Headache
© 2019 Eli Lilly and Company

ISSN 0017-8748
doi: 10.1111/head.13497
Published by Wiley Periodicals, Inc.

Research Submission

Psychometric Validation of the Role Function Restrictive Domain of the Migraine Specific Quality-of-Life Questionnaire Version 2.1 Electronic Patient-Reported Outcome in Patients With Episodic and Chronic Migraine

Rebecca M. Speck, PhD, MPH; Huda Shalhoub, PhD; Kathleen W. Wyrwich, PhD; Ren Yu, MA; David W. Ayer, PhD; Janet Ford, PhD, MPH; Elizabeth N. Bush, MHS; Richard B. Lipton, MD

Objectives.—To assess the measurement properties of the Migraine-Specific Quality of Life Questionnaire version 2.1 (MSQv2.1) electronic patient-reported outcome (ePRO) Role Function-Restrictive (RFR) domain to evaluate the functional impact of migraine in patients with episodic (EM) or chronic migraine (CM) enrolled in clinical trials.

Methods.—The 7-item MSQv2.1 ePRO RFR measures the functional impact of migraine on relationships with family and friends, leisure time, work or daily activities, productivity, concentration, tiredness, and energy. Measurement properties of the RFR were assessed using data from 2 EM (CGAG [n = 851] and CGAH [n = 909]) and 1 CM (CGAI [n = 1090]) Phase 3 galcanezumab clinical trials. Anchor- and distribution-based analyses were utilized to derive a responder threshold for clinical interpretation of change over time. The Migraine Disability Assessment (MIDAS), Patient Global Impression of Severity (PGI-S), Patient Global Impression of Improvement (PGI-I), and migraine headache days (MHD) served as anchors. Responsiveness and responder threshold analyses were completed from baseline to the average of months 4-6 for EM studies, and from baseline to month 3 for the CM study; timeframes selected were based on the primary endpoints in these studies.

Results.—Cronbach's alpha values for internal consistency reliability were 0.93, 0.92, and 0.92, for CGAG, CGAH, and CGAI, respectively. Test–retest reliability intra-class correlation coefficients were 0.82 and 0.84 for CGAG and CGAH, and 0.85 for CGAI in stable patients. Convergent validity was supported by moderate to strong correlations (≥0.30) between the RFR and both MIDAS and PGI-S. Known-groups validity was established between subgroups stratified by baseline PGI-S and MHD ($P < .05$; $\delta = 0.35$-1.96). For the EM studies, anchor variables suggested a change of ≥25 points (equivalent to 9 points/state changes on raw scale) in the RFR was an appropriate threshold to interpret a treatment benefit. For the CM study a change of ≥17.14 points (6 points/state changes on raw scale) was an appropriate threshold. In all 3 studies, significantly ($P < .01$) more galcanezumab patients achieved the responder definition thresholds, as compared to placebo (odds ratios of 1.98, 2.45, 2.27, 2.44, 1.64, and 1.66 for the 120 and 240 mg arms in the CGAG, CGAH, and CGAI trials, respectively).

Conclusion.—The MSQv2.1 ePRO RFR has sufficient reliability, validity, responsiveness, and appropriate interpretation standards for use in EM and CM clinical trials to assess the functional impact of migraine.

Key words: psychometric validation, Migraine-Specific Quality of Life Questionnaire, episodic migraine, chronic migraine

From Evidera, Bethesda, MD USA (R.M. Speck, H. Shalhoub, and R. Yu); Eli Lilly and Company, Indianapolis, IN USA (K.W. Wyrwich, D.W. Ayer, J. Ford, and E.N. Bush); Albert Einstein College of Medicine, Bronx, NY USA (R.B. Lipton).

Address all correspondence to R.M. Speck, Patient-Centered Research, Evidera, 7101 Wisconsin Avenue, Suite 1400, Bethesda, MD 20814, USA, email: rebecca.speck@evidera.com

Accepted for publication January 31, 2019.

Abbreviations: 15Q-MC-CGAG CGAG, 15Q-MC-CGAH CGAH, 15Q-MC-CGAI CGAI, ANCOVA analysis of covariance, AUC area under the curve, C concordance, EF Emotional Function, FDA Food and Drug Administration, HRQL health-related quality of life, ICC intra-class correlations coefficient, ICHD-3 IHS International Classification of Headache Disorders – 3rd edition, MHD migraine headache days, MID minimally important difference, MIDAS Migraine Disability Assessment, MSQ Migraine-Specific Quality of Life Questionnaire, PGI-I Patient Global Impression of Improvement, PGI-S Patient Global Impression of Severity, PRO patient-reported outcome, RFP Role Function-Preventive, RFR Role Function-Restrictive, ROC receiver operating characteristic, SEM standard error of mean, US United States, WHO World Health Organization, YI Youden Index, YLD years lived with disability

(*Headache* 2019;59:756-774)

INTRODUCTION

The International Headache Classification describes migraine as a recurrent, often life-long, disease characterized by migraine attacks with features such as: moderate or severe pain intensity, one-sided, pulsating in quality, aggravated by routine physical activity, a duration ranging from hours to 2-3 days, nausea or vomiting, and/or phonophobia and photophobia and varying attack frequencies.[1] Epidemiologic data indicate that the prevalence of migraine poses a notable burden to public health; in the United States the prevalence of migraine is nearly 12%, with 17% of the female population and 6% of the male population suffering from this disease.[2,3] Prevalence of migraine peaks during prime working ages of adulthood (ie, >18 years), with females between the ages of 18-44 years old experiencing the highest prevalence rate at 26.1%.[4] In addition, migraine is recognized as a leading cause of disability globally, expressed as years lived with disability (YLD)[5] and disability-adjusted life-years,[6] and is the number 1 cause of YLD of those under 50 years of age.[7] Health-related quality of life (HRQL) impairments due to migraine include work, family, social, and personal ramifications.[8]

As migraine disrupts physical and emotional well-being, assessment of HRQL is key to understanding patient burden and important changes over time.[9] Patient-reported outcome (PRO) instruments are often included as efficacy endpoints in clinical trials to provide insight into patients' perspectives of important HRQL impacts.[10] The Migraine-Specific Quality of Life Questionnaire (MSQ) is one such PRO instrument that was developed to measure the disease-specific HRQL impairments due to migraine.[11] The original MSQ v1.0, created in 1992 by Glaxo Wellcome Inc., was developed via a combination of a literature review and discussions with migraine specialists and patients, including one-on-one patient interviews. The 16-item instrument consisted of the Role Function-Restrictive (RFR), Role Function-Preventive (RFP), and Emotional Function (EF) domains.[12] Additional

Conflict of Interest: KWW, DWA, JF, and ENB are employed by Eli Lilly and Company, and are minority stockholders of Eli Lilly and Company. RBL is employed by Albert Einstein College of Medicine, and is the Edwin S. Lowe Professor of Neurology at the Albert Einstein College of Medicine in New York. He receives research support from the NIH: 2PO1 AG003949 (Program Director), 5U10 NS077308 (PI), RO1 NS082432 (Investigator), 1RF1 AG057531 (Site PI), RF1 AG054548 (Investigator), 1RO1 AG048642 (Investigator), R56 AG057548 (Investigator), K23 NS09610 (Mentor), K23AG049466 (Mentor), and 1K01AG054700 (Mentor). He also receives support from the Migraine Research Foundation and the National Headache Foundation. He serves on the editorial board of Neurology, senior advisor to Headache, and associate editor to Cephalalgia. He has reviewed for the NIA and NINDS, holds stock options in eNeura Therapeutics and Biohaven Holdings; serves as consultant, advisory board member, or has received honoraria from: American Academy of Neurology, Alder, Allergan, American Headache Society, Amgen, Autonomic Technologies, Avanir, Biohaven, Biovision, Boston Scientific, Dr. Reddy's, Electrocore, Eli Lilly, eNeura Therapeutics, GlaxoSmithKline, Merck, Pernix, Pfizer, Supernus, Teva, Trigemina, Vector, Vedanta. He receives royalties from Wolff's Headache 7th and 8th editions, Oxford Press University, 2009, Wiley and Informa. RMS, HS, and RY are employed by Evidera, which provides consulting and other research services to pharmaceutical, medical device, and related organizations. In their salaried positions, they work with a variety of companies and organizations, and are precluded from receiving payment or honoraria directly from these organizations for services rendered. Evidera received funding from Eli Lilly and Company to participate in the study and the development of this manuscript.
Funding: Eli Lilly and Company provided the funding for the study and for the manuscript.

developmental research lead to the creation of the MSQ v2.0 and later v2.1 based on sequential psychometric refinements.[13] Martin and colleagues evaluated the 14-item MSQ v2.1 for reliability (internal consistency and test–retest reproducibility at 4 weeks), construct validity, and the ability to detect between-group change, and reported the instrument demonstrated adequate performance on all evaluations.[9]

Subsequently, the pencil-and-paper MSQ v2.1 was migrated to tablet for electronic administration (MSQ v2.1 ePRO). Mean change in score over time for the MSQ v2.1 ePRO RFR domain was a key secondary endpoint in 3 recently completed Phase 3 placebo-controlled studies (I5Q-MC-CGAG (CGAG), I5Q-MC-CGAH (CGAH), and I5Q-MC-CGAI (CGAI)). The trials were designed to compare the efficacy and safety of galcanezumab (120 and 240 mg/month) to placebo in the prevention of episodic or chronic migraine.[14-16] In studies CGAG and CGAH, migraine headache days were significantly reduced by 4.3, 4.2, 2.3, and 4.7, 4.6, and 2.8 days for the 120 mg, 240 mg, and placebo arms, respectively.[14,15] In study CGAI, migraine headache days were significantly reduced by 4.8, 4.6, and 2.7 days for the 120 mg, 240 mg, and placebo arms, respectively.[16] The RFR domain, which includes 7 items that measure the functional impact of migraine on relationships with family and friends, leisure time, work or daily activities, productivity, concentration, tiredness, and energy, was selected as a key secondary endpoint to measure patient functioning for potential inclusion in labeling based on a literature review. A review of qualitative and quantitative literature, focused on the experience of migraine symptoms, demonstrated that, the RFR domain was most relevant to patients with episodic or chronic migraine. Specifically, patients with migraine have been described as working through their migraines;[8] therefore, the RFP domain was not selected and the EF domain alone was deemed insufficient in measuring the impact of migraine on both the physical and social aspects of daily living.

There is a body of research supporting the equivalence of paper-based and electronic administration of PRO instruments;[17,18] however, there are circumstances when demonstrating an instrument's psychometric properties administered in each mode is recommended.[19] Therefore, this study aimed to assess the reliability, validity, and ability to detect change of the MSQ v2.1 ePRO RFR domain using data from the 3 Phase 3 galcanezumab studies. In addition, the responder definition of the RFR domain for episodic and chronic migraine was determined for these 3 clinical trials.

METHODS

Patients.—The 3 clinical trials were conducted in accordance with the Declaration of Helsinki, local independent ethics committee/institutional review board requirements, and good clinical practice guidelines. Written informed consent was obtained from all individual participants included in the study.

CGAG and CGAH.—Studies CGAG (N = 858 patients enrolled) and CGAH (N = 915 patients enrolled) were Phase 3, multi-site randomized, double-blind, placebo-controlled studies to compare the efficacy and safety of 2 dosing regimens of galcanezumab with placebo for the prevention of migraine in adult patients with episodic migraine. Results for the primary endpoints are presented elsewhere.[14-16] A key secondary objective of the studies was to compare galcanezumab with placebo with respect to change in functioning, as measured by changes from baseline to the end of the 6-month double-blind treatment phase (average of Months 4, 5, and 6) in the MSQ v2.1 ePRO RFR domain.

Notable patient eligibility criteria included the following: 18-65 years of age at screening; have a diagnosis of migraine as defined by IHS International Classification of Headache Disorders – 3rd edition, beta (ICHD-3) (1.1 or 1.2),[1] with a history of migraine headaches of at least 1 year prior to Visit 1, and migraine onset prior to age 50; prior to Visit 1, have a history of 4-14 migraine headache days (MHD), and at least 2 migraine attacks per month on average within the past 3 months; from Visit 2 to Visit 3 (prospective baseline period), have a frequency of 4-14 MHDs, and at least 2 migraine attacks; from Visit 2 to Visit 3 (prospective baseline period), must achieve sufficient compliance with ePRO daily headache entries as demonstrated by completion of at least 80% of daily diary entries.

CGAI.—Study CGAI (N = 1113 patients enrolled) was a Phase 3, multi-site, randomized, double-blind, placebo-controlled study to compare the efficacy and safety of 2 doses of galcanezumab with placebo in the prevention of migraine in adult patients with chronic migraine. A key secondary objective of CGAI was also change in functioning, as measured by changes from baseline to end of the 3-month double-blind treatment phase in MSQ v2.1 ePRO RFR domain.

Eligibility criteria were similar to that of CGAG and CGAH, except for patients had to have a diagnosis of chronic migraine as defined by IHS ICHD-3 (1.3);[1] that is, a headache occurring on 15 or more days per month for more than 3 months that has the features of migraine on at least 8 days per month; prior to Visit 1, have a history of at least 1 headache-free day per month for the past 3 months; from Visit 2 to Visit 3 (prospective baseline period), have a frequency of at least 15 headache days, of which at least 8 must have the features of migraine; and from Visit 2 to Visit 3 (prospective baseline period), have at least 1 headache-free day.

Instruments.—*MSQ v2.1 ePRO.*—The 3 MSQ v2.1 ePRO domains are scored independently, with higher scores indicating better health status. Response options range from 1 (None of the time) to 6 (All of the time), and are reverse-recoded before the domain scores are calculated. The raw score for each domain is the sum of the final item value for all items in that domain. After the raw score for each MSQ v2.1 domain is computed, each domain score is linearly transformed to a 0-100 scale.[12] The MSQ v2.1 total score is calculated using the same process as for the domain scores. Because the MSQ v2.1 ePRO required a response to each item before advancing to the next item, no item-level missing data occurred.

The MSQ v2.1 ePRO was completed by patients at the clinical site visits on a tablet. For studies CGAG and CGAH, it was completed at baseline prior to treatment exposure, and at monthly clinic visits at Months 1-6 and Month 10 (last visit of the post-treatment period) or early termination visit.[14,15] For study CGAI, it was completed at baseline prior to treatment exposure, monthly for the 3-month double-blind treatment phase, monthly during the 9-month open-label treatment period, and every 2 months during the 4-month post-treatment period.[16] Copyright permission was obtained for use of the instrument.

Migraine Disability Assessment.—The Migraine Disability Assessment (MIDAS) was designed to quantify headache-related disability over a 3-month period. This instrument consists of 5 items that reflect the number of days reported as either missing completely or experiencing reduced productivity in school or work, household work, and/or social or leisure activities. A higher value on the MIDAS is indicative of more disability in units of lost days due to migraine.[20,21] This instrument is considered reliable and valid, and is correlated with clinical judgment regarding the need for medical care.[20,21] MIDAS has been used to stratify patients based on treatment needs and as an outcome in patient centered treatment trials.[22,23]

Patient Global Impression of Severity.—The Patient Global Impression of Severity (PGI-S) scale measures a patient's assessment of their level of illness for their current condition (ie, migraine).[24] This measure was obtained at baseline and monthly for each of the 3 studies. The PGI-S includes a 7-point scale: a score of 1 = "Normal, not at all ill," 2 = "Borderline ill," 3 = "Mildly ill," 4 = "Moderately," 5 = "Markedly," 6 = "Severely," and 7 = "Extremely ill."

Patient Global Impression of Improvement.—The Patient Global Impression of Improvement (PGI-I) scale is a patient-rated instrument that measures improvement of the patient's disease; the patient provides a self-evaluation of their disease (ie, migraine) since randomization to treatment.[24] It is on a 7-point scale; a score of 1 indicates the patient is "Very much better," 2 = "Much better," 3 = "A little better," 4 = "No change," 5 = "A little worse," 6 = "Much worse," and 7 indicates the patient is "Very much worse."

Statistical Analysis.—Distribution of Scores.—The distribution of scores for the MSQ v2.1 ePRO RFR domain was assessed using descriptive statistics at baseline, including mean (SD), median, range, and ceiling/floor effects.

Internal Consistency Reliability.—Internal consistency reliability was assessed using Cronbach's α

coefficient at baseline. Alpha coefficients of at least 0.70 in magnitude indicate acceptable internal consistency.[25,26]

Test–retest Reliability.—Test–retest reliability using paired *t*-tests and intra-class correlations coefficients (ICCs) was examined to assess the stability of the RFR domain score over time within a stable population. Among the placebo patients, stable patients were defined as patients who had either no change or a change of only 1 day in their number of migraine headache days per month during the last 2 time points of the treatment phase (months 5 and 6 for studies CGAG and CGAH, and months 2 and 3 for study CGAI) due to known placebo effects that occur early in migraine studies.[27] The hypothesis was that there would be no significant differences in the RFR domain scores when there is no change in disease status. Landis and Koch characterized ICC values of 0.01-0.20 as slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1.00 as almost perfect agreement.[28]

Convergent Validity.—The relationship between the RFR domain with the MIDAS, PGI-S, and MHD was examined using Spearman's rank correlation coefficients at baseline. A correlation coefficient >0.3 indicates moderate convergent validity, whereas a correlation coefficient >0.5 indicates strong convergent validity.[29] The RFR domain was hypothesized to have a moderate-to-strong relationship with the MIDAS and the PGI-S. The RFR domain and number of MHD were hypothesized to be moderately correlated.

Known-Groups Validity.—To evaluate known-groups validity, the RFR domain was analyzed by the PGI-S and number of MHD per month at baseline. The PGI-S has 7 response options; however, given the small sample size (<12) in 2 of the response option subgroups, subject PGI-S levels were collapsed into 5 PGI-S groups (combining "Normal" with "Borderline Ill," and "Severely Ill" with "Extremely Ill" groups). It was hypothesized that patients with the worse severity levels of illness as assessed by the PGI-S would have lower mean RFR domain scores.

Using the median number of MHD at baseline, 2 groups were created for episodic migraine (CGAG and CGAH trials) comparing patients with <8 and ≥8 MHD per month. For chronic migraine (CGAI trial) the groups were defined as 8 to 19 and ≥20 MHD per month at baseline. For MHD, it was hypothesized that the groups with fewer MHD per month would have higher mean RFR domain scores. All known-groups validity analyses used the analysis of covariance (ANCOVA) model, which included the RFR domain as the dependent variable and the known-group criterion variable as the independent variable, while age and sex were adjusted for as covariates. Bonferroni correction was used to adjust for multiple comparisons, and effect sizes examined using Cohen's *d*, a calculation of the difference of the means divided by the pooled standard deviation.

Responsiveness.—ANCOVA methods were also used to assess the mean change of the RFR domain scores between change groups based on relevant change levels in the MIDAS, PGI-I, PGI-S, and percent change in MHD. ANCOVA models were adjusted for baseline RFR domain scores.[30] Analyses were completed for the change during the blinded treatment period. For studies CGAG and CGAH, the RFR domain change score from baseline to months 4-6 average was evaluated. The 3-month average was used because the clinical course of episodic migraine is characterized by a high degree of variability of monthly MHD.[31] Therefore, a 3-month average would be more likely to capture a representative sample of MHD and thus provide more representative RFR domain score. For study CGAI, the RFR domain change score from baseline to month 3 was evaluated.

It was hypothesized that those patients in the MIDAS, PGI-S, PGI-I, and MHD groups that improved would have statistically significantly higher (that is, better) mean RFR domain change scores than those in the group that stayed the same or worsened. With adjustments for the baseline RFR score and sex, comparisons between least square (LS) means incorporated Bonferroni test corrections for multiple comparisons. In addition, effect size calculations using Cohen's *d* were conducted to aid in the interpretation

of the results comparing the LS means scores for these known-groups.

Responder Definition Thresholds.—A responder definition threshold, defined as the individual patient PRO score change over a predetermined time period that should be interpreted as a treatment benefit,[10] was estimated for the MSQ v2.1 ePRO RFR domain using both anchor- and distribution-based approaches. The anchor-based analyses first included dichotomizing the 4 anchor variables used to identify patients achieving a treatment benefit (responders) over the course of each trial: MIDAS (≥1 category improvement vs no improvement), PGI-S (≥1 unit improvement vs no improvement), PGI-I (≥ 1 unit improvement vs no improvement), and percent change in MHD (≥50% reduction in number of MHD vs <50% reduction). Secondly, a series of analyses to estimate the threshold value of RFR domain changes (baseline to the average of months 4-6 for studies CGAG and CGAH, and baseline to month 3 for study CGAI) that provided the greatest discriminative ability between patients in the responder/non-responder groups were completed. The Concordance (C) statistic, equivalent to the area under the receiver operating characteristic (ROC) curve, was used as a measure of goodness of each logistic model's fit, with $C = 0.50$ indicating that the model is no better than predicting an outcome than random chance while $C \geq 0.70$ indicates a good model and $C \geq 0.80$ indicates a strong model. The ROC curves were created by plotting sensitivity vs 1-specificity at all RFR domain change threshold possibilities. The best possible prediction would yield a point in the upper left corner of the ROC space (that is, coordinate [0,1]), representing 100% sensitivity and 100% specificity.[32] The Youden Index (YI) measures the effectiveness of a predictive marker and enables selection of an optimal threshold value/cut-point for that marker. The YI was calculated at each level of improvement in the RFR domain. [YI(c) = sensitivity(c) + (specificity(c) −1)], and the maximum YI value will represent the optimal cut-point in the ROC analysis. In addition, sensitivity and specificity for several change scores above and below the optimal cutoff point were provided to show the range of alternative RFR domain change score thresholds.

The responder definition threshold was also assessed using 2 distribution-based strategies. The first approach consisted of calculating a 0.5 SD unit at baseline for the RFR domain. It has been suggested that one-half SD of a measure represents a clinically meaningful change.[33] The second strategy estimated the responder definition as 1 standard error of measurement (SEM = baseline SD × [√1 − reliability]).[34] The SEM is expressed in the original metric of the instrument, and change beyond 1 SEM has demonstrated correspondence with an important change in several other chronic diseases.[34-37] In the SEM calculation, reliability of the RFR domain was assessed by the ICC, as noted above.[38] The distribution-based parameters were used to categorize changes over time, and were then compared with anchor-based estimates to provide confidence in the responder definition, with the distribution-based results considered as supportive to the anchor-based results.

Triangulation of the RD analyses results, that is, engaging in an iterative process examining results of the described anchor- and distribution-based methods to determine a single estimate of a responder definition threshold, was employed.[39-41] Consistent with the FDA PRO Guidance,[10] priority was placed on the anchor-based methods, while distribution-based methods are considered to play a supportive role. Therefore, triangulation involved examination of the range of anchor-based estimates, with stronger consideration given to the anchor-based estimates where the baseline correlations and change score correlations of the MSQ v2.1 ePRO RFR domain were highest among the 4 anchors, and examination of the actual change scores in the range identified in the ROC curves and ROC tables through cumulative distribution function (CDF) and probability density function (PDF) plots. The triangulation process also incorporated an awareness that the MSQ RFR domain's possible score change over time (increase or decrease) is in increments of ~2.857 points (also known as a state change[42]) on the 0-100 point RFR domain scale (100 possible points, divided by 7 items, divided by 5 levels of change per item). Therefore, the thresholds for baseline to month 3 change in study CGAI are increments of ~2.857. However, for studies CGAG and CGAH, which had an endpoint of RFR domain

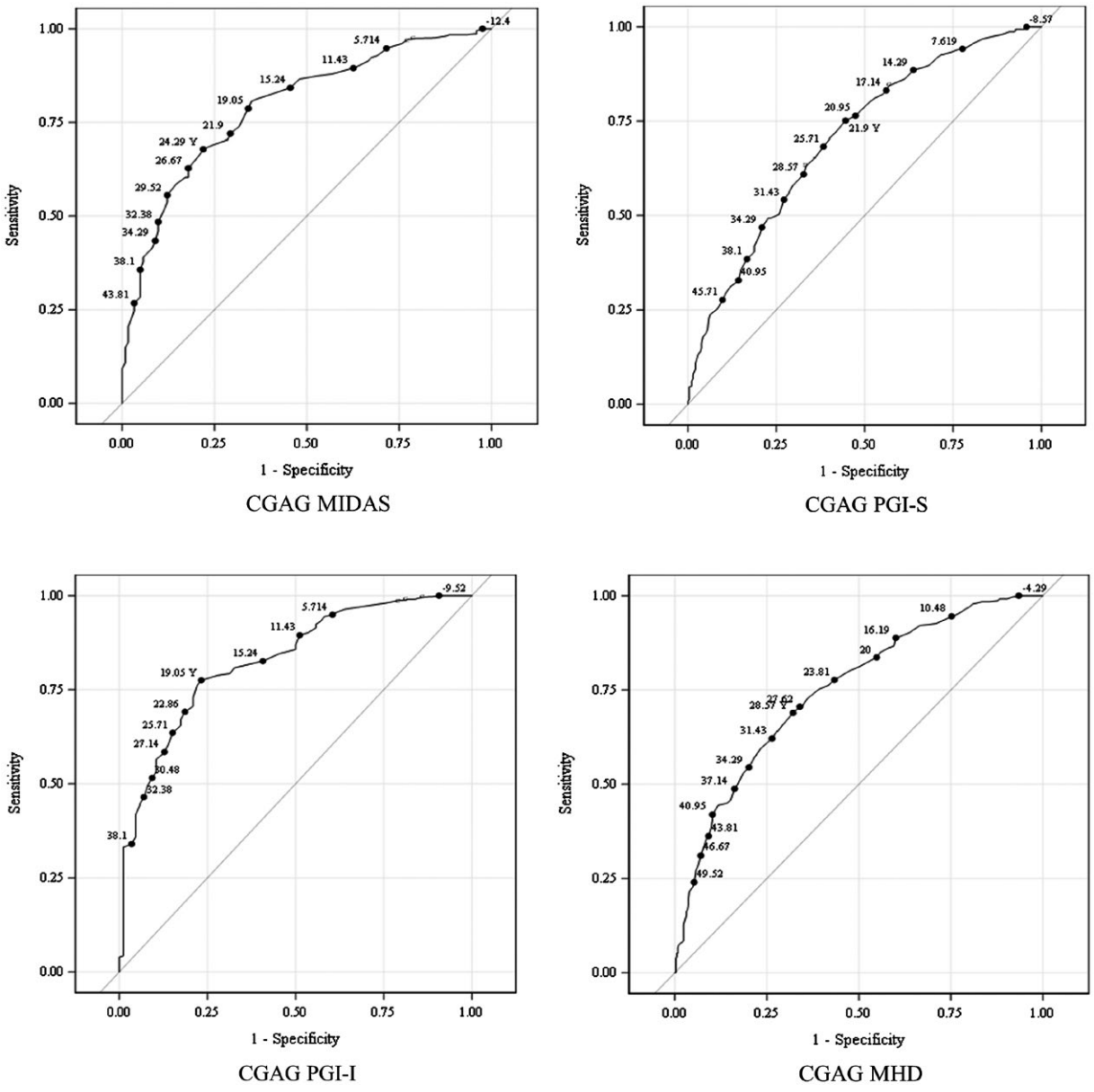

Fig. 1.—ROC curves for anchor improvement: change in MSQ v2.1 ePRO role function restrictive domain from baseline to average months 4-6 (CGAG).

change scores using months 4-6 average, the ~2.857 increment does not apply. As a result, integer levels of change (~2.857/3 months) guided by the ROC curves (Figs. 1–3) were generated to select the plausible and optimal change score cutoff points. All statistical analyses were completed using SAS version 9.4.

RESULTS

A summary of patients' baseline demographic and disease characteristics for each of the 3 Phase 3 studies are shown in Table 1 based on the PRO population (defined as all patients who had a baseline and post-baseline MSQ v2.1 ePRO total score at month 4, 5, or 6 for episodic studies CGAG and CGAH, and at month 3 for chronic study CGAI). No significant floor or ceiling effects for the RFR domain scores were observed. Descriptive statistics for the 2 episodic migraine trials were very similar for the RFR domain score, and the mean item level scores were nearly identical (Table 1). The mean RFR domain score was 51.5 for

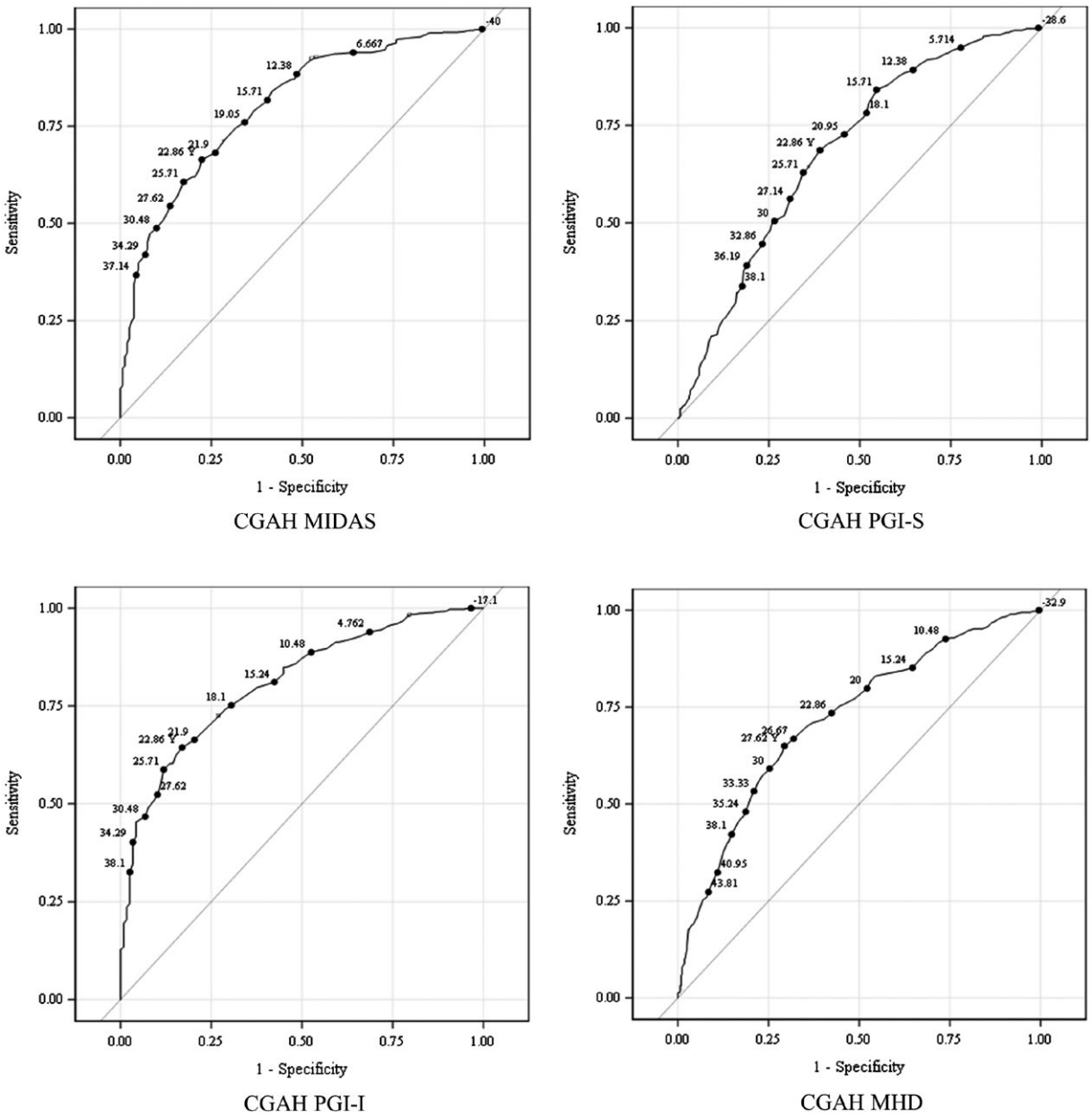

Fig. 2.—ROC curves for anchor improvement: change in MSQ v2.1 ePRO role function restrictive domain from baseline to average months 4-6 (CGAH).

CGAG (SD = 16.0; range: 0-94.3) and 51.7 for CGAH (SD = 15.6; range: 0-100), and the median RFR domain scores were the same for both studies (51.4). The RFR domain and item scores were notably lower in the chronic migraine trial than in the episodic trials, with a mean of 38.7 (SD = 17.2; range: 0-94.3) and median of 37.1 (Table 1).

Reliability.—*Internal Consistency Reliability.*—Cronbach's α estimates of internal consistency for the RFR domain for the 3 studies exceeded the recommended 0.70 threshold,[26] with values of 0.93, 0.92, and 0.92 for studies CGAG, CGAH, and CGAI, respectively.

Test–Retest Reliability.—Test–retest reliability was assessed among 105, 87, and 107 stable placebo patients in studies CGAG, CGAH, and CGAI, respectively. The ICC values were 0.82, 0.84, and 0.87 for the RFR domain score in studies CGAG, CGAH, and CGAI, respectively, demonstrating near perfect agreement.[28]

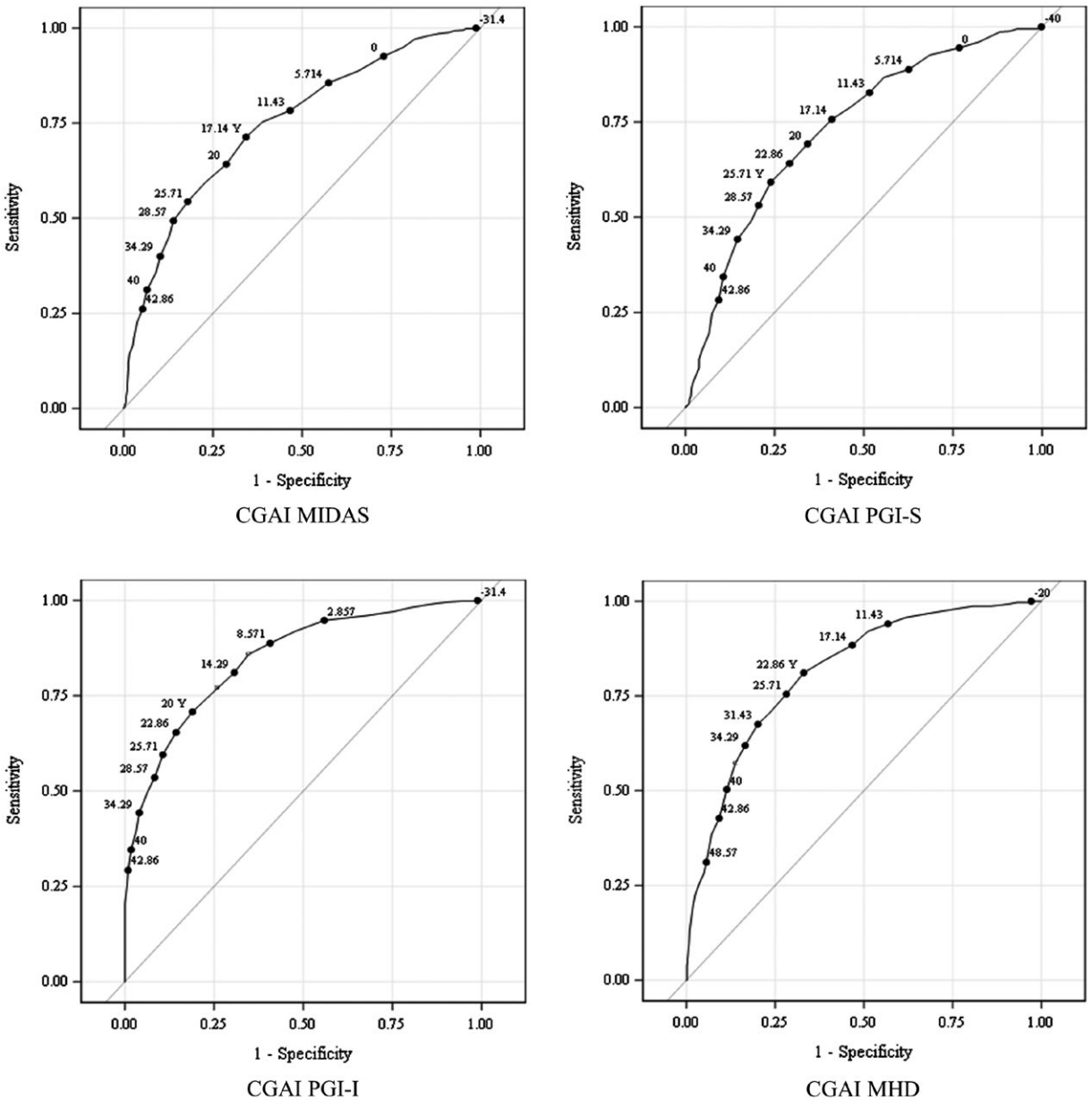

Fig. 3.—ROC curves for anchor improvement: change in MSQ v2.1 ePRO role function restrictive domain from baseline to month 3 (CGAI).

Validity.—*Convergent Validity*.—Results supporting convergent validity of the RFR domain in relation to MIDAS, PGI-S, and number of MHD are presented in Table 2. At baseline, moderate-to-large associations between the RFR domain scores and MIDAS (−0.57, −0.51, and −0.53) and PGI-S (−0.54, −0.46, and −0.52) scores were observed in CGAG, CGAH, and CGAI, respectively. These associations provide evidence for a strong relationship between the episodic and chronic migraine headache patient experiences of role function impairment, headache-related disability, and illness severity. Small correlations were observed between the RFR domain scores and number of MHD (r = −0.27 in study CGAG, −0.22 in study CGAH, and −0.27 in study CGAI).

Known-Groups Validity.—Patients with worse severity levels of illness as assessed by the PGI-S

Table 1.—Patient Demographic and Disease Characteristics, and MSQ v.2 1 ePRO Role Function Restrictive Domain Scores at Baseline: PRO Population

Characteristics	CGAG[†] PRO Population (n = 851)	CGAH[†] PRO Population (n = 909)	CGAI[‡] PRO Population (n = 1090)
Age (years)			
Mean (SD) [min, max]	40.6 (11.6) [17.0, 65.0]	41.8 (11.1) [18.0, 65.0]	41.0 (12.1) [17.0, 65.0]
Gender, n (%)			
Female	712 (83.7%)	776 (85.4%)	929 (85.2%)
Race, n (%)			
White	683 (80.3%)	638 (70.2%)	863 (79.2%)
Black or African American	94 (11.0%)	63 (6.9%)	69 (6.3%)
Asian	24 (2.8%)	102 (11.2%)	53 (4.9%)
American Indian or Alaska Native	3 (0.4%)	41 (4.5%)	6 (0.6%)
Native Hawaiian or Other Pacific Islander	3 (0.4%)	2 (0.2%)	1 (0.1%)
Multiple	44 (5.2%)	63 (6.9%)	97 (8.9%)
Missing	0 (0%)	0 (0%)	1 (0.1%)
Years since migraine diagnosis			
Mean (SD) [min, max]	20.0 (12.4) [0.2, 58.1]	20.6 (12.4) [0.1, 57.7]	21.2 (12.8) [0.1, 56.4]
Number of migraine headache days			
Mean (SD) [min, max]	9.1 (3.0) [4.0, 16.7]	9.1 (2.9) [4.0, 18.0]	19.4 (4.5) [8.0, 29.0]
Role Function Restrictive Domain			
Mean (SD)	51.5 (16.0)	51.7 (15.6)	38.7 (17.2)
Median (Q1-Q3)	51.4 (42.9-60.0)	51.4 (40.0-60.0)	37.1 (25.7-51.4)
Range (min-max)	(0.0, 94.3)	(0.0, 100.0)	(0.0, 94.3)
Missing n (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Floor n (%)[§]	4 (0.5%)	2 (0.2%)	11 (1.0%)
Ceiling n (%)[§]	0 (0.0%)	2 (0.2%)	0 (0.0%)

[†]Episodic migraine.
[‡]Chronic migraine.
[§]For all values, the floor effect is assessed based on minimum value, and the ceiling effect is assessed based on the maximum value possible for the range.
Max = maximum; min = minimum; n = number of patients within each specific category; PRO = patient-reported outcome; SD = standard deviation.

Table 2.—Convergent Validity: Spearman Correlation Between MSQ v2.1 ePRO Role Function Restrictive Domain, and MIDAS, PGI-S, and Number of Migraine Headache Days at Baseline

	Correlations		
Role Function Restrictive Domain/Study	MIDAS	PGI-S	Number of Migraine Headache Days[§]
CGAG[†] (N = 851)	−0.57	−0.54	−0.27
CGAH[†] (N = 909)	−0.51	−0.46	−0.22
CGAI[‡] (N = 1090)	−0.53	−0.52	−0.27

[†]Episodic migraine.
[‡]Chronic migraine.
[§]The number of migraine headache days ranged from 4 to 18 in the episodic population, and 8 to 29 in the chronic population.
Spearman's correlation coefficients reported: *P* value ≤ .0001 for all correlations.
ePRO = electronic patient-reported outcome; MIDAS = Migraine Disability Assessment; MSQ v2.1 = Migraine-Specific Quality of Life Questionnaire version 2.1; N = number of patients in the analysis population; PGI-S = Patient Global Impression of Severity.

Table 3.—Known-Groups Validity: MSQ v2.1 ePRO Role Function Restrictive Domain Scores at Baseline (PGI-S, 5-level)

	PGI-S at Baseline (5 Groups) LS Mean of MSQ v2.1 ePRO RFR at Baseline (SE), N								
	Normal & Borderline Ill	Mildly Ill	Moderately Ill	Markedly Ill	Severely Ill & Extremely Ill	Differences Between Means LS Mean (95% CI)	Effect Size	Overall F Value (*P* Value)[§]	Pairwise Comparison for Each Group[¶]
Groups	1	2	3	4	5				
CGAG[†]	67.8 (2.07), 44	63.4 (1.39), 104	56.9 (0.83), 346	47.8 (0.96), 237	36.8 (1.30), 120	1 v 2: 4.4 (−2.4, 11.2); 1 v 3: 10.9 (4.9, 17.0); 1 v 4: 20.1 (13.8, 26.3); 1 v 5: 31.1 (24.4, 37.7); 2 v 3: 6.5 (2.3, 10.7); 2 v 4: 15.6 (11.2, 20.1); 2 v 5: 26.6 (21.6, 31.7); 3 v 4: 9.1 (5.9, 12.3); 3 v 5: 20.1 (16.1, 24.1); 4 v 5: 11.0 (6.8, 15.2)	1v 2 = 0.36 1v 3 = 0.81 1 v 4 = 1.54 1 v 5 = 1.96 2 v 3 = 0.48 2 v 4 = 1.24 2 v 5 = 1.86 3 v 4 = 0.71 3 v 5 = 1.45 4 v 5 = 0.80	61.91 (<.0001)	*P* < .0001 (1 vs 3; 1 vs 4; 1 vs 5; 2 vs 4; 2 vs 5; 3 vs 4; 3 vs 5; 4 vs 5) *P* = .0002 (2 vs 3) *P* = .6649 (1 vs 2)
CGAH[†]	63.9 (1.66), 75	61.1 (1.35), 118	56.3 (0.91), 320	48.2 (0.92), 295	38.4 (1.44), 101	1 v 2: 2.8 (−3.0, 8.5); 1 v 3: 7.5 (2.5, 12.5); 1 v 4: 15.6 (10.6, 20.6); 1 v 5: 25.5 (19.6, 31.4); 2 v 3: 4.7 (0.5, 8.9); 2 v 4: 12.8 (8.6, 17.1); 2 v 5: 22.7 (17.4, 28.0); 3 v 4: 8.1 (5.0, 11.2); 3 v 5: 18.0 (13.5, 22.4); 4 v 5: 9.9 (5.4, 14.3)	1 v 2 = 0.19 1 v 3 = 0.53 1 v 4 = 1.10 1 v 5 = 1.49 2 v 3 = 0.35 2 v 4 = 0.98 2 v 5 = 1.56 3 v 4 = 0.63 3 v 5 = 1.31 4 v 5 = 0.73	43.49 (<.0001)	*P* < .0001 (1 vs 4; 1 vs 5; 2 vs 4; 2 vs 5; 3 vs 4; 3 vs 5; 4 vs 5) *P* = .0003 (1 vs 3) *P* = .0161 (2 vs 3) *P* = 1.0000 (1 vs 2)
CGAI[‡]	54.8 (2.00), 59	55.4 (2.25), 44	49.1 (1.03), 249	38.6 (0.86), 403	29.1 (0.93), 335	1 v 2: −0.5 (−8.9, 7.8); 1 v 3: 5.8 (−0.3, 11.8); 1 v 4: 16.2 (10.4, 22.0); 1 v 5: 25.7 (19.8, 31.6); 2 v 3: 6.3 (−0.5, 13.1); 2 v 4: 16.8 (10.2, 23.4); 2 v 5: 26.2 (19.6, 32.9); 3 v 4: 10.5 (7.1, 13.8); 3 v 5: 20.0 (16.5, 23.4); 4 v 5: 9.5 (6.4, 12.6)	1 v 2 = 0.04 1 v 3 = 0.36 1 v 4 = 1.09 1 v 5 = 1.64 2 v 3 = 0.43 2 v 4 = 1.21 2 v 5 = 1.81 3 v 4 = 0.71 3 v 5 = 1.32 4 v 5 = 0.66	65.76 (<.0001)	*P* < .0001 (1 vs 4; 1 vs 5; 2 vs 4; 2 vs 5; 3 vs 4; 3 vs 5; 4 vs 5) *P* = .0732 (1 vs 3) *P* = .0946 (2 vs 3) *P* = 1.0000 (1 vs 2)

Table 3.—(Continued)

	Migraine Headache Days at Baseline LS Mean of MSQ v2.1 ePRO RFR at Baseline (SE), N		Differences between means LS Mean (95% CI)	Effect Size	Overall F value (*P* value)[†]
	<8	≥8			
CGAG[†]	57.2 (1.0), 291	50.2 (0.8), 560	7.0 (4.7, 9.2)	−0.47	15.99 (<.001)
CGAH[†]	56.1 (1.0), 303	50.9 (0.8), 606	5.2 (3.1, 7.3)	−0.35	9.66 (<.001)
	<20	≥20			Overall F value (*P* value)[†]
CGAI[‡]	43.1 (0.9), 600	36.1 (0.9), 490	7.0 (5.0, 9.0)	−0.42	18.00 (<.001)

[†]Episodic migraine.
[‡]Chronic migraine.
[§]An analysis of covariance (ANCOVA) model adjusting for age and gender was used.
[¶]Pairwise comparisons between LS means were performed using Bonferroni test adjusting for multiple comparisons.
ePRO = electronic patient-reported outcome; LS Mean = least squares mean; MSQ v2.1 = Migraine-Specific Quality of Life Questionnaire version 2.1; N = number of patients in the analysis population; PGI-S = Patient Global Impression of Severity; RFR = Role Function-Restrictive; SE = standard error.

had lower RFR domain scores (Table 3). Statistically significant differences in mean RFR domain scores at baseline were observed between patients in nearly all of the PGI-S levels (δ = 0.35-1.96); however, in none of the 3 studies were there statistically significant differences between mean scores of patients with PGI-S levels of "Normal" or "Borderline," when compared to the "Mildly Ill" group. The groups with fewer MHD had higher mean RFR domain scores for all 3 studies ($P < .001$; $\delta = -0.35$ to -0.47; Table 3).

Responsiveness.—Patients who had improvements in MIDAS, PGI-S, PGI-I, and/or experienced at least 50% fewer MHD also demonstrated improvements in their RFR domain mean change score (Table 4). Mean differences in change scores over the pre-specified timespans on the RFR domain were statistically significantly different between patients with vs without categorical improvement in MIDAS, PGI-S, PGI-I, and percent change in number of MHD (all $P < .001$; δ = 0.67-1.40; Table 4). These results provide strong evidence to support the responsiveness (ability to detect change) of the RFR domain in patients with episodic and chronic migraine.

Responder Definition Threshold.—The ROC curves for each of the pre-specified anchors demonstrated good-to-strong model fit, with *C* statistics ranging between 0.69 and 0.85. The optimal ROC cutoff point was selected as the MSQ v2.1 ePRO RFR domain change score where the YI was highest across cutoff values for the pre-specified anchors.

In studies CGAG and CGAH, the YI was optimal at a RFR domain change score of ≥25 using months 4-6 average. The 2.857 RFR domain change score increment nearest 25 is ~25.71 points, which translates to a 9-point change on the MSQ v2.1 raw scale (range 7-42). These data suggest that a RFR domain change score of ≥25 from baseline to months 4-6 average is a clinically meaningful change threshold. In study CGAI, the YI was optimal at a change score of 17.14 at month 3, which translates to a 6-point change on the raw scale. These data suggest that a RFR domain change score of 17.14 from baseline to month 3 is a clinically meaningful change threshold. In all 3 studies, a statistically significantly greater proportion of galcanezumab patients, as compared to placebo, achieved the ascertained responder definition

Table 4.—Responsiveness: ANCOVA With MSQ v2.1 ePRO Role Function Restrictive Domain Change Among MIDAS, PGI-S, PGI-I, and Percent Change in Number of Migraine Headache Days Improvement Groups

	MIDAS			
	No Category Improvement LS Mean (SE), N	≥ 1 Category Improvement LS Mean (SE), N	*P* Value	Effect Size
CGAG[†]	14.2 (1.6), 123	32.6 (0.7), 572	<.0001	1.06
CGAH[†]	11.5 (1.3), 161	30.7 (0.7), 613	<.0001	1.15
CGAI[‡]	8.9 (1.1), 324	28.0 (0.8), 673	<.0001	0.92
	PGI-S			
	No Unit Improvement LS Mean (SE), N	≥ 1 Unit Improvement LS Mean (SE), N	*P* Value	Effect Size
CGAG[†]	20.6 (1.1), 287	34.2 (0.8), 463	<.0001	0.76
CGAH[†]	19.2 (1.0), 328	31.1 (0.8), 491	<.0001	0.67
CGAI[‡]	12.6 (1.0), 473	30.0 (0.9), 527	<.0001	0.84
	PGI-I			
	No Unit Improvement LS Mean (SE), N	≥ 1 Unit Improvement LS Mean (SE), N	*P* Value	Effect Size
CGAG[†]	10.5 (1.9), 86	32.0 (0.7), 609	<.0001	1.24
CGAH[†]	9.6 (1.6), 118	29.7 (0.7), 657	<.0001	1.19
CGAI[‡]	4.7 (1.0), 349	31.0 (0.7), 650	<.0001	1.40
	Percent Change in Number of Migraine Headache Days			
	<50% Improvement LS Mean (SE), N	≥ 50% Improvement LS Mean (SE), N	*P* Value	Effect Size
CGAG[†]	21.1 (0.9), 383	37.2 (0.9), 367	<.0001	0.94
CGAH[†]	19.9 (0.8), 439	33.9 (0.9), 377	<.0001	0.81
CGAI[‡]	14.4 (0.7), 696	39.0 (1.1), 302	<.0001	1.26

[†]Episodic migraine.
[‡]Chronic migraine.
[§]Comparisons between LS means were performed using Bonferroni test adjusting for multiple comparisons.
ANCOVA = analysis of covariance; ePRO = electronic patient-reported outcome; LS Mean = least squares mean; MIDAS = Migraine Disability Assessment; MSQ v2.1 = Migraine-Specific Quality of Life Questionnaire version 2.1; N = number of patients in the analysis population; PGI-I = Patient Global Impression of Improvement; PGI-S = Patient Global Impression of Severity; SE = standard error.

threshold (Table 5). Through CDF and PDF plots that depict the RFR domain change score by PGI-S change (improved, stayed the same, and worsened), it can be seen that a proportion of patients in all 3 groups experience score improvements and score worsening. In fact, over 50% of patients in all 3 trials experienced score improvements, despite their self-reported level of illness. However, in all trials it can be seen that a greater proportion of patients that self-reported improvement in their level of illness also experienced greater improvement in role functioning reflected through their higher RFR domain change scores (Fig. 4 series).

Distribution-based methods supported the ≥25 (~9 raw scale points) and 17.14 (6 raw scale points) identified via anchor-based methods for studies CGAG and CGAH, and study CGAI, respectively. The SD at baseline for the RFR domain ranged from 15.62 to 17.23, and 0.5 SD ranged from 7.81 to 8.62. Thus, the anchor-based ≥25 and 17.14-point thresholds for change in RFR domain are at least 2 times larger than the 0.5 SD estimate. SEM estimates ranged from 6.16

Table 5.—Proportion of Patients by Treatment Group Meeting the MSQ v2.1 ePRO Role Function Restrictive Domain Responder Definition Thresholds

		Placebo	Galcanezumab 120 mg	Galcanezumab 240 mg
CGAG[†]	Number of patients	377	189	184
	Months 4-6 Average	47.2%	63.5%[§]	69.6%[§]
	OR (95% CI)		1.98 (1.32, 2.97)[§]	2.45 (1.60, 3.73)[§]
CGAH[†]	Number of patients	396	213	210
	Months 4-6 Average	43.4%	58.2%[§]	60.0%[§]
	OR (95% CI)		2.27 (1.55, 3.33)[§]	2.44 (1.66,3.59)[§]
CGAI[‡]	Number of patients	494	252	253
	Months 3	54.1%	64.3%[¶]	64.8%[¶]
	OR (95% CI)		1.64 (1.18, 2,27)[¶]	1.66 (1.20, 2.30)[¶]

[†]Episodic migraine.
[‡]Chronic migraine.
[§]$P < .001$.
[¶]$P < .01$.
CI = confidence interval; OR = odds ratio.

to 6.84, making the ≥25 and 17.14-point item threshold 2-4 times larger than SEM, and therefore, conservative estimates that are beyond the measurement error associated with the measure.

DISCUSSION

The stability of the psychometric properties, including reliability, validity, and responsiveness, of the MSQ v2.1 ePRO RFR domain demonstrated in prior psychometric analyses of the MSQ v2.1 were supported using data from 3 Phase 3 clinical trials of galcanezumab. Specifically, test–retest reliability analyses revealed near-perfect levels of agreement in RFR domain values among patients considered stable across 2 assessment periods. Results were supportive of convergent validity with MIDAS and PGI-S, consistent with hypothesized relationships and prior analyses.[9,12,13,43,44] The lower correlation at baseline between MHDs and the RFR domain may be related to the clinical trials inclusion exclusion criteria on the number of MHDs with no restrictions for RFR scores, and may also suggest that HRQL measures capture aspects of the migraine experience not captured by MHD alone. These features, in addition to migraine attack frequency, could include attack duration and severity, profiles of associated symptoms, treatment effects, and comorbidities, among other factors. Known-groups validity was similarly supported as mean values of the RFR domain were significantly different based on known-groups using the PGI-S and MHD. Finally, the RFR domain was responsive to change using all 4 different anchors to define responders.

Ascertainment of the responder definition threshold of the MSQ v2.1 ePRO RFR domain using episodic and chronic migraine clinical trial data aids in the interpretability of RFR domain results. Triangulation of the methods executed herein resulted in responder definition thresholds of a ≥ 25 point change (~9 raw scale points) for episodic migraine and 17.14 point change (6 raw scale points) for chronic migraine. The difference in the responder thresholds between the patients with episodic and chronic migraine is equivalent to 3 additional state changes (3 × 2.857) in the RFR domain change score. The higher threshold for the episodic trials is reflected in the longer trial timespan for improvement (6 months for studies CGAG and CGAH vs 3 months for study CGAI), and the generally greater migraine burden of patients with chronic migraine (reflected in the RFR domain baseline summary scores reported above).

Two prior analyses have been conducted to determine meaningful change in the RFR domain.[45,46] Using data from a randomized, placebo-controlled, double-blind study of topiramate for the treatment

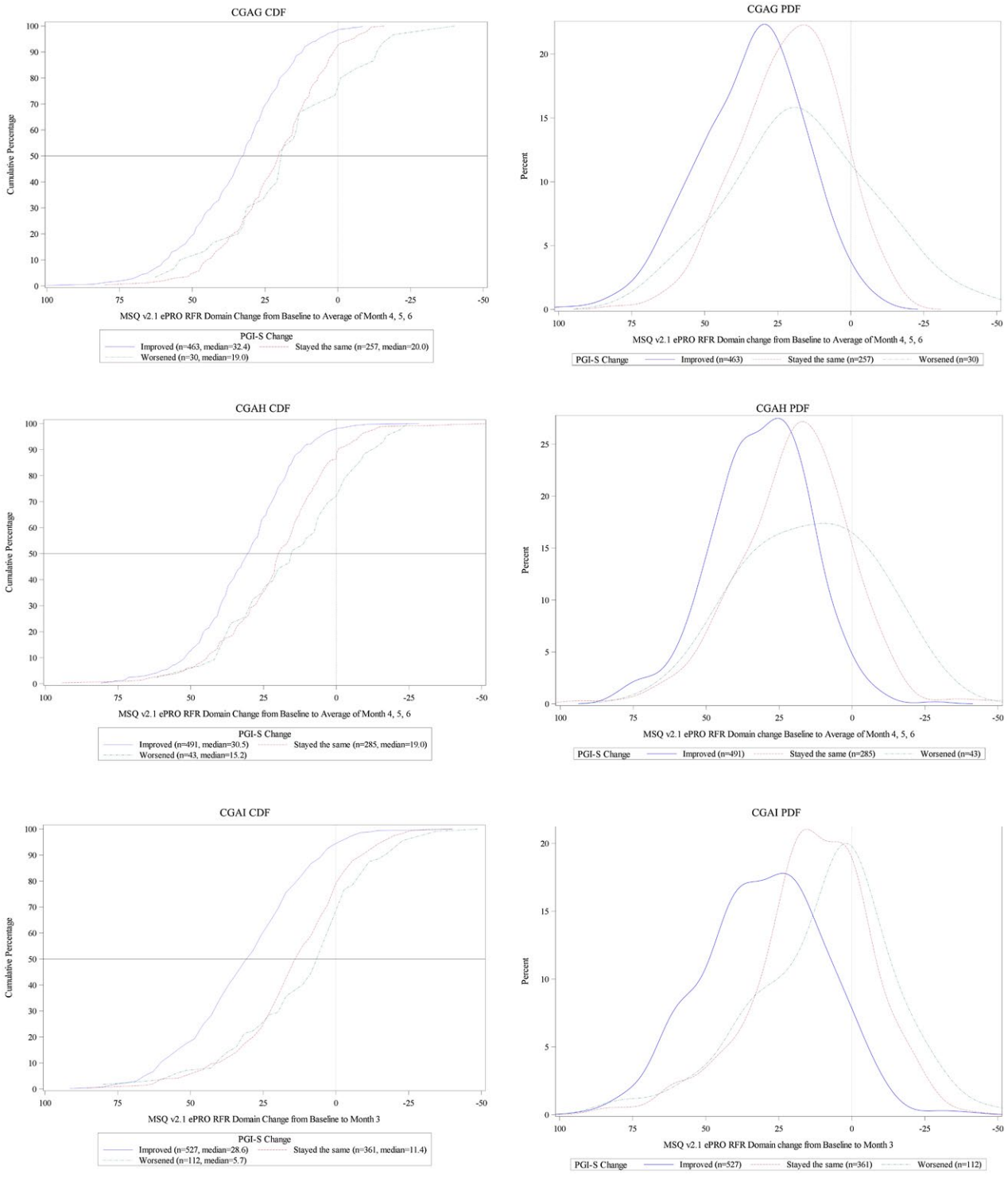

Fig. 4.—Cumulative distribution function and probability density function plots: change in MSQ v2.1 ePRO role function restrictive domain. [Color figure can be viewed at wileyonlinelibrary.com]

of chronic migraine, Dodick and colleagues used a 1-unit improvement in Subject's Global Impression of Change as an anchor to estimate the minimal important difference (MID) in the RFR domain via regression analysis.[46] They estimated the MID for the RFR domain to be 10.9 (95% CI = 9.4-12.4).

Anchor and distribution-based analyses completed by Cole and colleagues among patients with episodic migraine resulted in estimates of 5, and 8.5 (0.5 SD) and 4.8 (1 SEM), respectively.[45] The responder definition thresholds that emerged from CGAG, CGAH, and CGAI data were based on the use of multiple relevant anchors and ROC analyses, and yielded greater level of change as compared to earlier publications. As seen in the CDFs and PDFs, in addition to those that self-reported an improvement in their illness, a proportion of patients that self-reported no change or worsening in their illness also experienced RFR domain score improvements, though the changes were not sizable and the medians below the responder definition thresholds. Even with the higher threshold, the proportion of responders was statically significantly greater for patients treated with galcanezumab when compared to placebo across all 3 studies, with approximately 1.5-2.5 greater odds of being a responder.

The analytic methods reported herein are consistent with the FDA guidance on the use and interpretation of PRO scores in medical product development.[10] However, there are limitations in the generalizability of the findings to be considered. CGAH and CGAI enrolled patients in the United States (including Puerto Rico) and in 12 other countries. CGAG was limited to the United States (including Puerto Rico) and Canada. Interpretability in other patient populations is unknown. Additionally, the endpoint analysis time points differed for the episodic and chronic migraine trials. The average of months 4-6 as the endpoint for the episodic trials was chosen to increase the accuracy of measurement as migraine headaches per month may differ from month to month in patients with an episodic migraine diagnosis;[31] an adequate time on treatment to observe changes in patient functioning was needed. Overall, using the average of the final 3 months of data were believed to allow for more stable endpoint estimate, and was consistent with the methods used for the primary analyses on the change in monthly MHDs. However, a limitation was that the baseline measure of MHD was captured over a 1-month prospective period and does not capture an average over multiple months to account for potential variability at baseline. Using other or further longitudinal time periods to assess improvement in RFR should be considered in future trials.

CONCLUSION

The results presented herein are the first known validation study of the MSQ v2.1 ePRO and substantiate the RFR domain as a reliable and validated responsive measure of the impact of episodic and chronic migraine on function and performance of activities from the patient's perspective.[9,12,13,43,44] The instrument is psychometrically robust and appropriate for inclusion in future episodic and chronic migraine studies designed to measure the impacts of migraine on role functioning.[47] The findings of this analysis indicate that improvement on the MSQ v2.1 ePRO RFR domain of ≥25 over 4-6 months and 17.14 over 3 months in patients with episodic and chronic migraine, respectively, are reasonable and practical thresholds for identifying patients who have experienced a clinically significant change in the functional impact of migraine. When the thresholds were applied to the 3 galcanezumab trials, a greater proportion of responders was observed among patients treated with galcanezumab as compared to placebo. PROs can provide unique information on the effects of treatment; therefore, establishing a responder definition threshold for the MSQ v2.1 ePRO RFR domain provides the basis for its use in future clinical trials of episodic and chronic migraine.

STATEMENT OF AUTHORSHIP

Category 1

(a) Conception and Design

Rebecca M. Speck, Huda Shalhoub, Kathleen W. Wyrwich, David W. Ayer, Janet Ford, Elizabeth N. Bush, Richard B. Lipton

(b) Acquisition of Data

Kathleen W. Wyrwich, David W. Ayer, Janet Ford, Elizabeth N. Bush

(c) Analysis and Interpretation of Data

Rebecca M. Speck, Huda Shalhoub, Kathleen W. Wyrwich, Ren Yu, David W. Ayer, Janet Ford, Elizabeth N. Bush, Richard B. Lipton

Category 2

(a) Drafting the Manuscript

Rebecca M. Speck

(b) Revising It for Intellectual Content

Huda Shalhoub, Kathleen W. Wyrwich, Ren Yu, David W. Ayer, Janet Ford, Elizabeth N. Bush, Richard B. Lipton

Category 3

(a) Final Approval of the Completed Manuscript

Rebecca M. Speck, Huda Shalhoub, Kathleen W. Wyrwich, Ren Yu, David W. Ayer, Janet Ford, Elizabeth N. Bush, Richard B. Lipton

REFERENCES

1. IHS. The International Classification of Headache Disorders, 3rd edition (beta version). *Cephalalgia*. 2013;33:629.
2. Buse DC, Manack AN, Fanning KM, et al. Chronic migraine prevalence, disability, and sociodemographic factors: Results from the American Migraine Prevalence and Prevention Study. *Headache*. 2012;52:1456-1470.
3. Lipton RB, Bigal ME, Diamond M, et al. Migraine prevalence, disease burden, and the need for preventive therapy. *Neurology*. 2007;68:343-349.
4. Smitherman TA, Burch R, Sheikh H, Loder E. The prevalence, impact, and treatment of migraine and severe headaches in the United States: A review of statistics from national surveillance studies. *Headache*. 2013;53:427-436.
5. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet*. 2017;390:1211–1259.
6. GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet Neurol*. 2017;16:877–897.
7. Steiner TJ, Stovner LJ, Vos T, Jensen R, Katsarava Z. Migraine is first cause of disability in under 50s: Will health politicians now take notice? *J Headache Pain*. 2018;19:17.
8. Abu Bakar N, Tanprawate S, Lambru G, Torkamani M, Jahanshahi M, Matharu M. Quality of life in primary headache disorders: A review. *Cephalalgia*. 2016;36:67-91.
9. Martin BC, Pathak DS, Sharfman MI, et al. Validity and reliability of the Migraine-Specific Quality of Life Questionnaire (MSQ Version 2.1). *Headache*. 2000;40:204-215.
10. Food and Drug Administration. Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims. *Fed Reg*. 2009;74:65132-65133.
11. 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.
12. Jhingran P, Davis SM, LaVange LM, Miller DW, Helms RWMSQ. Migraine-Specific Quality-of-Life Questionnaire. Further investigation of the factor structure. *Pharmacoeconomics*. 1998;13:707-717.
13. Pathak D, Martin B, Kwong J, Batenhorst A. Evaluation of the Migraine-Specific Quality of Life Questionnaire (MSQ Version 2.0) using confirmatory factor analysis. *Qual Life Res*. 1998;7:647.
14. Skljarevski V, Matharu M, Millen BA, Ossipov MH, Kim BK, Yang JY. Efficacy and safety of galcanezumab for the prevention of episodic migraine: Results of the EVOLVE-2 Phase 3 randomized controlled clinical trial. *Cephalalgia*. 2018;38:1442-1454.
15. Stauffer VL, Dodick DW, Zhang Q, Carter JN, Ailani J, Conley RR. Evaluation of galcanezumab for the prevention of episodic migraine: The EVOLVE-1 randomized clinical trial. *JAMA Neurol*. 2018;75:1080-1088.
16. Eli Lilly and Company. Announces Positive Results for Three Phase 3 Studies of Galcanezumab for the Prevention of Episodic and Chronic Migraine. Available at: https://investor.lilly.com/static-files/959f9ca0-24ed-4968-b3b7-f8f9506ede0f. Accessed November 2, 2018.
17. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: A meta-analytic review. *Value Health*. 2008;11:322-333.
18. Muehlhausen W, Doll H, Quadri N, et al. Equivalence of electronic and paper administration of patient-reported outcome measures: A systematic review and meta-analysis of studies conducted between 2007 and 2013. *Health Qual Life Outcomes*. 2015;13:167.

19. Coons SJ, Gwaltney CJ, Hays RD, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health*. 2009;12:419-429.
20. 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.
21. Stewart WF, Lipton RB, Kolodner K, Liberman J, Sawyer J. Reliability of the Migraine Disability Assessment score in a population-based sample of headache sufferers. *Cephalalgia*. 1999;19:107-114; discussion 174.
22. Lipton RB, Stewart WF, Stone AM, Lainez MJ, Sawyer JP. Disability in Strategies of Care Study g. Stratified care vs step care strategies for migraine: The Disability in Strategies of Care (DISC) Study: A randomized trial. *JAMA*. 2000;284:2599-2605.
23. Matchar DB, Harpole L, Samsa GP, et al. The headache management trial: A randomized study of coordinated care. *Headache*. 2008;48:1294-1310.
24. Guy W. *ECDEU Assessment Manual for Psychopharmacology*. Rockville, MD: National Institute of Mental Health, Psychopharmacology Research Branch; 1976. Available at https://archive.org/details/ecdeuassessmentm1933guyw. Accessed January 24, 2018.
25. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297-334.
26. Nunnally JC, Bernstein IH. *Psychometric Theory*, 3rd ed. New York: McGraw-Hill; 1994.
27. Macedo A, Banos JE, Farre M. Placebo response in the prophylaxis of migraine: A meta-analysis. *Eur J Pain*. 2008;12:68-75.
28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
29. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
30. European Medicines Agency. *Guideline on adjustment for baseline covariates in clinical trials*. 26 February 2015. Accessed August 5, 2017. Available at: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2015/03/WC500184923.pdf.
31. 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.
32. Habibzadeh F, Habibzadeh P, Yadollahie M. On determining the most appropriate test cut-off value: The case of tests with continuous results. *Biochem Med (Zagreb)*. 2016;26:297-307.
33. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Med Care*. 2003;41:582-592.
34. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care*. 1999;37:469-478.
35. Cella D, Eton DT, Fairclough DL, et al. What is a clinically meaningful change on the Functional Assessment of Cancer Therapy-Lung (FACT-L) Questionnaire? Results from Eastern Cooperative Oncology Group (ECOG) Study 5592. *J Clin Epidemiol*. 2002;55:285-295.
36. Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol*. 1999;52:861-873.
37. Wyrwich KW, Tierney WM, Wolinsky FD. Using the standard error of measurement to identify important changes on the Asthma Quality of Life Questionnaire. *Qual Life Res*. 2002;11:1-7.
38. Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: Is there a connection? *J Biopharm Stat*. 2004;14:97-110.
39. Leidy NK, Wyrwich KW. Bridging the gap: Using triangulation methodology to estimate minimal clinically important differences (MCIDs). *COPD*. 2005;2:157-165.
40. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61:102-109.
41. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes*. 2006;4:70.
42. Wyrwich KW, Spertus JA, Kroenke K, et al. Clinically important differences in health status for patients with heart disease: An expert consensus panel report. *Am Heart J*. 2004;147:615-622.
43. Bagley CL, Rendas-Baum R, Maglinte GA, et al. Validating Migraine-Specific Quality of Life Questionnaire v2.1 in episodic and chronic migraine. *Headache*. 2012;52:409-421.

44. Rendas-Baum R, Bloudek LM, Maglinte GA, Varon SF. The psychometric properties of the Migraine-Specific Quality of Life Questionnaire version 2.1 (MSQ) in chronic migraine patients. *Qual Life Res*. 2013;22:1123-1133.
45. Cole JC, Lin P, Rupnow MF. Minimal important differences in the Migraine-Specific Quality of Life Questionnaire (MSQ) version. *Cephalalgia*. 2009;29:1180-1187.
46. Dodick DW, Silberstein S, Saper J, et al. The impact of topiramate on health-related quality of life indicators in chronic migraine. *Headache*. 2007;47:1398-1408.
47. Food and Drug Administration. EMGALITY (galcanezumab-gnlm) injection, for subcutaneous use. Initial U.S. Approval; 2018. Available at: https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/761063s000lbl.pdf. Accessed September 28, 2018.