A self-administered screener for migraine in primary care

The ID Migraine™ validation study

R.B. Lipton, MD; D. Dodick, MD; R. Sadovsky, MD; K. Kolodner, PhD; J. Endicott, PhD; J. Hettiarachchi, MD; and W. Harrison, MD

Abstract—*Background:* Migraine is a highly prevalent and disabling illness that remains substantially undiagnosed in primary care. Because of the potential value of a screening tool, the current study was designed to establish the validity and reliability of a brief, self-administered migraine screener in patients with headache complaints in the primary care setting. *Methods:* A total of 563 patients presenting for routine primary care appointments and reporting headaches in the past 3 months completed a self-administered migraine screener. All patients were then referred for an independent diagnostic evaluation by a headache expert, of whom 451 (80%) completed a full evaluation. Migraine diagnosis was assigned based on International Headache Society criteria after completing a semi-structured diagnostic interview. *Results:* Of nine diagnostic screening questions, a three-item subset of disability, nausea, and sensitivity to light provided optimum performance, with a sensitivity of 0.81 (95% CI, 0.77 to 0.85), a specificity of 0.75 (95% CI, 0.64 to 0.84), and positive predictive value of 0.93 (95% CI, 89.9 to 95.8). Test-retest reliability was good, with a kappa of 0.68 (95% CI, 0.54 to 0.82). The sensitivity and specificity of the three-item migraine screener was similar regardless of sex, age, presence of other comorbid headaches, or previous diagnostic status. *Conclusions:* The three-item ID Migraine™ migraine screener was found to be a valid and reliable screening instrument for migraine headaches. Its ease of use and operating characteristics suggest that it could significantly improve migraine recognition in primary care.

NEUROLOGY 2003;61:375–382

Despite the prevalence,[1-5] severity,[5,6] and burden of migraine,[7-12] recent surveys suggest that fewer than half of current migraine sufferers have ever received a medical diagnosis of migraine.[13] Only one-third of migraine sufferers currently receive treatment with prescription drugs.[4,14] The low rates of diagnosis and treatment have several causes, including low rates of medical consultation specifically for headache. Although healthcare utilization has increased for migraine,[4,13,15,16] fewer than half of migraine sufferers have consulted a doctor for a complaint of headache in the previous year.[13] Barriers to effective migraine diagnosis have been reviewed elsewhere.[13,17-23]

Migraine is one of the leading reasons for outpatient visits to neurologists in the United States.[17,18] Further, neurologists play a leadership role in efforts to improve migraine diagnosis and treatment in primary care. Recognition of migraine in primary care settings may lead to neurologic referral for the patients who require specialty care. To improve diagnosis in primary care settings, the US Headache Consortium Guidelines recommended the use of screening or case-finding instruments.[23] Although screening sometimes refers to detection of an illness early in its course during an asymptomatic phase, herein we use the term to refer to the detection of previously undiagnosed individuals with migraine. An ideal screening instrument for migraine should be brief and easy to use (physician focus groups had previously suggested that a three-item screener would be ideal). It should have sufficient sensitivity to detect most patients with migraine and sufficient specificity to minimize the number of patients who screen positive for migraine but do not have the condition.

There have been efforts to develop screening instruments for migraine.[24-30] Some studies were conducted in specialty referral settings,[27] or in the general population, and results may not be applicable to the setting of intended use—primary care.[30] Other studies did not have optimal expert confirmation of the migraine diagnosis, i.e., verification by a headache specialist applying International Headache Society (IHS) criteria.[25,29] Still others developed instruments that either had unfavorable performance characteristics or were too time-consuming or burdensome for routine use in primary care.[24-27]

Our goal was to establish the validity and reliability

From the Departments of Neurology, Psychiatry, Epidemiology, and Population Health (Dr. Lipton), Albert Einstein College of Medicine, New York, NY; IMR, a Division of AdvancePCS (Dr. Kolodner), Hunt Valley, MD; Department of Neurology (Dr. Dodick), Mayo Clinic, Scottsdale, AZ; Department Family Practice/SUNY Downstate Medical Center at Brooklyn (Dr. Sadovsky), NY; New York State Psychiatric Institute (Dr. Endicott), New York; and Pfizer Inc. (Drs. Hettiarachchi and Harrison), New York, NY.

Supported by a grant from Pfizer, Inc.

Received November 4, 2002. Accepted in final form May 6, 2003.

Address correspondence and reprint requests to Dr. Richard B. Lipton, Department of Neurology, Albert Einstein College of Medicine, 1165 Morris Park Avenue, Rousso Building, Room 332, Bronx, NY 10461; e-mail: rlipton@aecom.yu.edu

of a brief screening instrument that uses self-report by the patient and has both the sensitivity and specificity that would make it useful in the outpatient primary care setting. Because the setting of intended use was primary care, in our validation study screening occurred in the primary care setting. The gold standard diagnostic assessment was subsequently performed by headache experts applying IHS criteria.[31] We conducted reliability and validity studies using a nine-item screener. After an item reduction process, we developed a three-item screener that was a subset of the initial screener. We then conducted a second study to compare the reliability of responses on the three-item screener to the full nine-item scale.

Subjects and methods. *Study setting.* The validity and reliability studies were conducted in the United States at 27 primary care practice (PCP) sites and 12 headache specialty practice sites. Headache specialty sites were selected for their geographic diversity and for the availability of a diagnostician (neurologist or internist) with a national reputation for expertise in headache diagnosis, and experience in migraine diagnosis for research studies. In addition, as noted below, diagnostic procedures were standardized across headache expert sites. PCP sites were selected for their proximity to the headache specialist site; PCP sites that regularly referred to the headache specialist site were excluded to avoid selection bias. Patients were enrolled between April 2000 and August 2000.

Study subjects. All men and women aged 18 to 55 years old, inclusive, who were visiting a PCP office for any reason (including a specific medical complaint or routine follow-up care) were potentially qualified to participate in the initial screening phase of the study. Additional entry requirements included the ability to read and write English, lack of participation in a previous Pfizer-sponsored migraine study, and the report of two or more headaches in the previous 3 months. In addition, because the goal of the screener was to identify migraine among PCP attendees complaining of headache, eligible subjects had to indicate that they had experienced a headache that had limited their ability to work, study, or enjoy life, or simply that they might wish to speak with a healthcare professional about their headaches. After one-third of the sample had been enrolled, we added an additional entry criterion that excluded patients with a previous diagnosis of migraine to ensure that a high proportion of patients had not previously been diagnosed with migraine.

The study protocol was reviewed by a central Institutional Review Board (IRB) or, if available, by the local IRB at the individual site. The benefits and expectations of study participation were explained to each patient, and written informed consent was obtained. Patients were reassured that declining to participate would in no way affect the care or treatment that they had come to receive from their PCP.

Study procedures. The overall scale validation process was divided into three phases: In phase one (completed before the current study), a nine-item screener was developed by a consensus panel based on IHS criteria and empirically tested for patient acceptability in a study that evaluated the prevalence, clinical characteristics, associated disability, and treatment characteristics of headache sufferers in the primary care setting.[32] In phase two, reported on here, an independent sample of patients was recruited from a different set of sites for the current validity and reliability studies. Based on this study, we developed a three-item migraine screener. In phase three, we administered the nine-item phase two screener and the three-item subset to determine if formatting changes altered the performance of the screener. This final agreement study was conducted in a third independent sample of subjects.

The validation study, which is the primary focus of the current report, had two components: 1) a PCP-based migraine screening phase and 2) a headache expert diagnostic evaluation phase. In the first part, outpatients visiting their PCP office who met the entry criteria itemized above were asked to complete the migraine screener, a self-rated paper-and-pencil questionnaire.

The initial screener consisted of nine questions that were developed by a consensus panel based on IHS criteria and empirically tested for patient acceptability in a previous study designed to evaluate the prevalence, clinical characteristics, associated disability, and treatment characteristics of headache patients in the primary care setting.[32] None of the patients used in the previous study were included in the current study. Eight questions were rated by the patient on an ordinal severity scale whose four descriptors were "never," "rarely," "less than half the time," and "half the time or more." These response options had been shown previously to improve the accuracy of a diagnostic telephone interview for migraine, with false positive symptom reporting minimized by categorizing "rarely" along with "never" as a negative response.[33] Four of these questions focused on the IHS-defined pain features of migraine and three of the IHS-defined associated symptoms of migraine. One question on aura was included. The ninth question focused on the number of days the patient reported that the headaches limited his or her ability to work, study, or do what he or she needed to do. This question quantified the days of disability caused by a patient's headache over the previous 3 months. Our goal was to draft an overly inclusive questionnaire, and then identify the items with the greatest discriminative validity for migraine. These items were to be used to develop the simplest possible instrument with the best possible operating characteristics. Additional questions not used for case-finding obtained information on age, sex, race, previous diagnosis, and frequency of headaches.

Headache specialist diagnosis. In the validation study, the patient-based screener was not scored or discussed with the patient, but was reviewed for completeness by the PCP or a member of his or her staff. The patient was then referred to a headache specialist for a structured diagnostic headache evaluation, within 2 weeks of PCP screening. Results of the migraine screener were not available to the headache specialist.

The second part of the phase two validation study, the gold standard diagnosis by the headache specialist, included a medical history, physical examination, comprehensive neurologic history and examination (including additional diagnostic tests if clinically indicated), and a semi-structured interview that included the IHS features of migraine supplemented by additional questions relating to family history and medical treatment history. In addition to asking the required questions, the headache expert was encouraged to probe for clinical information necessary to clarify the differential diagnosis.

Based on the evaluation detailed above, the headache specialist completed a symptom checklist based on IHS criteria[31] and assigned a clinical diagnosis of migraine (IHS 1.1, 1.2), migrainous disorder (IHS 1.7), or other headache. A computer-based algorithm was run on the symptom checklist and on the IHS criteria, and compared with clinician diagnoses. A patient was judged to have met operationally defined criteria for a diagnosis of migraine based on either the algorithm-based approach or an IHS diagnosis.

If the patient requested it, the results of the specialist diagnostic evaluation were forwarded to his or her PCP.

Health-related quality of life (HRQoL) and functional measures. During the headache specialist evaluation, patients completed several self-administered HRQoL and functional measures, including the Migraine-Specific Quality-of-Life Questionnaire, version 2.1 (MSQ),[34] a 14-item scale designed to assess the effect of migraine on HRQoL. Patients also completed the Migraine Disability Assessment (MIDAS) questionnaire,[35-38] which asks patients to record lost time due to headache in three domains: paid work, chores (household work), and nonwork activities (social, family, and leisure activities). In addition, patients completed the Migraine-Related Work Productivity Questionnaire (WPQ-24),[39] a 24-item scale designed to assess the effect of migraine on work productivity and work functioning. Finally, they completed the Henry Ford Hospital Headache Disability Inventory (HDI),[40,41] a 25-item scale designed to assess the effect of migraine on daily activities and functioning.

Test-retest reliability sample. To assess test–retest reliability, a subset of 123 patients, approximately 25% of the sample at each PCP site, were randomly selected to complete the migraine screener for a second time at home, approximately 2 to 7 days after first completing the screener at the PCP office, but before visiting the headache specialist. Participants in the reliability study were asked to mail the second screener directly to an inde-

pendent study monitor to minimize the possibility that the headache expert might see it.

Short-form/long-form migraine screener agreement study. The original migraine screener was a nine-item scale, with a choice of four responses permitted for the first eight items, and a numerical response for the ninth item (disability days). Based on analysis of the initial data set, we developed a three-item screener that provided excellent operating characteristics and brevity. We were concerned that the three items embedded in a longer questionnaire might perform differently than the three items as a standalone scale. We therefore conducted a second study at nine PCP sites selected from the previous study participants, using a new patient sample. The goal of this study was to assess the level of agreement between the three-item screener used in isolation vs the three items embedded in the original screener. As a secondary aim, we also compared the original four-category response options with a binary (yes/no) response option in an attempt to further simplify the instrument. Patients were recruited in the PCP setting using the same entry criteria, and received, in random order and at 5 ± 2 day intervals, three versions of the screener: 1) the nine-item screener with ordinal responses, 2) the three-item screener with ordinal responses, and 3) the three-item Identification of Migraine (ID Migraine) Screener with binary responses.

Statistical analyses. Descriptive statistics were performed on the demographic and clinical variables, as well as on responses on the migraine screener.

Item selection and validity assessment. Each of the nine migraine-screener items was analyzed for item-total correlations and the influence of item elimination on internal consistency (Cronbach alpha). We planned to eliminate items with poor (<0.20) item-total correlations. We further planned to exclude highly correlated items. A combination of statistical analysis and clinical/medical judgment was used to identify redundant questions for deletion.

Validity assessment. Sensitivity and specificity of individual items were evaluated against the headache specialist's diagnosis. Based on prior work, we planned to treat the response to each item as a binary variable with a "no" assigned to responses of "never" or "rarely" and "yes" assigned to responses of "less than half the time" or "half the time or more."[33] We also performed exploratory analyses to examine how varying the cut-score for the "yes–no" assignment changed the level of agreement between individual items on the migraine screener and the equivalent physician-rated items.

To identify the optimal migraine screener, the following analyses were performed. First, using the total score of the nine-item migraine screener against the gold standard, a receiver-operator characteristic (ROC) curve was generated. The ROC curve is a graphical method for evaluating the trade-off between false positive and false negative rates for every possible cutoff. The ROC graph plots the false positive rate on the x-axis and the true positive rate (one minus the false negative rate) on the y-axis. The area under the curve is a measure of the correlation between the prediction of the screener and the gold standard diagnosis. Using the method proposed by Halpern et al.,[42] the optimal operating point (OOP) was determined from the ROC curve. The OOP serves as a cut-score for classifying subjects as likely migraine sufferers vs nonmigraine sufferers. Based on the OOP, sensitivity and specificity with their 95% two-sided confidence intervals were calculated.

Second, sensitivity and specificity were calculated using selected migraine screener items based on an algorithm that paralleled the operationalized IHS criteria for migraine.

Third, a logistic regression analysis was performed with the gold standard diagnosis as the dependent variable and individual screener items as predictor variables to determine which combination of items was independently associated with the gold standard diagnosis of migraine. Based on the results of the logistic regression, empirical combinations were chosen using the three items with the highest individual sensitivities and specificities. An OOP was then determined from the ROC curve for short versions of the scale in an attempt to identify the most efficient diagnostic screener.

Test-retest reliability assessment. The reliability assessment was performed using the test-retest sample. The test-retest reliability of the refined total score for the migraine screener was evaluated with an intraclass correlation coefficient using the Fleiss and Shrout formula.[43] Using the intraclass correlation coefficient, the screener-refined total scores at time one (initial screening visit) were compared to the time two (retest visit) scores.

The kappa statistic and percent agreement were calculated to compare the responses to each of the nine items as binary variables with a "no" assigned to responses of "never" or "rarely" and "yes" assigned to responses of "less than half the time" or "half the time or more" at time one compared with time two.

The migraine screener generated a score used for classifying subjects as likely migraine sufferers vs nonmigraine sufferers. Test-retest reliability was also calculated on this binary scoring of the migraine screening questionnaire. For this purpose, the kappa statistic was used to compare migraine diagnosis by the optimal migraine screener at time one compared with time two.

The adequacy of the kappa coefficient (K) was evaluated using the following descriptive ranges: a kappa of 0.40 to 0.59 is considered to be moderate, a kappa of 0.60 to 0.79 is considered to be substantial, and a kappa ≥0.80 is considered to be excellent agreement.[44]

Secondary analyses. Secondary analyses examined the impact of the following variables on the sensitivity and specificity of the optimal migraine screener: sex, age, history of a previous diagnosis of migraine headache, and presence or absence of other headache comorbidity. Likelihood ratios were calculated using the standard formula for a positive likelihood ratio (LR+, sensitivity/1 − specificity) and for a negative likelihood ratio (LR−, 1 − sensitivity/specificity).

Results. *Patient disposition across the PCP and headache-specialist phases of the study.* The inception cohort consisted of 563 patients who met study eligibility criteria and agreed to participate. Of these, 550 patients were screened and referred for evaluation by a headache expert, and 451 (80%) completed the structured diagnostic assessment, and they constitute the validation sample (figure). The validation sample had a mean age of 39.3 ± 10.1 years, and consisted of 341 women (75.6%) and 110 men (24.4%). A total of 312 patients (69%) were white, 105 (23%) were African American, 13 (2.9%) were Asian, and 21 (4.7%) reported other races. A total of 223 (50%) were currently married; 143 (32%) had never married; 70 (15%) were divorced, separated, or widowed; and 15 patients had uncertain marital status. In terms of current work status, 332 (74%) reported being employed full-time, 49 (11%) were employed part time, 30 (7%) were homemakers, 22 (5%) were unemployed, 13 (3%) were students, and 3 (1%) were retired. Among patients who eventually received a gold standard diagnosis of migraine, 124 (28%) said they had received a previous migraine diagnosis.

The 99 patients who completed the screener and were referred but did not complete the diagnostic evaluation by a headache expert showed only modest differences from the validation sample: they tended to be younger (mean ± SD, 31.0 ± 11.6 years vs 39.3 ± 10.1 years), reported less severe headache pain (moderate to severe, 82.7% vs 92.0%; $p < 0.01$), and reported less nausea (42.7% vs 52.5%; $p = 0.07$). These differences are consistent with previously reported differences among patients seeking, compared with those not seeking, medical help for headaches.[17]

Gold standard diagnosis of migraine. Headache experts assigned a diagnosis based on both a clinical evaluation and IHS criteria (using a semi-structured interview). The agreement between the two diagnostic methods used by headache specialists was very high: 97.6% (329/337) of patients receiving a clinical diagnosis of migraine were also diagnosed with migraine based on IHS criteria (checked for accuracy by computer scoring). Similarly, 92.2% (329/357) of patients who met IHS criteria for mi-

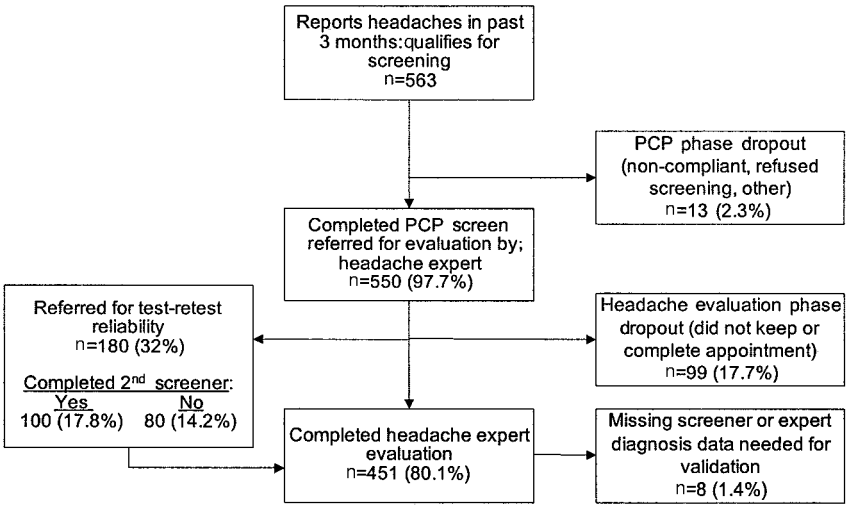

Figure. Subject disposition during study. PCP = primary care practice.

graine were also given a clinical diagnosis of migraine. An additional eight patients received a clinical diagnosis of migrainous disorder, although none of these patients received an IHS criteria-based diagnosis of either migrainous disorder or migraine. A consensus decision was reached, prior to performing the validation analysis, that these eight patients would not be classified as having a gold standard diagnosis of migraine.

Sensitivity and specificity of individual ID Migraine Screener items. We calculated the sensitivity and specificity of each individual item on the nine-item migraine screener. The results show that the individual items have sensitivities ranging from a low of 43% for aura to a high of 94% for pain intensity (table 1). Specificity was highest for nausea (81%), photophobia (74%), and aura (74%). Items providing levels of sensitivity and specificity greater than 50% each include one-sided pain, exacerbation with activity, nausea, photophobia, phonophobia, and functional impairment.

Exploratory analyses examined how varying the cut-score for "yes–no" assignment changed the level of agreement between individual items on the migraine screener and the equivalent physician-rated items. Consistent with previous reports,[33] we found that the "never/rarely" vs "less/more-than-half-the-time" dichotomization yielded the highest kappas, in the range of 0.23 to 0.47. These kappas for individual items are only fair, supporting the plan to combine items.

Developing the migraine screener. We employed three different approaches to optimizing the migraine screener. The first approach involved selecting migraine screener items based on an algorithm that paralleled the operationalized IHS criteria for migraine.[31] The results of this item definition yielded a screener with a sensitivity of 0.81 (95% CI, 0.77 to 0.85) and a specificity of 0.69 (95% CI, 0.58 to 0.79).

The second method involved simply summing the binary responses to the nine screener items. We generated an ROC curve by plotting the sensitivity of the screener total score against one minus the specificity.[45] The optimal operating point was determined to be a total score of six, which yielded a sensitivity of 0.77 (95% CI, 0.72, 0.81) and a specificity of 0.74 (95% CI, 0.62, 0.84). Compared to the

Table 1 *Sensitivity and specificity of individual screener items vs the gold standard migraine diagnosis**

Item	Sensitivity	95% CI	Specificity	95% CI
1. Pain is worse on just one side	0.75	0.70–0.79	0.50	0.39–0.61
2. Pain is pounding, pulsing, or throbbing	0.87	0.83–0.90	0.22	0.14–0.33
3. Pain is moderate or severe	0.94	0.91–0.96	0.16	0.09–0.25
4. Pain is made worse by activities such as walking or climbing stairs	0.67	0.62–0.72	0.57	0.46–0.68
5. You feel nauseated or sick to your stomach	0.60	0.55–0.65	0.81	0.71–0.89
6. You see spots, stars, zig-zags, lines, or gray areas for several minutes or more before or during your headaches (aura symptoms)	0.43	0.38–0.48	0.74	0.63–0.83
7. Light bothers you (a lot more than when you don't have headaches)	0.75	0.71–0.80	0.74	0.63–0.83
8. Sound bothers you (a lot more than when you don't have headaches)	0.83	0.78–0.86	0.56	0.45–0.67
9. Functional impairment due to headache in last 3 months†	0.87	0.83–0.90	0.52	0.40–0.63

* Sample size ranged from 438 to 448 owing to occasional missing values.
† Scored positive if disability reported on any 1 day in the past 3 months.

Table 2 *Results of bivariate and multivariate logistic regression models: unadjusted and adjusted odds: ratios for gold standard migraine diagnosis for each screener item*

Item	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
You felt nauseated or sick to your stomach	6.46 (3.61–11.59)	3.97 (1.92–8.20)
How many days did your headache limit you from working, studying, or doing what you needed to do?	6.98 (4.08–11.95)	3.82 (1.97–7.42)
Light bothered you (a lot more than when you don't have headaches)	8.68 (5.05–14.92)	3.30 (1.54–7.06)
The pain was worse on just one side	2.92 (1.78–4.79)	1.66 (0.88–3.14)
Sound bothered you (a lot more than when you don't have headaches)	6.03 (3.62–10.03)	1.56 (0.76–3.20)
The pain was made worse by activities such as walking or climbing stairs	2.69 (1.66–4.37)	1.24 (0.65–2.38)
The pain was moderate or severe	2.81 (1.35–5.85)	0.85 (0.32–2.28)
The pain was pounding, pulsing, or throbbing	1.91 (1.04–3.50)	0.82 (0.35–1.90)
You saw spots, stars, zig-zags, lines, or gray areas for several minutes or more before or during your headaches	2.12 (1.25–3.60)	0.69 (0.33–1.45)

IHS algorithm-scored screener described above, this empirical approach resulted in a small decrement in sensitivity, but a small increment in specificity.

The third approach used a logistic regression analysis to determine which combination of items was independently associated with the gold standard diagnosis of migraine. In univariate models each item was found to be significantly associated with migraine status (table 2). When all nine items were included into the model (see table 2), three variables had a strong and significant association with the diagnosis of migraine: nausea, photophobia, and headache-related disability (any day in the previous 3 months). Other IHS criteria did not yield significant additional predictive value when included in models containing these three variables. Empirical testing of combinations of these three items found that the optimal total score was two (i.e., scored positive on any two out of the three items). This yielded a sensitivity of 0.81 (95% CI, 0.77 to 0.85) and a specificity of 0.75 (95% CI, 0.64 to 0.84). In the primary care setting, the three-item ID Migraine Screener was found to have a positive predictive value of 93.3 (95% CI, 89.9 to 95.8).

Internal consistency reliability analyses. The reliabilities of both the long (nine-item) and short (three-item) versions of the migraine screener were evaluated. The internal consistency of the items in the long-form screener (binary scoring) was found to be good, with a Cronbach alpha for the total scale of 0.70. Deleting any single item attenuated the alpha, suggesting that all items contributed to internal consistency. The item-to-total score correlations ranged from 0.27 to 0.57, suggesting that no single item showed too high a correlation.

Test-retest reliability. For the 123 patients who completed the nine-item migraine screener a second time, test-retest reliability was found to be good. For the three-item ID Migraine Screener (binary response) the test-retest kappas for individual items ranged from 0.61 to 0.69. For the total score on the three-item ID Migraine Screener (binary response), the kappa was 0.68 (95% CI, 0.54 to 0.82). The test-retest interval (0 to 2 days vs 3 to 5 days vs 6 or more days) did not influence agreement.

Performance of three-item ID Migraine Screener in clinically relevant subgroups. We examined the performance of the three items in subgroups defined by age, sex, previous diagnostic status, and headache comorbidity. There were only modest differences in the performance of the ID Migraine Screener in these clinically relevant subgroups (table 3). Most notably, the screener appeared to have somewhat lower sensitivity in men compared to women (0.65 vs 0.86), but higher specificity.

Agreement study for the three-item ID Migraine Screener. An additional study, using an independent patient sample, was conducted in nine PCP sites to confirm the agreement between the original nine-item long form of the migraine screener and the three-item ID Migraine Screener. In a randomly assigned order, 161 patients received each of the migraine screeners at approximately 5-day intervals. Because of an a priori decision to exclude any patient who fell outside a 14-day test-retest window, data on 133 patients were included in the analysis. The results of the analysis found an 88% agreement between the nine-item long form and the three-item screener, using ordinal scale response options; the kappa was 0.68 (95% CI, 0.53, 0.82). The agreement using binary vs ordinal response options for the three-item scale was 87% with a corresponding kappa of 0.61 (95% CI: 0.44, 0.78).

Discussion. The current study demonstrates that the three-item ID Migraine Screener, consisting of questions on disability, nausea, and photophobia, is a valid and reliable screening instrument for migraine headaches in the primary care setting. In the PCP setting it had a sensitivity of 0.81 (95% CI, 0.77 to 0.85) and a specificity of 0.75 (95% CI, 0.64 to 0.84), relative to an IHS-based migraine diagnosis assigned by a headache specialist.

No additional gain in sensitivity or specificity was achieved by use of the longer nine-item version of the screener, whether using a quantitative scale or the application of an algorithm based on IHS criteria. The presence of aura is the defining feature for one

Table 3 *Sensitivity, specificity, and positive (LR+) and negative (LR−) likelihood ratios for the three-item ID Migraine screener in the total sample and in clinically relevant subgroups*

Patient group*	Sensitivity (95% CI)	LR+ (95% CI)	Specificity (95% CI)	LR− (95% CI)
Total validation sample, n = 443*	0.81 (0.77–0.85)	3.25 (2.69–3.93)	0.75 (0.64–0.84)	0.25 (0.22–0.28)
Men, n = 110	0.65 (0.53–0.76)	3.27 (2.31–4.63)	0.80 (0.63–0.92)	0.43 (0.36–0.51)
Women, n = 333	0.86 (0.81–0.89)	2.99 (2.38–3.75)	0.71 (0.57–0.83)	0.20 (0.17–0.25)
Ages 18–44 y, n = 274	0.83 (0.77–0.87)	3.44 (2.67–4.43)	0.76 (0.62–0.87)	0.23 (0.19–0.27)
Ages 45–55 y, n = 169	0.79 (0.71–0.86)	2.99 (2.24–3.99)	0.74 (0.56–0.87)	0.28 (0.23–0.34)
No previous diagnosis, n = 291	0.76 (0.70–0.82)	3.01 (2.45–3.70)	0.75 (0.63–0.84)	0.32 (0.28–0.37)
Previous diagnosis, n = 152	0.89 (0.83–0.94)	3.87 (2.33–6.43)	0.77 (0.46–0.95)	0.14 (0.11–0.19)
Migraine only, n = 184	0.80 (0.73–0.85)	NA	NA	NA
Migraine + other headache, n = 148	0.83 (0.76–0.89)	NA	NA	NA

* Validation data were missing for eight subjects: one subject could not be classified based on the physician assessment and an additional seven subjects were missing one of the key three items from the screener. The resulting sample size available for validation was 443.

NA = not applicable.

of the main clinical subtypes of migraine.[3,46,47] Physicians frequently rely on aura as a cardinal symptom of migraine, as suggested by the 1.9-fold higher rate of medical diagnosis in migraine with aura compared to those who have migraine without aura.[28,47] In the current validation study aura had good specificity but relatively low sensitivity (43%). Although aura is significantly associated with migraine status in univariate analyses, a logistic regression analysis found that it did not add significantly to the performance of the three-item screener. This occurs because most patients who have migraine with aura also meet criteria for migraine without aura, and specifically have at least two out of the three migraine screener items: i.e., disability, nausea, and photophobia.[48]

The performance of the three-item ID Migraine Screener is approximately equivalent to the performance of the Patient Health Questionnaire, which is the patient-rated version of the depression section of the PRIME-MD® (sensitivity, 85%; specificity, 75%).[49] Similarly, the ID Migraine Screener performs with a sensitivity and specificity in the same range as other commonly utilized medical screening tests such as PSA for prostate cancer (sensitivity, 75%; specificity, 74%),[50] Pap test for cervical cancer (sensitivity, 50% to 70%; specificity, 92% to 95%),[51] and electrocardiogram for myocardial infarction (sensitivity, 50% to 60%; specificity, 95%).[52] Because of the prevalence of migraine in the primary care setting, the high sensitivity and specificity of the ID Migraine Screener translated into a very high positive predictive value of 93.3%, suggesting that the Screener is very efficient in this setting. The predictive value of the Screener remains untested in broader community settings. It is important to emphasize that the ID Migraine Screener is a screening tool and not a diagnostic instrument. The ID Migraine Screener has a false positive rate (1 − sensitivity) of 19% (see table 3). Therefore, a thorough evaluation is essential to make the diagnosis of migraine.

The ID Migraine Screener also had very good test-retest reliability with a kappa of 0.68. In a separate study the three-item Screener agreed closely with the full questionnaire, indicating that question context does not substantially influence the operating characteristics of the Screener.

This kappa is quite good in light of the inherent uncertainty associated with the clinical diagnosis of migraine. Kappas range from 0.55 to 0.81 among neurologists asked to assign a headache diagnosis based on review of videotaped patient interviews.[53]

ID Migraine includes a disability item as one of the three Screener questions. Inclusion of this item distinguishes the ID Migraine Screener from some of the previous symptom-based migraine screeners with lower sensitivity and specificity.[24-30]

A large community survey[3] found that approximately 75% of migraineurs reported severe disability or complete bed rest as a functional outcome of their acute migraine. Despite this, the extent of disability associated with migraine is not only underestimated, but is frequently not even evaluated in patients with headache. The lack of physician-patient communication regarding migraine-related disability appears to contribute significantly to the underdiagnosis of the illness.[54]

A subsequent report will present more detailed analyses of the disability and HRQoL data obtained on patients in the current study. The high disability scores on the MIDAS and the low HRQoL in the migraine sufferers identified by screening in this study suggest that the ID Migraine Screener identifies migraine sufferers in need of medical care, as noted below.

The ID Migraine Screener, scored positive if two or more of the three items were endorsed, showed similar sensitivities and specificities across various

clinically relevant subgroups (see table 3). It performed well in a sample of previously undiagnosed migraine sufferers identified in primary care. These migraine sufferers had moderate to high levels of disability, losing a mean (± SD) of 6.2 (±8.6) days from work due to headache over the previous 3 months. These individuals also experienced significant decrements in HRQoL and functional status when compared to the general population. The mean MIDAS disability score among previously undiagnosed migraine sufferers was 30.2, somewhat lower than the score reported by patients who had previously been diagnosed (37.7), but notably worse than patients reporting only nonmigraine headaches (13.3). The ID Migraine Screener represents a potentially useful screening tool for primary care healthcare providers and their patients. Given the availability of effective treatment, the use of this screening tool might represent an important step toward reducing the burden of an illness that the World Health Organization has identified as being one of the top 20 most disabling disorders globally.[8,55] In most cases, effective initial migraine treatment can take place in the primary care setting followed by neurologic referral in selected cases.[23]

The strengths of the study include the large primary care sample and the use of an IHS gold standard diagnosis of migraine by headache specialists. It should be noted that expertise in migraine diagnosis is difficult to define operationally, so an IHS checklist-guided diagnosis was also included. Possible study limitations include uncertainty regarding the representativeness of both the participating PCP practices and the patient sample. Although attrition was relatively low (20%), the decision to participate, and to accept referral to a headache expert, might introduce bias, while study procedures may also, and unavoidably, have influenced results. Despite this possible limitation, the large number of geographically diverse primary care practice patients makes it likely that these results are broadly generalizable to US primary care sites.

Diagnostic rates of migraine in the primary care setting remain relatively low, despite its high prevalence and associated pain and disability. The results of this study confirm the validity and reliability of a brief, self-administered screening instrument, the three-item ID Migraine Screener, with excellent performance characteristics that recommend it as a simple method for increasing the recognition of migraine in the primary care setting.

Acknowledgment

The authors thank the investigators who participated in this study: J. Brandes, J. Asher, Nashville, TN; U. Patel, Murfreesboro, TN; J. Couch, M. Dees IL, J. Liprie, Oklahoma City, OK; M. Cutrer, Charleston, MA; F. Kronenberg, Chestnut Hills, MA; M. Williams, Boston, MA; M. Diamond, B. Bernheim, N. LaSala, Chicago, IL; J. Files, M. Ko, P. Shah, Scottsdale, AZ; D. Kudrow, Encino, CA; M. Rashti, Los Angeles, CA; C. Yang, Beverly Hills, CA; V. Martin, J. Fidelholtz, J. Konerman, J. Zerbe, Cincinnati, OH; A. Mauskop, M. Beer, R. Samuelson, L. Newman, G. Leeds, New York, NY; R. Sadovsky, Brooklyn, NY; J. Saper, E. Linkner, P. Scuccimarri, D. Wilson, Ann Arbor, MI; K. Wilks, R. Goodwin, Towson, MD; S. Miller, Baltimore, MD; P. Winner, J. Dresner, N. Vasquez, West Palm Beach, FL.

References

1. Stewart WF, Lipton RB, Celentano DD, Reed ML. Prevalence of migraine headache in the United States. Relation to age, income, race, and other sociodemographic factors. JAMA 1992;267:64–69.
2. Rasmussen BK. Epidemiology of headache. Cephalalgia 2001;21:774–777.
3. Lipton RB, Stewart WF, Diamond S, Diamond ML, Reed M. Prevalence and burden of migraine in the United States: data from the American Migraine Study II. Headache 2001;41:646–657.
4. Lipton RB, Scher AI, Kolodner K, Liberman J, Steiner TJ, Stewart WF. Migraine in the United States: epidemiology and patterns of health care use. Neurology 2002;58:885–894.
5. Linet MS, Stewart WF, Celentano DD, Ziegler D, Sprecher M. An epidemiologic study of headache among adolescents and young adults. JAMA 1989;261:2211–2216.
6. Stewart WF, Shechter A, Lipton RB. Migraine heterogeneity. Disability, pain intensity, and attack frequency and duration. Neurology 1994; 44(suppl 4):S24–S39.
7. Von Korff M, Stewart WF, Simon DJ, et al. Migraine and reduced work performance: a population-based diary study. Neurology 1998;50:1741–1745.
8. Menken M, Munsat TL, Toole JF. The global burden of disease study: implications for neurology. Arch Neurol 2000;57:418–420.
9. Lipton RB, Hamelsky SW, Kolodner KB, et al. Migraine, quality of life, and depression: a population-based, case-control study. Neurology 2000; 55:629–635.
10. Hu XH, Markson LE, Lipton RB, Stewart WF, Berger ML. Burden of migraine in the United States: disability and economic costs. Arch Intern Med 1999;159:813–815.
11. Terwindt GM, Ferrari MD, Tijhuis M, et al. The impact of migraine on quality of life in the general population: the GEM study. Neurology 2000;55:624–629.
12. Osterhaus JT, Gutterman DL, Plachetka JR. Health care resource and lost labor costs of migraine headache in the US. Pharmacoeconomics 1992;2:67–76.
13. Lipton RB, Diamond S, Reed M, Diamond M, Stewart WF. Migraine diagnosis and treatment: results from the American Migraine Study II. Headache 2001;41:646–657.
14. Celentano DO, Stewart WF, Lipton RB, Reed M. Medication use and disability among migraineurs. Headache 1992;32:223–228.
15. Clouse JC, Osterhaus JT. Healthcare resource use and costs associated with migraine in a managed healthcare setting. Ann Pharmacother 1994;28:659–664.
16. Rapoport AM, Adelman JU. Cost of migraine management: a pharmacoeconomic overview. Am J Manag Care 1998;4:531–545.
17. Bekkelund SI, Albretsen C. Evaluation of referrals from general practice to a neurological department. Fam Pract 2002;19:297–299.
18. Carson AJ, Ringbauer B, MacKenzie L, Warlow C, Sharpe M. Neurological disease, emotional disorder, and disability: they are related: a study of 300 consecutive new referrals to a neurology outpatient department. J Neurol Neurosurg Psychiatry 2000;68:202–206.
19. Lipton RB, Stewart WF, Simon D. Medical consultation for migraine: results from the American Migraine Study. Headache 1998;38:87–96.
20. Carr-Hill R, Jenkins-Clarke S, Dixon P, et al. Do minutes count? Consultation lengths in general practice. J Health Serv Res Policy 1998;3: 207–213.
21. Lipton RB, Goadsby PJ, Sawyer JPC, et al. Migraine: diagnosis and assessment of disability. Rev Contemp Pharmacother 2000;11:63–67.
22. MacGregor EA. The doctor and the migraine patient: improving compliance. Neurology 1997;48(suppl 3):S16–S20.
23. Matchar DB, Young WB, Rosenberg JH, et al. Evidence-based guidelines for migraine headache. Neurology 2000;55:754–762.
24. Rasmussen BK, Jensen R, Olesen J. Questionnaire versus clinical interview in the diagnosis of headache. Headache 1991;31:290–295.
25. Bensenor IJ, Lotufo PA, Pereira AC, et al. Validation of a questionnaire for the diagnosis of headache in an outpatient clinic at a university hospital. Arq Neuropsiquiatr 1997;55:364–369.
26. Michel P, Henry P, Letenneur L, Jogeix M, Corson A, Dartigues JF. Diagnostic screen for assessment of the IHS criteria for migraine by general practitioners. Cephalalgia 1993;13(suppl 12):54–59.
27. Russell MB, Rasmussen BK, Brennum J, Iversen HK, Jensen RA, Olesen J. Presentation of a new instrument: the diagnostic headache diary. Cephalalgia 1992;12:369–374.
28. Leone M, Filippini G, D'Amico D, Farinotti M, Bussone G. Assessment of International Headache Society diagnostic criteria: a reliability study. Cephalalgia 1994;14:280–284.
29. Gervil M, Ulrich V, Olesen J, Russell MB. Screening for migraine in the general population: validation of a simple questionnaire. Cephalalgia 1998;18:342–348.
30. Pryse-Phillips W, Aube M, Gawel M, Nelson R, Purdy A, Wilson K. A headache diagnosis project. Headache 2002;42:728–737.

31. Headache Classification Committee of the International Headache Society. Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. Cephalalgia 1988;8(suppl 7):1–96.
32. Taylor K, Couch J, Schultz J, Denaro J, Kelly M, Harrison W. Prevalence and impact of migraine in primary care. Headache 2000;40:434.
33. Stewart WF, Lipton RB, Liberman J. Variation in migraine prevalence by race. Neurology 1996;47:52–59.
34. 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.
35. 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.
36. Stewart WF, Lipton RB, Whyte J, et al. An international study to assess reliability of the Migraine Disability Assessment (MIDAS) score. Neurology 1999;53:988–994.
37. Stewart WF, Lipton RB, Kolodner KB, Sawyer J, Lee C, Liberman JN. Validity of the Migraine Disability Assessment (MIDAS) score in comparison to a diary-based measure in a population sample of migraine sufferers. Pain 2000;88:41–52.
38. 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(suppl 1):S20–S28.
39. Davies GM, Santanello N, Gerth W, Lerner D, Block GA. Validation of a migraine work and productivity loss questionnaire for use in migraine studies. Cephalalgia 1999;19:497–502.
40. Jacobson GP, Ramadan NM, Aggarwal SK, Newman CW. The Henry Ford Hospital Headache Disability Inventory (HDI). Neurology 1994;44:837–842.
41. Jacobson GP, Ramadan NM, Norris L, Newman CW. Headache Disability Inventory (HDI): short-term test-retest reliability and spouse perceptions. Headache 1995;35:534–539.
42. Halpern EJ, Albert M, Krieger AM, Metz CE, Maidment AD. Comparison of receiver operating characteristic curves on the basis of optimal operating points. Acad Radiol 1996;3:245–253.
43. Fleiss JL, Shrout PE. The effects of measurement errors on some multivariate procedures. Am J Public Health 1977;67:1188–1191.
44. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.
45. Kraemer H. Evaluating medical tests. Objective and quantitative guidelines. Thousand Oaks, CA: Sage Publications, 1992.
46. Lipton RB, Stewart WF, Celentano DD, Reed ML. Undiagnosed migraine headaches. Arch Intern Med 1992;152:1273–1278.
47. Stewart AL, Greenfield S, Hays RD, et al. Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. JAMA 1989;262:907–913.
48. Russell MB, Iversen HK, Olesen J. Improved description of the migraine aura by a diagnostic aura diary. Cephalalgia 1994;14:107–117.
49. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders Patient Health Questionnaire. JAMA 1999;282:1737–1744.
50. Bangma CH, Kranse R, Blijenberg BG, Schroder FH. Free and total prostate-specific antigen in a screened population. Br J Urol 1997;79:756–762.
51. Nanda K, McCrory DC, Myers ER, et al. Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. Ann Intern Med 2000;132:810–819.
52. Khaw K, Moreyra AE, Tannenbaum AK, Hosler MN, Brewer TJ, Agarwal JB. Improved detection of posterior myocardial wall ischemia with the 15-lead electrocardiogram. Am Heart J 1999;138:934–940.
53. Granella F, D'Alessandro R, Manzoni GC, et al. International Headache Society classification: interobserver reliability in the diagnosis of primary headaches. Cephalalgia 1994;14:16–20.
54. Holmes WF, MacGregor EA, Sawyer JP, Lipton RB. Information about migraine disability influences physicians' perceptions of illness severity and treatment needs. Headache 2001;41:343–350.
55. World Health Organization. World Health Report 2001, figure 2.3. Available at: www.who.int/whr/index.htm. Accessed July 18, 2002.