Validation of the web-based LUMINA questionnaire for recruiting large cohorts of migraineurs

4

W.P.J. (Ron) van Oosterhout[1], Claudia M. Weller[2]
Anine H. Stam[1], Floor Bakels[1], Theo Stijnen[3]
Michel D. Ferrari[1], Gisela M. Terwindt[1]

Departments of [1]Neurology,
[2]Human Genetics, [3]Medical Statistics
Leiden University Medical Center, Leiden

Cephalalgia 2011;31(13):1359–1367

Abstract

Objective – To assess validity and reliability of a self-administered web-based migraine-questionnaire in diagnosing migraine aura for the use of epidemiological and genetic studies.
Methods – Self-reported migraineurs enrolled via the LUMINA-website and completed a web-based questionnaire on headache and aura symptoms, after fulfilling screening criteria. Diagnoses were calculated using an algorithm based on the International Classification of Headache Disorders (ICHD-2), and semi-structured telephone-interviews were performed for final diagnoses. Logistic regression generated a prediction rule for aura. Algorithm-diagnoses and predicted diagnoses were subsequently compared to the interview-derived diagnoses.
Results – In one year, we recruited 2,397 migraineurs, of which 1,067 were included in the validation. A seven-question subset provided higher sensitivity (86%vs.45%), slightly lower specificity (75%vs.95%) and similar positive predictive value (86%vs.88%) in assessing aura when comparing with the ICHD-2 based algorithm.
Conclusions – This questionnaire is accurate and reliable in diagnosing migraine aura among self-reported migraineurs, and enables detection of more aura cases with low false-positive rate.

Key Words: migraine ■ headache ■ questionnaire validation ■ cohort study ■ screening in epidemiology

Introduction

MIGRAINE is a common brain disorder characterized by recurrent, disabling attacks of headache, autonomic features (migraine without aura; MO), and, in one third of patients, transient neurological aura symptoms (migraine with aura; MA). In western countries, the overall migraine prevalence in the general population is at least 12 percent, two-thirds of which concerns females[1-4]. Since no biomarker for migraine exists, diagnosis according to the headache classification of the International Headache Society (IHS)[5] relies exclusively on the headache history. A careful history taken by a headache specialist is the gold standard for making a valid migraine and aura diagnosis.

Large-scale studies with several thousands of participants are important to obtain information for epidemiological and genetic migraine research and may yield important insights in migraine pathophysiology. Migraine is a complex genetic disorder, i.e. multiple genetic and environmental factors contribute to migraine susceptibility.

Twin and population-based family studies showed that genetic factors play an important role in migraine susceptibility, especially in the MA subtype[6-12]. However, genetic linkage studies using migraine subtypes as an end diagnosis did not yield gene variants thus far. Clinical heterogeneity in migraine and aura diagnosis may have hampered the identification of such variants. Recently, in a large genome wide association analysis (GWA) with a large set of clinic-based migraineurs, a first-ever genetic risk factor was identified associated with common types of migraine, in patients that were largely recruited from specialist headache clinics with a clinic-based migraine diagnosis[13]. However, population-based large-scale studies exclude the possibility of a face-to-face examination, and, therefore, a less time-consuming and less costly diagnostic strategy has to be chosen. A web-based questionnaire represents an attractive and inexpensive alternative for a clinic interview. Several groups have reported on the use of internet to recruit headache and other patients for clinical research[14-18]. However, reliably diagnosing aura remains an issue.

The availability of a validated, aura-specific questionnaire is important when large numbers of cases are needed, especially in studies with self-reported migraineurs from the general population[19,20]. We developed the LUMINA (Leiden University MIgraine Neuro-Analysis) website and designed and validated a self-reporting, web-based questionnaire to reliably diagnose migraine headache and aura symptoms, using only a limited number of questions. In this paper, we will present the validation of this web-based migraine and aura questionnaire.

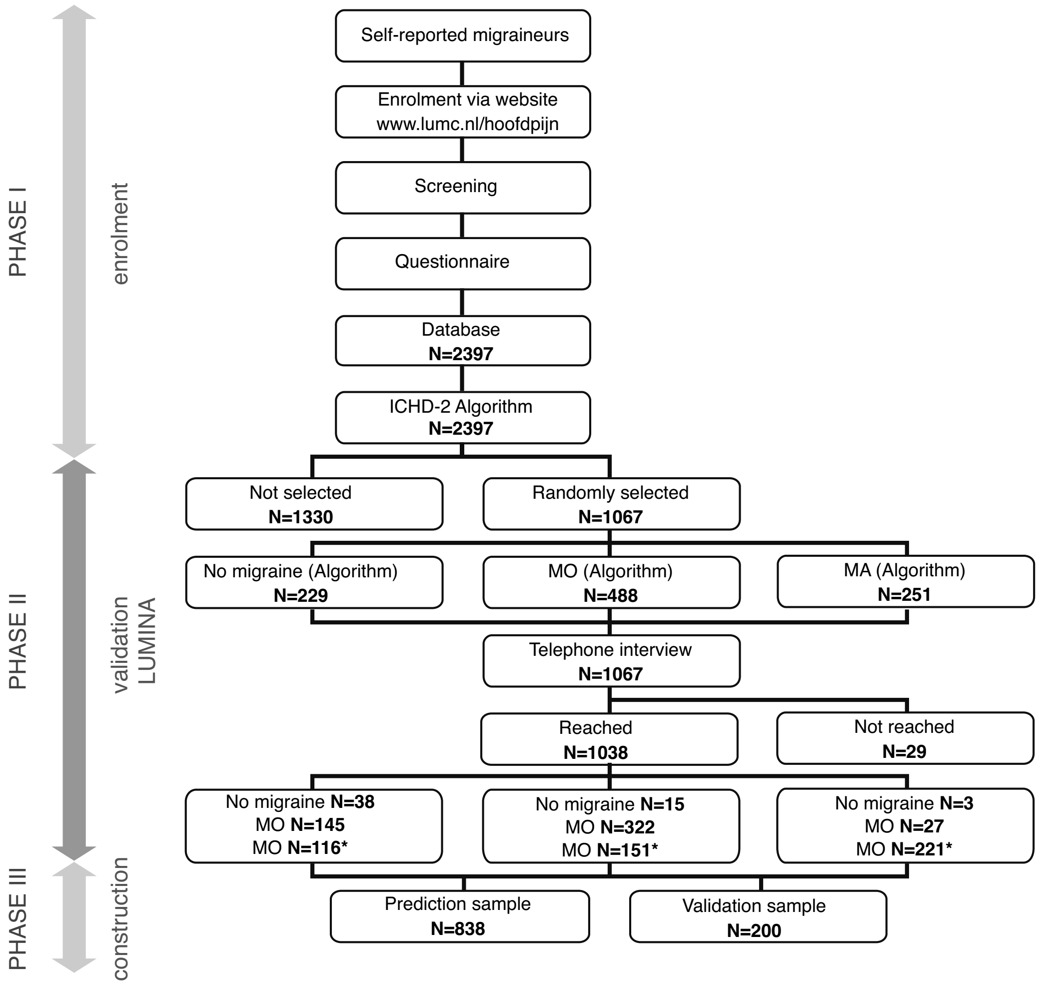

Figure 1 - Flowchart of (semi-)automated study flow. Screening = screening questionnaire; questionnaire = extended questionnaire; Alg.= ICHD-2 based algorithm diagnosis; Int.= interview diagnosis; MA = migraine with aura; MO = migraine without aura; *In the total MA group, 91.6% (447/488) reported visual aura symptoms.

Methods

Subjects

Participants were Dutch adults aged 18 to 74 years with migraine (MA and MO), who were informed via the lay press nationwide to enrol via the especially designed LUMINA website. Additionally, patients from our outpatient headache clinic were invited by a letter. In this clinic-based study, all participants were self-reporting migraineurs, of which approxiamtely 90% had previously been diagnosed with migraine by a physician.

Study flow

Study flow is depicted in Figure 1. Patients who visited the website were informed about the study and could enrol directly. The first step was to fulfil the screening criteria, using a simple screening questionnaire that was validated previously in the population-based GEM-study[3]. This screening questionnaire included five questions asking whether the patient i) had severe headaches in the past 12 months; ii) what the headache severity was; iii) had suffered from headaches which were preceded by visual disturbances; iv) had been diagnosed with migraine by a physician; and v) had ever used anti-migraine medication. After fulfilling these criteria, cases received a unique user ID-code via e-mail to log on to the study website, where they could participate in an extended, web-based questionnaire study. Having completed the extended questionnaire, a number of randomly selected participants were contacted by telephone by WPJvO, CMW, and AHS, who are experienced in diagnosing migraine. This semi-structured telephone interview detailed questions on headache and aura characteristics including ICHD-2 migraine and aura criteria[5] with special attention for visual, sensory, motor and speech aurasymptoms, was used as the gold standard. Median interview duration was 10-15 minutes, ranging up to 30 minutes if necessary. Afterwards, a final diagnosis was made: in case of ambiguity, a headache specialist (GMT) was consulted. Patients were excluded from the analysis if they could not be reached by telephone after five failed telephone contact attempts. The study was approved by the local medical ethics committee. All participants provided written informed consent.

Construction of questionnaire

The extended questionnaire (accessible via www.lumc.nl/hoofdpijn) was based on the ICHD-2 criteria[5] and incorporated 127 items on migraine headache and aura characteristics, premonitory symptoms, trigger factors, allodynia, and medication use and was presented to participants as a digital web-form. The questions were to be answered by choosing from categorical alternatives. On the web-form multicolour exemplary illustrations were shown with the most characteristic visual aura features (hemianopsia, scotoma, fortification spectra, visual blurring) and sensory aura features (anatomical distribution).

ICHD-2 based algorithm

After completion of the extended questionnaire, an algorithm based on ICHD-2 criteria[5] migraine criteria was run and individual diagnosis was determined. The algorithm had the following possible outcomes: 'no migraine'; 'migraine without aura'; and 'migraine with aura'. In the analysis, the algorithm outcomes were dichotomised into 'aura' and 'no aura' (Figure 2).

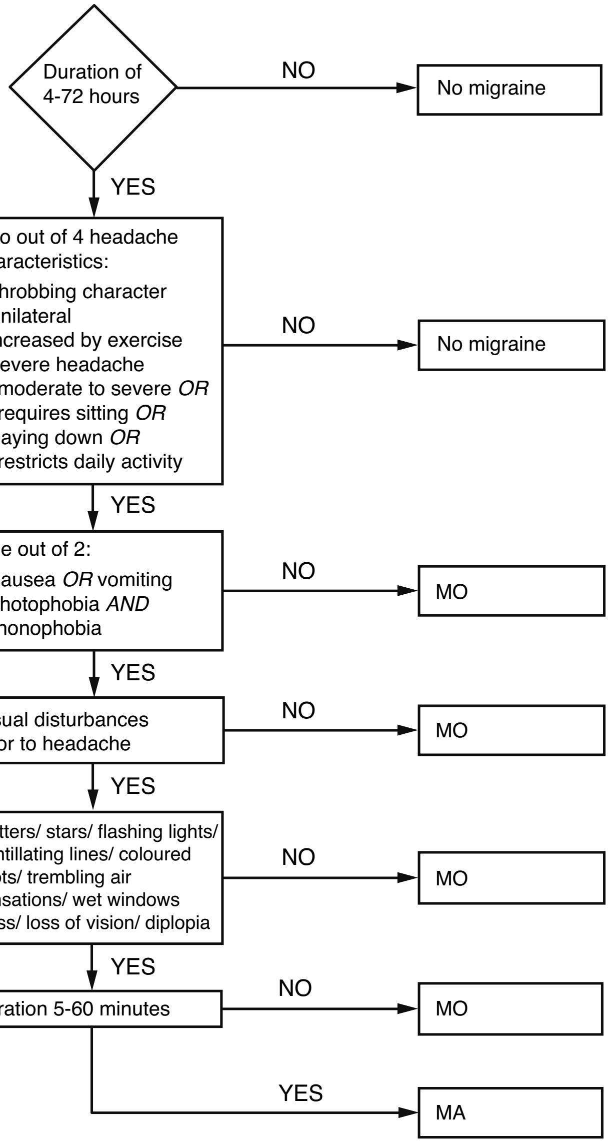

Figure 2 - Structure of ICHD-II based algorithm used in LUMINA study. MO = migraine without aura; MA = migraine with aura;

Statistical analysis

Descriptive statistics
Descriptive statistics were performed on demographic and clinical variables, on the algorithm based diagnoses and on the interview-derived diagnoses. Results are reported as mean ± SD or as percentage. Differences in between-groups means were analyzed with independent sample t-tests and ANOVAs. Proportions were compared using Chi-square tests. All items from the extended questionnaire that concerned ICHD-2 migraine criteria were evaluated separately. Likelihood ratios were calculated using standard formulas for positive likelihood ratio (LR+, sensitivity/[1 – specificity]) and negative likelihood ratio (LR-, [1 – sensitivity]/ specificity).

Questionnaire validation process
The questionnaire validation process was divided into two phases and was aimed at identifying a combination of items that were better predictors for diagnosing migraine aura than the ICHD-2 based algorithm, with the interview-derived diagnosis as the gold standard. In phase I, a sample of 838 self-reported migraineurs (approximately 80% of total group) was randomly selected and used as a training sample (see Figure 1) to derive a predictive model. These patients fulfilled set screening criteria from the five-item LUMINA screener before they could enter the extended questionnaire. Logistic regression (see below) was used to develop the predictive model that included questionnaire items most contributing to predict subcategories 'aura' and 'no aura'. Subsequently, we compared both the ICHD-2 based algorithm diagnoses and the diagnoses predicted by the logistic model, to the gold standard. In phase II, we validated this derived predictive model in an independent validation sample, consisting of 200 patients, approximately 20% of our sample (see Figure 1).

Phase I: Development of prediction rule
In phase I, a prediction rule for the aura subcategories 'aura' vs. 'no aura' was developed using a multivariate logistic regression analysis. Relevant, individual, dichotomized items (n=33) were selected from the extended questionnaire and were used as predictor variables for aura in the model. Selection of items was made by the authors (WPJvO; CMW; GMT) and was based on clinical relevance to migraine aura, and sensitivity, specificity, PPV, NPV en likelihood ratios of individual items. Inter-item correlation was assessed for relevant items using Spearman's rank coefficients and when items correlated with coefficients >0.9, one of these items was excluded from the analysis. A forward selection strategy using the likelihood ratio test was performed to identify items that were significant ($p<0.05$) predictors for the outcome of aura. For each subject in this sample (n=838), a prediction score was calculated using these items. Subsequently, a receiver operator

characteristics (ROC) curve was generated to assess the optimum cut off point for this prediction score. Using the method proposed by Halpern et al.[21], an optimum cut-off (highest sensitivity and specificity) was determined from the ROC curve. Therefore, the logistic model resulted in a selection of the 33 items with significant ($p<0.05$) contribution in the aura prediction.

Phase II: Validation of prediction rule

The derived predictive rule was subsequently validated in the second sample (validation sample; n=200; see Figure 1). Validity of this predictive model was assessed by checking whether the selected items contributed significantly ($p<0.05$) for the prediction in the second sample too. Subsequently, the sensitivity and specificity from the ROC optimum in the training sample were compared with these parameters in the validation sample, using the same cut-off value.

Overall outcome measures

Sensitivity, specificity, positive and negative predictive values were calculated to compare the fit of the three different models with the interview-derived aura diagnosis as the gold standard. These models were: 1) ICHD-2 based algorithm; 2) predictive model from phase I; and 3) validation of predictive rule in phase II.

All data analyses were performed using SPSS 16.0.2 (SPSS inc., IBM, USA). P values less than 0.05 were considered significant. When appropriate, categorical items were dichotomized into binary variables for the analysis in an attempt to simplify the instrument.

Table 1 - Baseline characteristics of total study population and separate study samples.

		Selection for study		Telephone interview		Sample	
Characteristic	**Total (n = 2397)**	**Not selected (n=1330)**	**Selected (n=1067)**	**Not reached (n=29)**	**Reached (n=1038)**	**Training (n=838)**	**Validation (n=200)**
Age, years	42.8±11.9	41.6±12.0	44.3±11.6	43.9±11.1	44.4±11.6	44.6±1.7	43.3±11.5
Female	84.8	83.9	85.8	89.7	85.6	85.0	88.5
Ever M diagnosis	88.9	87.8	90.2	100	89.9	90.2	89.0
Anti-M drug use	82.8	80.3	85.8	93.1	85.6	85.2	87.5
Algorithm diagnosis M	87.1	97.3*	71.4*	79.3	72.4	72.1	73.5

Values are mean ± SD or %. * $p<0.001$ (χ^2-test). M = migraine.

Table 2 - Sensitivity, specificity, and positive and negative predictive values as well as the corresponding likelihood ratios for diagnosis of migraine aura based on the IHCD-II-based algorithm (in both the total group and training sample) and the derived seven-item prediction model (in both the training sample and in the validation sample)

	ICHD-2 based algorithm		Model	
Characteristic	Total sample (n=1038)	Predictive sample (n=838)	Training sample (n=838)	Validation sample (n=200)
Sensitivity, %	45	44	83	86
Specificity, %	95	95	74	75
PPV MA, %	88	89	74	74
PPV MO (=NPV MA), %	70	64	83	86
Positive likelihood ratio	8.2	8.7	3.1	3.5
Negative likelihood ratio	0.6	0.6	0.2	0.2

MA = migraine with aura; MO = migraine without aura; NPV = negative predictive value; PPV = positive predictive value.

Receiver Operator Characteristics (ROC) curve

From the data in the training sample, we generated an ROC curve by plotting the sensitivity of the questionnaire against one minus the specificity. As a graphical representation of the trade-off between false negative and false positive rates for every possible cut-off point, the ROC curve reflects the trade-offs between sensitivity and specificity, and plots the false positive rate on the X axis and the true positive rate on the Y-axis. The area under the curve is a measure of correlation between the prediction of the questionnaire and the gold standard diagnosis. The closer the area under the curve (AUC) is to 1, the better the test is. To validate the derived logistic model, we compared the ROC from the prediction sample (n=838) to the ROC of the validation sample (n=200).

Results

General results

Over a 1-year period, from April 2008 until April 2009, 2,397 subjects fulfilled the set screening criteria and completed the extended questionnaire (Figure 1). During this time period, a total of 1,067 subjects (44.5%) were randomly selected for the semi-structured telephone interview, of which 1,038 (97.3%) were reached and could be used in the analysis. A total of 29 subjects (2.7%) were not included in the analysis because they could not be reached by telephone, after having tried at least five times. From these 1,038 subjects, 838 (79.4%) were randomly selected and used for the prediction model and the remaining sample of 200 subjects (18.9%) was used for validation (Figure 1).

Baseline characteristics of the total study population and separate prediction and validation samples are depicted in Table 1. Almost 90% of self-reported migraineurs had previously been diagnosed with migraine by a physician. Age, gender, prevalence of previous migraine diagnosis and use of anti-migraine medication did not differ significantly between selected subjects and non-selected subjects, nor between subjects that were reached compared to those that could not be reached for telephone interview (see Table 1). In the selected subjects (n=1,067; with special attention to patients which fulfilled ICHD-2 migraine criteria except for attack duration), the algorithm diagnosis of 'no-migraine' was more prevalent (28.6% [305/1,067] vs. 2.7% [36/1,330]; $p<0.001$) compared to non-selected subjects (n=1,330).

Screening questionnaire

In total, 94.6 percent of subjects (982/1,038) fulfilling the screening criteria, fulfilled ICHD-2 migraine criteria in the telephone interview. We considered everyone fulfilling the screening criteria to be migraineur. We used a logistic model to predict individual aura vs. no aura status.

Algorithm diagnosis

From the total sample of 1.038 subjects, the ICHD-2 based algorithm classified 488 subjects as MO patients, 251 as having MA, and 299 subjects as non-migraineurs (Figure 1). Of these, 243 were misclassified as non-migraineurs due to reporting of longer than actual attack duration. Table 2 summarizes the sensitivity, specificity, positive and negative predictive values as well as the corresponding likelihood ratios for the ICHD-2 based

Table 3 - Sensitivity, specificity, predictive values and likelihood ratios of individual questionnaire headache items vs. the interview diagnosis of migraine headache.

	Question									
	Yes		No							
Variable	**M.**	**No M.**	**M.**	**No M.**	**Sens.**	**Spec.**	**PPV**	**NPV**	**LR+**	**LR-**
Duration 4-72 hrs	721	19	249	49	0.74	0.72	0.97	0.16	2.64	0.36
Throbbing	670	232	40	96	0.94	0.29	0.74	0.71	1.32	0.21
Unilateral	863	57	46	72	0.95	0.56	0.94	0.61	2.16	0.89
Increase by activity	878	57	63	40	0.93	0.41	0.94	0.39	1.58	0.17
Severe	516	11	455	56	0.53	0.84	0.98	0.11	3.31	0.56
Nausea	867	63	36	72	0.96	0.53	0.93	0.67	2.04	0.08
Vomiting	627	87	64	260	0.91	0.75	0.88	0.80	3.64	0.12
Photophobia	859	91	25	63	0.97	0.41	0.90	0.72	1.64	0.07
Phonophobia	809	128	30	71	0.96	0.36	0.86	0.70	1.50	0.11

M = migraine; Sens. = sensitivity; Spec. = specificity; PPV = positive predictive value; NPV = negative predictive value; LR+ = positive likelihood ratio; LR- = negative likelihood ratio.

Table 4 - Sensitivity, specificity, predictive values and likelihood ratios of individual questionnaire aura items vs. the interview diagnosis of migraine aura.

	Question									
	Yes		No							
Variable	**M.**	**No M.**	**M.**	**No M.**	**Sens.**	**Spec.**	**PPV**	**NPV**	**LR+**	**LR-**
					Visual aura symptoms					
Suffer from visual disturbances?	436	235	42	278	0.91	0.54	0.65	0.87	1.98	0.17
Shitters	335	117	143	396	0.70	0.77	0.74	0.74	3.04	0.39
Stars	201	71	277	442	0.42	0.86	0.74	0.62	3.00	0.67
Flashes	178	42	300	471	0.37	0.92	0.81	0.61	4.63	0.68
Scintillating lines	223	25	255	488	0.47	0.95	0.90	0.66	9.40	0.56
Figures	111	29	367	484	0.23	0.94	0.79	0.57	3.83	0.82
Coloured spots	153	70	325	443	0.32	0.86	0.69	0.58	2.29	0.79
Trembling air sensations	488	412	25	66	0.14	0.95	0.73	0.54	2.80	0.91
Wet window glass	118	71	360	442	0.25	0.86	0.62	0.55	1.79	0.87
Loss of vision	283	62	195	451	0.59	0.88	0.82	0.70	4.92	0.47
Diplopia	146	72	332	441	0.31	0.86	0.67	0.57	2.21	0.80
Other specific visual disturbances	87	67	391	446	0.18	0.87	0.57	0.53	1.38	0.94
					Sensory aura symtpoms					
Sensory numbness/ tingling	114	268	13	623	0.90	0.70	0.30	0.98	3.00	0.14
Unilateral	111	236	16	655	0.87	0.73	0.32	0.98	3.22	0.18
5-60 min	49	50	78	841	0.39	0.94	0.50	0.92	6.50	0.65
Start before headache	94	154	33	737	0.74	0.83	0.38	0.96	4.35	0.31
					Motor aura symptoms					
Muscle weakness	20	203	6	802	0.77	0.80	0.09	0.99	3.85	0.29
Unilaterality	14	59	12	946	0.54	0.94	0.19	0.99	9.00	0.49
Duration 5-60 minutes	6	47	20	958	0.23	0.95	0.11	0.98	4.60	0.81
Starts prior to headache	14	128	12	877	0.54	0.87	0.10	0.99	4.15	0.53
Pinching	13	117	13	888	0.50	0.88	0.10	0.99	4.17	0.57
Arm lifting problem	10	62	16	943	0.39	0.94	0.14	0.98	6.50	0.65
Crippled walking	9	51	17	954	0.35	0.95	0.15	0.98	7.00	0.68
Facial asymmetry	8	26	18	979	0.31	0.97	0.24	0.98	10.33	0.71
					Speech disturbances					
Speech problems	132	366	8	489	0.94	0.57	0.27	0.98	2.19	0.11
Stiff mouth/ tongue	66	103	74	752	0.47	0.88	0.39	0.91	3.92	0.60
Wrong words	80	96	60	759	0.57	0.89	0.46	0.93	5.18	0.48
Expressive aphasia	119	311	21	544	0.85	0.64	0.28	0.96	2.36	0.23
Dysarthria	73	98	67	757	0.52	0.89	0.43	0.92	4.73	0.54
Prior to headache	102	154	38	701	0.73	0.82	0.40	0.95	4.06	0.33

Abbreviations as in Table 3.

Table 5 - Significantly correlated questions (n = 7) with regression coefficients, odds ratios and significance levels derived from the logistic regression model (training sample; n = 838).

	B	OR (95% CI)	p-value
Did you have visual disturbances before headache in the past 12 months?	0.729	2.07 (132-3.26)	0.002
Did the visual disturbances last 5-60 minutes?	1.658	5.25 (3.08-8.96)	<0.001
Have you had scintillating lines before or during your headache in the past 12 months?	1.210	3.35(2.06-5.45)	<0.001
Have you had loss of vision before or during your headache in the past 12 months?	0.913	2.49(1.63-3.80)	<0.001
Did you suffer from numbness or a tingling feeling in your face/ unilateral arm/ leg that started prior to headache in the past 12 months?	0.631	1.88(1.07-3.29)	0.027
Did you use nonsense words prior or during your headache in the past 12 months?	0.680	1.97(1.22-3.19)	0.005
Did you use a triptan in the past 12 months?	-0.561	0.57 (0.39-0.83)	0.003

B = regression coefficient; CI = confidence interval; OR = odds ratio..

algorithm diagnosis of migraine aura in the total sample (n=1,038). Similar values for this classification in the training sample (n=838) suggest this sample is a good representation of the whole group. In both the total group and the training sample, sensitivity for aura was approximately 0.45, specificity 0.95, positive predictive value (PPV) 0.88 and negative predictive value (NPV) 0.70 (Table 2). Additionally, we calculated characteristics of all individual questionnaire items that reflect migraine headache and migraine aura criteria and summarized those in Table 3 and 4. The results show individual sensitivity ranging up to 0.97 (photophobia; nausea) and PPV up to 0.98 (headache severity; headache duration).

Phase I: Derivation of predictive model

Using logistic regression, 7 questions (from the 33 included; none showed Spearman rank correlation >0.9) showed a significant impact on the likelihood of having a migraine aura in accordance to the gold standard derived from the telephone interview. These questions are summarized in Table 5, which also shows significance levels and regression coefficients derived from the logistic model. The questions show partial overlap with the questions used in the ICHD-2 based algorithm. This model explained between 35.4% (Cox and Snell R Square) and 47.3% (Nagelkerke adjusted R Squared) of variance, and correctly classified 651/838 (77.8%) of subjects.

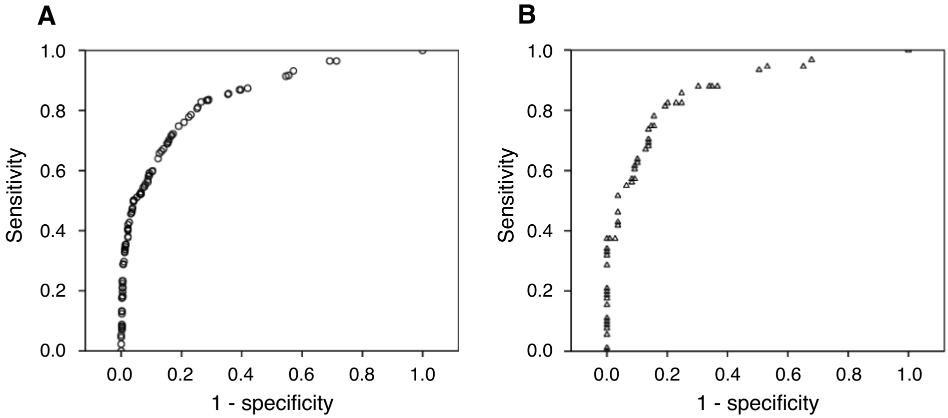

Figure 3 - Receiver operator characteristics (ROC) curves for (*A*) the derived prediction rule in the initial training sample (n=838) and (*B*) in the validation sample (n=200). The area under the ROC-curve (C-statistic) for the prediction rule was 0.85 (95% CI 0.83-0.88) in the training sample and 0.87 (95% CI 0.82-0.92) in the validation sample.

ROC curve

From the data in the predictive cohort, we generated an ROC curve by plotting the sensitivity of the questionnaire against one minus the specificity (Figure 3A). This analysis resulted in an optimal cut off point for the used logistic model at 0.35 with AUC of 0.85 (95% C.I. 0.83-0.88), yielding a 7 item questionnaire with a sensitivity of 0.83 and a specificity of 0.74. Compared to the ICHD-2 based algorithm outcome, this approach therefore resulted in a vast increment in sensitivity, with only small decrement of specificity (Table 2).

Phase II: Validation of derived prediction rule

Using the predictive model and cut-off point (0.35) derived from the training sample (n=838), we validated this model in a second, independent sample (n=200) of subjects who also fulfilled the set screening criteria. This analysis showed the model to have approximately similar sensitivity and specificity in this validation sample (Table 2). In the validation cohort, the ROC curve yielded an AUC of 0.87 (95% C.I. 0.82-0.92), which is comparable to the output from the training cohort (Figure 3B). When using this cut off from the training cohort, migraine aura diagnosis was predicted correctly in 160/200 (80.0%) of subjects.

Test-retest reliability

For a random selection of 44 patients who completed the extended questionnaire a second time, with a mean test-retest interval of 155 days (median 89 days, range 1-422 days), test-retest reliability was found to be good with a test-retest kappa for algorithm diagnostic group of 0.59 (95% CI 0.38-0.80). Test-retest interval did not influence agreement (linear regression, p=0.852).

Discussion

Our study has been the first one to validate a web-based questionnaire for purposes of diagnosing aura cases using a large sample of self-reported migraineurs. Few previous studies on migraine screeners and questionnaires have focussed on migraine aura, and the numbers of MA cases used to validate the questionnaire instruments in these studies were limited to n=8-186[17,19,22-24] respectively, in comparison to the large number of 488 aura cases in our study. Physicians frequently rely on aura as a cardinal symptom of migraine, as suggested by the 1.9 fold higher rate of medical diagnosis in interview settings when comparing MA cases to cases of MO[25]. Our study shows that, in self-reported migraineurs, a distinction between MA and MO can be made via a self-administered web-based questionnaire, with a focus on visual aura symptoms. The difficulty in diagnosing other aura types might be explained by the lack of perceptions and recognition of verbal and other non-visual auras by patients[26]. For diagnosing patients with these specific aura symptoms a clinical interview is needed. However, since the vast majority of the self-reported aura cases suffer from visual auras and only a small minority suffers from non-visual auras[27], we believe this number is neglectable when recruiting aura cases from a population of self-reported migraineurs. Perhaps the most helpful item identifying aura cases is the duration of the aura phenomena, since this question enables to distinguish visual aura symptoms from non-specific visual disturbances. Additionally, our data show aura patients are less likely to use triptans for rescue medication, which might be an indicator of lower headache severity.

We show that the question addressing the duration of the headache may hamper correct identification of migraine cases in a web-based questionnaire setting because some migraineurs overestimate the duration of an attack. Conversely, a question addressing headache severity should be included because this is helpful in distinguishing aura cases with migraineous headache from patients with non-specific headache.

The strength of our study includes the large samples of both the training (n=838) and validation sample (n=200), which are representative for the population studied. Both outclinic patients and other patients (most of whom are treated by their own GP or neurologist elsewhere) were included via the same web-based flow. We found no clinical or demographic differences between these populations that could have affected the predictive model. Secondly, the use of a telephone interview as a gold standard by well-trained physicians with consultation of a headache specialist assured precise categorisation of migraineurs. Although we did not had a face-to-face interview as gold standard, we feel that our thorough semi-structured telephone interview safeguarded a very reliable migraine and aura diagnosis. Thirdly, the use of a validated screening instrument prior to our new questionnaire resulted in a group of self-reported migraineurs in which 95% could in fact be diagnosed with migraine. Fourth, we used a web-based questionnaire that was easy to fill out and send in for participants. With this approach, we successfully recruited large samples of migraineurs and contributed to the identification of the first genetic risk factor for the common forms of migraine[13]. We included a selected population of self-reported migraineurs, that had already been diagnosed with migraine by a physician, or otherwise thought they suffered from migraine, in which our questionnaire shows a high reliability in diagnosing aura. Our study did not aim to validate the questionnaire as a screening instrument for migraine in a naïve, general population.

The World Wide Web as a tool for recruiting patients and conducting research has several advantages. First, a large and diverse subject population can be reached at low cost[16]. Secondly, internet research imposes fewer burdens on participants, compared to non-internet research[15]. Thirdly, available software permits data entry and analysis in a secure Web database. Fourth, investigators may be able to increase patient awareness and participation on clinical research. However, there might be certain challenges too[28]. Internet users tend to be younger and better educated than the patient population as a whole; visually impaired and minority groups may be underrepresented; and the symptoms expressed by participants may be more severe then is typical. We feel, however, these potential biases haven't pivotally influenced our data. Additionally, the so-called 'virtual Munchhausen syndrome', i.e. individuals referring themselves for studies for which they are not truly eligible, may compromise the validity of results[29]. In our study, we have no evidence that data have been influenced by subjects masquerading electronically as patients. This is in accordance with previous migraine research[15]. Even with such biases, altogether, the internet represents an appropriate aid to conduct research aimed at collecting clinical headache data from large numbers of patients.

We conclude that our web-based recruitment system in combination with an automated study flow is a very successful instrument to truly distinguish MA and MO in self-reported migraine patients. We propose to use our identified seven questions that have a higher accuracy in identifying aura cases from a population of self-reported migraineurs than an ICHD-2 based algorithm.

Acknowledgements

Source of funding

This work was supported by grants of the Netherlands Organization for Scientific Research (NWO) (903-52-291, M.D.F.; Vici 918.56.602, M.D.F; 907-00-217 GMT; Vidi 917-11-319 GMT), and by a grant from the Centre for Medical Systems Biology (CMSB) established by the Netherlands Genomic Initiative/ Netherlands Organisation for Scientific Research (NGI/NWO).

Reference list

1. Stewart WF, Shechter A, Rasmussen BK. Migraine prevalence. A review of population-based studies. *Neurology*. 1994 Jun;44:S17-S23.
2. Scher AI, Stewart WF, Liberman J, Lipton RB. Prevalence of frequent headache in a population sample. *Headache*. 1998 Jul;38:497-506.
3. Launer LJ, Terwindt GM, Ferrari MD. The prevalence and characteristics of migraine in a population-based cohort - The GEM Study. *Neurology*. 1999 Aug 11;53:537-542.
4. Stovner LJ, Zwart JA, Hagen K, Terwindt GM, Pascual J. Epidemiology of headache in Europe. *Eur J Neurol*. 2006 Apr;13:333-345.
5. Headache Classification Subcommittee of the International Headache Society. The International Classification of Headache Disorders: 2nd edition. *Cephalalgia*. 2004;24 Suppl 1:9-160.
6. Russell MB, Rasmussen BK, Thorvaldsen P, Olesen J. Prevalence and Sex-Ratio of the Subtypes of Migraine. *International Journal of Epidemiology*. 1995 Jun;24:612-618.
7. Ulrich V, Gervil M, Kyvik KO, Olesen J, Russell MB. Evidence of a genetic factor in migraine with aura: a population-based Danish twin study. *Ann Neurol*. 1999 Feb;45:242-246.
8. Ulrich V, Gervil M, Kyvik KO, Olesen J, Russell MB. The inheritance of migraine with aura estimated by means of structural equation modelling. *J Med Genet*. 1999 Mar;36:225-227.
9. Gervil M, Ulrich V, Kyvik KO, Olesen J, Russell MB. Migraine without aura: a population-based twin study. *Ann Neurol*. 1999 Oct;46:606-611.
10. Ulrich V, Gervil M, Fenger K, Olesen J, Russell MB. The prevalence and characteristics of migraine in twins from the general population. *Headache*. 1999 Mar;39:173-180.
11. Gervil M, Ulrich V, Kaprio J, Olesen J, Russell MB. The relative role of genetic and environmental factors in migraine without aura. *Neurology*. 1999 Sep 22;53:995-999.
12. Stewart WF, Staffa J, Lipton RB, Ottman R. Familial risk of migraine: a population-based study. *Ann Neurol*. 1997 Feb;41:166-172.
13. Anttila V, Stefansson H, Kallela M, et al. Genome-wide association study of migraine implicates a common susceptibility variant on 8q22.1. *Nat Genet*. 2010 Aug 29.
14. de Groen PC, Barry JA, Schaller WJ. Applying World Wide Web technology to the study of patients with rare diseases. *Ann Intern Med*.1998 Jul 15;129:107-113.

15. Lenert LA, Looman T, Agoncillo T, Nguyen M, Sturley A, Jackson CM. Potential validity of conducting research on headache in internet populations. *Headache*. 2002 Mar;42:200-203.
16. Strom L, Pettersson R, Andersson G. A controlled trial of self-help treatment of recurrent headache conducted via the Internet. *J Consult Clin Psychol*. 2000 Aug;68:722-727.
17. Hagen K, Zwart JA, Vatten L, Stovner LJ, Bovim G. Head-HUNT: validity and reliability of a headache questionnaire in a large population-based study in Norway. *Cephalalgia*. 2000 May;20:244-251.
18. Cady RK, Borchert LD, Spalding W, Hart CC, Sheftell FD. Simple and efficient recognition of migraine with 3-question headache screen. Headache. 2004 Apr;44:323-327.
19. Kirchmann M, Seven E, Bjornsson A, et al. Validation of the deCODE Migraine Questionnaire (DMQ3) for use in genetic studies. *Eur J Neurol*. 2006 Nov;13:1239-1244.
20. Hagen K, Stovner LJ, Zwart JA. Potentials and pitfalls in analytical headache epidemiological studies--lessons to be learned from the Head-HUNT study. *Cephalalgia*. 2007 May;27:403-413.
21. 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 Mar;3:245-253.
22. Valentinis L, Valent F, Mucchiut M, Barbone F, Bergonzi P, Zanchin G. Migraine in adolescents: validation of a screening questionnaire. *Headache*. 2009 Feb;49:202-211.
23. Kallela M, Wessman M, Farkkila M. Validation of a migraine-specific questionnaire for use in family studies. *Eur J Neurol*. 2001 Jan;8:61-66.
24. Gervil M, Ulrich V, Olesen J, Russell MB. Screening for migraine in the general population: validation of a simple questionnaire. *Cephalalgia*. 1998 Jul;18:342-348.
25. Leone M, Filippini G, D'Amico D, Farinotti M, Bussone G. Assessment of International Headache Society diagnostic criteria: a reliability study. *Cephalalgia*. 1994 Aug;14:280-284.
26. Facheris MF, Vogl FD, Hollmann S, et al. Adapted Finnish Migraine-Specific Questionnaire for family studies (FMSQ(FS)): a validation study in two languages. *Eur J Neurol*. 2008 Oct;15:1071-1074.
27. Rasmussen BK, Olesen J. Migraine with Aura and Migraine Without Aura - An Epidemiologic-Study. *Cephalalgia*. 1992 Aug;12:221-228.
28. Rothman KJ, Cann CI, Walker AM. Epidemiology and the internet. *Epidemiology*. 1997 Mar;8:123-125.
29. Soetikno RM, Mrad R, Pao V, Lenert LA. Quality-of-life research on the Internet: feasibility and potential biases in patients with ulcerative colitis. *J Am Med Inform Assoc*. 1997 Nov;4:426-435.