Development and validation of an Arabic-language headache questionnaire for population-based surveys

Naglaa A El-Sherbiny, Hatem Samir Shehata, Hanan Amer, Alaa Elmazny, Mohamed Masoud, Hanan Helmy, Nevin M Shalaby

Faculty of Medicine, Fayoum University, Faiyum, Egypt; Faculty of Medicine, Cairo University, Cairo, Egypt

Background: The reported prevalence of headache disorders in Arab regions varies considerably between countries. This may be due to a lack of standardized survey instruments that capture the prevalence.

Purpose of the study: Our goal was to construct and validate a structured headache questionnaire for Arabic-speaking headache patients to be used as an epidemiological survey instrument.

Methods: We developed a culturally adapted interviewer-administered questionnaire in Arabic language comprising two sets of questions. The first set included personal and sociodemographic data together with a screening question regarding the presence of headaches over the last year. The second set was designed to define the type and pattern of headaches according to the International Classification of Headache Disorders criteria (for subjects with “yes” answers on the screening question). Validation process took place in two phases through probability random sampling selected from 1,221 headache subjects collected in an epidemiological survey 3 (n=70) and 6 months (n=232) later. A detailed assessment of patients’ headaches was performed by neurologists (blinded from the questionnaire diagnosis) who clinically assessed the patients’ headache.

Results: The validity of the questionnaire was tested in 232 subjects with a mean age of 41.2±10.9 years, 72.8% of whom were females. The mean time to complete the questionnaire was 8.4±1.7 minutes. The intraclass correlation coefficient was 0.903 (95% confidence interval: 0.875–0.925), the Cronbach κ coefficient was 0.775 (95% confidence interval: 0.682–0.837), and the percentage of agreement was 84.5%.

Conclusion: Our results support the use of this comprehensive questionnaire as a valid tool for headache assessment among Arabic-speaking patients.

Keywords: Arabic, Egypt, headache, questionnaire, validation

Introduction

Headache disorders are rated among the ten most disabling conditions worldwide.1 Headaches impose a recognizable social and economic burden on individual sufferers and society, as they are one of the prime causes of reduced productivity in working adult patients.2 However, headache is still underdiagnosed and undertreated in developing countries.3 In Arab countries, and particularly Egypt, the most populous country in the Middle East, there are several challenges that affect proper headache management, including insufficient patients’ education, underestimating the significance of “headaches” by family members, and the availability of analgesics as over-the-counter self-medications without access to prescription drugs.4,5

Other principal barriers responsible for poor headache care include: the paucity of large-scale epidemiological studies in Arab countries, limited funding, large and often...
rural (less accessible) populations, and the low profile of headache disorders compared with communicable diseases; all these prevent the systematic collection of information.5,6 Furthermore, the results of the available cohorts are heterogeneous, with divergent prevalence rates that varied from 8% in Saudi Arabia7 to as high as 83.6% in Oman.8 These conflicting results are likely multifactorial, which are not only related to the differences in methodologies and tools employed for headache assessment in different studies but also are due to defects in good quality data9 and lack of standardized headache survey instruments for Arabic-speaking patients, which results in striking differences even with similar methodological designs.6

All this indicates a clear need for a culturally appropriate headache assessment tool based on the International Headache Society criteria to allow for comparable population-based studies of the prevalence and burden of headache disorders in Arab countries. However, in the process of developing a culturally neutral questionnaire, literal translation, word meanings and methods of administration should be linked to the diversity of societal norms and local realities among Arab regions.

Our goal was to develop a validated, interviewer-administered structured headache questionnaire in Arabic to be used for epidemiological surveys to assess the prevalence of primary headaches in Egypt.

Methods

Questionnaire development

An interviewer-administered structured questionnaire was developed after a review of the literature and was prepared in English by the authors HSS and NMS. It was then translated into Arabic and then back-translated to verify that the original meaning had not been lost; this was done according to a standardized protocol (Figure 1) adopted from Shaik et al.10 Translation and back-translation was carried out by two separate expert staff members from the English Department at Cairo University; they were fluent in both English and Arabic and were “blinded” to the source language version. A special panel (HSS, NMS, AE, and NAE) reviewed the forward- and back-translated versions to produce the final draft. The final survey document had a clear and concise title and an introduction explaining the purpose of the questionnaire together with clear instructions on how to complete it. It was organized into two sections; the first section included relevant personal and sociodemographic data (age, gender, education, marital status, occupation, and location of residence); history of medication use (oral contraceptive use, analgesic medications, and other drugs); and other related medical history (smoking, hypertension, diabetes, and other relevant clinical disorders).

The assessment of medical history began as recommended by previous studies,11 with a screening question about headaches in the last year (1-year prevalence), “Have you had headache during the last year not related to flu, cold, or head injury?” The second part of the questionnaire was completed by subjects who responded “yes” in the screening questionnaire. It included questions designed to define the nature of the headaches and to assess patterns of the headaches according to the International Classification of Headache Disorders (ICHD-IIIb).12

Types of questions

Four types of questions were used: binary questions (eg, yes/no or male/female); specific questions that do not specify options (eg, How old are you?); multiple choice questions (in which the options were mutually exclusive and covered all possible answers); and scaling questions (eg, on a visual analog scale).13 Examples of questions types are presented in the questionnaire sample (Supplementary material).
Wording of the questions
Simple and precise language was used, and abbreviations and undefined terms were avoided. For the newly developed questions, all efforts were made to keep questions direct, brief, and clear. The questions were assessed using the Flesch reading ease score, which identifies the number of words in sentences and the number of syllables in words, and then develops a readability formula (using the following link): http://www.readabilityformulas.com/freetests/six-readability-formulas.php. The Flesch reading ease score for the final form of the questionnaire was 73.2 (indicating that it was fairly easy-to-read).14

Training of interviewers
The initial interviewer training was done as a part of an epidemiological study5 in Fayoum Governorate, a developing city in Middle Egypt.15 This was followed by three consecutive sessions for recruitment and training of the physician interviewers. They received training sessions (from two neurologists, HSS and NMS) that consisted of 1) background information about headache diagnosis, types, and the art of history-taking, 2) skill-based training to reduce the errors related to subjectivity, 3) tips on introducing the research topic to the patients, and 4) methods to apply the questionnaire to capture all data and receive frank answers. The last part consisted of watching five videos of real headache patient interviews and interactive role play. Finally, the physician interviewers were instructed about maintaining confidentiality, presenting the questions clearly and exactly as written in a respectful manner, and remaining neutral as the answers were given.

Pilot surveys
A pilot study of the preliminary version of the questionnaire was carried out with 30 adult subjects in both rural and urban areas: 60% were female; mean age was 43.6 years (range 18–78). This was done to verify adequate simplicity, precision, acceptance, and appropriateness for the participants’ educational level and to minimize leading or confusing questions. Final modifications were done based on the response of such pilot survey.

Validation of the questionnaire
The final version of the questionnaire was constructed in two phases. During Phase I, two of the authors (neurologists HSS and NMS) performed three field visits to a randomly selected subsample of 70 subjects (from a total of 1,221) to evaluate the diagnoses of their headaches. Randomization was done using a computer-designed method by one of the interviewers (NME). The investigators were blinded to the subjects’ questionnaire responses, and they used ICHD-IIIb criteria12 to make their diagnoses. This validity testing was done within 3 months of the completion of the initial questionnaire, and all 70 subjects participated. Phase II was carried out 6 months after the finish of the epidemiological study and was done in a random sample of 250 of the 1,221 subjects who had a detailed assessment of their headaches by the same two neurologists who, again, were blinded to the questionnaire data. Eighteen headache subjects did not appear for the interview (a dropout rate of 7.2%) and were, therefore, excluded from the current analysis.

Ethical considerations
This study was designed according to the recommendations of the International Headache Society and was approved by the Review Board of the Neurology Department, Cairo University. Informed consent was obtained from all the subjects. For illiterate subjects, the interviewers read the consent form about the objectives of the study and the confidentiality of the information so as to obtain verbal consent.

Statistical analyses
Data were collected, coded, and analyzed using the Statistical Package for Social Sciences (SPSS) version 24.0 (IBM Corporation, Armonk, NY, USA). Continuous data were expressed as mean and SD, and categorical data were expressed as frequencies and percentages.

Test–retest reliability and construct validity
The test–retest reliability was quantified using the intraclass correlation coefficient (ICC) of absolute agreement, and Cronbach’s $\alpha$ coefficient was calculated to assess internal consistency.16 ICCs were classified as follows: “excellent” (≥0.81), “good” (0.61–0.80), “moderate” (0.41–0.60), and “poor” (≤0.40).17 The percentage agreement was calculated with criteria established as “excellent” (90%–100%), “good” (75%–89%), “moderate” (60%–74%), or “poor” (<60%).18

Results
Participant characteristics
Of the 250 subjects were invited to participate, 18 (7.2%) refused, and the validity of the questionnaire was, therefore, assessed in 232 subjects. The mean age was 41.2±10.9 years, and the sample included 169 (72.8%) females. None of the participants had any difficulties understanding or answering
any part of the questionnaire. The subjects’ characteristics are shown in Table 1.

Acceptability and word precision
None of our subjects reported difficulty in understanding the questionnaire at the time of completion. The mean time to complete the questionnaire was 8.4±1.7 minutes.

Strength of agreement
The questionnaire diagnosed migraines in 82 subjects, chronic migraine in 9 subjects, episodic tension-type headache (ETTH) in 105 subjects, chronic tension-type headache (CTTH) in 13 subjects, and cluster headache in 8 subjects. The neurologists were able to diagnose symptomatic headaches in 7 subjects (3%): 2 had errors of refraction, 2 had chronic sinusitis, 1 had marked anemia (hemoglobin ~5.7 g%), and 2 had myofascial pain and temporomandibular joint dysfunction. The difference between the interviewers’ diagnoses and neurologists’ diagnoses is shown in Table 2.

Validity and reliability
The results of blind reassessment of the randomly selected 232 headache participants are shown in Table 3; the ICC was 0.903 (95% CI: 0.875–0.925) “excellent”; the overall Cronbach’s $\kappa$ coefficient was 0.775 (95% CI: 0.682–0.837); and the percentage of agreement was 84.5%.

Discussion
This study presented an interviewer-administered structured questionnaire in Arabic in an epidemiological survey to assess the prevalence of primary headaches in a sample of 232 Egyptians with headaches; 169 (72.8%) of them were females, which reflects the fact that primary headaches are more common than in women, as was recorded by several authorities.3,7–9 The study sample was randomly selected from a community-based, cross-sectional survey designed to elucidate the prevalence of primary headache disorders in both urban and rural locations in Egypt.5

The interviewer administration of our questionnaire offered additional control over the quality of the measurement, and it is suitable to be applied to illiterate population, as 25.9% of general population in Egypt is illiterate.15

An ample interviewer training course was carried out to guarantee an effective interviewing strategy, and this helped us in achieving a higher response rate in the previous epidemiological study; it also contributed to a good concordance between the neurologists’ diagnoses and the developed headache questionnaire in the current study.5

We found that the ICC was 0.903 (95% CI: 0.875–0.925) ie, “excellent,” and the overall Cronbach’s $\kappa$ coefficient was 0.775 (95% CI: 0.682–0.837), and percentage of agreement was 84.5%. Traditionally, clinical diagnosis is known to be more flexible compared to structured criteria used for case identification for research protocols. In clinical practice, some patients do not necessarily “fit” into any particular diagnostic category, and the majority may have a combination of symptoms that need management regardless of their formal clinical labeling.19 The results of validity testing of the questionnaire showed that it has high sensitivity compared to clinical diagnoses, with the highest sensitivity for episodic tension headaches (0.93) and the lowest for chronic migraine (0.71), with intermediate values for episodic migraine (0.86) and chronic tension-type headaches (0.79).

Our sensitivity results for episodic migraine are comparable to a study that presented a computer-assisted telephone interviewing questionnaires for migraine, where a sensitivity of 81.5% and specificity of 91.3%, with 91.7%
positive predictive value, were found. In another Spanish cross-sectional, multicenter study that included 9,670 patients to validate Migraine Screen-Questionnaire, the achieved sensitivity was (0.82) and the specificity was (0.97). Nevertheless, both studies focused on migraine and did not aim to identify types of headaches or differentiate migraine subtypes (“episodic” and “chronic”). Sensitivity and specificity are usually higher when an instrument addressed migraine only. Lipton et al presented a 3-item screening questionnaire for migraine and achieved a sensitivity of (0.81) and a specificity of (0.75), with a κ coefficient of 0.68, whereas, in a more detailed migraine-specific questionnaire, Kallela et al obtained a sensitivity of (0.99) and a specificity of (0.96).

The relatively low sensitivity for detection of chronic migraine by the current questionnaire demonstrated the difficulties in detecting this condition by a simple screening tool. Chronic migraine remains underdiagnosed and undertreated worldwide; in the Chronic Migraine Epidemiology and Outcomes study, only 13.6% of the subjects consulted a specialist and even among those who were evaluated by an expert clinician, only 36% reported receiving a diagnosis of chronic migraine. Lipton et al developed and validated a self-administered case-finding tool called “Identify Chronic Migraine (ID-CM),” which entailed an extensive, four-stage process with strong psychometric properties to help clinicians identify chronic migraine. The final 12-item ID-CM achieved a sensitivity of 80.6%, a specificity of 88.6%, a negative predictive value of 75.0%, and a positive predictive value of 91.5%. However, the sample used for this study was highly selected. Therefore, the results may not be fully generalizable to other settings.

In the current questionnaire, the sensitivity for ETTH was 0.93, but for CTTH it was 0.79. Several studies were conducted to develop and validate diagnostic questionnaires that could identify different headache types including TTH. In a Pakistani study, researchers validated an interviewer-administered headache questionnaire developed by Lifting The Burden that had an overall κ value of 0.77 but was relatively insensitive for TTH (0.6). Ayzenberg et al tested the same questionnaire and reported the TTH as 64%. Relatively few other questionnaires for the diagnosis of TTH have been developed, but Rasmussen et al reported a sensitivity of only 43% (albeit with 96% specificity). The investigators all agree that the difficulties arise from the lack of specific characteristics of TTH, making it difficult to design questions to easily diagnose it.

In our study, the sensitivity of chronic headaches was less than that of episodic types. It was reported that questionnaires cannot distinguish reliably between headache disorders characterized by headache on more days than not (eg, chronic migraine and CTTH), but can identify presumptive medication overuse headache from the reported frequency and type of medication taken for the headache.

The main strengths of this study are the fact that it is the first study to validate an interviewer-administered headache questionnaire in Arabic language among Egyptians in spite of the burden of the disease in our country’s population. Hence, the questionnaire can be used in population-based surveys. Another strength is the concern with literal translations, word precision, meanings and how they are linked to the local realities and culture. The appropriateness of the wording of the questions was assessed by the Flesch reading ease score, which assesses the number of words in sentences and the number syllables in words. Another strength of this study is the nonsignificant differences between the diagnoses provided by nonspecialist physician interviewers and those made by the neurologists.

Possible limitations of our study protocol included presenting headache as a unidimensional disorder, which may explain the low sensitivity for detecting chronic migraine. Although unidimensional data are more understandable, easier to explain, and can maximize Cronbach’s ω, yet, it can present an oversimplified paradigm for headache diagnosis and dismiss other contributors such as cervicogenic elements, temporomandibular joint disorders, or other neuralgias. In addition to that, an intense skill-based training program for the

### Table 3

<table>
<thead>
<tr>
<th>Type of headache</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Positive predictive value (95% CI)</th>
<th>Negative predictive value (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Episodic migraine</td>
<td>0.86 (0.78–0.97)</td>
<td>0.94 (0.86–0.98)</td>
<td>0.86 (0.69–0.98)</td>
<td>0.92 (0.87–0.99)</td>
</tr>
<tr>
<td>Chronic migraine</td>
<td>0.71 (0.52–0.99)</td>
<td>0.98 (0.82–1.00)</td>
<td>0.78 (0.63–0.96)</td>
<td>0.96 (0.82–1.00)</td>
</tr>
<tr>
<td>ETTH</td>
<td>0.93 (0.79–0.98)</td>
<td>0.93 (0.86–1.00)</td>
<td>0.95 (0.85–1.00)</td>
<td>0.92 (0.82–0.99)</td>
</tr>
<tr>
<td>CTTH</td>
<td>0.79 (0.64–1.00)</td>
<td>0.96 (0.89–0.97)</td>
<td>0.85 (0.69–0.99)</td>
<td>0.98 (0.94–1.00)</td>
</tr>
<tr>
<td>Unclassified</td>
<td>0.66 (0.12–0.81)</td>
<td>0.97 (0.92–1.00)</td>
<td>0.67 (0.31–0.89)</td>
<td>0.97 (0.91–1.00)</td>
</tr>
</tbody>
</table>

Abbreviations: CTTH, chronic tension-type headache; ETTH, episodic tension-type headache.
interviewers is needed, including both theoretical and practical education in addition to interactive role playing models, which is a resource-dependent and time-consuming task.

Conclusion
Our study produced the first Arabic questionnaire for evaluation of headache symptoms to be used in population-based surveys in the Arab regions using standard methods for validity and reliability.

Disclosure
The authors report no conflicts of interest in this work.
Supplementary material
A sample of the Arabic-language headache questionnaire for population-based surveys in Egypt

عينة من استبيان لقياس مدى انتشار الصداع النصفي في جمهورية مصر العربية

Interviewer name and signature (to be completed by the interviewer)

 Participant identification: Initials, Number of ID (to be completed by the interviewer)

First part: personal, sociodemographic data

How old are you? ........................ year

What is your gender?  Male □ Female □

Screening question: Have you had headache during the last year not related to flu, cold, or head injury?

Yes □ No □

Second part (completed by subjects who responded “yes” in the screening questionnaire)

Do you have the following symptoms during headache?

□ Nausea □ Vomiting □ Phonicophobia □ Photophobia □ Conjunctival injection □ Lacrimation □ Eyelid edema □ Ptosis □ Facial sweating □ Rhinorrhoea

Scaling question: Visual Analog Scale for Headache severity

Notes: We welcome the reader to inquire with the corresponding author for more details through the following e-mail (samirhatem@hotmail.com).