Original Article

Cephalalgia
An International Journal of Headache

International Headache Society

Cephalalgia
0(0) 1–12


Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/03331024211018138
journals.sagepub.com/home/cep
SAGE

A brief diagnostic screen for cluster headache: Creation and initial validation of the Erwin Test for Cluster Headache

Randika Parakramaweera[1], Randolph W Evans[2], Larry I Schor[3], Stuart M Pearson[3], Rebecca Martinez[1], Jacob S Cammarata[1], Amisha J Amin[1], Seung-Hee Yoo[4], Wei Zhang[5], Yuanqing Yan[1] and Mark J Burish[1]

Abstract

Objective: To use 1) newly generated data, 2) existing evidence, and 3) expert opinion to create and validate a new cluster headache screening tool.

Methods: In phase 1 of the study, we performed a prospective study of an English translation of an Italian screen on 95 participants (45 with cluster headache, 17 with other trigeminal autonomic cephalalgias, 30 with migraine, and 3 with trigeminal neuralgia). In phase 2, we performed a systematic review in PubMed of all studies until September 2019 with diagnostic screening tools for cluster headache. In phase 3, a 6-person panel of cluster headache patients, research coordinators, and headache specialists analyzed the data from the first two phases to generate a new diagnostic screening tool. Finally, in phase 4 this new screen was validated on participants at a single headache center (all diagnoses) and through research recruitment (trigeminal autonomic cephalalgias only, as recruitment was essential but was otherwise low).

Results: In total, this study included 319 unique participants including 109 cluster headache participants (95 total participants/45 cluster headache participants in phase 1, and 224 total participants/64 cluster headache participants in phase 4). It also found 123 articles on potential screening tools in our systematic review. In phase 1, analysis of the English translation of an Italian screen generated 7 questions with high sensitivity and specificity against migraine, trigeminal neuralgia, and other trigeminal autonomic cephalalgias, but had grammatical and other limitations as a general screening tool. In phase 2, the systematic review revealed nine studies that met inclusion criteria as diagnostic screening tools for cluster headache, including four where sensitivity and specificity were available for individual questions or small groups of questions. In phase 3, this data was reviewed by the expert panel to generate a brief (6-item), binary (yes/no), written screening test. In phase 4, a total of 224 participants completed the new 6-item screening test (81 migraine, 64 cluster headache, 21 other trigeminal autonomic cephalalgias, 35 secondary headaches, 7 neuralgias, 5 probable migraine, and 11 other headache disorders). Answers to the 6 items were combined in a decision tree algorithm and three items had a sensitivity of 84% (confidence interval or 95% confidence interval 73–92%), specificity of 89% (95% confidence interval 84–94%), positive predictive value of 76% (95% confidence interval 64-85%), and negative predictive value of 93% (95% confidence interval 88–97%) for the diagnosis of cluster headache. These three items focused on headache intensity, duration, and autonomic features.

Conclusion: The 3-item Erwin Test for Cluster Headache is a promising diagnostic screening tool for cluster headache.

[1]Department of Neurosurgery, University of Texas Health Science Center at Houston, Houston, TX, USA
[2]Department of Neurology, Baylor College of Medicine, Houston, TX, USA
[3]Department of Psychology, University of West Georgia, Carrollton, GA, USA
[4]Department of Biochemistry and Molecular Biology, University of Texas Health Science Center at Houston, Houston, TX, USA
[5]Department of Biostatistics and Data Science, UTHealth School of Public Health, Houston, TX, USA

Corresponding author:
Mark J Burish, Department of Neurosurgery, University of Texas Health Science Center at Houston, Houston, TX, USA.
Email: mark.j.burish@uth.tmc.edu

Keywords
trigeminal autonomic cephalalgia, diagnostic questionnaire, screening tool, systematic review, sensitivity and specificity, Erwin Test for Cluster Headache (ETCH)

Date received: 6 November 2020; revised: 5 April 2021; accepted: 8 April 2021

Introduction

Delays in diagnosis are a widespread issue in cluster headache (CH): in a recent meta-analysis, diagnostic delay was 1–8 years across 13 countries (1). Patients consulted multiple specialties before the correct diagnosis is made, including dentistry, otolaryngology, ophthalmology, neurosurgery, and psychiatry (2–8). Cluster headache was often diagnosed by a neurologist, but in one study 41.3% of patients were misdiagnosed by neurologists as well (2). The most common misdiagnoses were migraine, trigeminal neuralgia, sinusitis, and dental or jaw disease (1), and patients received unnecessary procedures such as tooth extractions and nasal septum surgery (2–5,7–9). Misdiagnoses and delays in diagnoses not only lead to increased morbidity due to incorrect treatments, but they may also contribute to the high costs for emergency room visits and radiology services (10).

There is thus a great need for an accurate and efficient method of diagnosing CH. Currently there is no molecular or imaging diagnostic biomarker; CH is diagnosed using clinical criteria from the International Classification of Headache Disorders (11). Given the diagnostic issues mentioned above, a diagnostic aid may be useful. A previously employed method to aid in clinical diagnoses has been the use of self-administered screening tools. In CH, tools have been developed in several languages (12–18) but self-administered screening tools are limited in English. We performed a four year study to develop a new self-administered screening tool in English. We first explored the most suitable questions for a new tool by testing a translated version of an Italian questionnaire and by performing a systematic review. We then used an expert consensus to create a new CH diagnostic screening tool and performed initial validation of this new tool.

Methods

This study was designed in 4 phases: 1) we tested the English translation of an Italian screen with high sensitivity and specificity for CH; 2) we performed a systematic review of currently published CH diagnostic screening tools, 3) an expert panel used data from phases 1 and 2 to create a new diagnostic screening tool, and 4) we performed initial validation of the new screen. The phases of the study related to participants (phases 1 and 4) were approved by the Institutional Review Board at the University of Texas Health Science Center at Houston.

Phase 1: a previous diagnostic screening tool for CH was established in Italian by Torelli et al (17) and an English version of the screen is available in their publication (though it was not used by participants). The previous study consisted of a self-administered screen of 16 items with responses of "yes", "no", or "don't know" and was tested on 71 Italian-speaking patients with either CH, migraine, or tension-type headache. The authors found that a combination of eight of the items had a sensitivity of 100% and specificity of 95.1% for the diagnosis of CH.

The goal of phase 1 is to investigate the English version of this screen in an American population, to evaluate its effectiveness in screening for CH, and (because of several issues in the translation and specific question topics) to use this data to generate a new screening test. Phase 1 was performed prospectively at a single site between July 2016 and June 2019. Participants were recruited by one of four methods: from inquiries to a local headache foundation, from information on our website, from our clinicaltrials.gov posting (NCT02910323), or during in-person clinic visits at a single site (patients enrolled nonconsecutively). Individuals residing in the United States who were interested in participating but who could not travel to the clinic were enrolled and interviewed by phone. The single clinical site is that of author MJB and is located in the Texas Medical Center. Inclusion criteria were: 1) at least 18 years of age, and 2) a diagnosis of CH, migraine, tension-type headache, or trigeminal neuralgia. In all cases, the gold standard diagnosis of an interview with a neurologist was made by a single neurologist (study author MJB) according to the International Classification of Headache Disorders (ICHD)-3 beta criteria (19), with the definitions of episodic and chronic CH updated to the ICHD-3 criteria (11). The neurologist was not always blind to the results of the screen. Participants were excluded if they had two or more headache

disorders associated with severe pain. This exclusion criterion was created because some patients could not distinguish between multiple headache types with severe headache intensity, such as migraine and CH attacks. However, in our experience patients could distinguish mild or moderate headache attacks, such as tension-type headache, from CH attacks. Thus a participant with episodic CH and tension-type headache would be included in the episodic CH group, but a participant with episodic CH and migraine would be excluded from the study.

Before starting our study, several of our authors noted that some of the translated English questions from the Italian screening test had complicated wording that could easily be explained verbally if necessary. A verbal interaction also provided useful input for creating a new screen. Thus, instead of a self-administered format like Torelli et al, our screen was administered verbally either in-person or over the phone by one of the study authors. All participants completed all 16 questions, and clarification on the questions was provided by study authors when needed.

Phase 2: A systematic literature review was performed to identify previously published headache screens using PRISMA guidelines (20) with one exception: PRISMA guidelines request that the title of the article mention that it is a systematic review. We did not follow this request because a systematic review was not the final goal of our research. No systematic review protocol was registered prior to conducting our search. The search string was created in two steps. In step A, a non-systematic review of screening tests was performed using PubMed, Google Scholar, Google, and a review of references in relevant articles. This step identified seven articles (12–18). In step B, a systematic review of screening tests was performed. Search criteria were created in PubMed and modified until all seven articles from step A were included among the results. The final search string was as follows: *("cluster headache" or "trigeminal autonomic") AND (screen or questionnaire or survey) AND (sensitivity or specificity or validation or prevalence)*. The search was performed on 4 September 2019 by author RM and 6 September 2019 by author MJB. Of note we used what is now considered the legacy version of PubMed, which uses a different search algorithm (21,22) and provided different results than those currently available on PubMed. Two authors (RM and MJB) reviewed the articles independently in 3 rounds (titles, then abstracts, then full text), stopping at the end of each round to compare results, discuss disagreements, and come to a consensus before proceeding to the next round. Articles were included if: 1) a questionnaire was administered, 2) the diagnosis was confirmed by interview with a neurologist or a provider working with a neurologist (e.g. nurse, medical student, or non-neurology physician), and 3) a diagnostic screen for CH was part of the questionnaire. Articles were excluded if they were duplicates, including the use of the same screening tool in the same language. We also excluded non-English articles. No methods were performed to assess risk of bias. After the review was complete, one author (MJB) examined each article and extracted title, authors, year, journal, country, population type, total number of participants completing the questionnaire, total number interviewed by a neurologist, specific screening questions (if available), and sensitivity/specificity/positive and negative predictive values (if available). Specific questions were grouped by one author (MJB) into general topics such as pain intensity, pain duration, or specific autonomic features.

Phase 3: An expert panel of six people was created: two CH patients, two research coordinators with experience administering surveys to CH participants (authors RM and SMP), and two headache specialists (authors RWE and MJB). This expert panel reviewed the results of phases 1 and 2 and created a new 6-item diagnostic screening tool that was examined in phase 4.

Phase 4: The new diagnostic screening tool was administered to participants. Data was collected retrospectively between October 2019 and August 2020, and the gold standard diagnosis of an interview with a neurologist was again made by a single neurologist (author MJB) according to the ICHD-3 (11); of note medication overuse headaches were rarely included as the strict definition of regular overuse for over 3 months was not asked for most patients. The neurologist was blind to the results of the screening test. Inclusion and exclusion criteria were the same as for Phase 1 with two exceptions: 1) we accepted all ICHD-3 diagnoses, as well as participants with multiple severe headaches; and 2) all participants included in phase 1 were excluded from phase 4. All participants filled out a paper screen, as this was thought to be the format that would be most commonly used in clinical practice. We consecutively enrolled headache patients from in-person visits at the same clinic as phase 1. Of note during the study, the clinic started utilizing telemedicine in March 2020 due to the coronavirus pandemic; telehealth patients were not enrolled as the form was completed on paper. Also, given the low prevalence of trigeminal autonomic cephalalgia (TAC) patients in the clinic, similar to phase 1 we also contacted TAC participants recruited from other means (inquiries to a local headache foundation, from information on our website, from our clinicaltrials.gov posting), as long as they had not been included in phase 1. All patients contacted remotely were required to print out the screening tool, fill it out on paper, and then submit either by mail or photograph/scan. STARD guidelines

for reporting studies of diagnostic accuracy (23) were followed.

Statistics: All statistical analyses were performed a priori. Sample size was based on a previous study (17). In phase 1, a confusion matrix was made to evaluate the sensitivity, specificity, positive predictive value, and negative predictive value for each item. To obtain a subset of the items with the best sensitivity and specificity, we performed a variable selection procedure through lasso regression with 6-fold cross validation. The procedure was repeated 100 times to select the best lambda value to maximize the area under the curve (AUC) value. A receiver operating characteristic (ROC) curve was built, and the optimal threshold was determined based on Youden's J statistic to maximize the distance to the identity line, followed by the calculation of sensitivity and specificity. In phase 4, two analyses of the screening tool were performed. In the first, CH (ICHD-3 codes 3.1, 3.1.1, and 3.1.2) was compared to all other headache disorders. Participants with multiple diagnoses were placed in the CH group if one of their diagnoses was CH. In the second, the same analysis was performed but we removed participants with other TACs (ICHD-3 codes 3.2, 3.3, 3.4, 3.5, and their subheadings such as 3.2.1). We combined all 6 questions to generate a decision tree algorithm to maximize the classification accuracy. The sensitivity, specificity, PPV, NPV, and their 95% confidence intervals were calculated to assess the performance of the decision tree algorithm. All statistical analysis was conducted in R software version 3.4.2, and in phase 4 the "rpart" package was used.

Results

This study involved two initial phases (1. testing of a screening tool with some limitations and 2. a systematic review), a third phase where an expert panel created a new screening tool, and a fourth phase where the new screening tool was tested on participants. Each of the four phases will be presented separately.

Phase 1 (English translation of prior screen)

A total of 186 participants were enrolled into phase 1, and 95 met inclusion and exclusion criteria (Figure 1a). There was low enrollment for paroxysmal hemicrania (n=5), SUNCT (n=2), SUNA (n=0), and trigeminal neuralgia (n=3).

Analysis of the screen for CH versus all other headaches is shown in Table 1. Questions with sensitivity above 90% were presence of headaches, pain severity, unilaterality, location, conjunctival injection, lacrimation, and pain occurring at night. However, these questions had some of the lowest specificities. There were no questions with specificity above 90%, but questions with specificity above 70% were headache duration, clock-like circadian pattern, and previous use of verapamil or lithium. Like the Torelli et al paper, we also evaluated if a particular subset of questions gave high sensitivity and specificity. We found 100% sensitivity and 96.8% specificity with the combination of pain severity, conjunctival injection, lacrimation, nasal congestion, headache duration, clock-like circadian pattern, and previous use of lithium or verapamil.

Two questions from the screen were problematic in creating a new screen: 1) a question about being headache-free for many months may exclude individuals with chronic CH, and 2) a question about trialing verapamil or lithium may exclude individuals who had not been seen by medical providers. Therefore we removed those 2 questions and again evaluated if a particular subset of questions gave high sensitivity and specificity (Table 1). We found 100% sensitivity and 93.9% specificity with the combination of pain severity, conjunctival injection, lacrimation, nasal congestion, rhinorrhea, headache duration, and clock-like circadian pattern.

Phase 2 (systematic review)

A total of 123 articles were identified with our search term (Figure 1b). Ultimately 9 articles met our inclusion criteria (Supplemental Table 1). Questions from these screens focused on ICHD criteria, in particular headache duration, cranial autonomic features, and restlessness/agitation. Sensitivity and specificity for specific questions or groups of questions were available from four articles (Table 2). The screens were performed in multiple languages, limiting the ability to compare question structure across studies. Question topics with the highest sensitivity asked about unilaterality, restlessness, lacrimation, duration of pain, and multiple headache attacks per week. The questions with the highest specificity asked about conjunctival injection and duration of pain. Interestingly, a German screen consisting of nine items identified 182 suspected cases of CH, only four of which were confirmed by a neurology interview to have CH: the majority had migraine (24). Like the others screens, the German screen included questions about localization, intensity, autonomic features, and restlessness. Unlike the screens in Table 2, the German screen did not include questions on headache frequency or duration. This difference suggests that duration and/or frequency questions increase specificity for CH.

ICHD criteria not mentioned in any of the screens included several autonomic features: eyelid edema, forehead and facial sweating, and miosis/ptosis. While the addition of these features may improve

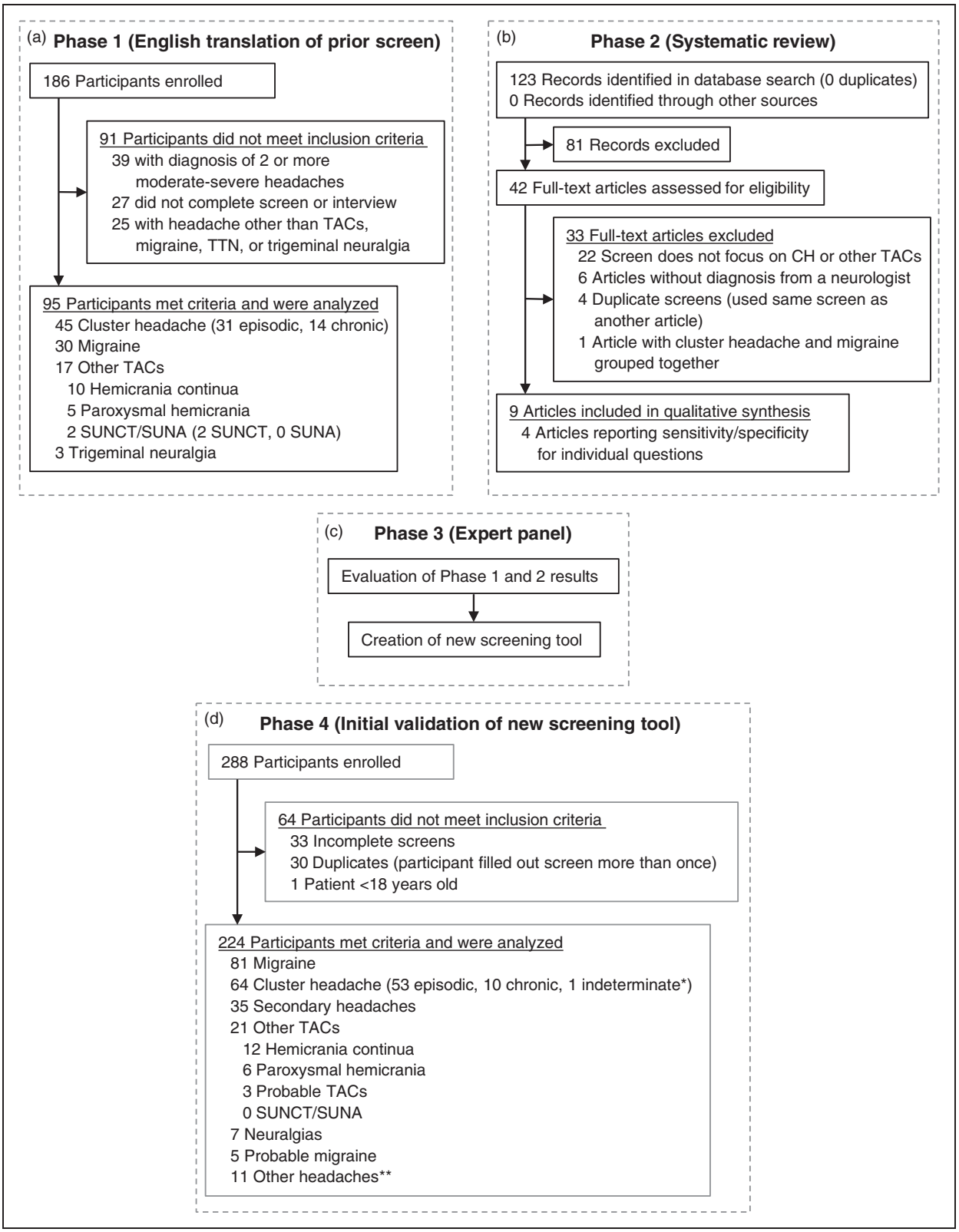

Figure 1. Flow diagrams for each phase of this study. In phase 2, within excluded full-text articles there were a total of 7 articles without diagnosis from a neurologist, but 2 of them also had a questionnaire that did not focus on cluster headache or other TACs, so were included there and not double-counted. *One participant was indeterminate for episodic or chronic because the headaches started less than 1 year ago and were without a remission period. **11 Other headaches: 4 new daily persistent headache, 2 tension-type headache, 2 headache unspecified, 1 primary stabbing headache, 1 persistent idiopathic facial pain, 1 primary headache associated with sexual activity. Abbreviations: TACs, trigeminal autonomic cephalalgias; TTH, tension-type headache.

Table 1. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for cluster headache in the English translation of the Torelli et al screen (full questions in English are available in the original study [17]). All values shown as percentages. Survey responses for participants with cluster headache (n=31 episodic, n=14 chronic) were compared to survey responses for migraine (n=30), hemicrania continua (n=10), paroxysmal hemicrania (n=5), trigeminal neuralgia (n=3), and short-lasting unilateral neuralgiform headache with conjunctival injection and tearing (n=2). The English translations were copied verbatim from the original paper with the exception of the last question: Carbolithium was changed to lithium, and Isoptin was changed to verapamil. Numbers represent percentage. Optimal combination of questions with high sensitivity and specificity is shown at bottom, first including questions 15 and 16 (which may exclude chronic cluster headache and patients without medical care), and second excluding questions 15 and 16.

	Sensitivity	Specificity	PPV	NPV
Question 1 (presence of headache)	100.0	2.0	47.3	100.0
Question 2 (severe pain)	100.0	14.0	50.6	100.0
Question 3 (unilateral location)	95.5	32.0	55.3	88.9
Question 4 (location around eye)	100.0	10.2	50.0	100.0
Question 5 (conjunctival injection)	92.5	58.1	67.3	89.3
Question 6 (lacrimation)	97.7	45.8	62.3	95.7
Question 7 (nasal congestion)	72.1	65.9	67.4	70.7
Question 8 (rhinorrhea)	86.4	60.0	67.9	81.8
Question 9 (restlessness)	84.1	63.8	68.5	81.1
Question 10 (duration <4 hours)	88.6	73.5	75.0	87.8
Question 11 (>1 attack per day without interictal pain)	78.6	52.1	58.9	73.5
Question 12 (frequency of daily for 7 days)	88.6	42.0	57.4	80.8
Question 13 (clock-like circadian pattern)	75.0	80.0	76.9	78.3
Question 14 (nocturnal circadian pattern)	93.2	35.4	56.9	85.0
Question 15 (remission periods lasting months)	74.4	68.1	68.1	74.4
Question 16 (use of lithium or verapamil for headaches)	78.6	80.9	78.6	80.9
Optimal combination of questions 1–16 = Questions 2 + 5 + 6 + 7 + 10 + 13 + 16	100.0	96.8	95.7	100
Optimal combination of questions 1–14 = Questions 2 + 5 + 6 + 7 + 8 + 10 + 13	100.0	93.9	93.6	100

specificity, it should be noted that they may not result in large improvements in sensitivity: a patient requires only one feature of restlessness or cranial autonomic features for a CH diagnosis, and the features omitted are much less common than the ones that were included (25,26).

Phase 3 (expert panel)

The screening tool in Phase 1 had two major limitations: grammatical issues (as a translation from another language) as well as two questions that potentially exclude chronic CH patients and patients who do not see medical providers. Therefore an expert panel was convened to create a new screening tool. The expert panel used information from Phase 1 and Phase 2 to create a new screening tool (Figure 1c). The new screen consisted of 6 yes-or-no questions concerning pain intensity, circadian pattern, duration, restlessness, laterality, and autonomic features.

Phase 4 (initial validation of new screening tool)

A total of 287 participants were enrolled into phase 4, and 224 met inclusion criteria (Figure 1d). Participants had a wide range of diagnoses: in the ICHD-3, which has 14 major sub-headings, only 3 sub-headings were not represented: headache attributed to infection, headache attributed to disorder of homoeostasis, and headache attributed to psychiatric disorder. Supplemental Table 2 lists diagnoses and screening tool responses for individual participants. The most common primary headache was migraine (n=81) followed by CH (n=64). CH patients were primarily male with the episodic version of the disease, while migraine patients were primarily female but had a higher prevalence of chronic migraine than the general population (Table 3). The most common secondary headaches were persistent headache attributed to traumatic injury to the head (n=8) and headache attributed to idiopathic intracranial hypertension (n=8). The most common neuralgia or facial pain was trigeminal neuralgia (n=3).

The full 6-item screening tool, as well as sensitivity and specificity for each individual question, is shown in Table 4. While many questions had high sensitivity, no single question had high specificity. However, a decision tree algorithm identified 3 questions that had 85% sensitivity and 89% specificity for CH compared to all other diagnoses in our cohort (Figure 2). Those 3 questions are presented in Figure 3 and called the Erwin Test for Cluster Headache (ETCH).

Of note the other TACs appear to drive some of the specificity. When we analyzed the data after removing

Table 2. Sensitivity and specificity (listed as "sensitivity %/specificity %") of similar questions across multiple screening tests for cluster headache. In Chung et al and Dousset et al, conjunctival injection and/or tearing were included as a single question; in Chung et al., nasal congestion and/or rhinorrhea was also included as a single question. Also in Chung et al. there was one question on pain frequency (in this case more than 3 times per week) similar to the other studies on multiple attacks per week, as well as another question that combined intensity and duration. The final column refers to data from Phase 1 of this study, namely an English translation of the Torelli et al. screen that was provided in their article. *Duration was listed as less than 4 hours for Torelli et al. and within 3 hours for Chung et al. The specific duration was not reported for Dousset et al., though their discussion notes that other authors have suggested that durations longer than 3 hours may be needed.

Question (paraphrased as each is in a different language)	Chung et al (9)	Dousset et al (10)	Torelli et al (5)	Wilbrink et al (14).	Phase 1 of this article
Pain is severe			100/34.1		100/14.0
Pain is unilateral	90.5/56.4	94.6/44.1	100/61.0		95.5/32.0
Pain is in or near the eye			100/58.5		100/10.2
Conjunctival injection	73.8/94.0	89.2/82.5	63.3/90.2		92.5/58.1
Lacrimation	73.8/94.0	89.2/82.5	80/75.6		97.7/45.8
Nasal congestion	38.1/97.7		63.3/90.2		72.1/65.9
Rhinorrhea	38.1/97.7		70.0/90.2		86.4/6.0.0
Restlessness/Agitation	83.3/84.6		90.0/92.7		84.1/63.8
Duration of attack*	83.3/86.9	91.9/91.4	100/90.2		88.6/73.5
Multiple attacks per day			73.3/73.2		78.6/52.1
Multiple attacks per week	85.7/51.3		96.7/68.3		88.6/42.0
Headache at specific times of day (i.e., clock-like)			63.3/78.0		75.0/80.0
Headache occurs at night			63.3/78.0		93.2/35.4
Headaches remit or disappear for months			56.7/95.1		74.4/68.1
Have used verapamil or lithium			66.7/97.6		78.6/80.9
Headache is disabling	97.6/47.7				
Nausea	73.8/34.6				
Photophobia	61.9/55.7				
Multiple intense attacks for over a week	88.1/70.8				
Combined unilateral + conjunctival injection/lacrimation + duration		78.4/100			
Combined attack duration 15–180 min + headaches remit or disappear for months + male sex				53.8/88.9	

Table 3. Baseline demographics of study population from phase 4 (initial validation of new screening tool). Age is presented as mean (standard deviation). Participants are organized by diagnosis code though some have multiple diagnoses. For this table, participants with multiple diagnoses were arbitrarily placed in highest category as follows: 1st cluster headache, 2nd other trigeminal autonomic cephalalgia, 3rd secondary headache, 4th neuralgia, 5th other headache, 6th migraine, 7th probable migraine. Age in years presented as mean (standard deviation). *One participant was indeterminate for episodic or chronic because the headaches started less than 1 year ago and were without a remission period. **Several participants had indeterminate episodic/chronic patterns. Abbreviations: TBI headache, persistent headache attributed to traumatic injury to the head; IIH headache, Headache attributed to idiopathic intracranial hypertension (IIH).

	Age in years	Percent male	Episodic/chronic
Total			
Primary Headaches			
Migraine (n=81)	41.1 (13.8)	14.8% (12 M/69 F)	27/54
Cluster headache (n=64)	49.7 (11.9)	64.1% (41 M/23 F)	54/10*
Other Trigeminal Autonomic Cephalalgias (n=21)	47.1 (10.7)	4.8% (1 M/20 F)	1/12**
All other primary headaches (n=13)	42.5 (19.3)	38.5% (5 M/8 F)	
Secondary headaches			
TBI headache (n=8)	46.6 (18.0)	62.5% (5 M/3 F)	
IIH headache (n=8)	44.4 (12.2)	0% (0 M/8 F)	
All other secondary headaches (n=19)	47.5 (13.7)	36.8% (7 M/12 F)	
Neuralgias, Facial Pains, and Other headaches (n=10)	54.6 (18.4)	60.0% (6 M/4 F)	

Table 4. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy for cluster headache in the 6-item diagnostic screen (answer choices are yes/no). All values shown as value (95% confidence interval). The 6-item questionnaire in this Table is © 2020, The University of Texas Health Science Center at Houston. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

	Sensitivity	Specificity	PPV	NPV	Accuracy
1. Is this the worst pain you have ever experienced?	0.94 (0.85–0.98)	0.53 (0.44–0.60)	0.44 (0.36–0.53)	0.95 (0.89–0.99)	0.64 (0.58–0.71)
2. Does the headache generally start at the same time each day?	0.89 (0.79–0.95)	0.57 (0.49–0.65)	0.45 (0.36–0.54)	0.93 (0.86–0.97)	0.66 (0.59–0.72)
3. Imagine setting a timer. Does the headache last less than 4 hours?	0.92 (0.83–0.97)	0.63 (0.55–0.71)	0.50 (0.41–0.59)	0.95 (0.89–0.98)	0.71 (0.65–0.77)
4. Do you feel the need to rock, move about, or bang your head during a headache?	0.84 (0.73–0.92)	0.66 (0.58–0.74)	0.50 (0.40–0.60)	0.91 (0.85–0.96)	0.71 (0.65–0.77)
5. During a headache, is the pain on only one side (right or left)?	0.97 (0.89–1.00)	0.44 (0.36–0.52)	0.41 (0.33–0.49)	0.97 (0.90–1.00)	0.59 (0.52–0.65)
6. During a headache, do one or more of these happen to you?	0.94 (0.85–0.98)	0.58 (0.50–0.66)	0.47 (0.38–0.56)	0.96 (0.90–0.99)	0.68 (0.62–0.74)
your eye turns red on only one side					
your eye waters on only one side					
your nose runs on only one side					
your nose gets congested on only one side					
OPTIMAL COMBINATION OF QUESTIONS 1–6 = QUESTIONS 1 + 3 + 6	0.84 (0.73–0.92)	0.89 (0.84–0.94)	0.76 (0.64–0.85)	0.93 (0.88–0.97)	0.88 (0.83–0.92)

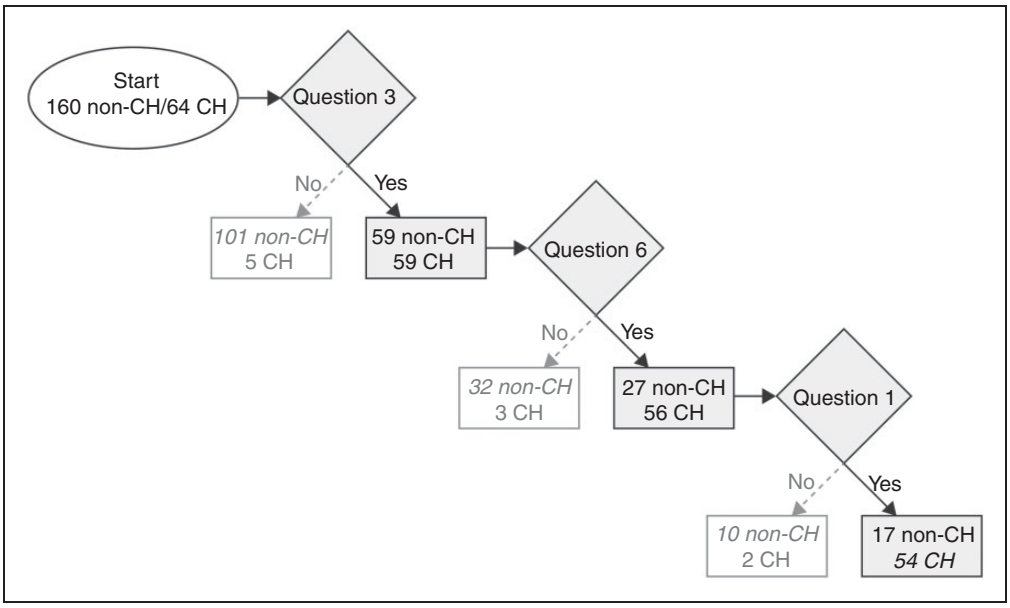

Figure 2. Decision tree for cluster headache classification with all 224 patients. The participants who answered "yes" to a question were then evaulated in the subsequent question. Text that is italicized represents accurate classifications. Three out of 6 questions, Q3, Q6, and Q1, were selected in the decision tree algorithm. Abbreviations: CH, cluster headache.

The Erwin Test for Cluster Headache

1. Is this the worst pain you have ever experienced?
   ☐ Yes ☐ No

2. Imagine setting a timer. Does the headache last less than 4 hours?
   ☐ Yes ☐ No

3. During a headache, do one or more of these happen to you?
   …your eye turns red on only one side
   …your eye waters on only one side
   …your nose runs on only one side
   …your nose gets congested on only one side
   ☐ Yes ☐ No

Figure 3. The 3-item Erwin Test for Cluster Headache (ETCH). A yes to all 3 questions has 85% sensitivity and 89% specificity for cluster headache based on the analysis of 224 participants in this study. This Figure copyright 2020, The University of Texas Health Science Center at Houston. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

the 21 participants with paroxysmal hemicrania, SUNCT, SUNA, hemicrania continua, and probable TACs, the same 3 items were identified by a decision tree algorithm but had 5% more specificity (88% sensitivity and 94% specificity) for CH (Supplemental Table 3).

Discussion

Using in-depth testing of a prior screen, a systematic review of previous screening tests, and a consensus among experts, we created a new diagnostic screening tool and found that three questions had an 85% sensitivity and 89% specificity for CH in 224 headache participants. The first question focuses on the extreme intensity of pain in CH, which is significantly higher than other intensely painful disorders such as labor pain, pancreatitis, and nephrolithiasis (27). The second question focuses on the shorter duration of pain as many other headache disorders, in particular migraine and secondary headaches, are often longer than 4 hours. And the third focuses on several cranial autonomic features, which are not unique to CH or even the TACs as a whole (28) but are particularly prominent in these disorders (29). Patients with other TACs have very similar features to CH, and when those 21 participants were removed from the analysis of 224 participants, the specificity increased by 5%. This fact suggests that the other TACs are an important differential diagnosis to consider when a participant tests positive for our screening tool. However, the possible confounder of other TACs is likely negligible in many settings because paroxysmal hemicrania (30,31), SUNCT/SUNA (32,33), and hemicrania continua (30,34) are even less common than CH (which has a prevalence of 1 in 1000 [35]). A strength of our study is that our validation step included a variety of headache and facial pain disorders: it included 58 participants with diagnoses other than migraine or TACs.

Thus the high specificity suggests that this screen has promise in the general population.

After our systematic review, a screening test in the United Kingdom was published with 86.4% sensitivity and 92.0% specificity for CH compared to migraine (36). This screening test includes two written components (the visual analog scale and a selection of six pain images depicting headache) and two verbal components (a description of pain and several questions differentiating migraine from CH, which are then categorized based on the responses). Like other screening tests, their 12-item screening tool focuses on pain (intensity, location, duration, and quality), associated cranial autonomic features, and a circadian pattern of attacks. Their screen differs from ours in that their screen uses a structured interview, and only migraine and CH were included. Their screen is similar to ours in that it was tested at a single headache center, and we agree with their assessment that single site studies (including ours) require further evaluation in larger settings.

Our study has several limitations, specific to each phase. In phase 1, patients were not enrolled consecutively, and our study was administered verbally instead of self-administered. The goal was to ensure that the wording of the screen was understood and to gain additional insight into participants' interpretation of the questions. However, a verbal version that allows clarification with a study author introduces variation that could bias results. We also excluded most co-morbid headache disorders (with the exception of tension-type headaches) and all secondary headaches. Thus the sensitivity and specificity of our phase 1 findings might be higher than expected because of our verbal explanations and because we excluded many other headache types. In phase 2, we searched only one database (PubMed), which could give an incomplete list of articles. As we started with seven articles to seed the search term, our search only identified an additional two articles (37,38). The primary issue with all the limitations in phases 1 and 2 is that they could bias the data analyzed in phase 3 that was used to generate the new screening tool. Specifically, the new screening tool was based on data that lacked patients with co-morbid primary headache disorders and had a low number of other trigeminal autonomic cephalalgias, neuralgias, and secondary headaches. In phase 4, the first limitation is that the study was a single site with all participants diagnosed by a single headache specialist. Second, test-retest reliability was not evaluated. Third, we enrolled a low number of secondary headaches listed in the differential diagnosis of CH, such as tooth impaction, maxillary sinusitis, and headaches due to neoplasms such as pituitary tumors (34). We also did not characterize medication overuse headaches, though in most headache disorders medication overuse headaches do not resemble CH attacks (39). However, a small study in CH suggests that CH patients with medication overuse may have headaches much different than their CH attacks (40). Fourth, patients with TACs were recruited through multiple methods while other disorders were only recruited through the clinic. While this was done for practical reasons (the TACs were essential for the study but also uncommon), the TAC population may be different because of their recruitment. Finally, in both phases 1 and 4, some respondents were taking preventive medications at the time of the screen, thus their frequency or duration of headache may not be accurate.

In summary, our findings suggest that our screening test may be a useful tool for the identification of CH. Its utilization in primary care settings as well as specialty clinics involved in facial pain (e.g. dentistry, neurology, otolaryngology, and pain medicine) could produce quicker diagnoses and consequently faster delivery of appropriate treatments. A similarly brief 3-item screening tool, the ID migraine screener with 81% sensitivity and 75% specificity for migraine in a study of primary care (41), has been used extensively in migraine epidemiologic and translational research. Thus our screening tool may also be useful in research settings.

Key findings

- A brief 3-item diagnostic screening tool for cluster headache has 84% sensitivity and 89% specificity in 224 headache patients.
- This screening tool has the potential to be used clinically to decrease diagnostic delay, and scientifically to screen research participants.

Acknowledgments

We thank our two cluster headache expert panel members who assisted in phase 3 of the study, as well as all of the participants in phases 1 and 4.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study received support from the Will Erwin Headache Research Foundation. The Erwin Test for Cluster Headache (ETCH) is named in honor of the late Will Erwin.

ORCID iD

Mark J Burish https://orcid.org/0000-0002-8931-6436

References

1. Buture A, Ahmed F, Dikomitis L and Boland JW. Systematic literature review on the delays in the diagnosis and misdiagnosis of cluster headache. *Neurol Sci* 2019; 40: 25–39.
2. Vikelis M and Rapoport AM. Cluster headache in Greece: an observational clinical and demographic study of 302 patients. *J Headache Pain* 2016; 17: 88.
3. Del Rio MS, Leira R, Pozo-Rosich P, et al. Errors in recognition and management are still frequent in patients with cluster headache. *Eur Neurol* 2014; 72: 209–212.
4. Bahra A and Goadsby PJ. Diagnostic delays and mismanagement in cluster headache. *Acta Neurol Scand* 2004; 109: 175–179.
5. van Vliet JA, Eekers PJE, Haan J, et al. Features involved in the diagnostic delay of cluster headache. *J Neurol Neurosurg Psychiatry* 2003; 74: 1123–1125.
6. Maytal J, Lipton RB, Solomon S, et al. Childhood Onset Cluster Headaches. *Headache J Head Face Pain* 1992; 32: 275–279.
7. Bittar G and Graff-Radford SB. A retrospective study of patients with cluster headaches. *Oral Surg Oral Med Oral Pathol* 1992; 73: 519–525.
8. Van Alboom E, Louis P, Van Zandijcke M, et al. Diagnostic and therapeutic trajectory of cluster headache patients in Flanders. *Acta Neurol Belg* 2009; 109: 10–17.
9. Frederiksen H-H, Lund NL, Barloese MC, et al. Diagnostic delay of cluster headache: A cohort study from the Danish Cluster Headache Survey. *Cephalalgia* 2020; 40: 49–56.
10. Choong CK, Ford JH, Nyhuis AW, et al. Health care utilization and direct costs among patients diagnosed with cluster headache in U.S. health care claims data. *J Manag Care Spec Pharm* 2018; 24: 921–928.
11. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition. *Cephalalgia* 2018; 38: 1–211.
12. Chung PW, Cho SJ, Kim BK, et al. Development and validation of the cluster headache screening questionnaire. *J Clin Neurol* 2019; 15: 90–96.
13. Dousset V, Laporte A, Legoff M, et al. Validation of a brief self-administered questionnaire for cluster headache screening in a tertiary center. *Headache* 2009; 49: 64–70.
14. Fritsche G, Hueppe M, Kukava M, et al. Validation of a German language questionnaire for screening for migraine, tension-type headache, and trigeminal autonomic cephalgias. *Headache* 2007; 47: 546–551.
15. Kukava M, Dzagnidze A, Janelidze M, et al. Validation of a Georgian language headache questionnaire in a population-based sample. *J Headache Pain* 2007; 8: 321–324.
16. Maizels M and Wolfe WJ. An expert system for headache diagnosis: The computerized headache assessment tool (CHAT). *Headache* 2008; 48: 72–78.
17. Torelli P, Beghi E and Manzoni GC. Validation of a questionnaire for the detection of cluster headache. *Headache* 2005; 45: 644–652.
18. Wilbrink LA, Weller CM, Cheung C, et al. Stepwise web-based questionnaires for diagnosing cluster headache: LUCA and QATCH. *Cephalalgia* 2013; 33: 924–931.
19. Headache Classification Committee of the International Headache Society (IHS). The International Classification of Headache Disorders, 3rd edition (beta version). *Cephalalgia* 2013; 33: 629–808.
20. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; 151(4): 264–259.
21. Canese K. An Updated PubMed Is on Its Way. NLM Tech Bull. 4 March 2019. Available from: https://www.nlm.nih.gov/pubs/techbull/ma19/ma19_pubmed_update.html (Accessed 10 August 2020)
22. NIH NL of M. A New PubMed: Highlights for Information Professionals: Questions and Answers. 2019. Available from: https://www.nlm.nih.gov/oet/ed/pubmed/events/2019_09_faq.html (Accessed 10 August 2020)
23. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. BMJ 28 October 2015. Available from: https://pubmed.ncbi.nlm.nih.gov/26511519/ (Accessed 5 November 2020)
24. Katsarava Z, Obermann M, Yoon M-S, et al. Prevalence of cluster headache in a population-based sample in Germany. *Cephalalgia* 2007; 27: 1014–1019.
25. Moon H-S, Cho S-J, Kim B-K, et al. Field testing the diagnostic criteria of cluster headache in the third edition of the International Classification of Headache Disorders: A cross-sectional multicentre study. *Cephalalgia*. 2019; 39: 900–907.
26. de Coo I, Wilbrink L, Haan J, et al. Evaluation of the new ICHD-III beta cluster headache criteria. *Cephalalgia*. 2016; 36: 547–541
27. Burish MJ, Pearson SM, Shapiro RE, et al. Cluster headache is one of the most intensely painful human conditions: Results from the International Cluster Headache Questionnaire. Headache 2021; 61: 117–124.
28. Riesco N, Pérez-Alvarez AI, Verano L, et al. Prevalence of cranial autonomic parasympathetic symptoms in chronic migraine: Usefulness of a new scale. *Cephalalgia* 2016; 36: 346–350.
29. Goadsby PJ and Lipton RB. A Review of paroxysmal hemicranias, SUNCT syndrome and other short-lasting headaches with autonomic feature, including new cases. *Brain* 1997; 120: 193–209.

30. Eller M and Goadsby P. Trigeminal autonomic cephalalgias. *Oral Dis* 2014; 22: 1–8.
31. Benoliel R and Sharav Y. Paroxysmal hemicrania. *Case studies and review of the literature*. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 1998; 85: 285–292.
32. Cohen A. SUN: Short-Lasting Unilateral Neuralgiform Headache Attacks. *Headache J Head Face Pain* 2017; 57: 1010–1020.
33. Lambru G and Matharu MS. SUNCT and SUNA: medical and surgical treatments. *Neurol Sci* 2013; 34: 75–81.
34. McGeeney BE. Cluster headache and other trigeminal autonomic cephalalgias. *Semin Neurol*. 2018; 38: 603–607.
35. Fischera M, Marziniak M, Gralow I, et al. The incidence and prevalence of cluster headache: A meta-analysis of population-based studies. *Cephalalgia* 2008; 28: 614–618.
36. Buture A, Boland JW, Dikomitis L, et al. Development and evaluation of a screening tool to aid the diagnosis of a cluster headache. *Brain Sci* 2020; 10: 77.
37. Evers S, Fischera M, May A, et al. Prevalence of cluster headache in Germany: Results of the epidemiological DMKG study. *J Neurol Neurosurg Psychiatry* 2007; 78 (11): 1289–1290
38. Haimanot T, Seraw B, Forsgren L, et al. Migraine, chronic tension-type headache, and cluster headache in an Ethiopian rural community. *Cephalalgia* 1995; 15: 482–488.
39. Diener HC, Holle D, Solbach K, et al. Medication-overuse headache: Risk factors, pathophysiology and management. *Nat Rev Neurol* 2016; 12(10): 575–583
40. Paemeleire K, Bahra A, Evers S, et al. Medication-overuse headache in patients with cluster headache. *Neurology* 2006; 67: 109–113.
41. Lipton RB, Dodick D, Sadovsky R, et al. A self-administered screener for migraine in primary care: The ID migraine^{TM} validation study. *Neurology* 2003; 61: 375–382.