Computerized migraine diagnostic tools: a systematic review

Yohannes W. Woldeamanuel (iD) **and Robert P. Cowan**

Abstract

Background: Computerized migraine diagnostic tools have been developed and validated since 1960. We conducted a systematic review to summarize and critically appraise the quality of all published studies involving computerized migraine diagnostic tools.

Methods: We performed a systematic literature search using PubMed, Web of Science, Scopus, snowballing, and citation searching. Cutoff date for search was 1 June 2021. Published articles in English that evaluated a computerized/automated migraine diagnostic tool were included. The following summarized each study: publication year, digital tool name, development basis, sample size, sensitivity, specificity, reference diagnosis, strength, and limitations. The Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool was applied to evaluate the quality of included studies in terms of risk of bias and concern of applicability.

Results: A total of 41 studies (median sample size: 288 participants, median age = 43 years; 77% women) were included. Most (60%) tools were developed based on International Classification of Headache Disorders criteria, half were self-administered, and 82% were evaluated using face-to-face interviews as reference diagnosis. Some of the automated algorithms and machine learning programs involved case-based reasoning, deep learning, classifier ensemble, ant-colony, artificial immune, random forest, white and black box combinations, and hybrid fuzzy expert systems. The median diagnostic accuracy was concordance = 89% [interquartile range (IQR) = 76–93%; range = 45–100%], sensitivity = 87% (IQR = 80–95%; range = 14–100%), and specificity = 90% (IQR = 77–96%; range = 65–100%). Lack of random patient sampling was observed in 95% of studies. Case–control designs were avoided in all studies. Most (76%) reference tests exhibited low risk of bias and low concern of applicability. Patient flow and timing showed low risk of bias in 83%.

Conclusion: Different computerized and automated migraine diagnostic tools are available with varying accuracies. Random patient sampling, head-to-head comparison among tools, and generalizability to other headache diagnoses may improve their utility.

Keywords: automated migraine diagnosis, computerized migraine diagnosis, digital health, migraine, systematic review

Introduction

The rise of disruptive digital health technologies is revolutionizing healthcare.[1,2] The increasing prevalence of chronic health burden such as headache disorders and the growing digitally native population are drivers for innovative modes of disease management.[3,4] The application of digital health tools (e.g. computerized expert systems/software, mobile health, wearables, health information technology, Internet of Medical Things, telehealth and telemedicine, robotic surgery) is known to reduce diagnostic delays and inaccuracies, improve inefficiencies, improve access, and facilitate remote management and continuity of care – all of which reduce healthcare-related costs.[1,2,5] In addition, digital health platforms empower patients' self-efficacy by increasing patient's shared decision-making within the clinician–patient relationship as well as by facilitating behavioral activation and internal locus of

Correspondence to:
Yohannes W. Woldeamanuel
Division of Headache & Facial Pain, Department of Neurology & Neurological Sciences, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, USA.
ywoldeam@stanford.edu; yohannes.woldeamanuel@gmail.com

Robert P. Cowan
Division of Headache & Facial Pain, Department of Neurology & Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA

control.[3,4] Patient-centered digital care can be enhanced by promoting patient engagement and redefining patient–doctor relationship.[3,4] In the United States, healthcare delivery is suboptimal and inefficient as compared with economically similar countries that spend half annual per capita cost.[6,7] Digital health tools can promote value-based care to improve the health of populations (scalability), improve individual patient outcomes, and lower per capita costs.[1,8,9] The ongoing COVID-19 pandemic demonstrates the utility of digital health platforms such as telehealth.[10]

That the diagnosis of headache disorders is primarily based on patient history provides a valuable opportunity for developing and evaluating digital tools to diagnose and manage headache types. Headache disorders remain largely underdiagnosed and misdiagnosed, and thus are undertreated and mistreated.[11,12] Delays in accurate diagnosis lasting more than a decade lead to a higher risk for chronification and increased complexity in headache treatment.[13–16] Digital tools have the potential to close the gap between accurate diagnosis and increasing headache burden. Daily monitoring of headache incidence and related attributes using e-diary reduces diagnostic errors due to recall bias.[17,18] Automated prediction of headache attacks and medication overuse can be made by using data-driven machine learning approaches.[19,20] Wearables can identify real-time digital biomarkers for early detection of headache attacks (e.g. sleep time,[21] galvanic skin response,[21–24] heart rate variability[21,22]) and can also be used to improve adherence and deliver treatment such as smartphone-based biofeedback,[22–24] behavioral therapy,[25] and smartphone-controlled electronic pulses.[26]

The strong relationship between psychiatric comorbidities, traumatic brain injury, and headache disorders such as migraine is an important aspect of the growing headache burden.[27,28] Comprehensive evaluation of psychiatric disorders (e.g. depression, anxiety, post-traumatic stress disorder) is a critical component of proper headache management which can be optimized using digital health tools.[27–29] Digital health tools can solve unmet needs in these complex migraine comorbidities by providing early and accurate real-time identification of clinical symptoms or signs (e.g. quantitative electroencephalography/ EEG wearable headsets, transcranial Doppler, photoplethysmography for temporal artery pulse wave), as well as by aiding in monitoring symptom management.[29] Currently, there is active research involving development and validation of digital tools for traumatic brain injury and headaches by utilizing physiological biomarkers such as nystagmus, vestibular-ocular function, saccades, intracranial pressure, optic nerve sheath diameter, pupillary reflex, and reaction time.[29]

The diagnostic classification of headache disorders (International Classification of Headache Disorders, 3rd edition or ICHD-3[30]) is well suited for building artificial intelligence automated diagnoses. Patient-centered care and patient–physician shared decision-making are important features that are characteristic of digital headache tools.[31,32] Headache clinical trials and population research can benefit by utilizing digital health tools for screening, improved oversight, and real-time analysis of patient outcomes.[33,34] Personalized headache care, behavioral, and lifestyle-based management approaches can be enhanced using digital health applications.[35,36] The number of headache-trained healthcare providers is relatively low and poorly accessible to most headache sufferers.[37] With nearly a billion people estimated to suffer from migraine worldwide,[38] digital health technologies have the scalability potential to expand and improve access and outcomes for patients.[39] Migraine stigmatization, prevalent in headache care,[40] and migraine-specific disabilities present an opportunity for digital tools to serve patient populations not comfortable in seeking care through traditional face-to-face interviews[41] or unable to access care due to disease-specific limitations.

Although multiple computerized headache diagnostic tools have been developed and evaluated since the 1960s,[42] migraine ranks last among the eight most prevalent conditions globally in terms of available mobile health applications.[43] Migraine causes the highest disability in the productive age group.[44] Despite the development of many computerized migraine diagnostic and management digital tools, there is no systematic review that specifically summarized the accuracy, performance, and quality of published computerized migraine diagnostic digital tools. In this study, we conducted a systematic review to summarize and critically appraise the quality of published studies that developed and evaluated computerized migraine diagnostic tools. We hypothesized that there will be a significant variation in the accuracy

performance of the published migraine diagnostic tools.

Methods

Search methodology

The methods of systematic review followed the standard guidelines of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[45]. Literature search was performed using PubMed Advanced Search Builder, Web of Science Advanced Search, and Scopus Advanced Search using text-word searching. For PubMed Advanced Search Builder, the following search terms were combined using the Boolean operator 'AND': ((computerized) AND (diagnosis)) AND (migraine); ((algorithm) AND migraine)) AND (diagnosis); ((automated) AND (migraine)) AND (diagnosis). For Web of Science Advanced Search, the Web of Science Core Collection database was employed using the following search terms, Boolean operators, and field tags: TS=(computerized* AND migraine diagnosis); TS=(computerized* AND migraine); TS= (automated* AND migraine); TS=(algorithm* AND migraine). The field tag TS stands for topics. The asterisk * represents a wildcard in Web of Science for any group of characters, including no character. For Scopus Advanced Search, the following search terms were used: 'migraine AND computerized AND diagnosis'.

We employed PubMed, Web of Science, and Scopus because these three databases are the most comprehensive and most optimum resources for conducting systematic review of medical literature.[46,47] We did not include ScienceDirect as it covers publications by Elsevier only[48] – which are all available in Web of Science/PubMed/Scopus. Embase was not employed as its coverage is similar to PubMed's MEDLINE. PsycInfo is for specialized articles on behavioral studies;[47] we attempted PsycInfo search but could not find relevant articles. Additional search was done by snowballing and citation searching. Cutoff date for search was 1 June 2021.

The inclusion criterion was published articles and translated abstracts in English that evaluated a computerized/automated diagnostic tool for migraine. Some of the included studies involved evaluation of computerized digital tools for other headache types in addition to migraine – in such studies, only the migraine data were extracted for our systematic review. The study protocol was registered on PROSPERO (International Prospective Register of Systematic Reviews) [PROSPERO CRD42021033196]. The search results were uploaded to Covidence (http://covidence.org) for deduplication and screening. Two authors (YWW, RPC) screened references and reviewed full-text articles independently.

Data extraction

After reviewing the identified articles and deciding on inclusion, data were extracted and entered into a summary table by first author and publication year, digital tool name, development basis, sample size, concordance rate, sensitivity, specificity, reference diagnosis, strength, and limitations.

Summary analysis

Descriptive statistics [median, interquartile range (IQR), range, percentages] were used to summarize the following parameters: age, sex ratio, proportion of tools developed based on ICHD criteria, proportion of self-administered and physician/provider-administered, study settings, that is, percentage of tools validated in headache clinical centers *versus* nonclinical settings, percentage of reference diagnosis conducted using face-to-face interviews, and median diagnostic accuracy (concordance, sensitivity, specificity). Studies were categorized by decades to assess whether there was progressive increment in the number of digital tools developed. Median group differences (where group size imbalance is not larger than 1:3) were compared using the Mann–Whitney test. Spearman's nonparametric correlation between year of publication and accuracy performance measurands was examined to assess whether there was improvement in diagnostic accuracy over the successive years. A value of $p < 0.05$ was considered to be statistically significant.

Quality Assessment of Diagnostic Accuracy Studies

The Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2[49] tool was applied to critically appraise the quality of the included studies in terms of risk of bias and concern of applicability

Identification of studies via databases and registers

Identification of studies via other methods

Identification

Records identified from:
  Databases
  PubMed Advanced Search
  Builder (n = 434)
  Web of Science (n = 617)
  Scopus (n = 1813)

Records removed *before screening*:
  Duplicate records removed
  (n = 832)

Records identified from:
  Snowballing (n = 6)
  Citation searching (n = 15)
  ResearchGate, Google Scholar

Screening

Records screened
(n = 2032)

Records excluded
(n = 2008)

Reports sought for retrieval
(n = 24)

Reports not retrieved
(n = 0)

Reports sought for retrieval
(n = 21)

Reports not retrieved
(n = 0)

Reports assessed for eligibility
(n = 20)

Reports excluded:
  Diagnostic evaluation not
  included (n = 4)

Reports assessed for eligibility
(n = 21)

Reports excluded:
  (n = 0)

Included

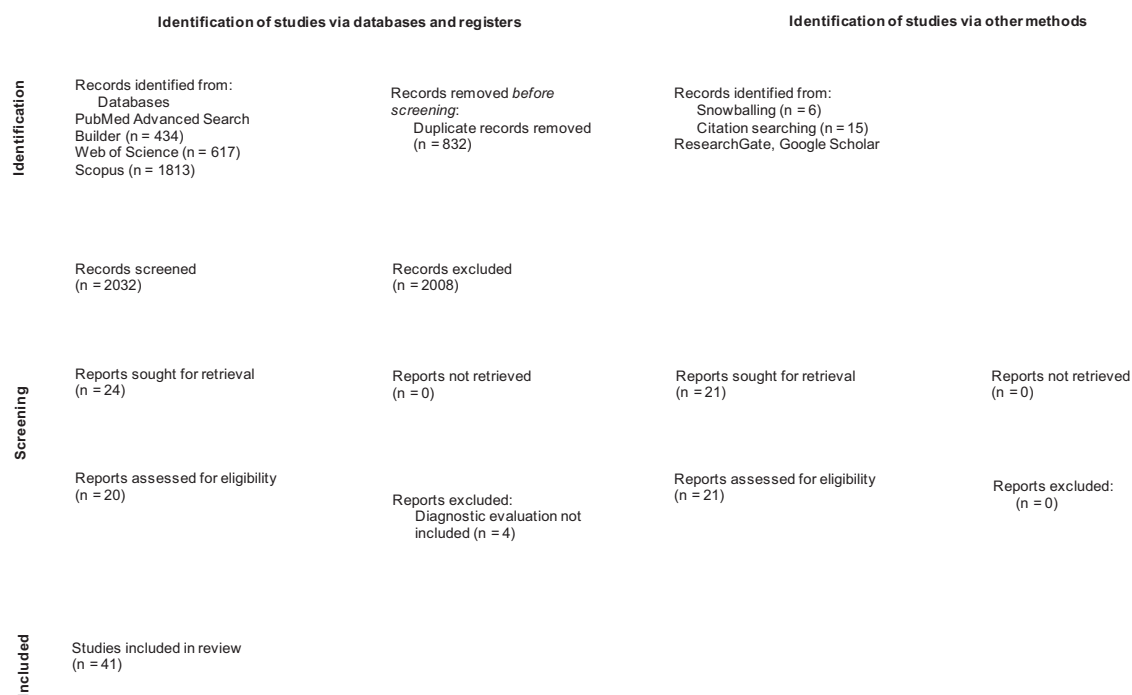Studies included in review
(n = 41)

Figure 1. PRISMA flowchart.
PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

over four key domains: patient selection, index test, reference standard, flow, and timing. Risk levels are classified as high, low, and unclear. See Supplemental File 1 for details on QUADAS assessment.

Results

Search results

The number of published studies identified by PubMed Advanced Search Builder on ((computerized) AND (diagnosis)) AND (migraine); ((algorithm) AND (migraine)) AND (diagnosis); ((automated) AND (migraine)) AND (diagnosis) was 155, 218, and 97, respectively. The number of published studies identified through Web of Science Core Collection Advanced Search database using the following search terms TS=(computerized* AND migraine diagnosis); TS=(computerized* AND migraine); TS=(automated* AND migraine); TS=(algorithm* AND migraine) was 58, 199, 138, and 294, respectively. The number of published studies identified through Scopus Advanced Search database using the terms 'migraine AND computerized AND diagnosis' was 1813. There were 6 studies identified through snowballing and 15 studies

through citation searching on ResearchGate and Google Scholar (Figure 1 PRISMA flowchart).

Included studies and summary results

After screening records, removing duplicates, and reviewing full-length articles, 41 studies (median age = 43 years; 77% women) were included in the systematic review (Figure 1 PRISMA flowchart; Supplemental File 2 and 3) as per the inclusion criteria. The median sample size of the studies was 288. Sixty percent of the digital tools were developed based on ICHD criteria. The remaining tools were developed based on local or national criteria (8), Wolff 1962 criteria (2), 1962 Ad Hoc criteria (1), case-based reasoning following ICHD-3-beta (1), artificial immune algorithm following ICHD-2 (1), deep learning framework (1), and ID-CM or Identify Chronic Migraine (1). Half (50%) of the tools were self-administered. Eighty-eight percent of tools were evaluated in headache clinical centers. The remaining were validated in nonclinical or community care settings, that is, employees (1), primary care (2), and students (1).

Face-to-face interviews were used in 82% of the reference diagnosis, phone interviews in two
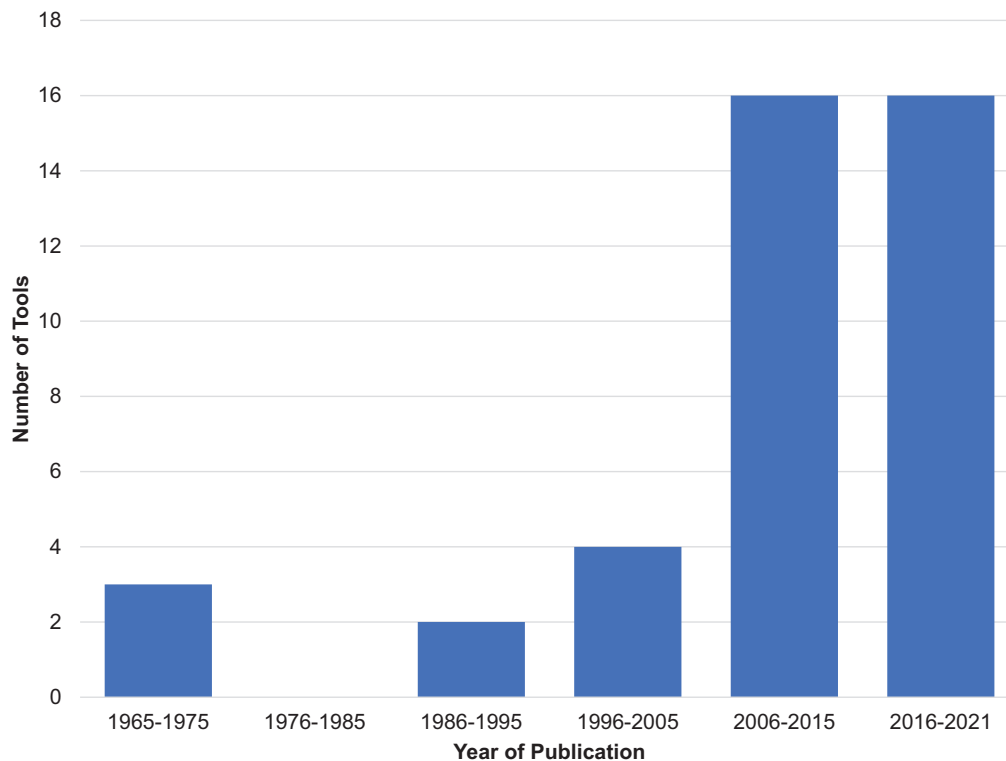
Figure 2. A decade-by-decade timeline of the number of computerized migraine diagnostic tools developed and validated since 1965. There was a 4.5-time increment in the number of computerized migraine diagnostic tools since 2005 compared with before 2005. There were six new tools developed and evaluated since the beginning of 2020.

studies, and questionnaire-based diagnosis in four studies. The majority (70%) of the digital tools were available in English. Five were available in Italian, four in Turkish, three in Chinese, one in Polish, one in Persian, two in Dutch, one in Serbian, one in Spanish, and one in Japanese. Some of the automated algorithms and machine learning programs involved case-based reasoning, deep learning, classifier ensemble, ant-colony, artificial immune, random forest, white and black box combinations, and hybrid fuzzy expert systems. Ten (25%) studies compared multiple machine learning programs to identify the most accurate tool in diagnosing migraine. Two of the digital tools are available as open-source software, and two other tools contained a daily electronic diary (e-diary). Management options were embedded in one of the digital tools.

There was a 4.5-time increase in the number of digital tools after 2005 compared with before (Figure 2). There were six new digital tools since the COVID-19 pandemic started in the beginning of 2020. The median diagnostic accuracy was concordance = 89% (IQR = 76–93%; range = 45–100%), sensitivity = 87% (IQR = 80–95%; range = 14–100%), and specificity = 90% (IQR = 77–96%; range = 65–100%). Five (12%) of the digital tools achieved diagnostic accuracy of 100% in one of the accuracy performance measurands, that is, sensitivity, specificity, and concordance. Lack of random patient sampling was observed; only 2 (5%) studies applied random sampling, while the majority 23 (56%) of the studies used convenience or consecutive sampling. Patient sampling method was unclear in 16 (39%) studies. There was no description of age or sex ratio of participants in 25 (61%) studies. There was no statistically significant correlation between year of publication and accuracy performance of the digital tools.

Twelve (30%) of the digital tools were developed specifically for migraine diagnosis (i.e. binary classification into 'migraine' or 'no migraine'), while the remainder 29 (70%) were developed to diagnose migraine as well as other primary headache disorders [i.e. multiclass classification into
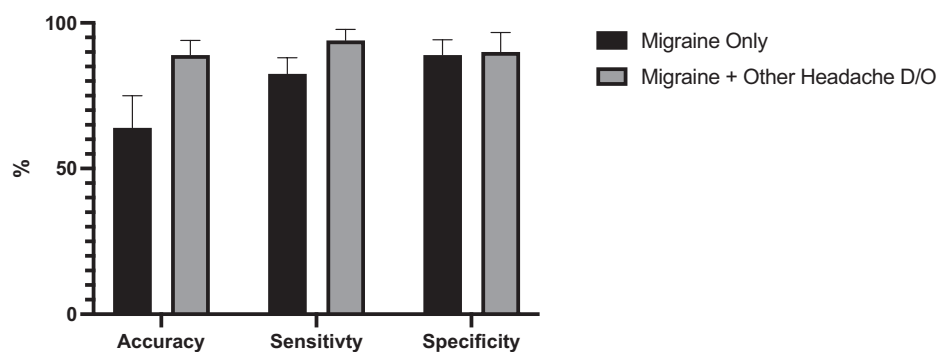
Figure 3. Comparison of diagnostic accuracy performance between binary and multiclass classification. The digital tools that were specifically designed for diagnosis of migraine only ('migraine' *versus* 'no migraine', i.e. binary classification) performed observably lower than the digital tools that were developed for multiclass classification of migraine and other headache disorders. The median values of sensitivity, specificity, and accuracy for the migraine-only classification digital tools (black bars) were 75%, 85%, and 61%, while for tools that diagnosed migraine and other headache disorders were 89%, 88%, and 88%. The differences did not reach statistical significance on the Mann–Whitney test.
D/O, disorders.

'migraine' or 'tension-type headache (TTH)' or 'trigeminal autonomic cephalalgias (TAC)' or 'other headache']. Although the accuracy performance of the migraine-only diagnostic tools (binary) was observably lower than the tools that diagnose migraine plus other headache disorders (multiclass), the difference was not statistically significant (Figure 3). We are currently conducting a separate systematic review appraising accuracy performance of digital diagnostic tools for other headache disorders (e.g. TTH, TACs) in addition to comparing accuracy performance of migraine diagnosis *versus* diagnosis of other headache disorders (e.g. TTH, TACs).

Brief description of performance of selected migraine diagnostic digital tools – a timeline

Since the 1960s, there have been efforts to create computer-generated diagnostic tools in headache.[42,50] One of the early computer-based self-administered diagnostic tools developed in 1968 correctly identified all nine headache patients as vascular headache, as compared with diagnoses made by neurologists with interest in headache.[42] A self-administered computer-based diagnostic tool developed in 1972 performed with sensitivity and specificity of 74% and 75% of diagnosing migraine, respectively, while correctly differentiating 36 of 50 patients to common migraine, classical migraine, cluster migraine, muscle contraction, and other headache types, as compared with diagnoses made by neurologists with interest in headache.[50] However, in 1974, another computerized

diagnostic tool performed considerably lower with sensitivity and specificity of 14% and 98%, respectively; the reason for its lower performance was considered to be due to lesser power of algorithms involving 19 headache types in its diagnostic capacity.[51] In 1980, a voice-based, self-administered, computer diagnostic study involving 40 headache patients[52] showed that 38 (95%) patients found the computer interview easy to use, while 22 (55%) patients thought that the computer took a complete headache history; 55% of the headache clinicians found the computer interview to be useful.[52] While these early studies on computer-based tools laid the foundation for self-administered headache diagnostic tools, the lack of standard headache classification and absence of widely available computers made it challenging to adopt their use for daily practice. Headache classification has evolved greatly through the decades; phrases such as vascular headache, common migraine, cluster migraine, and muscle contraction are currently outdated.

In the 2000s, the Italian Neurological Association for Headache Research developed and validated the Archivio Informatico con funzioni di Diagnosi Assistita per Cefalee – Informatic Register with Assisted Diagnosis for Headaches (AIDA CEFALEE) – computerized ICHD-2-based expert software that is filled by physicians and automatically generates a textual patient history, diagnosis, and treatment recommendations.[53,54] Its diagnostic performance was evaluated in 200 headache patients from five headache centers

against an ICHD-2-based structured clinical interview conducted by medical staff.[53,54] The AIDA CEFALEE[34,35] and the structured clinical interview showed 100% sensitivity and 76% specificity in diagnosing migraine.[53,54] In 2008, US researchers developed the self-administered Computerized Headache Assessment Tool (CHAT) based on the 1988 ICHD-1 and evaluated its diagnostic accuracy in 135 headache patients[55] using a reference of phone interview diagnoses by headache specialist nurse. The reference diagnosis was based on the 1994 UCSD (University of California San Diego) Migraine Questionnaire[56] and Silberstein–Lipton criteria.[57] The CHAT had an overall accuracy of 92.8% (91/98), that is, 91.6% (77/84) in migraine [100% (35/35) in episodic migraine, 85.7% (42/49) in transformed migraine] and 83.3% (5/6) in probable migraine. The reference used to validate the CHAT, that is, the Silberstein–Lipton criteria are currently outdated.

In 2014, Chinese researchers developed and validated the first ICHD-3-based, physician-administered, computerized clinical decision support systems (CDSS) in 543 patients against a reference diagnosis by 2 headache specialists.[58] The CDSS was designed to assist general practitioners based on the ICHD-3.[58] Its diagnostic sensitivity and specificity were 94% and 95%, respectively.[58] That the AIDA and CDSS were only available in Italian and Chinese, respectively, and had to be physician administered reduce their utility for direct patient use and self-management in anglophone populations. In 2015, a four-item migraine diagnostic algorithm, that is, the Migraine-4, was found to have a sensitivity and specificity of 94% and 92%, respectively, compared with a reference diagnosis made by structured self-administered questions.[59] The performance assessment of the Migraine-4 was conducted in a nonclinical population and its accuracy was not compared with diagnosis made by live interviews.[59]

In the last 5 years, Turkish,[60] Serbian,[61] Iranian,[62,63] German,[64] US,[65] and Dutch[66] researchers have evaluated different algorithms, machine learning expert systems in diagnosing migraine and other headache types. These algorithms involved ant-colony, artificial immune, classifier ensembles, white and black box combinations, and hybrid fuzzy expert systems. Their accuracy ranged from 65% to 100%. Some tools are available as open source, include electronic headache diary, and provide treatment options.

Limitations of these newly developed tools include the fact that they were based on non-ICHD criteria, retrospective analysis, and that they were not tested against live interviews. Some of these tools contained contamination where the reference diagnostician was exposed to diary data before interviews. The ongoing COVID-19 pandemic may have contributed to the rise in digital tools and telemedicine research in headache including mobile phone applications. In 2020, the M-sense app was developed based on ICHD-3 criteria and 3-monthly, patient-administered electronic headache diary.[64] The diagnostic accuracy of M-sense was validated in 102 patients against a headache specialist diagnosis using the same data collected by M-sense.[64] The M-sense showed 96.9% (62/64) sensitivity and 68.4% (26/38) specificity for migraine.[64]

QUADAS results

Consecutive or convenience sampling was used in 56% of the studies (Table 1, Supplemental File 4). The sampling method was unclear in 39% of the studies. Only 5% of the studies applied random sampling. Most studies (61%) mentioned consecutive or random sampling resulting in low risk of bias in patient selection (domain 1) according to QUADAS-2. Case–control design was avoided in all studies. The index tests (domain 2) showed low risk of bias (interpreted without knowledge of the reference test result) and low concern of applicability (the index test, its conduct, or interpretation did not differ from the review questions) in all the studies. Similarly, the majority (76%) of reference standard tests (domain 3) exhibited low risk of bias (the reference standard, its conduct, or its interpretation did not introduce bias) and low concern of applicability (the target condition as defined by the reference standard, i.e. migraine matches the review question). The majority, 83% of the studies, featured low risk of bias in the flow and timing (domain 4) of participating patients, that is, most patients received both the index and reference tests within appropriate interval. Percentage results are displayed in Supplemental File 4. Sample size estimation was lacking in all included studies.

Discussion

Our systematic review shows the progressive increment in the quality and quantity of computerized migraine diagnostic tools. The rise of digital health in headache medicine deserves special

Table 1. QUADAS assessment of included studies.

Study	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Freemon[42]	Low	Low	Unclear	Low	Low	Low	Unclear
Stead[50]	Unclear	Low	Low	Low	Low	Low	Low
Toole et al.[51]	Unclear	Low	Unclear	Low	Low	Low	Unclear
Penzien et al.[67]	Unclear	Low	Low	Low	Low	Low	Low
Andrew et al.[68]	Unclear	Low	Low	Low	Low	Low	Low
Gallai et al.[69]	Low	Low	Low	Low	Low	Low	Low
Pryse-Phillips et al.[70]	Unclear	Low	Low	Low	Low	Low	Low
Kopec et al.[71]	Unclear	Low	High	High	Low	Low	High
Sarchielli et al.[72]	Low	Low	Low	Low	Low	Low	Low
Sarchielli et al.[73]	Low	Low	Low	Low	Low	Low	Low
De Simone et al.[54]	Low	Low	Low	Low	Low	Low	Low
Maizels and Wolfe[55]	Low	Low	Low	High	Low	Low	Low
Mendes et al.[74]	Low	Low	Low	High	Low	Low	Low
Porta-Etessam et al.[75]	Low	Low	Low	Low	Low	Low	Low
van Oosterhout et al.[76]	Low	Low	Low	Low	Low	Low	Low
Tezel and Köse[77]	Unclear	Low	Unclear	High	Low	Low	Unclear
Yurtay et al.[78]	Low	Low	Low	High	Low	Low	Low
Krawczyk et al.[79]	Unclear	Low	Low	Low	Low	Low	Low
Eslami et al.[80]	Low	Low	Low	Low	Low	Low	Low
Dong et al.[58]	Low	Low	Low	Low	Low	Low	Low
Yanping and Huilong[81]	Unclear	Low	High	Low	Low	Low	High
Yin et al.[82]	Low	Low	Low	Low	Low	Low	Low
Jackowski et al.[83]	Unclear	Low	High	Low	Low	Low	High
Çelik et al.[60]	Unclear	Low	Low	Low	Low	Low	Low
Walters and Smitherman[59]	Low	Low	Low	Low	Low	Low	Low
Lipton et al.[65]	Low	Low	Low	Low	Low	Low	Low
Çelik et al.[84]	Low	Low	Low	Low	Low	Low	Low
Çelik and Yurtay[85]	Low	Low	Low	Low	Low	Low	Low

(Continued)

Table 1. (Continued)

Study	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Keight et al.[86]	Unclear	Low	High	Low	Low	Low	High
Vandewiele et al.[87]	Unclear	Low	Low	Low	Low	Low	Low
Kaiser et al.[88]	Low	Low	Low	High	Low	Low	Low
Khayamnia et al.[62]	Low	Low	Unclear	Low	Low	Low	Unclear
Sacco et al.[89]	Low	Low	Low	Low	Low	Low	Low
Qawasmeh et al.[63]	Unclear	Low	Low	Low	Low	Low	Low
Roesch et al.[64]	Low	Low	Low	Low	Low	Low	Low
Kwon et al.[90]	Low	Low	Unclear	Low	Low	Low	Unclear
Katsuki et al.[91]	Low	Low	Low	Low	Low	Low	Low
van Casteren et al.[66]	Low	Low	Low	High	Low	Low	Low
Simić et al.[61]	Unclear	Low	Unclear	Low	Low	Low	Unclear
Simić et al.[92]	Low	Low	Low	Low	Low	Low	Low
Groccia et al.[93]	Unclear	Low	Low	Low	Low	Low	Low

QUADAS, Quality Assessment of Diagnostic Accuracy Studies.
The risk of bias in the domain of patient selection was low in 56% (23) of the studies that used consecutive/convenience sampling and 5% (2) of the studies that applied random sampling. The sampling method was unclear in 39% (16) of the studies. The index test results were all interpreted without knowledge of the results of the reference standard. The conduct or interpretation of the index test did not introduce bias – hence resulting in 'low risk' and 'low concern' for 'Risk of Bias' and 'Applicability'. Likewise, most studies (76%) featured 'low risk' and 'low concern' in the conduct and interpretation of the reference tests. Similarly, most (83%) studies showed 'low risk' of bias in the flow and timing of patients, that is, all patients received both the index and reference tests at appropriate interval.
For each included study, Risk of bias and Applicability concerns are tabulated as "Low" (green shade), "Unclear" (blue shade), and "High" (orange shade).

attention because of shortage and/or inaccessibility of headache specialists for a growing headache burden worldwide. As artificial intelligence and machine learning tools are advancing rapidly, we anticipate a continued research in and adoption of computerized headache diagnostic and management tools. This is evidenced by the 4.5-time increase in the number of digital tools developed and validated in the last decade. The high median values of 87–90% for concordance, sensitivity, and specificity reflect the high level of accuracy and precision of computerized migraine diagnostic tools. That the ICHD criteria were used by most of the tools to build the automation algorithm is a worthy feature in terms of applying standardized migraine diagnosis.

That half of the computerized diagnostic tools were self-administered was a useful indicator that the tools can be completed at patient's convenience, saving time and empowering patient self-efficacy. The majority of the computerized tools were evaluated in a headache clinical center – a barrier for the generalizability of the results to community or primary care settings. Future studies may develop and conduct field testing of such tools designed for community level or nonclinical populations and for primary care settings.

Our findings show that the majority of the validation studies did not apply probability patient sampling (e.g. random patient selection). Nonrandom

patient recruitment (e.g. consecutive or convenience sampling) leads to a type of selection bias known as spectrum bias. Spectrum bias reflects the varying accuracy of diagnostic tools in different patients.[94] That case–control study design was avoided by all included studies and reflects the strong quality of the results.[49,95] By including extreme populations, that is, cases of chronic migraine *versus* controls of fully healthy participants, the performance of diagnostic accuracies can be more inflated in case–control designs than when including the average migraine patient *versus* typically healthy individuals.[95] The majority of the included studies measured the migraine diagnostic accuracy by enrolling patients with migraine and control participants that included patients with other headache types (e.g. TTH) as well as healthy individuals. Future studies can improve presentation of their results by providing sensitivity (e.g. subgroup) analysis to address heterogeneous patient and control populations.

The application of face-to-face interviews by the majority of the reference diagnosis was a strong feature, as a direct comparison between machine automation and human can be allowed. Some diagnostician interviewers possess naturally engaging personality and utilize memory jogging techniques to reduce recall bias during interviews.[96] Embedding e-diary in the digital tools can reduce patients' recall bias, as was seen in two of the included studies.[64,66] One of the digital tools includes a management option[58] – indicating the potential of the tools for headache care in terms of both diagnosis and management. Digital tools have been validated in several clinical trials as self-management therapies (e.g. neurofeedback, behavioral therapy) for migraine and other headache disorders.[22,23] Overall, the QUADAS quality assessments showed low risk of bias and low concern of applicability in all categories, which indicated the quality of the validation studies.

The expert systems utilized in the computerized migraine diagnostic tools involved purely data-driven approaches and hybrid expert diagnostic systems. Hybrid systems involved combination of different automation and algorithm approaches as well as combination of data- and knowledge-driven modalities. This indicates the versatility of the digital platform in incorporating multiple approaches. The marginal improvement we found in the accuracy performance of multiclass over binary classification digital tools may indicate that multiclass classification may capture multiple response variables of clinical features, thereby matching with refined headache phenotypes. However, this speculation involving potential differences in machine learning migraine diagnostic digital tools needs to be appropriately tested and validated.

Our systematic review was focused on migraine because migraine is the most commonly studied headache diagnosis. Some of the computerized diagnostic tools included (Table 1) are capable to simultaneously perform diagnosis of multiple headache disorders. It will be interesting to compare computerized migraine diagnostic accuracy with computerized diagnosis of other headache types.

Limitations

Because the studies included in this systematic review evaluated different digital tools, we were not able to conduct a weighted meta-analysis and derive a combined performance for diagnostic accuracy. Similarly, because of the interstudy heterogeneity of diagnostic automation software, it was not possible to undertake detailed sensitivity analysis such as publication bias, confounding or moderation, small-study bias, outliers, and cumulative meta-analysis among the different computerized migraine diagnostic tools. That patient selection method was missing from 39% of the included studies is another limitation. Studies that involved inappropriate exclusion criteria (e.g. overexclusion of difficult cases) may inflate the accuracy performance of a digital migraine diagnostic tool – only because confirmed or known migraine cases were enrolled, whereas 'difficult to diagnose' patients were excluded.[97,98] Vice versa, studies that involved overinclusion of 'difficult to diagnose' migraine cases may underestimate the accuracy performance of a digital migraine diagnostic tool – only because 'easy to diagnose' migraine patients were excluded.[97,98]

By virtue of incorporating the same ICHD question items in the index test (digital tools) and reference tests (interviews), validation of computerized migraine diagnostic tools is inherently faced with incorporation bias which can overestimate accuracy performance of the digital tools tested.[99–101] This circularity problem can only be solved by testing digital/computerized migraine diagnostic tools

that rely on digital biomarkers (e.g. heart rate variability, photoplethysmography) instead of self-report symptoms and signs, that is, when the index and reference tests contain parameters that are completely independent to each other (e.g. physiological/objective *versus* patient-reported parameters).[99–101] Our review was limited to studies published in English, although we have included translated abstracts. As such, we may have missed computerized migraine diagnostic tools published in journals with other languages.

Conclusion

The emergence of digital health in headache medicine is evident. Computerized migraine diagnostic tools have been shown to perform with diagnostic accuracy as high as 100% compared with traditional interview-based diagnosis by headache experts. Increased testing in different clinical settings and patient populations as well as random patient sampling and *a priori* sample size estimations will enhance generalizability and robustness of computerized migraine diagnostic tools for implementation toward daily use. By virtue of being a chronic condition whose episodic attacks can be triggered by lifestyle-related changes (e.g. sleep disruption, skipped mealtime) and featuring several lifestyle-related multimorbidities (e.g. depression, anxiety),[14,17,38] migraine care can benefit from digital health applications helping sufferers self-monitor and better manage their headache and related problems by reinforcing healthy lifestyle behavior.

Compared with traditional headache care model that is not capable of reaching the growing burden of a billion of migraine sufferers worldwide,[102] computerized migraine diagnostic tools have the potential to provide efficient, patient-centered, and improved headache care by delivering early diagnosis and management, enhancing diagnostic accuracy, saving time, boosting accessibility, enabling remote care with reduced costs, and decreasing travel to hospital/clinic care setting thereby reducing the exposure to communicable diseases in healthcare settings.[103] People with migraine can increase their self-efficacy and get a greater control of their health by using digital tools.[104] Personalized migraine care can be facilitated by digital tools.[105] Headache care providers can obtain a comprehensive objective overview of the patient including real-time monitoring via sensors and diary recording.[103] Virtual or augmented

reality treatment modalities can be tested using migraine digital health tools, for example, to reduce head pain perception[106] or to provide biofeedback sessions for associated psychological comorbidities.[107] Scalability as well as data security of digital migraine tools can be increased by using blockchain technology.[108] Ingestible sensors can be tested to measure gastroparesis, gut metabolomics, or gut–brain axis involvement in migraine.[109]

Author contributions

Yohannes W. Woldeamanuel: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Resources; Validation; Visualization; Writing – original draft; Writing – review & editing.
Robert P. Cowan: Data curation; Methodology; Writing – review & editing.

Conflict of interest statement

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

ORCID iD

Yohannes W. Woldeamanuel https://orcid.org/0000-0003-4879-6098

Supplemental material

Supplemental material for this article is available online.

References

1. Meskó B, Drobni Z, Bényei É, *et al.* Digital health is a cultural transformation of traditional healthcare. *mHealth* 2017; 3: 38.

2. Mitchell M and Kan L. Digital technology and the future of health systems. *Health Syst Reform* 2019; 5: 113–120.

3. Petersen CL, Weeks WB, Norin O, *et al*. Development and implementation of a person-centered, technology-enhanced care model for managing chronic conditions: cohort study. *JMIR mHealth uHealth* 2019; 7: e11082.

4. Finkelstein J, Knight A, Marinopoulos S, *et al*. Enabling patient-centered care through health information technology. *Evid Rep Technol Assess* 2012; 206: 1–1531 (Prepared by Johns Hopkins University Evidence-based Practice Center under Contract No 290- 2007-10061-I. AHRQ Publication No 12-E005-EF. Rockville, MD: Agency for Healthcare Research and Quality), http://www.ncbi.nlm.nih.gov/pubmed/24422882 (accessed 8 September 2021).

5. Ekman B. Cost analysis of a digital health care model in Sweden. *PharmacoEcon Open* 2018; 2: 347–354.

6. Doty MM, Tikkanen RS, FitzGerald M, *et al*. Income-related inequality in affordability and access to primary care in eleven high-income countries. *Health Aff* 2021; 40: 113–120.

7. Centers for Medicare & Medicaid Services. Historical, https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical (accessed 20 June 2021).

8. Rosalia RA, Wahba K and Milevska-Kostova N. How digital transformation can help achieve value-based healthcare: Balkans as a case in point. *Lancet Reg Health Eur* 2021; 4: 100100.

9. Putera I. Redefining health: implication for value-based healthcare reform. *Cureus* 2017; 9: e1067.

10. Monaghesh E and Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. *BMC Public Health* 2020; 20: 1193.

11. Viana M, Khaliq F, Zecca C, *et al*. Poor patient awareness and frequent misdiagnosis of migraine: findings from a large transcontinental cohort. *Eur J Neurol* 2020; 27: 536–541.

12. Cevoli S, D'Amico D, Martelletti P, *et al*. Underdiagnosis and undertreatment of migraine in Italy: a survey of patients attending for the first time 10 headache centres. *Cephalalgia* 2009; 29: 1285–1293.

13. Viticchi G, Falsetti L, Bartolini M, *et al*. Migraine: incorrect self-management for a disabling disease. *Neurol Int* 2018; 10: 24–27.

14. Buse DC, Rupnow MFT and Lipton RB. Assessing and managing all aspects of migraine: migraine attacks, migraine-related functional impairment, common comorbidities, and quality of life. *Mayo Clin Proc* 2009; 84: 422–435.

15. Adams AM, Serrano D, Buse DC, *et al*. The impact of chronic migraine: the Chronic Migraine Epidemiology and Outcomes (CaMEO) study methods and baseline results. *Cephalalgia* 2015; 35: 563–578.

16. Peres MFP, Swerts DB, de Oliveira AB, *et al*. Migraine patients' journey until a tertiary headache center: an observational study. *J Headache Pain* 2019; 20: 88.

17. Bandarian-Balooch S, Martin PR, McNally B, *et al*. Electronic-diary for recording headaches, triggers, and medication use: development and evaluation. *Headache* 2017; 57: 1551–1569.

18. Park J-W, Chu MK, Kim J-M, *et al*. Analysis of trigger factors in episodic migraineurs using a smartphone headache diary applications. *PLoS ONE* 2016; 11: e0149577.

19. Ferroni P, Zanzotto FM, Scarpato N, *et al*. Machine learning approach to predict medication overuse in migraine patients. *Comput Struct Biotechnol J* 2020; 18: 1487–1496.

20. Tasdelen B, Helvaci S, Kaleagasi H, *et al*. Artificial neural network analysis for prediction of headache prognosis in elderly patients. *Turk J Med Sci* 2009; 39: 5–12.

21. Siirtola P, Koskimäki H, Mönttinen H, *et al*. Using sleep time data from wearable sensors for early detection of migraine attacks. *Sensors* 2018; 18: 1374.

22. Minen MT, Corner S, Berk T, *et al*. Heartrate variability biofeedback for migraine using a smartphone application and sensor: a randomized controlled trial. *Gen Hosp Psychiatry* 2021; 69: 41–49.

23. Bembalgi V and Naik KR. Galvanic skin resistance (GSR) biofeedback in tension-type headache – Auditory, visual or combined feedback: which is beneficial? A randomized controlled trial. *Adv Physiother* 2012; 14: 123–131.

24. Bembalgi V and Naik KR. Comparative study on the efficacy of electromyography and galvanic skin resistance biofeedback in tension

type headache: a single blinded randomized controlled trial. *Int J Disabil Hum Dev* 2013; 12: 353–361.

25. Minen MT, Adhikari S, Seng EK, *et al.* Smartphone-based migraine behavioral therapy: a single-arm study with assessment of mental health predictors. *NPJ Digit Med* 2019; 22: 46.

26. Yarnitsky D, Dodick DW, Grosberg BM, *et al.* Remote electrical neuromodulation (REN) relieves acute migraine: a randomized, double-blind, placebo-controlled, multicenter trial. *Headache* 2019; 59: 1240–1252.

27. Dresler T, Caratozzolo S, Guldolf K, *et al.* Understanding the nature of psychiatric comorbidity in migraine: a systematic review focused on interactions and treatment implications. *J Headache Pain* 2019; 20: 51.

28. Lambru G, Benemei S, Andreou AP, *et al.* Position paper on post-traumatic headache: the relationship between head trauma, stress disorder, and migraine. *Pain Ther* 2021; 10: 1–13.

29. Schmid W, Fan Y, Chi T, *et al.* Review of wearable technologies and machine learning methodologies for systematic detection of mild traumatic brain injuries. *J Neural Eng* 2021; 18: 041006.

30. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. *Cephalalgia* 2018; 38: 1–211.

31. Dodick DW, Tepper SJ, Lipton RB, *et al.* Improving medical communication in migraine management: a modified Delphi study to develop a digital migraine tracker. *Headache* 2018; 58: 1358–1372.

32. Minen MT, Gumpel T, Ali S, *et al.* What are headache smartphone application (app) users actually looking for in apps: a qualitative analysis of app reviews to determine a patient centered approach to headache smartphone apps. *Headache* 2020; 60: 1392–1401.

33. Bensink M, Shah S, Shah N, *et al.* Tracking migraine digitally: the future of migraine management. *J Nurse Pract* 2021; 17: 462–470.

34. Diener HC, Ashina M, Durand-Zaleski I, *et al.* Health technology assessment for the acute and preventive treatment of migraine: a position statement of the International Headache Society. *Cephalalgia* 2021; 41: 279–293.

35. Kleiboer A, Sorbi M, Mérelle S, *et al.* Utility and preliminary effects of online digital assistance (ODA) for behavioral attack prevention in migraine. *Telemed J E Health* 2009; 15: 682–690.

36. Crawford MR, Luik AI, Espie CA, *et al.* Digital cognitive behavioral therapy for insomnia in women with chronic migraines. *Headache* 2020; 60: 902–915.

37. Mauser ED and Rosen NL. So many migraines, so few subspecialists: analysis of the geographic location of United Council for Neurologic Subspecialties (UCNS) certified headache subspecialists compared to United States headache demographics. *Headache* 2014; 54: 1347–1357.

38. Woldeamanuel YW and Cowan RP. Migraine affects 1 in 10 people worldwide featuring recent rise: a systematic review and meta-analysis of community-based studies involving 6 million participants. *J Neurol Sci* 2017; 372: 307–315.

39. Demaerschalk BM, Blegen RN and Ommen SR. Scalability of telemedicine services in a large integrated multispecialty health care system during COVID-19. *Telemed J E Health* 2021; 27: 96–98.

40. Young WB, Park JE, Tian IX, *et al.* The stigma of migraine. *PLoS ONE* 2013; 8: e54074.

41. Grist R, Porter J and Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. *J Med Internet Res* 2017; 19: e176.

42. Freemon FR. Computer diagnosis of headache. *Headache* 1968; 8: 49–56.

43. Martínez-Pérez B, de la Torre-Díez I and López-Coronado M. Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis. *J Med Internet Res* 2013; 15: e120.

44. Steiner TJ, Stovner LJ, Vos T, *et al.* Migraine is first cause of disability in under 50s: will health politicians now take notice? *J Headache Pain* 2018; 19: 17.

45. Page MJ, McKenzie JE, Bossuyt PM, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372: n71.

46. Singh VK, Singh P, Karmakar M, *et al.* The journal coverage of Web of Science, Scopus and Dimensions: a comparative analysis. *Scientometrics* 2021; 126: 5113–5142.

47. Bramer WM, Rethlefsen ML, Kleijnen J, *et al*. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst Rev* 2017; 6: 245.

48. https://www.elsevier.com/en-au/solutions/sciencedirect (accessed 20 October 2021).

49. Whiting PF. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; 155: 529–536.

50. Stead WW. Computer-assisted interview of patients with functional headache. *Arch Intern Med* 1972; 129: 950–955.

51. Toole JF, Brady WA, Cochrane CM, *et al*. Use of computerized questionnaire in the etiologic diagnosis of headache. *Headache* 1974; 14: 73–76.

52. Bana DS, Leviton A, Swidler C, *et al*. A computer-based headache interview: acceptance by patients and physicians. *Headache* 1980; 20: 85–89.

53. De Simone R, Marano E and Bonavita V. Towards the computerisation of ANIRCEF Headache Centres. Presentation of AIDA CEFALEE, a computer assisted diagnosis database for the management of headache patients. *Neurol Sci* 2004; 25(Suppl. 3): S218–S222.

54. De Simone R, Coppola G, Ranieri A, *et al*. Validation of AIDA Cefalee, a computer-assisted diagnosis database for the management of headache patients. *Neurol Sci* 2007; 28(Suppl. 2): S213–S216.

55. Maizels M and Wolfe WJ. An expert system for headache diagnosis: the Computerized Headache Assessment tool (CHAT). *Headache* 2008; 48: 72–78.

56. Tom T, Brody M, Valabhji A, *et al*. Validation of a new instrument for determining migraine prevalence: the UCSD Migraine Questionnaire. *Neurology* 1994; 44: 925–928.

57. Siberstein SD, Lipton RB, Solomon S, *et al*. Classification of daily and near-daily headaches: proposed revisions to the IHS criteria. *Headache* 1994; 34: 1–7.

58. Dong Z, Yin Z, He M, *et al*. Validation of a guideline-based decision support system for the diagnosis of primary headache disorders based on ICHD-3 beta. *J Headache Pain* 2014; 15: 40.

59. Walters AB and Smitherman TA. Development and validation of a four-item migraine screening algorithm among a nonclinical sample: the migraine-4. *Headache* 2016; 56: 86–94.

60. Çelik U, Yurtay N, Koç ER, *et al*. Diagnostic accuracy comparison of artificial immune algorithms for primary headaches. *Comput Math Methods Med* 2015; 2015: 465192.

61. Simić S, Banković Z, Villar JR, *et al*. A hybrid fuzzy clustering approach for diagnosing primary headache disorder. *Logic J IGPL* 2021; 29: 220–235.

62. Khayamnia M, Yazdchi M, Heidari A, *et al*. Diagnosis of common headaches using hybrid expert-based systems. *J Med Signals Sens* 2019; 9: 174–180.

63. Qawasmeh A, Alhusan N, Hanandeh F, *et al*. A high performance system for the diagnosis of headache via hybrid machine learning model. *Int J Adv Comput Sci Appl* 2020; 11: 655–663.

64. Roesch A, Dahlem MA, Neeb L, *et al*. Validation of an algorithm for automated classification of migraine and tension-type headache attacks in an electronic headache diary. *J Headache Pain* 2020; 21: 75.

65. Lipton RB, Serrano D, Buse DC, *et al*. Improving the detection of chronic migraine: development and validation of identify chronic migraine (ID-CM). *Cephalalgia* 2016; 36: 203–215.

66. van Casteren DS, Verhagen IE, de Boer I, *et al*. E-diary use in clinical headache practice: a prospective observational study. *Cephalalgia* 2021; 41: 1161–1171.

67. Penzien DB, Andrew ME, Knowlton GE, *et al*. Computer-aided system for headache diagnosis with the IHS headache diagnostic criteria: development and validation. *Cephalalgia* 1991; 11(Suppl. 11): 325–326.

68. Andrew ME, Penzien DB, Rains JC, *et al*. Development of a computer application for headache diagnosis: the headache diagnostic system. *Int J Biomed Comput* 1992; 31: 17–24.

69. Gallai V, Sarchielli P, Alberti A, *et al*. Application of the 1988 International Headache Society Diagnostic Criteria in Nine Italian Headache Centers using a computerized structured record. *Headache* 2002; 42: 1016–1024.

70. Pryse-Phillips W, Aube M, Gawel M, *et al*. A headache diagnosis project. *Headache* 2002; 42: 728–737.

71. Kopec D, Shagas G, Selman J, *et al*. Development of an expert system for aiding migraine diagnosis. *J Inf Technol Healthc* 2004; 2: 355–364.

72. Sarchielli P, Pedini M, Alberti A, *et al.* Application of ICHD 2nd edition criteria for primary headaches with the aid of a computerised, structured medical record for the specialist. *J Headache Pain* 2005; 6: 205–210.

73. Sarchielli P, Pedini M, Coppola F, *et al.* Application of the ICHD-II criteria to the diagnosis of primary chronic headaches via a computerized structured record. *Headache* 2007; 47: 38–44.

74. Mendes KB, Fiuza RM and Steiner MTA. Diagnosis of headache using artificial neural networks. *IJCSNS Int J Comput Sci Netw Secur* 2010; 10: 172–178.

75. Porta-Etessam J, López-Gil A, Matías-Guiu J, *et al.* [Validation of a diagnostic questionnaire for migraine adapted for a computer assisted telephone interviewing]. *Neurologia* 2010; 25: 163–167.

76. van Oosterhout WP, Weller CM, Stam AH, *et al.* Validation of the web-based LUMINA questionnaire for recruiting large cohorts of migraineurs. *Cephalalgia* 2011; 31: 1359–1367.

77. Tezel G and Köse U. Headache disease diagnosis by using the clonal selection algorithm. In: *6th international advanced technologies symposium (IATS'11), Elazig*, 16–18 May 2011, pp. 144–148. Elazig: Fırat University.

78. Yurtay N, Celik U and Yurtay Y. Headache diagnosis with K-means algorithm. In: AWER Procedia information technology & Computer Science, Antalya, November 2012, pp. 1074–1081. Istanbul: Academic World Education & Research Center.

79. Krawczyk B, Simić D, Simić S, *et al.* Automatic diagnosis of primary headaches by machine learning methods. *Open Med* 2013; 8: 157–165.

80. Eslami V, Rouhani-Esfahani S, Hafezi-Nejad N, *et al.* A computerized expert system for diagnosing primary headache based on International Classification of Headache Disorder (ICHD-II). *SpringerPlus* 2013; 2: 199.

81. Yanping W and Huilong D. An application of ant colony optimization clustering approach for primary headache diagnosis. In: Park J, Pan Y, Chao HC, *et al.* (eds) *Ubiquitous computing application and wireless sensor*. Dordrecht: Springer, 2015, pp. 643–648.

82. Yin Z, Dong Z, Lu X, *et al.* A clinical decision support system for the diagnosis of probable migraine and probable tension-type headache based on case-based reasoning. *J Headache Pain* 2015; 16: 29.

83. Jackowski K, Jankowski D, Simić D, *et al.* Migraine diagnosis support system based on classifier ensemble. In: Bogdanova A and Gjorgjevikj D (eds) *ICT innovations 2014*. Cham: Springer, 2015, pp. 329–339.

84. Çelik U, Yurtay N, Koc ER, *et al.* Migraine, tension-type and cluster-type of headaches classification by using immunos algorithms. *J Med Imaging Health Inform* 2016; 6: 1173–1177.

85. Çelik U and Yurtay N. An ant colony optimization algorithm-based classification for the diagnosis of primary headaches using a website questionnaire expert system. *Turk J Electr Eng Comput Sci* 2017; 25: 4200–4210.

86. Keight R, Aljaaf AJ, Al-Jumeily D, *et al.* An intelligent systems approach to primary headache diagnosis. In: Huang DS, Jo KH and Figueroa-García J (eds) *Intelligent computing theories and application*. Cham: Springer, 2017, pp. 61–72.

87. Vandewiele G, de Backere F, Lannoye K, *et al.* A decision support system to follow up and diagnose primary headache patients using semantically enriched data. *BMC Med Inform Decis Mak* 2018; 18: 98.

88. Kaiser EA, Igdalova A, Aguirre GK, *et al.* A web-based, branching logic questionnaire for the automated classification of migraine. *Cephalalgia* 2019; 39: 1257–1266.

89. Sacco S, Benemei S, Cevoli S, *et al.* Development and validation of the ID-EC – the ITALIAN version of the identify chronic migraine. *J Headache Pain* 2019; 20: 15.

90. Kwon J, Lee H, Cho S, *et al.* Machine learning-based automated classification of headache disorders using patient-reported questionnaires. *Sci Rep* 2020; 10: 14062.

91. Katsuki M, Narita N, Matsumori Y, *et al.* Preliminary development of a deep learning-based automated primary headache diagnosis model using Japanese natural language processing of medical questionnaire. *Surg Neurol Int* 2020; 11: 475.

92. Simić S, Villar JR, Calvo-Rolle JL, *et al.* An application of a hybrid intelligent system for diagnosing primary headaches. *Int J Environ Res Public Health* 2021; 18: 1890.

93. Groccia M, Guido R, Conforti D, *et al.* A knowledge-based clinical decision support system for headache disorders management. In: *Proceedings of the 14th international joint conference on Biomedical Engineering Systems*

and Technologies, Vienna, 11–13 February, 2021.

94. Willis BH. Spectrum bias – why clinicians need to be cautious when applying diagnostic test studies. *Fam Pract* 2008; 25: 390–396.

95. Rutjes AW, Reitsma JB, Vandenbroucke JP, *et al*. Case–control and two-gate designs in diagnostic accuracy studies. *Clin Chem* 2005; 51: 1335–1341.

96. Bowling A. Mode of questionnaire administration can have serious effects on data quality. *J Public Health* 2005; 27: 281–291.

97. Whiting P, Rutjes AWS, Reitsma JB, *et al*. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Ann Intern Med* 2004; 140: 189–202.

98. Lijmer JG, Mol BW, Heisterkamp S, *et al*. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999; 282: 1061–1066.

99. Mower WR. Evaluating bias and variability in diagnostic test reports. *Ann Emerg Med* 1999; 33: 85–91.

100. Rutjes AWS, Reitsma JB, di Nisio M, *et al*. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ* 2006; 174: 469–476.

101. Worster A and Carpenter C. Incorporation bias in studies of diagnostic tests: how to avoid being biased about bias. *CJEM* 2008; 10: 174–175.

102. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017; 390: 1211–1259.

103. Coravos A, Goldsack JC, Karlin DR, *et al*. Digital medicine: a primer on measurement. *Digit Biomark* 2019; 3: 31–71.

104. Newby K, Teah G, Cooke R, *et al*. Do automated digital health behaviour change interventions have a positive effect on self-efficacy? A systematic review and meta-analysis. *Health Psychol Rev* 2021; 15: 140–158.

105. Cancela J, Charlafti I, Colloud S, *et al*. Digital health in the era of personalized healthcare: opportunities and challenges for bringing research and patient care to a new level. In: Syed-Abdul S, Zhu X and Fernandez-Luque L (eds) *Digital health*. Amsterdam: Elsevier, 2021, pp. 7–31.

106. de Tommaso M, Ricci K, Laneve L, *et al*. Virtual visual effect of hospital waiting room on pain modulation in healthy subjects and patients with chronic migraine. *Pain Res Treat* 2013; 2013: 515730.

107. Kim H, Kim DJ, Kim S, *et al*. Effect of virtual reality on stress reduction and change of physiological parameters including heart rate variability in people with high stress: an open randomized crossover trial. *Front Psychiatry* 2021; 12: 614539.

108. Bell L, Buchanan WJ, Cameron J, *et al*. Applications of blockchain within healthcare. *Blockchain Healthc Today* 2018; 1: 1–7.

109. Kalantar-Zadeh K, Ha N, Ou JZ, *et al*. Ingestible sensors. *ACS Sens* 2017; 2: 468–483.