

Preliminary Diagnostic Accuracy and Usability of a Rule-Based mHealth Application for Differentiating Migraine and Tension-Type Headache: A Pilot Study

Abdelhameed M. A. Elshenawy^{1,2*}, Amine Hajaj³, Laith Mohammed³, Mariam Mohammed³, Fatima Aldanhany³, Farida M. Habib³

¹Fujairah University, 2202, Fujairah, United Arab Emirates

²Cairo University, Giza, 12613, Egypt

³College of Dentistry and Health Sciences, Fujairah University, 1207, Fujairah, United Arab Emirates

*Corresponding Author Email id: Abdelhameed_mahros@cu.edu.eg

ABSTRACT

Background: Headache disorders are common causes of disability, and differentiating migraine from Tension-Type Headache (TTH) can be challenging because of overlapping clinical features. Mobile Health (mHealth) applications may support structured symptom assessment, but diagnostic performance should be established before such tools are incorporated into clinical triage. **Objectives:** This pilot study evaluated the preliminary diagnostic accuracy and inter-method agreement of a rule-based mHealth application for differentiating migraine from TTH using International Classification of Headache Disorders, 3rd edition (ICHD-3) criteria and assessed application usability. In this study, diagnostic accuracy was assessed against an independently established clinical reference diagnosis, while agreement between the application and the clinical reference diagnosis was assessed using Cohen's kappa. **Methods:** A quantitative, cross-sectional pilot study included 73 participants recruited at the University of Fujairah. The application used a deterministic, logic-based weighted algorithm derived from ICHD-3 symptom criteria. Diagnostic accuracy was evaluated in a subset of 25 participants. After completing the application assessment, these participants attended clinical assessments by physicians from neurology and medicine, who independently applied ICHD-3 criteria without access to the application classification. Sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), overall accuracy, and Cohen's kappa were calculated. Reliability in this pilot study refers to agreement with the independent clinical classification as quantified by Cohen's kappa; test-retest reproducibility was not assessed. Usability was assessed using the System Usability Scale (SUS). **Results:** In the validation subset, overall accuracy was 88.0% (95% CI: 68.8%-97.5%), sensitivity was 80.0% (95% CI: 44.4%-97.5%), specificity was 93.3% (95% CI: 68.1%-99.8%), PPV was 88.9% (95% CI: 51.8%-99.7%), and NPV was 87.5% (95% CI: 61.7%-98.4%). Agreement with the clinical reference diagnosis was substantial (Cohen's kappa = 0.74). The mean SUS score was 78.6 (SD = 12.3). The wide confidence intervals indicate limited precision because of the small validation subset. **Conclusion:** The application demonstrated promising preliminary diagnostic accuracy and inter-method agreement with an independently established clinical diagnosis, together with satisfactory usability in this pilot sample. Because the clinical validation subset was small and confidence intervals were wide, these findings should be interpreted cautiously and should not be generalized beyond similar populations. Larger studies are required to confirm diagnostic accuracy and to evaluate test-retest and other forms of reproducibility before routine clinical implementation.

Keywords: Diagnostic Accuracy; Headache Diagnosis; Migraine; Mobile Health; Nursing; Tension-Type Headache; Usability

INTRODUCTION

Headache disorders are among the most prevalent neurological conditions and contribute substantially to disability worldwide. Migraine accounts for a particularly high burden of years lived with disability, while TTH

Received: January 10, 2026 Received in revised form: August 17, 2026 Accepted: August 27, 2026

is highly prevalent across populations (Mokdad, 2021; Sait *et al.*, 2021). Updated global estimates indicate that TTH affected approximately 2.0 billion people and migraine affected approximately 1.2 billion people in 2021, with migraine associated with disproportionately greater disability despite lower prevalence (Wijeratne *et al.*, 2025). Although migraine and TTH are classified using established clinical criteria, overlapping symptoms can complicate differentiation in routine assessment (Arnold, 2018; Robbins, 2021).

Nurses contribute to headache history taking, symptom assessment, patient education, monitoring, and coordination of care. Recent European consensus work emphasizes a structured nursing role within headache services, and contemporary nursing literature similarly highlights the importance of standardized assessment and patient-centered support (Rasmussen *et al.*, 2024; Pini *et al.*, 2024). Beyond headache-specific services, broader nursing evidence indicates that nurse-led remote digital support can improve self-management and satisfaction with care across chronic conditions (Kilfoy *et al.*, 2025), and that nurses play a key role in helping patients interpret and act on digitally collected symptom data (Wathne *et al.*, 2025). Digital decision-support tools may be particularly useful when they help organize symptom information without replacing clinical judgment.

Mobile health applications are increasingly used for headache diaries, symptom tracking, behavioral support, and communication. Earlier reviews found that mHealth approaches for headaches were feasible but supported by limited evidence, while more recent studies have examined telemedicine, digital self-management, and app usability in headache care (Stubberud & Linde, 2018; Young *et al.*, 2025). Computerized diagnostic tools have also shown variable accuracy, with methodological concerns regarding sampling, reference standards, and incorporation bias remaining important (Woldeamanuel & Cowan, 2022; Kim *et al.*, 2022). Broader audits of consumer-facing symptom-checker tools have likewise reported inconsistent diagnostic and triage accuracy, emphasizing the necessity of independent validation of algorithm-based classification tools before clinical use (Semigran *et al.*, 2015).

The application evaluated in this study is not an artificial intelligence (AI) or machine-learning system. It uses a deterministic, logic-based weighted algorithm that applies predefined symptom rules derived from ICHD-3 criteria. This distinction is important because current AI-based headache research involves learned models and predictive methods that require different development and validation procedures (Petrušić *et al.*, 2025; Espinoza-Vinces *et al.*, 2025).

Aim of the Study

The study aimed to evaluate the preliminary diagnostic accuracy and inter-method agreement (as defined above) of a developed rule-based mHealth application for differentiating migraine from TTH, and to assess its usability in a pilot sample. Test-retest reliability and longitudinal reproducibility were not assessed.

Research Questions

What is the preliminary diagnostic accuracy of the application for differentiating migraine from TTH, as measured by sensitivity, specificity, PPV, NPV, and overall percent agreement against the independent clinical reference diagnosis?

What is the level of agreement between the application classification and the independent clinical reference diagnosis, as measured by Cohen's kappa?

How do participants rate the usability and acceptability of the application using the System Usability Scale?

METHODOLOGY

Study Design

A quantitative, cross-sectional pilot design was used to obtain preliminary estimates of diagnostic performance and usability. Cross-sectional designs are appropriate for assessing variables at a defined point, whereas pilot studies are primarily intended to assess feasibility and generate preliminary parameter estimates rather than provide definitive evidence of diagnostic performance (Setia, 2016; Hertzog, 2008). Consensus-

based frameworks similarly distinguish pilot studies, which test the feasibility of study procedures, from feasibility studies more broadly (Eldridge *et al.*, 2016). Diagnostic-accuracy reporting followed STARD principles, particularly the need to define the index test, reference standard, participant selection, and uncertainty around accuracy estimates (Cohen *et al.*, 2016).

Sample Size and Rationale

A total of 73 participants were recruited for the pilot study; sample size followed published pilot-study guidance rather than a definitive diagnostic-validation calculation (Julious, 2005; Teare *et al.*, 2014). Diagnostic-accuracy studies commonly use a partial-verification design, in which the reference standard is applied to a subset of participants rather than the full sample, because independent clinical assessment is more resource-intensive than the index test itself (Begg & Greenes, 1983; de Groot *et al.*, 2011). All 73 participants completed the application-based index assessment; diagnostic performance was evaluated in the 25 participants for whom an independent, clinician-confirmed reference diagnosis was obtained. This subset is not necessarily representative of the full sample, and the resulting estimates should be treated as exploratory rather than definitive, pending future verification of the complete sample.

Setting and Participants

The study was conducted at the University of Fujairah using convenience sampling from nursing and health sciences settings. Participants were primarily university students and health professionals. Eligibility criteria included age 18 years or older, ownership of a smartphone with internet access, and a history of recurrent headaches consistent with migraine or TTH during the preceding three months. Individuals with suspected secondary headaches or incomplete study data were excluded. Because recruitment was concentrated in a young, health-educated population, the sample was not intended to represent the broader population of patients with migraine or TTH.

Instrumentation

The study used a purpose-built smartphone application developed by information-technology experts in collaboration with clinicians. The application included the following components:

Diagnostic Algorithm

Classification was performed by a deterministic, logic-based weighted algorithm incorporating predefined ICHD-3 symptom criteria, including pain quality, duration, laterality, aggravation by physical activity, and associated symptoms (Arnold, 2018). The application did not use artificial intelligence, machine learning, model training, or adaptive prediction.

Usability Measure

The System Usability Scale (SUS) is a widely used ten-item instrument for evaluating perceived usability across diverse products and systems, and normative score bands supported its interpretation in this study (Bangor *et al.*, 2008). The SUS was administered immediately after use to provide a standardized measure of perceived usability. Similar usability approaches have been used in recent migraine app development and implementation studies (Young *et al.*, 2023; Young *et al.*, 2025).

Reference Standard

For the 25 participants in the validation subset, the clinical reference diagnosis was established after completion of the application assessment. Following use of the application, participants attended the clinics and were assessed by physicians from neurology and medicine. The physicians were not informed of the application-generated classification. During the clinical assessment, the physicians independently evaluated the headache presentation and independently applied ICHD-3 diagnostic criteria to establish the reference diagnosis. Accordingly, the reference diagnosis was determined after the index application assessment and independently of the application result, reducing the risk of direct diagnostic incorporation bias.

Data Collection Procedure

Data collection was completed in late March 2025. After providing informed consent, participants

completed a demographic profile and entered headache symptoms in real time or shortly after an episode. The application generated a provisional classification that was stored in a secure central database for analysis. Digital headache diaries and mHealth platforms have been used to capture symptom and self-management data in headache populations (Minen *et al.*, 2020; Mosadeghi-Nik *et al.*, 2016).

A partial blinding approach was used for the validation subset. The application-generated classification was not visible to participants at the time of the clinical assessment, and the clinician establishing the reference diagnosis was not informed of the application classification. Data used for statistical analysis were coded without participant identifiers. This design reduced the potential for direct diagnostic contamination, although the limitations of the reference standard remain relevant.

Ethical Considerations

The research proposal was approved by the University of Fujairah Research, United Arab Emirates, with Ethics Approval No. EA#12 and committee meeting reference CHS-CRP-SA-IW #5, dated 12th February 2025. The approval period was from 5th February 2025 to 5th February 2027. Written informed consent was obtained from all participants prior to participation, and confidentiality was maintained through data de-identification and restricted access to study data.

Data Analysis

Data was analyzed using PSPP software. Diagnostic accuracy in the validation subset was summarized using sensitivity, specificity, PPV, NPV, and overall percent agreement with exact 95% binomial confidence intervals. Agreement between the application classification and the independently established clinical diagnosis was assessed using Cohen's kappa, interpreted using established benchmarks for interrater agreement (McHugh, 2012). Descriptive statistics were used for participant characteristics and SUS scores. The confirmed 2 x 2 validation counts were 8 true-positive, 14 true-negative, 1 false-positive, and 2 false-negative classifications; these counts were used to derive PPV and NPV. The symptom variables used in the application's diagnostic algorithm are structural components of its classification rule rather than statistically independent predictors of its output, so predictor-level regression modelling of these variables was outside the scope of this analysis.

RESULTS

Participant Characteristics

A total of 73 participants completed the study. The mean age was 22.1 years (SD = 7.4). Table 1 presents demographic counts and percentages for all 73 participants. Gender and participant-role items were collected as optional fields in the self-report demographic questionnaire, consistent with standard practice for non-mandatory demographic disclosure, and non-response did not affect eligibility, application use, or inclusion in the diagnostic or usability analyses. One participant did not disclose gender and is categorized as not reported. Thirteen participants did not identify their role as university students or health professional; because no further role category was offered on the questionnaire, these responses are categorized as other/not reported.

Table 1: Participant Demographics

Characteristic	Value (N = 73)
Age (years), mean ± SD	22.1 ± 7.4
Gender: Female	46 (63.0%)
Gender: Male	26 (35.6%)
Gender: Not reported	1 (1.4%)
Participant role: University student	52 (71.2%)
Participant role: Health professional	8 (11.0%)
Participant role: Other/not reported	13 (17.8%)
Family history of headache	19 (26.0%)

Preliminary Diagnostic Performance

Diagnostic performance was evaluated in the 25-participant validation subset. The confirmed validation counts were 8 true-positive, 14 true-negative, 1 false-positive, and 2 false-negative classifications. These

counts are consistent with the reported sensitivity, specificity, and overall accuracy and were used to calculate PPV and NPV. Table 2 presents the preliminary diagnostic-accuracy estimates and corresponding confidence intervals. The wide intervals, particularly for sensitivity and PPV, indicate substantial statistical uncertainty; therefore, the findings should be interpreted as preliminary evidence rather than definitive validation.

Table 2: Preliminary Diagnostic Performance in the Validation Subset ($n = 25$)

Metric	Estimate (95% CI)
Overall accuracy	88.0% (68.8%-97.5%)
Sensitivity	80.0% (44.4%-97.5%)
Specificity	93.3% (68.1%-99.8%)
Positive predictive value	88.9% (51.8%-99.7%)
Negative predictive value	87.5% (61.7%-98.4%)
Cohen's kappa	0.74 (substantial agreement)

Usability and User Acceptance

The mean SUS score was 78.6 (SD = 12.3). In the post-use survey, 93.2% of participants reported that the application was easy to use, 89.0% indicated willingness to use it frequently for symptom tracking, and 8.2% reported a need for technical support. These findings describe acceptability in the study sample and may not generalize to users with lower digital or health literacy.

DISCUSSION

This pilot study provides preliminary evidence supporting the diagnostic accuracy and inter-method agreement of a deterministic ICHD-3-based mHealth application for differentiating migraine from TTH. Accuracy was reflected in diagnostic accuracy estimates against an independently established clinical reference diagnosis, while inter-method agreement between the application and the clinical classification was substantial. The point estimates for accuracy and specificity were encouraging; however, the confidence intervals were wide, and the sensitivity interval was particularly imprecise. Accordingly, the findings should be interpreted as preliminary rather than definitive evidence. This cautious interpretation is consistent with systematic reviews showing that computerized migraine diagnostic tools can achieve useful accuracy but remain sensitive to study design, reference-standard quality, and sampling methods (Woldeamanuel & Cowan, 2022; Kim *et al.*, 2022).

The reference standard is a central methodological issue in diagnostic-accuracy studies. In the present study, clinical diagnosis was established after the application assessment by physicians from neurology and medicine who were blinded to the application classification and independently applied ICHD-3 criteria. This temporal sequence and blinding strengthened the independence of the reference diagnosis. Nevertheless, diagnostic-accuracy studies require careful interpretation whenever the index test and reference assessment draw on related clinical symptom criteria (Cohen *et al.*, 2016). Reviews of computerized migraine diagnosis similarly identify incorporation and selection biases as important methodological considerations (Woldeamanuel & Cowan, 2022).

The usability findings support continued evaluation of the application as a structured symptom-assessment tool. Recent mHealth research has shown that app-based care plans and digital tools can be feasible and acceptable when they are designed around user needs and integrated into clinical workflows (Young *et al.*, 2023; Young *et al.*, 2025). Broader evidence also supports the feasibility of telemedicine and digital self-management in headache care, while emphasizing that implementation quality, adherence, and clinical context influence outcomes (Noser *et al.*, 2022; Huang *et al.*, 2024). In nursing practice, digital tools may assist structured history taking and documentation, but they should complement rather than replace clinical assessment, education, and referral pathways (Rasmussen *et al.*, 2024; Pini *et al.*, 2024).

The application uses a rule-based algorithm; no machine-learning model was trained or tested in this study. This distinguishes the present application from emerging AI systems for headache classification, prediction, and clinical decision support, which involve learned models and different development and validation procedures. Recent reviews describe rapid growth in AI applications for headache medicine but also

emphasize the need for external validation, transparent methodology, representative datasets, and careful evaluation before clinical deployment (Petrušić *et al.*, 2025; Espinoza-Vinces *et al.*, 2025).

Generalizability is limited by convenience sampling and the demographic composition of the sample. Participants were predominantly young university students and health professionals, groups likely to have greater health literacy and familiarity with symptom terminology than many patients encountered in routine practice. Consequently, both usability and diagnostic performance may differ in older adults, individuals with lower health or digital literacy, and patients recruited from neurology or primary-care clinics. Digital headache research increasingly emphasizes representative sampling and real-world evaluation as tools that move from feasibility studies to clinical implementation (Young *et al.*, 2025; Zhang *et al.*, 2023).

Limitations

Several limitations should be considered. First, only 25 of the 73 participants contributed to the diagnostic comparison; this partial-verification design, common in diagnostic pilot studies, resulted in wide confidence intervals, limited precision, and the possibility that the validation subset is not fully representative of the broader sample. Second, the convenience sample was predominantly young and health educated, limiting external validity. Third, although the clinical reference diagnosis was established after the application assessment by physicians from neurology and medicine who independently applied ICHD-3 criteria while blinded to the application result, the study did not assess agreement between different clinicians or repeat the reference assessment. Fourth, the cross-sectional design assessed a single diagnostic comparison and did not evaluate test-retest reliability, longitudinal stability, or other forms of reproducibility. Fifth, the study did not include an independent external validation sample. Finally, the symptom variables underlying the rule-based algorithm are structural components of its classification rule rather than statistically independent predictors, which precludes predictor-level regression modelling of the application's output.

Future Scope

Future research should evaluate the application prospectively in larger and more diverse clinical populations using standardized, independently applied ICHD-3 reference diagnoses. Repeated assessments across multiple headache episodes would permit formal evaluation of test-retest reliability, longitudinal stability, and reproducibility beyond the inter-method agreement examined in the present pilot study. Subsequent development could explore longitudinal digital phenotyping and optional integration of wearable-derived measures, but these extensions should be evaluated separately and should not be presented as validated features of the current application.

CONCLUSION

This pilot study provides preliminary evidence of diagnostic accuracy and inter-method agreement for a deterministic, ICHD-3-based mHealth application differentiating migraine from tension-type headache, together with satisfactory usability, in a young, health-educated convenience sample. Because the clinical validation subset was small ($n = 25$) and the corresponding confidence intervals were wide, these findings should be interpreted as preliminary rather than as evidence of validated diagnostic performance suitable for routine clinical use. On this basis, the application is best regarded as a research-stage decision-support tool rather than a validated stand-alone diagnostic instrument. Progression toward clinical use requires an adequately powered diagnostic-accuracy study with a larger and more clinically representative sample, an independently established reference diagnosis, and formal assessment of test-retest reliability and other forms of reproducibility beyond the single-comparison design used here. This pilot evaluated the algorithm's diagnostic performance and general usability rather than a nurse-delivered intervention or a nursing-sensitive outcome; because nurses are central to headache history-taking, symptom assessment, and patient education, future studies should evaluate the application when used directly by nurses during triage or symptom assessment, so that its relevance to nursing workflow and decision-making can be established directly rather than inferred from algorithm performance alone. If future versions incorporate machine-learning or predictive components, these should undergo separate development and both internal and external validation before any clinical claims are made.

CRediT Authorship Contribution Statement

A.M.A.E.: Conceptualization, Methodology, Writing - original draft. **A.H.:** Investigation, Writing - original draft. **L.M.:** Investigation, Methodology. **M.M.:** Investigation, Writing - original draft. **F.A.:** Software. **F. M. H.:** Formal analysis, Methodology, Writing - review and editing.

AI Assistance Declaration

During preparation of this manuscript, OpenAI ChatGPT was used to assist with language enhancement, grammar correction, academic phrasing, structural editing, and reference cross-checking. Before submission, the authors reviewed all AI-assisted content, verified the accuracy of the scientific content and references, and took full responsibility for the final manuscript.

Conflict of Interest

The authors declare no conflict of interest.

ACKNOWLEDGEMENT

The authors acknowledge the study participants for their contributions and acknowledge individuals who provided language support, proofreading, and statistical advice during manuscript preparation.

REFERENCES

- Arnold, M. (2018). Headache classification committee of the international headache society (IHS) the international classification of headache disorders. *Cephalalgia*, 38(1), 1-211. <https://doi.org/10.1177/0333102417738202>
- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24(6), 574-594. <https://doi.org/10.1080/10447310802205776>
- Begg, C. B., & Greenes, R. A. (1983). Assessment of diagnostic tests when disease verification is subject to selection bias. *Biometrics*, 39(1), 207-215. <https://doi.org/10.2307/2530820>
- Cohen, J. F., Korevaar, D. A., Altman, D. G., Bruns, D. E., Gatsonis, C. A., Hooft, L., ... & Bossuyt, P. M. (2016). STARD 2015 guidelines for reporting diagnostic accuracy studies: Explanation and elaboration. *BMJ Open*, 6(11), e012799. <https://doi.org/10.1136/bmjopen-2016-012799>
- de Groot, J. A. H., Bossuyt, P. M. M., Reitsma, J. B., Rutjes, A. W. S., Dendukuri, N., Janssen, K. J. M., & Moons, K. G. M. (2011). Verification problems in diagnostic accuracy studies: consequences and solutions. *BMJ*, 343, d4770. <https://doi.org/10.1136/bmj.d4770>
- Eldridge, S. M., Lancaster, G. A., Campbell, M. J., Thabane, L., Hopewell, S., Coleman, C. L., & Bond, C. M. (2016). Defining feasibility and pilot studies in preparation for randomised controlled trials: Development of a conceptual framework. *PloS One*, 11(3), e0150205. <https://doi.org/10.1371/journal.pone.0150205>
- Espinoza-Vinces, C., Martínez, M. C., Atorrasagasti-Villar, A., Rodríguez, M. D. M. G., Ezpeleta, D., & Irimia, P. (2025). Artificial intelligence in headache medicine: Between automation and the doctor-patient relationship. A systematic review. *The Journal of Headache and Pain*, 26(1), 192. <https://doi.org/10.1186/s10194-025-02143-8>
- Hertzog, M. A. (2008). Considerations in determining sample size for pilot studies. *Research in Nursing & Health*, 31(2), 180-191. <https://doi.org/10.1002/nur.20247>
- Huang, Y. B., Lin, L., Li, X. Y., Chen, B. Z., Yuan, L., & Zheng, H. (2024). An indirect treatment comparison meta-analysis of digital versus face-to-face cognitive behavior therapy for headache. *NPJ Digital Medicine*, 7(1), 262. <https://doi.org/10.1038/s41746-024-01264-9>
- Julious, S. A. (2005). Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry*, 4(4), 287-291. <https://doi.org/10.1002/pst.185>

- Kilfoy, A., Chu, C., Krisnagopal, A., Mcatee, E., Baek, S., Zworth, M., ... & Jibb, L. (2025). Nurse-led remote digital support for adults with chronic conditions: A systematic synthesis without meta-analysis. *Journal of Clinical Nursing*, 34(3), 715-736. <https://doi.org/10.1111/jocn.17226>
- Kim, K. M., Kim, A. R., Lee, W., Jang, B. H., Heo, K., & Chu, M. K. (2022). Development and validation of a web-based headache diagnosis questionnaire. *Scientific Reports*, 12(1), 7032. <https://doi.org/10.1038/s41598-022-11008-y>
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276-282. <https://hrcak.srce.hr/89395>
- Minen, M. T., Gumpel, T., Ali, S., Sow, F., & Toy, K. (2020). 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: The Journal of Head and Face Pain*, 60(7), 1392-1401. <https://doi.org/10.1111/head.13859>
- Mokdad, A. H. (2021). GBD 2016 Headache Collaborators. Global, regional, and national burden of migraine and tension-type headache, 1990-2016: A systematic analysis for the Global Burden of Disease Study 2016, *Lancet Neurology*, 17(11), 954-976. [https://doi.org/10.1016/S1474-4422\(18\)30322-3](https://doi.org/10.1016/S1474-4422(18)30322-3)
- Mosadeghi-Nik, M., Askari, M. S., & Fatehi, F. (2016). Mobile health (mHealth) for headache disorders: A review of the evidence base. *Journal of Telemedicine and Telecare*, 22(8), 472-477. <https://doi.org/10.1177/1357633X16673275>
- Noser, A. E., Gibler, R. C., Ramsey, R. R., Wells, R. E., Seng, E. K., & Hommel, K. A. (2022). Digital headache self-management interventions for patients with a primary headache disorder: A systematic review of randomized controlled trials. *Headache: The Journal of Head and Face Pain*, 62(9), 1105-1119. <https://doi.org/10.1111/head.14392>
- Petrušić, I., Chiang, C. C., Garcia-Azorin, D., Ha, W. S., Ornello, R., Pellesi, L., ... & Wells-Gatnik, W. (2025). Influence of next-generation artificial intelligence on headache research, diagnosis and treatment: the junior editorial board members' vision—part 2. *The Journal of Headache and Pain*, 26(1), 2. <https://doi.org/10.1186/s10194-024-01944-7>
- Pini, L. A., Cottafavi, K., & Ferri, P. (2024). The nursing role in the management of medication overuse headache: realities and prospects. *Brain Sciences*, 14(6), 600. <https://doi.org/10.3390/brainsci14060600>
- Rasmussen, A. V., Jensen, R. H., Gantenbein, A., Sumelahti, M. L., Braschinsky, M., Lagrata, S., ... & Mose, L. S. (2024). Consensus recommendations on the role of nurses in headache care: A European e-Delphi study. *Cephalalgia*, 44(5), 03331024241252161. <https://doi.org/10.1177/03331024241252161>
- Robbins, M. S. (2021). Diagnosis and management of headache: A review. *Jama*, 325(18), 1874-1885. <https://doi.org/10.1001/jama.2021.1640>
- Sait, A., Mitsikostas, D. D., Ji, L. M., Nooshin, Y., Shuu-Jiun, W., Messina, R., ... & Lipton, R. B. (2021). Tension-type headache (Primer). *Nature Reviews. Disease Primers*, 7(1). <https://doi.org/10.1038/s41572-021-00257-2>
- Semigran, H. L., Linder, J. A., Gidengil, C., & Mehrotra, A. (2015). Evaluation of symptom checkers for self diagnosis and triage: Audit study. *BMJ*, 351. <https://doi.org/10.1136/bmj.h3480>
- Setia, M. S. (2016). Methodology series module 3: Cross-sectional studies. *Indian Journal of Dermatology*, 61(3), 261. <https://doi.org/10.4103/0019-5154.182410>
- Stubberud, A., & Linde, M. (2018). Digital technology and mobile health in behavioral migraine therapy: A narrative review. *Current Pain and Headache Reports*, 22(10), 66. <https://doi.org/10.1007/s11916-018-0718-0>
- Teare, M. D., Dimairo, M., Shephard, N., Hayman, A., Whitehead, A., & Walters, S. J. (2014). Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: A simulation

- study. *Trials*, 15(1), 264. <https://doi.org/10.1186/1745-6215-15-264>
- Wathne, H., Storm, M., Morken, I. M., & Lunde Husebø, A. M. (2025). Nurse-assisted remote patient monitoring for self-management support to patients with long-term illness—a qualitative multimethod study. *Journal of Advanced Nursing*, 81(9), 5996-6010. <https://doi.org/10.1111/jan.16736>
- Wijeratne, T., Oh, J., Kim, S., Yim, Y., Kim, M. S., Shin, J. I., ... & Kumar, R. (2025). Global, regional, and national burden of headache disorders, 1990–2021, with forecasts to 2050: A Global Burden of Disease study 2021. *Cell Reports Medicine*, 6(10). <https://doi.org/10.1016/j.xcrm.2025.102348>
- Woldeamanuel, Y. W., & Cowan, R. P. (2022). Computerized migraine diagnostic tools: a systematic review. *Therapeutic Advances in Chronic Disease*, 13, 20406223211065235. <https://doi.org/10.1177/20406223211065235>
- Young, N. P., Ridgeway, J. L., Haddad, T. C., Harper, S. B., Philpot, L. M., Christopherson, L. A., ... & Ebbert, J. O. (2023). Feasibility and usability of a Mobile app–based interactive care plan for migraine in a community neurology practice: Development and pilot implementation study. *JMIR Formative Research*, 7, e48372. <https://doi.org/10.2196/48372>
- Young, N. P., Stern, J. I., Steel, S. J., & Ebbert, J. O. (2025). Mobile App-Based Interactive Care Plan for Migraine: Survey Study of Usability and Improvement Opportunities. *JMIR Formative Research*, 9(1), e66763. <https://doi.org/10.2196/66763>
- Zhang, Y., Huang, W., Pan, S., Shan, Z., Zhou, Y., Gan, Q., & Xiao, Z. (2023). New management strategies for primary headache disorders: Insights from P4 medicine. *Heliyon*, 9(11). <https://doi.org/10.1016/j.heliyon.2023.e22285>