**STUDY INFORMATION**

Title: Multicenter evaluation of headache-associated symptom clusters in youth using multiple correspondence analysis.

Proposed authors:

Carlyn Patterson Gentile

Andrew D. Hershey

Christina L. Szperka

Description:

The primary objective of this study is to determine if similar headache-associated symptom clusters are seen across two tertiary pediatric care centers.

The secondary objective is to determine if headache-associated symptom clusters change over time.

Primary Hypothesis: Similar headache-associated symptom clusters will be seen across two tertiary pediatric care centers. Such a finding would support canonical representations of headache-associated symptoms.

Secondary Hypothesis: Fewer headache-associated symptoms will be reported with improvement in headache burden.

Background: We developed a novel analysis to determine underlying associations between headache-associated symptoms in youth using a large single center headache registry (Patterson Gentile et al, 2023). However, it remains unclear if these findings will generalize to other populations of youth with headache. To address this question, we will compare results of MCA from two pediatric tertiary care centers. Statistical simulations of clinical risk prediction have demonstrated that center-specific predictions offer the best predictive performance when data are heavily clustered by the center (Wynants et al, 2018). We will take this observation into account for statistical analysis.

**DESIGN PLAN**

This is a pre-registered multi-site cross-sectional study with a longitudinal arm conducted from patient questionnaires collected at the Children’s Hospital of Philadelphia general neurology and headache outpatient clinics and the Cincinnati Children’s Hospital outpatient headache clinic. The study aims to assess the generalizability of findings from

Inclusion criteria: Youth must meet the following inclusion criteria. (1) ages 6 to 17 years old at the time the intake questionnaire was filled out; (2) any sex, any race/ethnicity; (3) filled out the CHOP intake headache questionnaire between November 2022 and December 2023 or the available Cincinnati intake headache questionnaire up through June 2023; (5) any headache diagnosis

The timing of data capture differs because data missingness decreased substantially after quality improvements were enacted at CHOP after November 2022.

Exclusion criteria: (1) data outside of the designated age (2) outside the date collection range

Participant demographics including age (in years), legal sex, race (White, Black, Asian, Other/unknown), and ethnicity (Hispanic, non-Hispanic, Unknown) will be included in statistical analysis. Additional information available on race and gender identity will be reported in a table.

Headache burden metrics will include the following:

1. Baseline pain severity (mild, moderate, severe scale) – for those missing mild, moderate, or severe who have numeric pain scores (0 – 10), mild will be defined as 1-3, moderate will be defined as 4-6, and severe will be defined as 7-10.
2. Frequency of headache exacerbations (<1/month, 1-3/month, 1/week, 2-3/week, >3/week, daily, always – this is worded differently as ‘bad headaches’ in the CHOP dataset, and ‘headache frequency’ in the Cincinnati dataset. The category ‘never’ in the CHOP dataset will be re-categorized as <1/month
3. Pain pattern (constant vs. intermittent) – this is determined by reporting an ‘always’ headache for Cincinnati dataset, and based on a series of pattern questions from the CHOP dataset
4. PedMIDAS grade (no disability, mild disability, moderate disability, severe disability)

The following 11 headache-associated symptoms will be considered: nausea, vomiting, photophobia, phonophobia, osmophobia, lightheadedness, spinning, difficulty thinking, vision changes, ear ringing, neck pain. This differs from the original analysis, which had 13 associated symptoms, to account for questionnaire discrepancies between the two centers. Specifically, ‘double vision’ and ‘blurry vision’ will be combined into the broader category of ‘vision changes’ used in the Cincinnati questionnaire. ‘Balance problems’ will be omitted because it was not asked in the Cincinnati dataset and had overlap with room spinning in the CHOP dataset.

Identified differences between the two centers include the following:

1. Geographic location
2. Differing demographics
3. The CHOP questionnaire was patient-reported, the Cincinnati questionnaire was questionnaire-guided history taking from a headache specialist provider.
4. Cincinnati data were collected from a headache clinic by headache specialists, CHOP questionnaires were collected from general neurology and headache clinics.
5. ICHD-3 diagnosis was made by headache specialists at Cincinnati Children’s, while it was made by an algorithm applied to patient questionnaires at CHOP.
6. TOCs from other neurologists at the same institution occurred in the CHOP dataset, but not the Cincinnati dataset. Both clinics saw second opinions for headache from other institutions.
7. ICHD3 diagnosis was exclusively migraine in the Cincinnati dataset (and this included participants with NDPH or PTH onset), where these diagnoses were separate in the CHOP dataset. Presence of aura was differentiated in the Cincinnati dataset, but not in the CHOP dataset.

To address differences in diagnosis, headache phenotype will be used for the CHOP dataset. The following primary headache disorder phenotypes based on ICHD-3 diagnostic criteria will be used (migraine/probable migraine, chronic migraine/chronic probable migraine, tension-type headache, trigeminal autonomic cephalalgia, other primary headache disorder).

**ANALYSIS PLAN**

All analyses will be carried out through Matlab®.

Primary analysis: MCA will be conducted on the 11 headache-associated symptoms including data from the two centers, compared to MCA performed on the individual center data. The first two dimensions will be plotted and compared. A mixed effects linear regression model will be used to determine if center (primary predictor) predicted differences in first- and second-dimension MCA scores from the analysis that included both datasets (primary outcome, respectively). Covariates will be selected if they have a significant association (defined as p-value<0.1) with either the primary predictor or the primary outcome. Covariates that will be evaluated include demographic information, headache phenotype, and headache burden metrics as detailed above. Non-significant covariates will be removed by backwards elimination. Any covariate that is significant (p<0.05) or impacts the primary predictor by >15% will remain in the final analysis.

Secondary analysis: For Cincinnati participants with a follow up visit, we will determine how MCA scores change across visits using a linear mixed effects longitudinal analysis. The primary predictor will be overall perceived change (better, the same, or worse) and the primary outcomes will be MCA dimension 1 scores, and MCA dimension 2 scores, respectively. Covariates that will be evaluated include demographic information, headache phenotype, and headache burden metrics as detailed above. Non-significant covariates will be removed by backwards elimination. Any covariate that is significant (p<0.05) or impacts the primary predictor by >15% will remain in the final analysis.

**REFERENCES**

Patterson Gentile, C., Aguirre, G. K., Hershey, A. D., & Szperka, C. L. (2023). Symptoms associated with headache in youth. *Cephalalgia*, *43*(7), 03331024231187162.

Wynants, L., Vergouwe, Y., Van Huffel, S., Timmerman, D., & Van Calster, B. (2018). Does ignoring clustering in multicenter data influence the performance of prediction models? A simulation study. *Statistical methods in medical research*, *27*(6), 1723-1736.