MiCOAS

SYSTEMATIC LITERATURE REVIEW OF OUTCOMES AND ENDPOINTS USED IN ACUTE MIGRAINE CLINICAL TRIALS

VECTOR PSYCHOMETRIC GROUP, LLC
CHAPEL HILL, NORTH CAROLINA U.S.A.

EXECUTIVE SUMMARY

This document provides a summary of the systematic literature review of acute clinical trials for adult patients with migraine completed by the Migraine Clinical Outcome Assessment System (MiCOAS) team in partial fulfillment of the objectives of the grant provided by the US Food and Drug Administration (FDA; 1 UG3 FD006795-01) to develop a standardized set of patient-centered outcomes and endpoints with a goal of using these endpoints in migraine clinical trials. This report focuses on the acute migraine treatment outcomes and endpoints found in the peer-reviewed literature summarizing clinical trials. A second report focuses on outcomes and endpoints for preventive migraine treatments.

We conducted a systematic literature review of English language published clinical trials of acute migraine therapies in adults using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist. We engaged in a two phase process: an initial review of every manuscript identified by key words and other criteria in PubMed to be a candidate for inclusion yielding 1,567 articles, followed by a review in which 705 of the publications were determined to be eligible and appropriate for the data extraction process. For the data extraction process, we preidentified five broad categories of outcome variable types and three potential endpoint timing possibilities. This report includes the results of the complete publication list and a more in-depth analysis of the 451 publications (64.0%) published in 1988 and later (when the International Classification of Headache Disorders [ICHD] first published diagnostic criteria for migraine), which were randomized and blinded, and focused on pharmacological or medical device interventions.

Among the pre-specified types of outcomes, 95.3% of the publications examined at least one pain-related outcome, 67.2% examined at least one non-pain symptom or most bothersome symptom (MBS), 41.2% examined at least one disability/impairment outcome, and 35.3% examined one or more patient reported outcome (PRO) (headache-related or non-headache focus).

As demonstrated in the analysis of data extracted from the articles summarizing acute clinical trials, the outcomes used to define endpoints vary substantially across trials, ranging from pain relief or freedom, use of acute/rescue medication, and various headache-related and non-headache PROM measures, such as those related to the impact migraine has on the patient's life or more general health-related quality of life (HRQoL). The definition of the endpoints used (e.g., change from baseline, fixed-time point comparisons, categorization of "responders" to treatment based on wide variety of "responder definitions") also differs substantially across publications. Endpoint timing varied across publications but key outcomes such as pain relief/freedom, non-pain symptoms, and disability/impairment were commonly assessed at 1 hours and 2 hours post-treatment.

While some of this inconsistency is attributable to the wide range of publication dates and changes in criteria, the treatment landscape, and the fields understanding of migraine, even within our focused subset of more recent publications, a large amount of variability exists in the outcomes and endpoints used and how those outcomes were operationalized. The results from examining the full set of selected articles demonstrated even more variability and lack of standardization across trials.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

Migraine is a chronic and potentially disabling neurological disease (Headache Classification Committee of the IHS [IHS], 2018). Migraine is highly prevalent, potentially severely disabling, and has a broad impact on individuals with the disease, their families, and society as a whole. The 2016 Global Burden of Disease (GBD) analysis (GBD 2016 Headache Collaborators, 2018) reported that migraine is a leading cause of years lived with disability. It is estimated that worldwide 1.04 billion people have migraine, corresponding to a prevalence of 14.4% overall, 18.9% in women, and 9.8% in men (GBD 2016 Headache Collaborators, 2018). Analyses from the US, population-based American Migraine Prevalence and Prevention Study (e.g., Buse et al., 2012; Buse et al., 2013; Lipton et al., 2007) found that approximately 12% of respondents, including 17.4% of females and 5.7% of males, met criteria for migraine and 0.91% met criteria for chronic migraine (1.29% of females; 0.48% of males). A migraine attack is frequently characterized by intense, debilitating headache but can also include associated symptoms in various combinations such as nausea, vomiting, sensitivity to light and sound during the headache phase as well as prodrome, aura and postdrome phases (IHS, 2018). Considering the substantial burden and impact of migraine, additional funding is needed to identify and test products that may improve outcomes for migraine patients.

Therapeutic approaches for migraine fall under 2 broad categories: preventive and acute treatments (American Headache Society, 2019). Preventive treatments, which include both pharmacological and non-pharmacological approaches, aim to reduce frequency, severity, and duration of attacks, improve responsiveness to treatment of acute attacks, and reduce level of disability (American Headache Society, 2019; Tassorelli et al., 2018). Acute migraine treatments aim to resolve migraine pain and symptoms when an attack occurs and return individuals to a “normal” level of functioning as quickly as possible (Marmura, Silberstein, & Schwedt, 2015).

The global standard for migraine classification is the International Classification of Headache Disorders (ICHD). With editions in 1988 (ICHD-1; IHS, 1988), 2004 (ICHD-2; IHS, 2004), 2013 (ICHD-3 beta; IHS, 2013) and 2018 (ICHD-3; IHS, 2018), criteria are provided for migraine and its subtypes often assessed in clinical trials. Some subtypes of migraine often studied in clinical trials and defined by the ICHD classification system include migraine with aura, migraine without aura, and chronic migraine (CM). Though criteria for migraine with and without aura have been relatively stable, criteria for CM emerged in ICHD-2, evolved in the ICHD-3 (beta) and were carried into ICHD-3. CM refers to people with migraine and 15 or more headache days per month for at least 3 months of which at least eight days are linked to migraine. In ICHD-3 beta and ICHD-3, forms of CM with and without medication overuse are recognized. The term episodic migraine (EM) refers to persons with migraine and fewer than 15 headache days per month and has only recently been added to the ICHD system (Goadsby & Evers, 2020) but is widely used in clinical trials, research, and clinical care.

Almost 30 years ago, the IHS first published guidelines to help improve the quality of clinical trials in migraine (International Headache Society Committee on Clinical Trials in Migraine, 1991). These guidelines were updated in 2000 (second edition), 2012 (third edition), and, most recently, 2019 (fourth edition) (Tfelt-Hansen et al., 2000; Tfelt-Hansen et al., 2012; Diener et al., 2019). The IHS guidelines address several topics including subject selection (migraine definition, attack frequency, duration of migraine, age of onset), trial design (blinding, randomization, placebo-control, study designs, number of treated attacks, rescue medication), evaluation of results (headache diaries, (co)primary endpoints, secondary endpoints, adverse events), and statistical analyses (hierarchy of endpoints, power analyses, alpha corrections, statistical analysis plans). The US Food and Drug Administration (FDA) has also recently provided guidance for the design and conduct of

acute migraine trials, including non-binding recommendations for outcomes and endpoints to be assessed and response scales for assessing those outcomes (US Food and Drug Administration (FDA), 2018).

In the current review, we extracted data on trial design components (e.g., blinding, randomization, placebo-control, designs) and endpoint definitions (pain relief, pain freedom, disability/pain scales). Where possible, this review captured whether publications were IHS-compliant in their definitions of migraine and IHS-compliant for key endpoints such as pain relief and pain freedom. Additionally, when pain and disability were assessed, this review captured whether IHS-recommended pain and disability scales were used. These elements provide indicators for how well the examined publications align with current recommended clinical trial guidelines.

The goal of this work is to provide an overview of the acute migraine clinical trial literature to aid in future endpoint, outcome assessment, and treatment developments. The purpose of this document is to summarize the findings from a systematic literature review that provides a comprehensive picture of concepts, endpoints, and associated outcomes used in clinical trials of acute treatments for adults (defined as 18 years or older) with migraine published in English in peer-reviewed scientific journals.

METHODS

A systematic literature review was conducted to understand the frequency of utilization for specific concepts, endpoints, and associated outcome measures used in clinical trials assessing acute treatments in adults with migraine. PRISMA provides a checklist related to consensus recommendations for the development and execution of high-quality systematic literature reviews (Moher et al, 2009). This checklist includes recommendations for the conduct of the literature search and review, including: pre-specification of eligibility criteria for located publications, the database to be used for the search as well as draft search terms, the standardized process used to review located publications including record tracking/data management systems to be used, the data planned to be extracted from each publication meeting inclusion criteria, and the plan for summarizing the extracted information. The protocol developed for this literature review adhered to PRISMA recommendations.

IDENTIFICATION OF PUBLICATIONS

PubMed, a search engine maintained by the National Center for Biotechnology Information at the U.S. National Library of Medicine, located at the National Institutes of Health was used as the primary database queried to identify initial articles for review. PubMed filters were used to limit results to human clinical trials and to articles published in English. No time frame restrictions were imposed on the results and the date of the final search was 10/28/2019.

The PubMed search term used to identify the initial articles was:

(((migraine[MeSH Terms]) AND acute AND Clinical Trial[ptyp] AND Humans[Mesh]) AND English[lang])

The title and abstract of each publication returned from the search were screened by two Vector Psychometric Group, LLC (VPG) methodologists, using the Covidence online systematic review tool, for relevance to the stated goals. Specifically, the inclusion criteria as specified in the Covidence system were:

- The screening reviews were based on the inclusion of an interventional, adult acute migraine trial description in the title, abstract, or keywords
- Interventions could be pharmacological (e.g., pills, injections), physical (acupuncture, massage, exercise, etc.), dietary, or other novel treatment intended to treat attacks
- Open-label studies and Phase 4 trials were included.
- Subtypes of migraine (e.g., menstrual migraine; medication overuse if sample is specified as migraine patients) were included
- Pilot studies with migraine patients were included

Exclusion criteria were as follows:

- Preventive migraine trials were excluded (mixed trials with preventive and acute outcomes included)
- Observational studies, surveys (not Post-Marketing Phase 4), epidemiological studies, etc. were excluded
- Trials with ONLY healthy volunteers given an acute intervention were excluded (mixed healthy/migraine samples were included)
- Given the limited information and difficulty in obtaining full documents for such references, exclude:

  - Peer-reviewed, stand-alone abstracts
  - Letters to the editor describing trials
  - Abstracts/papers from conference proceedings
  - Case studies
- Trials using only pediatric patients were excluded (mixed adult and pediatric trials included)

Once the initial list of screen-pass references was compiled, a review of the reference section in each located publication was undertaken to locate any potentially relevant publications that were previously undiscovered. Newly located articles were added to the "initial" list and title and abstract submitted to the screening review (as detailed above) for inclusion/exclusion in the final version of the initial list.

With the candidate reference list finalized, a brief review of each full publication was undertaken by two Vector Psychometric Group, LLC (VPG) doctoral level methodologists to confirm the relevance of the article to the current goals. With an agreed-upon positive assessment from the brief review, the publication was included in the final references list. All agreed upon negative reviews resulted in the exclusion of the publication from this list. Disagreements on the status of an article were reviewed by a third doctoral-level study team member and a discussion among all three reviewers determined the final status of an article regarding inclusion/exclusion in the final list of publications slated for extraction.

All articles in the final list of publications were fully reviewed by a VPG doctoral-level study team member and, if information relevant to the goal of the review was found in the publication during data extraction, it was included in the literature synthesis section of this literature review report.

Data Extraction

For all located publications included in the final list of publications, pre-identified salient key features of each acute publication were extracted. This included extracting all available information related to year of publication, journal name, ClinicalTrials.gov identifier(s), trial name, phase of trial (I - IV), general description of the trial design, sample size, patient sociodemographic descriptives (age, gender, race), salient migraine subtypes (e.g., migraine with aura only, menstrually-related migraine (MRM), only, EM vs. CM, etc.), and type of treatment investigated (pharmacologic, neurostimulation, behavioral, complimentary and integrative treatments, etc.). Additionally, data extraction from the articles included the concepts examined (e.g., pain freedom, pain relief, disability/impairment HRQoL), the endpoints used, and any specific outcome measures used.

Data related to the descriptive trial information was extracted by trained research assistants. A second research assistant independently extracted the same data for approximately 5% of candidate publications and rater/extractor agreement kappas were calculated. Data related to the concepts, outcomes, and endpoints examined were extracted by one of four VPG doctoral-level methodologists into a pre-coded, standardized, structured Excel worksheet.

Synthesis of Extracted Information

To synthesize the sizeable amount of information collected during the data extraction from the large number of articles on acute migraine treatment, numerous tables and figures were planned to present summary information in a digestible fashion. These included summary tables focused on the study design characteristics, demographics of study participants, and outcomes (pain, associated symptoms, migraine-

related PROMs, etc.), endpoint type (change from baseline, fixed timepoint comparisons, responder definitions), and endpoint timing (1hr, 2hr, 4hr, etc.) used.

Outcomes are presented within four broad categories:

1. Pain-related outcomes (pain relief, pain freedom, headache recurrence, rescue medication use, etc.)
2. Non-pain associated symptoms (most bothersome symptom, nausea, photophobia, etc.)
3. Disability/Impairment
4. Patient Reported Outcome Measures
    a. Migraine/headache related PROMs (24h Migraine Specific Quality of Life Questionnaire [24hr MSQoL], etc.)
    b. Non-headache related PROMs (includes patient global impression of change (PGIC), patient global impression of severity (PGIS), and treatment efficacy/satisfaction/preference items)
        i. Measurement of patient global impression, treatment efficacy, satisfaction, and preference varied across manuscripts (e.g., different response scales, varied verbal labels) but were grouped into general categorizes for tabled results.

Endpoint type was categorized within 3 broad categories:

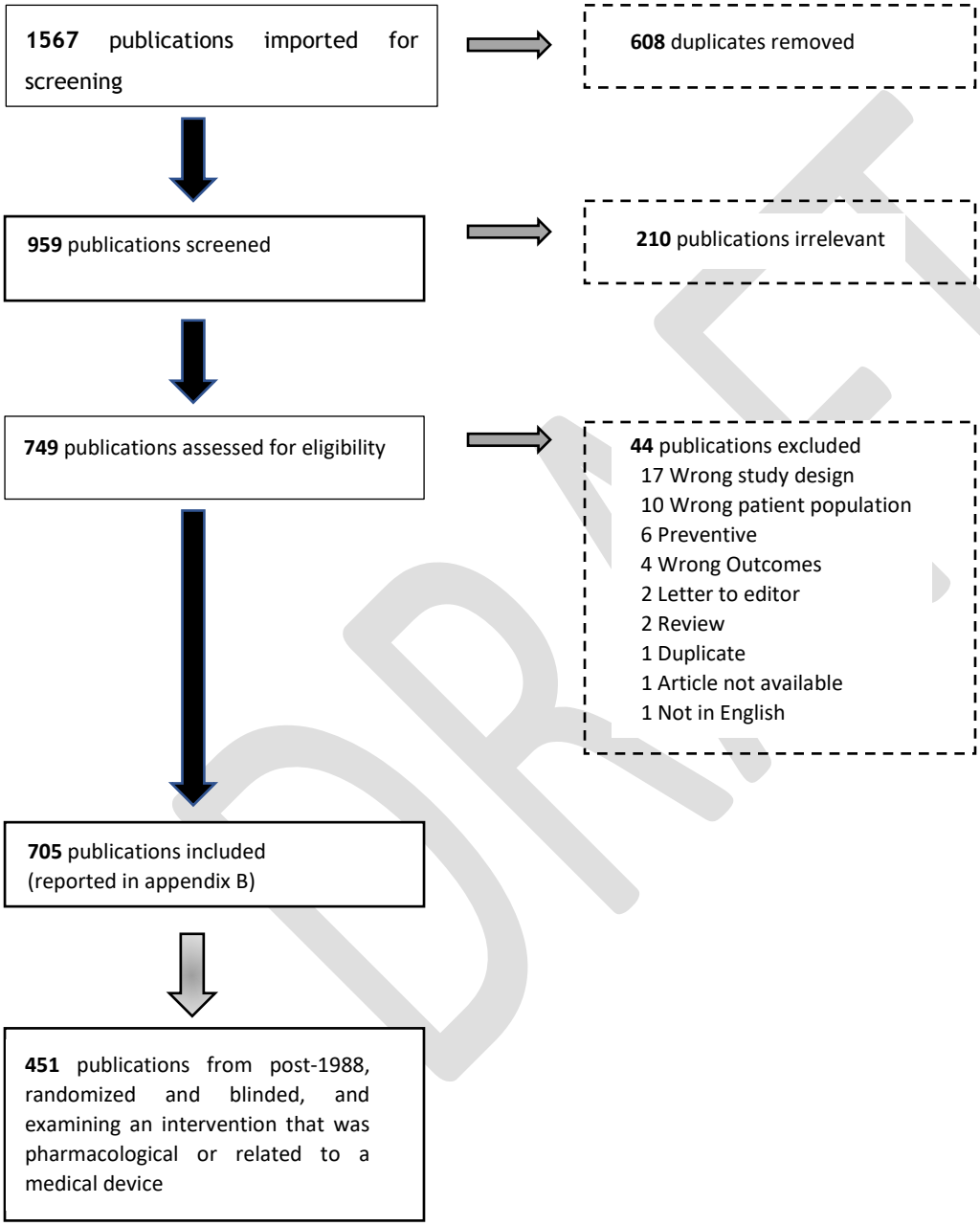
1. Change from baseline
2. Fixed timepoint
3. Responder definitions (e.g., 50% reduction, 75% reduction, 100% reduction, Other definitions)

Additionally, the specific timings of endpoint definitions (e.g., 15m, 1h, 2h, 24h) were also tracked.

RESULTS

Of the 1,567 publications found through the initial search and reference section reviews, 705 publications were included for data extraction. Figure 1 provides a more detailed break-down of the review and selection process outcomes. Appendix A provides a complete list of all publications located from the PubMed search and their ultimate status regarding inclusion/exclusion in the final selection of articles.

Figure 1. PRISMA diagram of article flow through the systematic literature review of acute migraine trials.



With respect to data extraction from the 705 publications, inter-rater agreement kappas for the descriptive variables extracted (age, sex, migraine characteristics, trial design features, etc.) had an average kappa estimate of .86. Given the somewhat inconsistent nature of reporting in the examined articles and the varied

age and quality of reporting in the publications, the observed level of inter-rater agreement was considered acceptable (and was above the recommended lower bound of .6 (McHugh, 2012)).

General Study Characteristics

Of the 705 publications included for review, just over 60.0% reported on publications that were placebo/sham controlled, over two-thirds were blinded and randomized (70.1%), more than 95% assessed a pharmacological or medical device treatment (96.3 %), and over three-fourths of publications used one of the iterations of the ICHD criteria for migraine (79.4 %). Many publications published prior to 1988 used the 1962 Ad Hoc Committee criteria for migraine (Ad Hoc Committee on Classification of Headache, 1962) which used similar criteria to the subsequent ICHD criteria for migraine.

Table 1. General Publication Characteristics (n=705)

Study Characteristic	Percent	N
Study Purpose(s)		
Efficacy Assessed	96.60	681
Safety Assessed	85.82	605
Pharmacokinetic Study	4.40	31
Study/Design Features		
Study 1988 or Later	94.04	663
Randomized	77.59	547
Blinded	73.33	517
Randomized and Blinded	70.07	494
ICHD Migraine Criteria Used	79.43	560
Placebo/Sham Controlled	60.57	427
Crossover Design	22.98	162
Intervention Information		
Drug /Medical Device	96.31	679
Multiple Active Treatments	36.74	259
Open-Label Study	21.84	154
Multiple Dose Levels Assessed	23.83	168

Of the 705 publications, most publications examined at least one efficacy outcome (96.6%) or a safety outcome (85.8%). While planned for extraction, phase of study was not reliably reported – over 88% of publications were not clearly marked as Phase I through IV (data not shown).

With respect to the interventions investigated in the 705 reviewed publications, nearly 95% of publications investigated pharmacological/medication treatments (pills, injections, etc.) and 2.3% examined medical devices (e.g., neurostimulation devices, dental devices). Table 2 provides a detailed break-down of the interventions examined.

Table 2. Acute Migraine Treatments Investigated (n=705)

Treatment	Percent	N
Pharmacological/medication	94.04	663
Medical device (electrical stimulation, dental plate)	2.27	16
Other/Multiple Categories	2.13	15
Alternative (acupuncture, osteopathic, herbal, etc.)	1.56	11

Given the presumed interest in outcomes and endpoints used in the publications that employed ICHD criteria to identify participants with migraine, from this point on we focus on the 451 publications that were published in 1988 or later (following the ICHD-1 publication), that were also randomized and blinded, and included interventions that were pharmacological or related to a medical device (64% of articles). (Results from the full sample of 705 articles are provided in Appendix B.)

Demographic and Descriptive Variables of Patients

Available demographic characteristics for the subjects from the selected publications (pooled over all treatment groups) are summarized in Table 3. The median total sample size was n= 243 (25th percentile: 80; 75th percentile: 640). Of publications that reported age, gender, and/or race descriptives, the average age was found to be 39.1 (SD = 3.9), with 82.9% of patients identifying as female, and 85.1% of patients reported as White/Caucasian.

Table 3. Demographic Characteristics of the Samples in Select Recent Publications (n=451)

Variable	N	Mean	SD	Minimum	25th Percentile	Median	75th Percentile	Maximum
Total N	451	570.58	916.83	10	80	243	640	8473
Mean Age	410	39.13	3.85	24.45	37.45	40.00	41.20	56.10
Percent Female	427	82.85	8.50	35.00	80.00	84.12	87.00	100.00
Percent White	178	85.10	17.25	2.12	80.30	88.39	94.85	100.00

Note. SD = Standard deviation. Descriptive statistics are based on all available, non-missing data.

Of note in these demographic summary values is that publications conducted exclusively outside of the United States (e.g., Chinese, Indian, or Iranian studies) often did not report the breakdown of patients into race/ethnicity categories and, therefore, did not contribute data to the summary Percent White value; the "typical" patient in an acute migraine is a middle-aged white female, but there is slightly more racial/ethnic diversity in the overall trial population than indicated by the reported values.

Tables 4 and 5 summarize migraine and aura group characteristics of the publications that reported such features. The majority (90.9%) of the publications looked at general migraine (unspecified/multiple types) and mixed aura types (82.5%).

Table 4. Migraine Group Characteristics (n=451)

Patient group characteristics	Percent	N
General migraine (classical/common migraine, unspecified, multiple types)	90.91	410
MRM	4.66	21
EM only	3.77	17
CM/TM only	0.67	3

Note. MRM = Menstrually-related migraine. EM = Episodic migraine. CM = Chronic migraine. TM = Transformed migraine.

Table 5. Aura Group Characteristics (n=451)

Aura Characteristics	Percent	N
Mixed (with and without aura)	82.48	372
Not specified	9.53	43
Without aura only	6.87	31
With aura only	1.11	5

OUTCOMES AND ENDPOINTS USED IN SELECTED SUBSET OF MORE RECENT PUBLICATIONS (N=451)

As seen in Table 6, over 90% (95.3%) of the publications looked at one or more pain-related outcome while 67.2% examined one or more non-pain symptoms outcome. A little over 40% of publications looked at disability/impairment outcomes, a little over one-third of publications used PROMs; of publications using PROMs, 94.3% used one or more non-headache specific PROM and only 17.6% of publications looked at migraine/headache-related PROMs.

Table 6. Outcomes Assessed Across Publications (n=451)

Outcome grouping	Percent	N
Pain-related	95.34	430
Headache Recurrence / Rescue Med Use	76.51	329
Pain Relief	72.09	310
Pain Free	64.88	279
Pain General	33.95	146
Meaningful Relief	8.60	37
Non-Pain Symptoms	67.18	303
Associated Symptoms (Nausea, Vomiting, Photophobia, phonophobia, etc.)	98.68	299

Most Bothersome Symptom	5.28	16
Disability/impairment	41.24	186
Patient Report Outcome Measures (PROMs)	35.25	159
Non-headache specific PROM	94.34	150
Headache-related PROM	17.61	28

In examining the various combinations of outcomes used (Table 7), over one-fifth of publications looked at one or more pain-related outcome combined with associated symptom outcomes. Under 20% of publications looked at only pain-related outcomes (18%), and a little over 17% of publications examined pain-related, non-pain symptoms, disability/impairment, and used one or more PRO. The 11 publications (2.4%) listed as using none of the outcomes in our constructed groupings reported primarily on safety studies (examining only adverse events), pharmacokinetic studies (examining assorted laboratory-provided values), or health economics outcomes.

Table 7. Combinations Assessed Across Publications (n=451)

Pain-related	Non-Pain Symptoms/MBS	Disability/ impairment	PROMs (headache and non-headache specific)	Percent	N
Yes	**Yes**	No	No	23.28	105
Yes	No	No	No	17.96	81
Yes	**Yes**	**Yes**	**Yes**	17.74	80
Yes	**Yes**	**Yes**	No	16.85	76
Yes	**Yes**	No	**Yes**	8.65	39
Yes	No	No	**Yes**	5.54	25
Yes	No	**Yes**	No	2.88	13
No	No	No	No	2.44	11
Yes	No	**Yes**	**Yes**	2.44	11
No	No	**Yes**	No	0.67	3
No	No	**Yes**	**Yes**	0.67	3
No	**Yes**	No	No	0.67	3
No	No	No	**Yes**	0.22	1

Pain-related Outcomes (n=430)

As noted earlier, pain-related outcomes were the most commonly encountered outcomes in the reviewed acute migraine literature, with over 95% of the publications using one or more pain-related outcome in the reported publication. In assessing headache pain intensity, the IHS acute trial guidelines (Diener et al., 2019) recommend three response scales: a four-category ordinal scale, an 11-point NRS, or a 100mm Visual Analog Scale (VAS). Of the 430 publications that assessed headache pain intensity in some manner, 89.1% used an IHS-recommended pain scale.

There were several commonly encountered specific outcomes that we have classified under the overarching pain-related outcome heading: pain relief, pain freedom, general pain, meaningful relief, headache recurrence, and rescue medication use.

Pain relief (n=310)

Pain relief in the context of clinical migraine trials refers to a reduction in headache pain that is not a complete resolution of the pain. The IHS acute clinical trial guidelines (Diener et al., 2019) define Headache Relief as a decrease in headache pain from moderate to severe at baseline to mild or moderate; while the current guidelines were recently published a majority of the publications assessing pain relief did define relief in a manner consistent with this recommendation (88.7%; top section of Table 8). In looking at the publications examining pain relief, a large majority (94.2%) used an ordinal response scale, with continuous response scales (such as a VAS or NRS) being the second most commonly used. Of the publications that stated they used a VAS and/or NRS, over 70% reported using a VAS. Of note is that the term "VAS" was applied to a wide variety of ratings scales, ranging from a true VAS (which asks patients to provide a tick mark on a line, the position of which is then measured), to NRSs (typically ranging from 0-10), to ordinal scales with as few as five response options.

Table 8. Rating Scale Features of Publications Assessing Pain Relief (n=310)

Rating Scale Features	Percent	N
IHS-recommended Definition of Relief		
No/Unknown	11.29	35
Yes	88.71	275
Rating Scale Features		
Binary	0.32	1
Ordinal	94.19	292
4 categories	97.95	286
5 categories	2.05	6
Continuous	3.87	12
Other/Multiple	1.61	5
VAS/NRS	4.84	14
NRS	28.57	4
VAS	71.43	10

With respect to the endpoint definitions used for pain relief, Table 9 provides a detailed breakdown of the time used in defining a pain relief endpoint, with observed timings ranging from 10 minutes to 24 hours with publications typically examining more than one endpoint related to pain relief. However, the most commonly used endpoint definition for pain relief was 2 hours (86.8%).

Table 9. Endpoint Definitions Used in Assessing Pain Relief (n=310)

Endpoint Definition	Percent	N
10m	5.48	17
15m	10.65	33
20m	3.55	11
25m	0.65	2
30m	48.06	149
45m	6.77	21
1hr	64.19	199
90m	27.42	85
2hr	86.77	269
3hr	17.74	55
4hr	39.03	121
24hr	15.81	49
Other	16.13	50

Building on the pain relief outcomes used, several different outcomes related to pain relief were also commonly examined, including sustained response (defined as meeting the criteria for pain relief at a given point and having no headache pain increases through a set later timepoint) and the consistency of obtaining pain relief across multiple attacks. Table 10 provides a breakdown of the outcomes that are variations of pain relief. Of the 310 publications that examined pain relief, one-third (33.9%) looked at sustained pain relief and about 10% of publications looked at consistency of pain relief across attacks (12.6%) and time to pain relief (9.7%). Of the 105 publications that examined sustained pain relief, 70.5% looked at sustained pain relief at 24 hours, 8.6% looked at sustained relief at 48 hours, and 16.2% looked at sustained relief at 24 and 48 hours.

Table 10. Additional Outcomes Derived from Pain Relief (n=310)

Additional outcomes derived from pain relief	Percent	N
Sustained Response	33.87	105
24hr	70.48	74
48hr	8.57	9
24hr and 48hr	16.19	17
Other	4.76	5
Consistency across attacks	12.58	39
Time to relief	9.68	30

Pain freedom (n=279)

Pain freedom in the context of clinical migraine trials refers to a reduction in headache pain that is a complete resolution of the pain. The IHS acute trial guidelines (Diener et al., 2019) define Pain Freedom as patients

who become free from headache pain following treatment. As noted previously, multiple pain severity rating scales are allowed by IHS acute migraine trial guidelines; however, a comment in the guidelines document does note that the four-category response scale is preferred. For pain freedom, therefore, we tracked the number of publications that used the four-category response scale for rating headache intensity and also defined freedom as the complete absence of pain (e.g., a response of "None" on the four-category scale). The majority of recently published articles defined relief in a manner consistent with the definition (90.3%; top section of Table 11). Of the publications focused on pain freedom, the majority (95.0%) used an ordinal response scale. Of the publications that stated they used a VAS and/or NRS (only 2.9% of pain freedom publications), 25% used the NRS alone, 62.5% used VAS alone, and 12.5% used the NRS and VAS.

Table 11. Rating Scale Features of Publications Assessing Pain Freedom (n=279)

Variable	Percent	N	
IHS-recommended Definition of Freedom			
No/Unknown	9.68	27	
Yes	90.32	252	
Binary	1.08	3	
Ordinal	94.98	265	
4 categories	99.62		264
Other	0.38		1
Continuous	2.51	7	
Other/Multiple	1.43	4	
VAS/NRS	2.87	8	
NRS	25.00		2
VAS	62.50		5
NRS and VAS	12.50		1

With respect to the endpoint definitions used for pain freedom, Table 12 provides a detailed breakdown of the time used in defining a pain freedom endpoint, with observed timings ranging from 10 minutes to 24 hours and publications often defining multiple endpoints for pain freedom. The most used endpoint definition was 2 hours (92.1%), conforming to the IHS acute trial guidelines and as was also seen with pain relief. Additionally, pain freedom was also commonly assessed at 30 minutes (44.8%) and 1 hour (58.1%).

Table 12. Endpoint Definitions Used in Assessing Pain Relief (n=279)

Endpoint Definition	Percent	N
10m	4.66	13
15m	9.32	26
20m	3.58	10
25m	0.36	1

30m	44.80	125
45m	6.09	17
1hr	58.06	162
90m	28.67	80
2hr	92.11	257
3hr	12.90	36
4hr	36.56	102
24hr	18.64	52
Other	12.90	36

Like pain relief, variant outcomes related to pain freedom were also commonly examined. Table 13 provides a breakdown of the outcomes that are variations of pain freedom. Of the 279 publications that examined pain freedom, almost half (47.0%) evaluated sustained pain freedom and under 10% of publications looked at consistency of pain freedom across attacks (7.9%) and time to pain freedom (9.32%). Of the 131 publications that examined sustained pain freedom, three-quarters assessed sustained pain freedom at 24 hours (74.1%) and 16.8% looked both 24 and 48 hours.

Table 13. Additional Outcomes Derived from Pain Freedom (n=279)

Additional outcomes derived from pain freedom	Percent	N
Sustained Response	46.95	131
24hr	74.05	97
48hr	7.63	10
24hr and 48hr	16.79	22
Other	1.53	2
Time to freedom	9.32	26
Consistency	7.89	22

General pain (n=146)

General pain is used to describe general assessments of pain that did not conform to pain freedom, pain relief, or meaningful relief. Given the somewhat broader nature of the category, more variability was seen in the response scales used to assess pain and the types of analyses examined. Table 14 provides an overview of the features of the utilized response scales for assessing pain. Continuous response scales (59.6%, of which about 70% are either 11 or 100-point scales) were most often used in assessing general pain, followed by ordinal scales (39%, of which 89.5% were four-category response scales). About half of 146 general pain publications (54.1%) examined either a VAS or NRS. The VAS was used in about 3 times more publications than the NRS (72.2% vs. 25.3% of publications). However, it is important to note again that response scales that were termed “visual analog scale” by authors covered a wide range of response scales, many that would not be considered a true VAS.

Table 14. Rating Scale Features of Publications Assessing General Pain (n=146)

Rating Scale Features	Percent	N
Binary	1.37	2
Ordinal	39.04	57
3 categories	1.75	1
4 categories	89.47	51
5 categories	5.26	3
Other	3.51	2
Continuous	59.59	87
10 point	16.09	14
11 point	35.63	31
100 point	34.48	30
Other range	13.79	12
VAS/NRS	54.11	79
NRS	25.32	20
VAS	72.15	57
NRS and VAS	2.53	2

While the previously examined pain-related outcomes of pain freedom and pain relief are, by definition, change from baseline analyses, general pain outcomes could be examined as change from baseline and/or fixed timepoint analyses (e.g., comparing treatment groups on mean headache intensity values at 2 hours post-treatment). Table 15 provides a breakdown of the type of endpoints that were seen in analyses examining general pain. Of the 146 examining a general pain outcome, about two-thirds examined change from baseline (62.3%), one-quarter examined a fixed timepoint (27.4%), and 10.3% examined fixed timepoint and change from baseline.

Table 15. Endpoint Types Used in Publications Assessing General Pain (n=146)

Endpoint type	Percent	N
Change from Baseline	62.33	91
Fixed Timepoint	27.40	40
Fixed and Change from Baseline	10.27	15
Responder definition	23.97	35
50%	45.71	16
Multiple	14.29	5
Other	40.00	14

Within the change from baseline analyses, a subgroup of publications examined responder definitions (or within-person meaningful change thresholds). These analyses set a specific threshold for minimum change

from baseline for the subject to be considered a "responder" at a certain time point and then compare proportions of responders across treatment groups. Of the publications that utilized responder definition analyses (24% of all general pain publications), the most commonly used threshold was a 50% reduction in pain from baseline (45.7%), followed by "Other" (40%) which included an assortment of other values used to define responders (e.g., a 4 point reduction on an 11 point scale, 75% reduction from baseline).

Finally, the endpoint definitions used in general pain analyses demonstrated more variability than the previously described pain relief and pain freedom endpoints. As can be seen in Table 16, the most commonly used endpoints were 30 minutes (46.6%), 1 hours (59.6%), and 2 hours (56.9%).

Table 16. Endpoint Definitions Used in Assessing General Pain (n=146)

Endpoint definition	Percent	N
10m	8.90	13
15m	19.86	29
20m	10.27	15
25m	2.05	3
30m	46.58	68
45m	15.75	23
1hr	59.59	87
90m	13.01	19
2hr	56.85	83
3hr	10.96	16
4hr	22.60	33
24hr	14.38	21
Other	26.03	38

Meaningful relief (n=37)

Meaningful relief is generally described as a subjective concept in which each trial subject interprets what "meaningful relief" is to them and provides responses relative to their own subject-specific definition. The rating scale most often used in assessing meaningful relief (Table 17) is binary (59.5%), which allows patients to indicate Yes/No to a question such as, "Have you obtained meaningful relief from your migraine?"

Table 17. Rating Scale Features of Publications Assessing Meaningful Relief (n=37)

Rating Scale Features	Percent	N
Binary	59.46	22
Ordinal	13.51	5
4 categories	100.00	5
Continuous	27.03	10

While most publications that investigated meaningful relief used a "Time to" analysis (67.6%) and often employed a stopwatch or timestamp approach, allowing subjects to report the exact duration of time from treatment to achieving meaningful relief, publications also described Time to analyses based on meaningful relief items asked at fixed timepoints. The endpoints/timing of the meaningful relief assessments used are summarized in Table 18. As can be seen, and mirroring all previous endpoints, two-hours post-treatment (64.9%) was the most commonly used timepoint at which meaningful relief was assessed.

Table 18. Endpoint Definitions Used in Assessing Meaningful Relief (n=37)

Endpoint Definition	Percent	N
10m	2.70	1
15m	16.22	6
30m	29.73	11
45m	8.11	3
1hr	45.95	17
90m	21.62	8
2hr	64.86	24
3hr	16.22	6
4hr	32.43	12
24hr	2.70	1
Other	18.92	7

Headache recurrence and rescue medication use (n=329)

The return of headache pain after it was resolved is termed headache recurrence. Recurrence was previously defined as the achieving pain freedom and then experiencing a return of moderate to severe headache pain. The most recent version of the IHS acute trial guidelines (Diener, et al., 2019) has moved to the term "relapse" which they define as the occurrence of a headache of any severity within 24 or 48 hours after the initial treatment; FDA recommendations suggest relapse should be assessed through 48 hours after treatment. Given the variability in the definition and the relatively recent publication of the new IHS "relapse" guidelines, any publication which specifically stated that they examined headache recurrence/relapse is included here under the broad categorization of headache recurrence.

Headache recurrence/relapse was examined in 44.79% of the examined publications (Table 19), with most publications using a 24-hour cut-off alone to define the recurrence window (81.7%). Only 8.9% used the 48 hours window alone which is preferred in the current IHS acute migraine guidelines.

Table 19. Additional Pain-related Outcomes Used in Acute Migraine Publications (n=329)

Additional Pain-related Outcomes	Percent	N
Headache Recurrence	44.79	202
24hr	81.68	165
48hr	8.91	18
24hr and 48hr	3.47	7
Other	5.94	12
Rescue Medication	61.64	278

The use of rescue medication or additional doses of treatment medication was also a commonly used outcome in acute migraine trials, with over 60% of the examined publications tracking subjects' use of additional medication to attempt to alleviate experience migraine attacks.

Other Non-Pain Symptoms and Most Bothersome Symptom (n=303)

As noted previously, other non-pain symptoms (often collectively referred to as associated migraine symptoms) were often examined in acute migraine trials with two-thirds of the publications examining at least one non-pain symptom or MBS (67.2%). Historically, the most commonly assessed associated symptoms of acute migraine attacks were nausea/vomiting, photophobia, and phonophobia but other symptoms (such as aura, allodynia, osmophobia, neck pain, or dizziness) were also found in publications.

Nausea, Vomiting, Photophobia, Phonophobia, and Others (n=299)

Of the previously mentioned "core" associated symptoms, nausea was the most commonly assessed (Table 20), 96% of the publications that examined associated symptoms including an assessment of nausea. Photophobia was next most common (84.3%) followed by phonophobia (76.9%). With respect to the rating scales used, across associated symptoms the most commonly used response scale was binary (Presence/Absence; ranging from 45% to 86% across specific symptoms); the use of a binary response scale for associated symptoms aligns with FDA current recommendations (FDA, 2018).

Table 20. Rating Scale Features of Publications Assessing Non-Pain Symptoms (n=299)

Associated Symptom	Percent	N
Nausea	95.99	287
Binary	77.35	222
Ordinal	18.82	54
Continuous	3.14	9
Other	0.70	2
Vomit	57.19	171
Binary	85.96	147
Ordinal	12.28	21
Continuous	0.58	1

Other	1.17	2
Photophobia	84.28	252
Binary	83.73	211
Ordinal	15.08	38
Continuous	0.79	2
Other	0.40	1
Phonophobia	76.92	230
Binary	84.35	194
Ordinal	14.78	34
Continuous	0.43	1
Other	0.43	1
Aura	3.68	11
Binary	72.73	8
Ordinal	27.27	3
Continuous	0.00	0
Other	0.00	0
Other Symptoms	13.38	40
Binary	45.00	18
Ordinal	37.50	15
Continuous	15.00	6
Other	2.50	1

With respect to the types of endpoints used in assessing associated symptoms, the majority (57.2%) of publications used fixed-timepoint analyses (often comparing across treatment groups the proportion of subjects with presence/absence of a symptom at specific timepoints) (Table 21).

Table 21. Endpoint Types Used in Publications Assessing Non-Pain Symptoms (n=299)

Endpoint type	Percent	N
Change from Baseline	35.45	106
Fixed Timepoint	57.19	171
Fixed and Change from Baseline	7.36	22

For analyses that investigated non-pain migraine attack symptoms, the most common endpoint used was at two hours post-treatment (82.6%), followed by analyses at one hour post-treatment (47.8%) (Table 22).

Table 22. Endpoint Definitions Used in Non-Pain Symptoms (n=299)

Endpoint Definition	Percent	N
10m	4.01	12
15m	6.69	20
20m	4.01	12
25m	0.67	2
30m	34.45	103
45m	4.01	12
1hr	47.83	143

90m	20.07	60
2hr	82.61	247
3hr	13.04	39
4hr	32.11	96
24hr	11.37	34
Other	16.72	50

Most Bothersome Symptom (n=16)

A relatively recent introduction to the assessment of non-pain migraine attack symptoms is the measurement of the most bothersome migraine-associated symptom, which is currently recommended by the FDA as a co-primary endpoint in acute migraine trials (e.g., Diener, 2019; FDA, 2018). The definition of most bothersome symptom (MBS) requires that patients designate their most bothersome (non-pain) migraine symptom from the choices of nausea, photophobia, or phonophobia; the MBS may be designated prior to randomization (and patients only then treat attacks in which MBS is present) or can by designated at the start of each attack, but prior to administration of study drug. Given the recent introduction of MBS, a limited number of publications were located that assessed it (n = 16). In those publications that did assess MBS, 93.8% of them used a binary (Present/Absent) response scale. The majority of these 16 publications (62.5%) used a fixed timepoint endpoint type for analyses, while 37.5% used a change from baseline formulation. As seen in Table 23, all 16 publications that assessed MBS (100.0%) used a two-hour post-treatment endpoint definition, with publications also commonly investigating MBS at one hour (43.8%) and 90 minutes (both 37.5%)

Table 23. Endpoint Definitions Used in Most Bothersome Symptom (n=16)

Endpoint Definition	Percent	N
10m	0.00	0
15m	0.00	0
20m	6.25	1
25m	0.00	0
30m	37.50	6
45m	6.25	1
1hr	43.75	7
90m	37.50	6
2hr	100.00	16
3hr	6.25	1
4hr	25.00	4
24hr	6.25	1
Other	12.50	2

DISABILITY/IMPAIRMENT (N=186)

Disability/impairment refers to the decrement in a subject's ability to function normally in a wide range of possible domains, such as daily life activities, self-care, mobility, or in employment/work-related contexts. We note that the majority of publications assessing disability did so with a single item, we tracked the number of publications using the IHS-recommended functional disability item or one very similar to it. Of the 186 publications that assess disability/impairment in some way, 67.7% of them used the IHS functional disability item or one substantially similar (Table 24).

Table 24. Rating Scale Features of Publications Assessing Disability/Impairment (n=186)

Variable	Percent	N
IHS Recommended Functional Disability Item		
No/Unknown	32.26	60
Yes	67.74	126
Binary	9.14	17
Ordinal	82.80	154
3 categories	1.95	3
4 categories	88.96	137
5 categories	7.14	11
Other	1.95	3
Continuous	5.38	10
Other/Multiple	2.69	5

With respect to the response scales used for assessing disability, a large majority used an ordinal response scale with four possible response categories (often corresponding to the IHS recommended item response options). Within the continuous category, there were continuous response scales (such as NRSs or VASs) but this could also include such outcomes as time lost to disability or estimated efficiency (as a percent of normal capacity) at work.

Table 25 provides a breakdown of the type of endpoints that were seen in analyses examining disability. As can be seen, the most common endpoint type was fixed timepoint analyses (51.6%), although 44.6% of the publications examined change from baseline.

Table 25. Endpoint Types Used in Publications Assessing Disability/Impairment (n=186)

Endpoint type	Percent	N
Change from Baseline	44.62	83
Fixed Timepoint	51.61	96
Fixed and Change from Baseline	2.69	5
Other	1.08	2

Finally, the endpoint definitions used in disability analyses demonstrated relatively limited variability across publications. As can be seen in Table 26, the most commonly used endpoints were measured at two hours (83.9%) and one hour post-treatment (54.3%), followed by 30 minutes post-treatment (38.2%).

Table 26. Endpoint Definitions Used in Assessing Disability (n=186)

Endpoint definition	Percent	N
10m	4.30	8
15m	7.53	14
20m	1.61	3
25m	0.00	0
30m	38.17	71
45m	7.53	14
1hr	54.30	101
90m	25.81	48
2hr	83.87	156
3hr	17.20	32
4hr	33.33	62
24hr	13.44	25
Other	20.97	39

Patient-reported Outcome Measures (n=159)

Combining all PROMs (headache-related and non-headache specific PROMs), 35.3% of publications examined one or more PROM. Most of the 159 publications assessing a PROM examined one or more non-headache specific PROM (94.3%) and publications less frequently assessed one or more headache-related PROM (17.6%).

Headache-related PROMs (n=28)

Compared to the preventive literature, the use of headache/migraine-related PROMs in acute migraine trials was much less frequent. As seen in Table 27, of the 28 publications that examined one or more headache-related PRO, about 60% used the 24-hour Migraine-Specific Quality of Life Questionnaire (24hr MSQoL; Hartmaier et al., 1995) and 17.9% assessed the Patient Perception of Migraine Questionnaire-Revised (PPMQr; Revicki et al., 2006). Other headache-related item/scales that were not used in 5 or more publications are not reported here but a full list (from all n=705 publications) is available in the appendix.

Table 27. Headache/migraine related PROMs Used in Acute Migraine Publications (n=28)

Headache/Migraine-specific PROM Used	Percent	N
24hr MSQoL	60.71	17
PPMQr	17.86	5

Note. 24hr MSQoL = 24-hour Migraine-Specific Quality of Life Questionnaire. PPMQr = Patient Perception of Migraine Questionnaire-Revised.

Table 28 shows that almost three-quarters of the publications assessing one or more headache-related endpoint used change from baseline (71.4%) and the remaining publications examined fixed timepoints.

Table 28. Endpoint Types Used for Headache-related PROMs (n=28)

Endpoint Type	Percent	N
Change from Baseline	28.57	8
Fixed Timepoint	71.43	20

Non-headache specific PROMs (n=150)

Non-headache specific PROMs are scales/items that are not directly related to headache and often are used in a variety of disease areas; 150 publications used at least one non-headache specific PROM (see Table 6). The most commonly used non-headache specific PROM measures/items seen in the examined acute migraine trials were related to treatment satisfaction (27.3%), treatment efficacy (38.0%), and treatment preference (28.7%) (Table 29). Many other scales/items did not fit within these existing categories and the full list of other occurring scales is available in the appendix (n=705 publications). Treatment satisfaction was often measured using an ordinal scale with four categories (19.5%) or seven categories (41.5%). Treatment efficacy was often measured using an ordinal scale with four categories (36.8%) or five categories (43.9%). Treatment preferences was most frequently measured using a binary scale (60.5%) or ordinal scale with three categories (16.3%) or five categories (11.6%).

Table 29. Non-headache-specific PROMs Used in Acute Publications and Their Response Scales (n=150)

Rating scale features of non-headache specific PROMs	Percent	N
Patient Global Impression of Change	6.67	10
7 categories	80.00	8
Unknown	20.00	2
Patient Global Impression of Severity	1.33	2
7 categories	100.00	2
Treatment Satisfaction	27.33	41

2 categories	9.76	4
3 categories	0.00	0
4 categories	19.51	8
5 categories	12.20	5
6 categories	2.44	1
7 categories	41.46	17
Other	14.63	6
Treatment Efficacy	38.00	57
2 categories	3.51	2
3 categories	3.51	2
4 categories	36.84	21
5 categories	43.86	25
6 categories	1.75	1
7 categories	5.26	3
Other	5.26	3
Treatment Preference	28.67	43
2 categories	60.47	26
3 categories	16.28	7
4 categories	2.33	1
5 categories	11.63	5
6 categories	2.33	1
7 categories	2.33	1
Other	4.65	2

Like the headache-related PROMs, the non-headache specific PROM endpoint type was primary based on fixed timepoints (94.7%) and change from baseline was less often observed (6.0%) (Table 30).

Table 30. Endpoint Types Used for Non-headache specific PROMs (n=150)

Endpoint Type	Percent	N
Change from Baseline	4.67	7
Fixed Timepoint	93.33	140
Fixed Timepoint and Change from Baseline	1.33	2
Unknown	0.67	1

DISCUSSION

We conducted a systematic literature review of clinical trials for migraine acute therapies in adults published in peer-reviewed scientific outlets. Of the complete set of 705 publications, almost 94% were published in 1988 or later, three quarters utilized randomization, 73.3% utilized a blinded approach, and 79.4% used ICHD criteria to identify migraine or other headache diagnosis. About 60% of the publications were placebo/sham controlled and more than 95% assessed a pharmacological or medical device treatment (96.3%). Nearly all (96.6%) assessed efficacy and 85.8% assessed safety.

In our focused subset of 451 publications that were published in 1988 or later and adhered to certain design characteristics (e.g., blinded, randomized), participants were largely around 40 years old (mean age across publications 39.1) and tended to be white (mean across publications 85.1%) women (mean across publications 82.9%). These results align broadly with previously reported epidemiological studies of people with migraine (e.g., Buse et al., 2012; Buse et al., 2013; Lipton et al., 2007).

With respect to the outcomes used in the subset, most publications examined one or more pain-related outcome (95.3%). These publications tended to focus on pain relief (72.1%), pain freedom (64.9%), and headache recurrence/rescue medication use (76.5%). Many of these publications focused on assessing pain-related outcomes at 2 hours post-treatment. Non-pain symptoms and MBS were also often evaluated (67.2%). Associated symptoms such as nausea, photophobia, and phonophobia were frequently measured (299 publications) at 2 hours post-treatment and MBS was assessed far less frequently (only 16 publications total), due to its recent introduction to the migraine research paradigm. Over one-third of eligible publications looked at disability/impairment and one or more PROM (41.2% and 35.3%, respectively). Of the 159 publications assessing one or more PRO, 94.3% examined a non-headache specific PROM (often single item measures of treatment satisfaction, efficacy, or preference) while only 17.6% examined one or more headache-related PROM (most commonly the 24hr MSQOL and PPMQr measures). Both headache-related and non-headache specific PROMs were often evaluated using a fixed time-point endpoint type; 71.4% (headache-related) and 93.3% (non-headache).

When examining if publications followed current acute clinical trial guidelines (Diener et al., 2019), of 430 publications that assessed headache pain intensity 89.1% used an IHS recommended pain scale (a four-category ordinal scale, an 11-point NRS, or a 100mm VAS). Of 186 publications that assessed disability or impairment, 67.7% used the IHS guideline recommended single functional disability item or something similar. The majority (81.7%) of publications that assessed pain relief or headache relief defined it in a manner consistent with IHS acute clinical trial guidelines recommendations and of the 375 publications that studied pain freedom, 82.1% defined it consistent with guideline recommendations.

Nonetheless, as can be seen in this report, there was wide variety in study design, endpoints included, and how endpoints were measured across publications. The endpoints and outcomes used in acute migraine treatment trials, even when “common” outcomes are used, had inconsistent operationalizations across publications. We use the term publication because in some cases several or many publications came from a single study. As demonstrated in summaries of data extracted from the articles reporting on acute clinical trials, the outcomes used to define endpoints in such trials vary substantially, ranging from changes in pain and associated symptoms (relief, freedom, sustained relief, etc.), use of acute/rescue medication frequency

(days, doses), and various headache-related and general PROM measures, such as those related to the impact migraine has on the patient's life or more general HRQoL. The definition of the endpoints used (e.g., change from baseline, fixed-time point comparisons, categorization of "responders" to treatment based on wide variety of "responder definitions") also differs substantially across publications.

We narrowed our focus to a subset of publications from 1988 or later that reported on randomized and blinded trials of pharmacologic and medical device interventions to attempt to summarize the field as it currently stands and to focus on clinical trials likely to be similar to future trials designed to obtain FDA approval for new acute interventions. Even within this subset, a large amount of variability exists in the outcomes and endpoints used and how those outcomes were operationalized. The results from examining the full set of 705 articles (see Appendix B), with an even wider range of publications and including changes in diagnostic criteria, the treatment landscape, and the fields understanding of migraine, demonstrated even more variability and lack of standardization across trials.

This report has limitations. It is likely that we also did not identify every acute migraine trial publication ever published; however, our sample of 705 fully reviewed and extracted publications is likely large enough to be representative of the field. We did not group publications by the parent study, so there may be unequal weighting for some studies or data sets which have several publications (e.g., efficacy and PRO-related analyses reported in separate publications) and, in a small number of cases, publications did not provide enough information to reliably report on specific attributes of endpoints (e.g. rating scale, timing of endpoints, whether scale was IHS-compliant). Additionally, as with any review of published literature there is the possibility that publication bias of positive studies has affected the results and that, within individual publications, outcomes and endpoints that were supportive of efficacy were selectively reported as well.

Findings from this report showed that while there were some common elements that often aligned with current recommendations from guidelines (e.g., assessing pain and associated symptoms, focusing on the 2 hour post-treatment timepoint, using IHS-recommended scales), there were examples of studies not following recommended guidelines (e.g., few PROMs were used, especially headache-related) and inconsistencies (e.g., specific pain-related outcomes, associated symptoms, and endpoint types/timing varied). Based upon these results, we recommend developing a uniform set of PRO-based endpoints to facilitate comparability across study reports and investing in qualitative work to confirm that the endpoints used represent what matters most to patients. The proposed tool should be reliable, valid, and sensitive to group level and within person change. Disease-specific measures should be optimally sensitive to within person change to facilitate treatment across classes. In acute migraine clinical trials, there are currently few disease-specific measures available. The limited number of publications that did assess headache-related PROMs tended to use the 24hr MSQOL or PPMQr instruments, which, to our knowledge, have not been demonstrated to be fit-for-purpose and lack the qualitative and quantitative support needed to draw inferences regarding treatment efficacy.

Future work to rigorously evaluate the existing scales and subscales against a prespecified set of criteria, and to explore other domains, such as physical and cognitive function, may be useful and may lead to the development of PROM instruments that fulfill the current area of opportunity in assessing outcomes in migraine trials with patient-centered and patient-relevant PROMs. Based on prior qualitative work and feedback from the FDA, the development of measures of cognitive function may also be informative. Future work should both

capture and distinguish between the ictal and interictal burden of migraine. Regardless of the specific endpoints and outcomes supported by patient feedback and future psychometric development, we believe that the need for a standardized approach to study design in migraine trials has been demonstrated. Current trials exhibit a large amount of variability in outcomes and endpoints used, in addition to the variability in operationalization of outcomes when "common" outcomes are used across trials. Standardization of the outcomes and endpoints used in acute migraine trials will facilitate cross-trial comparisons and allow for patients to develop a framework for understanding the possible outcome differences in the wide variety of acute migraine treatment options that are available.

REFERENCES

Ad Hoc Committee on Classification of Headache. (1962). Classification of headache. *Journal of the American Medical Association, 179*, 717–718.

American Headache Society. (2019). The American Headache Society position statement on integrating new migraine treatments into clinical practice. *Headache, 59*(1), 1-18.

Buse, D. C., Loder, E. W., Gorman, J. A., Steward, W. F., Reed, M. L., Fanning, K. M., Serrano, D., & Lipton, R. B. (2013). Sex differences in the prevalence, symptoms, and associated features of migraine, probable migraine and other severe headache: Results of the American Migraine Prevalence and Prevention (AMPP) study. *Headache, 53*(8), 1278-1299.

Buse, D. C., Manack, A. N., Fanning, K. M., Serrano, D., Reed, M. L., Turkel, C. C. & Lipton, R. B. (2012). Chronic migraine prevalence, disability, and sociodemographic factors: results from the American Migraine Prevalence and Prevention Study. *Headache, 52*(10), 1456-1470.

Diener, H-C., Tassorelli, C., Dodick, D. W., Silberstein, S. D., Lipton, R. B., Ashina, M., Becker, W. J., Ferrari, M. D., Goadsby, P. J., Pzo-Rosich, P., Wang, S-J., & Mandrekar, J., (2019). Guidelines of the International Headache Society for controlled trials of acute treatment of migraine attacks in adults: Fourth edition. *Cephalagia, 39(6),* 687-710.

GBD 2016 Headache Collaborators (2018). Global, regional, and national burden of migraine and tension-type headache, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology, 17*(11), 954-976.

Goadsby, P.J., & Evers, S. (2020) International Classification of Headache Disorders- ICHD-4 alpha. *Cephalalgia, 40,* 887-888.

Hartmaier, S. L., Santanello, N. C., Epstein, R. S., & Silberstein, S. D. (1995). Development of a brief 24-hour migraine-specific quality of life questionnaire. *Headache, 35*(6), 320-329.

Headache Classification Committee of the International Headache Society. (1988). Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. *Cephalalgia, 8*(S7), 1–96.

Headache Classification Committee of the International Headache Society. (2004). The International Classification of Headache Disorders: 2nd edition. *Cephalalgia, 24*(S1), 9-160.

Headache Classification Committee of the International Headache Society. (2013). The International Classification of Headache Disorders: 3rd edition (beta version). *Cephalalgia, 33*(9), 629-808.

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

International Headache Society Committee on Clinical Trials in Migraine (1991). Guidelines for controlled trials of drugs in migraine. *Cephalalgia, 11*, 1–12.

Kawata, A. K., Hsieh, R., Bender, R., Shaffer, S., Revicki, D. A., Bayliss, M., Buse, D. C., Desai, P., Sapra, S., Ortmeier, B. & Hareendran, A. (2017). Psychometric evaluation of a novel instrument assessing

the impact of migraine on physical functioning: The Migraine Physical Function Impact Diary. *Headache 57*(9), 1385-1398.

Lipton, R. B., Bigal, M. E., Diamond, M., Freitag, F., Reed, M. L., & Steward, W. F. (2007). Migraine prevalence, disease burden, and the need for preventive therapy. *Neurology, 68*(5), 323-349.

Marmura, M. J., Silberstein, S. D., & Schwedt, T. J. (2015). The acute treatment of migraine in adults: The American Headache Society evidence assessment of migraine pharmacotherapies. *Headache 55*(1), 3-20.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica, 22*(3), 276-282.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group. (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses:* The PRISMA Statement. *Journal of Clinical Epidemiology, 62, 1006-1012. doi:10.1016/j.jclinepi.2009.06.005*

Revicki, D. A., Kimel, M., Beusterien, K., Kwong, J. W., Varner, J. A., Ames, M. H., Mahajan, S. & Cady, R. K. (2006). Validation of the Revised Patient Perception of Migraine Questionnaire: Measuring satisfaction with acute migraine treatment. *Headache, 46*(2), 240-252.

Tassorelli, C., Diener, H. C., Dodick, D. W., Silberstein, S. D., Lipton, R. B., Ashina, M., Becker, W. J., Ferrari, M. D., Goadsby, P. J., Pozo-Rosich, P., & Wang, S. J. (2018). Guidelines of the International Headache Society for controlled trials of preventive treatment of chronic migraine in adults. *Cephalalgia, 38*(5), 815-832.

Tfelt-Hansen, P., Block, G., Dahlöf, C., Diener, H. C., Ferrari, M. D., Goadsby, P. J., Guidetti, V., Jones, B., Lipton, R. B., Massiou, H., Meinert, C., Sandrini, G., Steiner, T., & Winter, P. O (2000). Guidelines for controlled trials of drugs in migraine: second edition. *Cephalalgia, 20*(9), 765-786.

Tfelt-Hansen, P., Pascual, J., Ramadan, N., Dahlöf, C., D'Amico, D., Diener, H., Hansen, J. M., Lanteri-Minet, M., Loder, E., McCrory, D., Plancade, S., & Schwedt, T. (2012). Guidelines for controlled trials of drugs in migraine: A guide for investigators. *Cephalalgia, 32*(1), 6-38.

US Food and Drug Administration. (2018). Migraine: Developing drugs for acute treatment Guidance for Industry. https://www.regulations.gov/document?D=FDA-2014-D-1540-0008.

APPENDIX A: PUBLICATION TRACKING MATRIX OF IDENTIFIED CANDIDATE PUBLICATIONS

Appendix Acute Literature Review Arti

APPENDIX B: RESULTS FROM ALL 705 EXAMINED PUBLICATIONS

DEMOGRAPHIC AND DESCRIPTIVE VARIABLES OF PATIENTS

Available demographic characteristics for the patients from the included publications (pooled over all treatment groups) are summarized in Table A-1. The median total sample size was n=150 (25th percentile: 51, 75th percentile: 557). Of publications that reported age, gender, and/or race descriptives, the average age was found to be 39.0 (SD = 4.2), with 82.4% of patients being female, and 84.5% of patients reported as White/Caucasian.

Table A-1. Demographic and descriptive (n=705)

Variable	N	Mean	SD	Minimum	25th Percentile	Median	75th Percentile	Maximum
Total N	703	567.88	1671.71	6	51	150	557	33147
Mean Age	627	38.95	4.11	24.10	36.80	39.80	41.38	56.10
Percent Female	652	82.44%	9.51	35.00%	78.87%	84.00%	87.00%	100.00%
Percent White	233	84.47%	19.42	0.00%	80.30%	88.95%	95.20%	100.00%

Note. SD = Standard deviation.

Tables A-2 and A-3 summarize migraine and aura group characteristics of the publications that reported such features. Over 92% of the publications looked at general migraine (unspecified/multiple types). More than three-fourths of the publications examined mixed aura types (76.7%), 14.0% were not specified, 7.2% were without aura, and 2.0% were with aura only.

Table A-2. Migraine Group Characteristics (n=705)

Patient group characteristics	Percent	N
General migraine (classical/common migraine, unspecified, multiple types)	92.48	652
MRM	3.97	28
EM only	3.12	22
CM/TM only	0.43	3

Note. MRM = menstrually-related migraine. EM = Episodic migraine. CM = chronic migraine. TM = transformed migraine.

Table A-3. Aura Group Characteristics (n=705)

Aura Characteristics	Percent	N
Mixed (with and without aura)	76.74	541
Not specified	14.04	99
Without aura only	7.23	51
With aura only	1.99	14

Outcomes and Endpoints Used

As seen in Table A-4, over 90% (91.1%) of the publications looked at one or more pain-related outcomes while 57.5% examined one or more associated symptoms outcomes. A little over one-third of publications looked at disability/impairment outcomes, and a little over one-third of publications used a PRO. Most PROMs were non-headache specific PROMs (93.6%) and only 20% of publications that used a PROM utilized a migraine/headache-related PRO.

Table A-4. Outcomes Assessed Across Publications (n=705)

Outcome grouping	Percent	N
Pain-related	91.06	642
Pain Relief	67.29	432
Pain Free	58.41	375
Pain General	35.36	227
Meaningful Relief	7.63	49
Headache Recurrence/Rescue Med Use	69.16	444
Non-Pain Symptoms	57.49	406
Associated Symptoms (Nausea, Vomiting, Photophobia, phonophobia, etc.)	98.52	400
Most Bothersome Symptom	4.43	18
Disability/impairment	33.90	239
Patient Reported Outcomes (PROMs)	35.18	248
Non-headache specific PRO	93.55	232
Migraine/headache-related PRO	19.76	49

In examining the various combinations of outcomes used (Table A-5), over one-fifth of publications looked only at one or more pain-related outcome combined with non-pain symptom outcomes. Similarly, about 20% of publications looked at only pain-related outcomes (21.4%), and a little over 14% of publications examined pain-related, associated symptoms, disability/impairment, and used one or more PRO. The 35 publications (5.0%) listed as using none of the outcome in our constructed groupings reported primarily on safety studies (examining only adverse events), pharmacokinetic studies (examining assorted laboratory-provided values), or health economics outcomes.

Table A-5. Combinations Assessed Across Publications (n=705)

Pain-related	Non-pain symptoms	Disability/ impairment	PROMs (headache and non-headache specific)	Percent	N
Yes	Yes	No	No	21.56	152
Yes	No	No	No	21.42	151
Yes	Yes	Yes	Yes	14.18	100
Yes	Yes	Yes	No	13.33	94
Yes	No	No	Yes	7.80	55

Yes	Yes	No	Yes	7.80	55
No	No	No	No	4.82	34
Yes	No	Yes	Yes	2.70	19
Yes	No	Yes	No	2.27	16
No	No	No	Yes	1.99	14
No	No	Yes	No	0.71	5
No	No	Yes	Yes	0.71	5
No	Yes	No	No	0.71	5

Pain-related Outcomes (n=642)

As noted earlier, pain-related outcomes were the mostly commonly encountered outcomes in the reviewed acute migraine literature, with over 90% of the publications using one or more pain-related outcome in the reported study. In assessing headache intensity, IHS acute trial guidelines (Diener et al., 2019) allow for four possible response scales to be utilized: a four-category ordinal scale, an 11-point NRS, or a 100mm VAS. Of the publications that assessed headache pain intensity in some manner, 81.8% used an IHS recommended scale.

There were several commonly encountered specific outcomes that we have classified under the overarching pain-related outcome heading.

Pain Relief (n=432)

Pain relief in the context of clinical migraine trials refers to a reduction in headache pain that is not a complete resolution of the pain. The IHS acute trial guidelines (Diener et al., 2019) defines Headache Relief as a decrease in headache pain from moderate to severe at baseline to mild or moderate; while the current guidelines were recently published a majority of the publications assessing pain relief did define relief in a manner consistent with this recommendation (81.7%; top section of Table A-6). In looking at all publications examining pain relief, a large majority (89.4%) used an ordinal response scale, with continuous response scales (such as a VAS or NRS) being the second most commonly used. Of the publications that stated they used a VAS and/or NRS, over 75% reported using a VAS. Of note is that the term “VAS” was applied to a wide variety of ratings scales, ranging from true VAS (which asks patients to provide a tick mark on a line, the position of which is then measured), to NRSs (typically ranging from 0-10), to ordinal scales with as few as five response options.

Table A-6. Rating Scale Features of Publications Assessing Pain Relief (n=432)

Rating Scale Features	Percent	N
IHS-recommended Definition of Relief		
No/Unknown	18.29	79
Yes	81.71	353
Rating Scale Features		
Binary	1.85	8

Ordinal	89.35	386
4 categories	97.41	376
5 categories	1.81	7
Other	0.78	3
Continuous	6.02	26
Other/Multiple	2.78	12
VAS/NRS	6.48	28
NRS	21.43	6
VAS	78.57	22

With respect to the endpoint definitions used for pain relief, Table A-7 provides a detailed breakdown of the time used in defining a pain relief endpoint, with observed timings ranging from 10 minutes to 24 hours with publications typically examining more than one endpoint related to pain relief. However, the most commonly used endpoint definition for pain relief was 2 hours (82.4%).

Table A-7. Endpoint Definitions Used in Assessing Pain Relief (n=432)

Endpoint Definition	Percent	N
10m	4.40	19
15m	10.19	44
20m	3.47	15
25m	0.69	3
30m	40.74	176
45m	6.48	28
1hr	58.33	252
90m	24.07	104
2hr	82.41	356
3hr	13.89	60
4hr	34.03	147
24hr	14.58	63
Other	19.21	83

Building on the pain relief outcomes used, several different outcomes related to pain relief were also commonly examined, including sustained response (defined as meeting the criteria for pain relief at a given point and having no headache pain increases through a set later timepoint) and the consistency of obtaining pain relief across multiple attacks. Table A-8 provides a breakdown of the variation among general and specific outcomes derived from pain relief.

Table A-8. Additional Outcomes Derived from Pain Relief (n=432)

Additional outcomes derived from pain relief	Percent	N
Sustained Response	28.47	123
24hr	73.98	91
48hr	7.32	9
24hr and 48hr	14.63	18
Other	4.07	5
Consistency across attacks	12.96	56
Time to relief	12.96	56

Pain Free (n=375)

Pain freedom in the context of clinical migraine trials refers to a reduction in headache pain that is a complete resolution of the pain. The IHS acute trial guideline (Diener et al., 2019) defines Pain Freedom as patients who become freed from headache pain following treatment. As noted previously, multiple pain severity rating scales are allowed by IHS acute migraine trial guidelines; however, a comment in the guideline document does note that the four-category response scale is preferred. For pain freedom, therefore, we tracked the number of publications that used the four-category response scale for rating headache intensity and also defined freedom as the complete absence of pain (e.g., a response of "None" on the four-category scale). Of the 375 publications that examined pain freedom, 82.1% defined pain freedom in this way (top section of Table A-9).

Table A-9. Rating Scale Features of Publications Assessing Pain Freedom (n=375)

Variable	Percent	N
IHS-recommended Definition of Freedom		
No/Unknown	17.87	67
Yes	82.13	308
Binary	2.67	10
Ordinal	88.80	333
4 categories	99.40	331
Other	0.60	2
Continuous	5.33	20
Other/Multiple	3.20	12
VAS/NRS	5.60	21
NRS	19.05	4
VAS	76.19	16
NRS and VAS	4.76	1

In looking at all publications examining pain freedom, a large majority (88.8%) used an ordinal response scale, with continuous response scales (such as a VAS or NRS) being the second most commonly used. Of the publications that stated they used a VAS and/or NRS, over three-quarters (76.2%) reported using a VAS. Of note is that term "visual analog scale" was applied to a wide variety of ratings scales, ranging from true VAS (which asks patients to provide a tick mark on a line that is measured), to NRSs (typically ranging from 0-10), to ordinal scales with as few as 5 response options.

With respect to the endpoint definitions used for pain freedom, Table A-10 provides a detailed breakdown of the time used in defining a pain freedom endpoint, with observed timings ranging from 10 minutes to 24 hours and publications often defining multiple endpoints for pain freedom. However, the most commonly used endpoint definition was 2 hours (90.7%), conforming to the IHS acute trial guidelines and as was also seen with pain relief.

Table A-10. Endpoint Definitions Used in Assessing Pain Freedom (n=375)

Endpoint Definition	Percent	N
10m	4.27	16
15m	9.87	37
20m	3.20	12
25m	0.53	2
30m	38.93	146
45m	6.13	23
1hr	53.87	202
90m	25.07	94
2hr	90.67	340
3hr	10.40	39
4hr	31.73	119
24hr	17.33	65
Other	13.87	52

Building on the pain freedom outcomes used, several related outcomes were also commonly examined, including sustained response (defined as meeting the criteria for pain freedom at a given timepoint and having no headache pain recurrence through a set later timepoint) and consistency of obtaining pain freedom from headache pain across multiple attacks. Table A-11 provides a breakdown of the variation among general and specific outcomes derived from pain freedom.

Table A-11. Additional Outcomes Derived from Pain Freedom (n=375)

Additional outcomes derived from pain freedom	Percent	N
Sustained Response	42.40	159
24hr	76.73	122

48hr	6.92	11
24hr and 48hr	15.09	24
Other	1.26	2
Time to freedom	13.07	49
Consistency	10.40	39

General Pain (n=227)

General pain is used to describe general assessments of pain that did not conform to pain freedom, pain relief, or meaningful relief. Given the somewhat broader nature of the category, more variability was seen in the response scales used to assess pain and the types of analyses examined. Table A-12 provides an overview of the features of the utilized response scales for assessing pain. As can be seen, continuous response scales (most typical 11-point scales) were most often used in assessing general pain, followed by ordinal scales (within which, a four-category response scale was most common). Of the publications that specifically called the rating they used a VAS or NRS, almost 75% used a VAS; as with the previous pain-related outcomes, however, response scales that were termed "visual analog scale" by authors covered a wide range of scales, many of which that would not be considered a true VAS.

Table A-12. Rating Scale Features of Publications Assessing General Pain (n=227)

Rating Scale Features	Percent	N
Binary	1.32	3
Ordinal	40.09	91
3 categories	7.69	7
4 categories	80.22	73
5 categories	7.69	7
Other	4.40	4
Continuous	56.83	129
10 point	17.83	23
11 point	39.53	51
100 point	28.68	37
Other range	13.95	18
Other/multiple response scales	1.76	4
VAS/NRS	51.98	118
NRS	23.73	28
VAS	74.58	88
NRS and VAS	1.69	2

While the previously examined pain-related outcomes of pain freedom and pain relief are, by definition, change from baseline analyses, general pain outcomes could be examined as change from baseline and/or fixed timepoint analyses (e.g., comparing treatment groups on mean headache intensity values at 2 hours

post-treatment). Table A-13 provides a breakdown of the type of endpoints that were seen in analyses examining general pain. As can be seen, the most common endpoint type was change from baseline (58.6%).

Table A-13. Endpoint Types Used in Publications Assessing General Pain (n=227)

Endpoint type	Percent	N
Change from Baseline	58.59	133
Fixed Timepoint	31.72	72
Fixed and Change from Baseline	9.69	22
Responder definition	21.59	49
50%	44.90	22
MCID/MID	2.04	1
Multiple	16.33	8
Other	36.73	18

Within the change from baseline analyses, a subgroup of publications examined responder definitions (or within-person meaningful change thresholds). These analyses set a specific threshold for minimum change from baseline for the subject to be considered a "responder" at a certain time point and then compare proportions of responders across treatment groups. Of the publications that utilized responder definition analyses, the most commonly used threshold was a 50% reduction in pain from baseline, followed by "Other" which included an assortment of other values used to define responders (e.g., a 4 point reduction on an 11 point scale, 75% reduction from baseline).

Finally, the endpoint definitions used in general pain analyses demonstrated more variability than the previously described pain relief and pain freedom endpoints. As can be seen in Table A-14, the most commonly used endpoints were one hour and two hours post-treatment (both 52.4%), followed by 30 minutes post-treatment (37.9%).

Table A-14. Endpoint Definitions Used in Assessing General Pain (n=227)

Endpoint definition	Percent	N
10m	7.49	17
15m	15.42	35
20m	9.69	22
25m	1.76	4
30m	37.89	86
45m	11.89	27
1hr	52.42	119
90m	12.33	28
2hr	52.42	119
3hr	8.81	20

4hr	20.70	47
24hr	14.98	34
Other	27.75	63

Meaningful Relief (n=49)

Meaningful relief is generally described as a subjective concept in which each trial subject interprets what "meaningful relief" is to them and provides responses relative to their own subject-specific definition. The rating scale most often used in assessing meaningful relief (Table A-15) is binary (59.2%), which allows patients to indicate Yes/No to a question such as, "Have you obtained meaningful relief from your migraine?"

Table A-15. Rating Scale Features of Publications Assessing Meaningful Relief (n=49)

Rating Scale Features	Percent	N
Binary	59.18	29
Ordinal	12.24	6
4 categories	100.00	6
Continuous	28.57	14
VAS/NRS	4.08	2
NRS	50.00	1
VAS	50.00	1

While most publications that investigated meaningful relief used a "Time to" analysis (69.4%) and often employed a stopwatch or timestamp approach, allowing subjects to report the exact duration of time from treatment to achieving meaningful relief, publications also described Time to analyses based on meaningful relief items asked at fixed timepoints. The endpoints/timing of the meaningful relief assessments used are summarized in Table A-16. As can be seen, and mirroring all previous endpoints, two hours post-treatment (69.4%) was the most commonly used timepoint at which meaningful relief was assessed.

Table A-16. Endpoint Definitions Used in Assessing Meaningful Relief (n=49)

Endpoint Definition	Percent	N
10m	2.04	1
15m	12.24	6
30m	26.53	13
45m	6.12	3
1hr	38.78	19
90m	16.33	8
2hr	69.39	34
3hr	12.24	6
4hr	28.57	14

24hr	2.04	1
Other	14.29	7

Headache Recurrence and Rescue Medication Use (n=444)

The return of headache pain after it was resolved is termed headache recurrence. Recurrence was previously defined as the achieving pain freedom and then experiencing a return of moderate to severe headache pain. The most recent version of the IHS acute trial guidelines (Diener, et al., 2019) has moved to the term "relapse" which they define as the occurrence of a headache of any severity within 24 or 48 hours after the initial treatment. Given the variability in the definition and the relatively recent publication of the new IHS "relapse" guidelines, any publication which specifically stated that they examined headache recurrence/relapse are included here.

Headache recurrence/relapse was examined in 38.2% of the examined publications (Table A-17), with the majority of publications within that subset using a 24-hour cut-off to define the recurrence window. Fewer publications used a 48-hour recurrence/relapse window, which is preferred in the current IHS acute migraine guidelines.

Table A-17. Additional Pain-related Outcomes Used in Acute Migraine Publications (n=705)

Additional Pain-related Outcome	Percent	N
Headache Recurrence	38.16	269
24hr	81.41	219
48hr	7.43	20
24hr and 48hr	3.72	10
Other	7.43	20
Rescue Medication	52.06	367

The use of rescue medication or additional doses of treatment medication was also a commonly used outcome in acute migraine trials, with over 50% of the examined publications tracking subjects' use of additional medication to attempt to alleviate experience migraine attacks.

Other Non-Pain Symptoms and Most Bothersome Symptom (n=406)

As noted previously, other non-pain symptoms (often collectively referred to as associated migraine symptoms) were often examined in acute migraine trials with over 50% of the publications examining at least one non-pain symptom. Historically, the most commonly assessed associated symptoms of acute migraine attacks were nausea/vomiting, photophobia, and phonophobia but other symptoms (such as aura, allodynia, osmophobia, neck pain, or dizziness) were also found in publications.

Nausea, Vomiting, Photophobia, Phonophobia and Others (n=400)

Of the previously mentioned "core" associated symptoms, nausea was the most often included (Table A-18), with almost 95% of the publications that examined associated symptoms including an assessment of nausea. Photophobia was next most common (78%) followed by phonophobia (69.5%). With respect to the rating scales used, across associated symptoms the most commonly used response scale was binary (Presence/Absence; ranging from 39.3% to 83.5% across specific symptoms).

Table A-18. Rating Scale Features of Publications Assessing Non-Pain Symptoms (n=400)

Associated Symptom	Percent	N
Nausea	94.00	376
Binary	75.60	285
Ordinal	19.63	74
Continuous	3.45	13
Other	1.06	4
Vomit	56.00	224
Binary	83.11	187
Ordinal	14.22	32
Continuous	0.89	2
Other	1.34	3
Photophobia	78.00	312
Binary	81.79	256
Ordinal	16.29	51
Continuous	0.96	3
Other	0.64	2
Phonophobia	69.50	278
Binary	83.51	233
Ordinal	15.41	43
Continuous	0.36	1
Other	0.36	1
Aura	4.50	18
Binary	52.63	10
Ordinal	31.58	6
Continuous	10.53	2
Other	0.00	0
Other Symptoms	15.25	61
Binary	39.34	24
Ordinal	34.43	21
Continuous	21.31	13
Other	4.92	3

With respect the types of endpoints used in assessing associated symptoms, the majority (59.75%) of publications used fixed-timepoint analyses (often comparing across treatment groups the proportion of subjects with presence/absence of a symptom at specific timepoints).

Table A-19. Endpoint Types Used in Non-Pain Symptoms (n=400)

Endpoint type	Percent	N
Change from Baseline	33.75	135
Fixed Timepoint	59.75	239
Fixed and Change from Baseline	6.50	26

For analyses that compared non-pain migraine attack symptoms, the most common endpoint used was at two hours post-treatment (78.3%), followed by analyses at one-hour post-treatment (45.5%).

Table A-20. Endpoint Definitions Used with Non-Pain Migraine Symptoms (n=400)

Endpoint Definition	Percent	N
10m	3.75	15
15m	7.25	29
20m	5.00	20
25m	0.75	3
30m	30.50	122
45m	4.00	16
1hr	45.50	182
90m	18.25	73
2hr	78.25	313
3hr	10.50	42
4hr	29.50	118
24hr	12.25	49
Other	20.00	80

Most Bothersome Symptom (n=18)

A relatively recent introduction to the assessment of non-pain migraine attack symptoms is the measurement of the most bothersome migraine-associated symptom, which is currently recommended by the FDA as a co-primary endpoint in acute migraine trials (e.g., Diener, 2019). The definition of MBS requires that patients designate their most bothersome (non-pain) migraine symptom from the choices of nausea, photophobia, or phonophobia; the MBS may be designated prior to randomization (and patients only then treat attacks in which MBS is present) or can by designated at the start of each attack.

Given the recent introduction of MBS, a limited number of publications were located that assessed it (n = 18). In those publications that did assess MBS, 94.4% of them used a binary (Present/Absent) response scale. The

majority of these 18 publications (61.1%) used a fixed timepoint endpoint type for analyses, while 38.9% used a change from baseline formulation. As seen in Table A-21, all 18 publications that assessed MBS (100.0%) used a two hour endpoint definition, with publications also commonly investigating MBS at one hour (38.9%) and 30 and 90 minutes (both 33.3%)

Table A-21. Endpoint Definitions Used in Publications Assessing MBS (n=18)

Endpoint Definition	Percent	N
20m	5.56	1
30m	33.33	6
45m	5.56	1
1hr	38.89	7
90m	33.33	6
2hr	100.00	18
3hr	5.56	1
4hr	22.22	4
24hr	5.56	1
Other	16.67	3

DISABILITY/IMPAIRMENT (N=239)

Disability/impairment refers to the decrement in a subject's ability to function normally in wide range of possible domains, such as daily life activities, self-care, mobility, or in employment/work-related contexts. The current IHS acute trial guidelines (Diener et al., 2019) recommend that functional disability be assessed via a single item, "How well can you function right now?" with four possible response options (suggested response labels are "No disability (i.e., able to function normally)" to "Severe disability (i.e., unable to perform most to all activities of daily living or requiring best rest") or that scales (such as the Migraine Physical Function Impact Diary; Kawata et al., 2017) be used. Given that the majority of publications assessing disability did so with a single item, we tracked the number of publications using the IHS-recommended functional disability item or one very similar to it. Of the 239 publications that assess disability/impairment in some way, over 60% of them used the IHS item or one substantially similar (Table A-22).

Table A-22. Rating Scale Features of Publications Assessing Disability/Impairment (n=239)

Rating Scale Features	Percent	N
IHS Recommended Functional Disability Item		
No/Unknown	36.40	87
Yes	63.60	152
Binary	8.79	21
Ordinal	81.17	194
3 categories	4.64	9
4 categories	86.60	168

5 categories	6.70	13
Other	2.06	4
Continuous	7.53	18
Other	2.51	6

With respect to the response scales used for assessing disability, a large majority used an ordinal response scale with four possible response categories (often corresponding to the IHS recommended item response options). Within the continuous category, there were continuous response scales (such as NRSs or VASs) but this could also include such outcomes as time lost to disability or estimated efficiency (as a percent of normal capacity) at work.

Table A-23 provides a breakdown of the type of endpoints that were seen in analyses examining disability. As can be seen, the most common endpoint type was fixed timepoint analyses (56.1%), although 41% of the publications examined changed from baseline.

Table A-23. Endpoint Types Used in Publications Assessing Disability/Impairment (n=239)

Endpoint Type	Percent	N
Change from Baseline	41.00	98
Fixed Timepoint	56.07	134
Change from Baseline and Fixed Timepoint	2.09	5
Other	0.84	2

Finally, the endpoint definitions used in disability/impairment analyses demonstrated relatively limited variability across publications. As can be seen in Table A-24, the most commonly used endpoints were one hour (50.6%) and two hours post-treatment (77.8%), followed by 30 minutes post-treatment (33.1%).

Table A-24. Endpoint Types Used in Publications Assessing Disability/Impairment (n=239)

Endpoint Definition	Percent	N
10m	3.35	8
15m	7.53	18
20m	2.09	5
25m	0.42	1
30m	33.05	79
45m	6.69	16
1hr	50.63	121
90m	23.85	57
2hr	77.82	186
3hr	14.23	34
4hr	30.96	74
24hr	13.81	33

Other	22.59	54

PATIENT REPORTED OUTCOME MEASURES (N=248)

Combining all PROMs (Headache-related and non-headache specific), 35.2% of the 705 publications examined one or more PRO. Most of the 248 publications assessing a PRO, examined one or more non-headache specific PROM (93.6%) and publications less frequently assessed one or more headache-related PROM (19.8%).

Headache-related PROMs (n = 49)

Compared to the preventive literature, the use of headache/migraine-related PROMs in acute migraine trials was much less frequent. As seen in Table A-25, of the 49 publications that examined one or more headache-related PRO, 42.9% used the 24hr MSQoL and 20.4% assessed the PPMQr. Other headache-related item/scales that were not used in 5 or more publications were also present, highlighting again the lack of consistency in outcomes across publications and the high degree of variability in PROMs used across trials. These other "named" headache-related PROMs are provided in Table A-26; with longer recall periods than are typically usable in acute trials, some PROMs in Table A-26 (e.g., the 6-item short-form Headache Impact Test [HIT-6; Kosinski et al., 2003]; Migraine Disability Assessment [MIDAS; Stewart et al. 1999]) were most often seen in open-label studies that examined both acute endpoints (e.g., pain relief at 2 hours) and longer-term effects (e.g., HRQoL over 12 weeks).

Table A-25. Headache/migraine related PROMs Used in Acute Migraine Publications (n=49)

Headache/Migraine-specific PROM Used	Percent	N
24hr MSQOL	42.86	21
PPMQr	20.41	10
Migraine Specific Quality of Life Questionnaire (MSQ; e.g., Jhingran et al., 1998; Martin et al., 2000)	10.2	5

Note. Scales encountered in fewer than 5 publications are included in the "Other scales/items" category. 24hr MSQoL = 24-hour Migraine-Specific Quality of Life Questionnaire. PPMQr = Patient Perception of Migraine Questionnaire-Revised.

Table A- 26. Breakdown of other "named" headache-related PROMs used in acute migraine publications (n = 49)

Named Headache-related PROM	Percent	N
MIDAS	8.2	4
HIT-6	6.1	3
Completeness of Response Survey (Coon et al., 2012)	2.0	1
Hunter Headache Scale	2.0	1
Migraine Specific Quality of Life (MSQoL) measure (Wagner, Patrick, Galer, & Berzon, 1996)	2.0	1

Patient Perception of Migraine Questionnaire (PPMQ; Davis, Black, & Sleath, 2002)	2.0	1
Qualité de Vie et Migraine (Richard, et al., 1993)	2.0	1

Note. MIDAS = Migraine Disability Assessment. HIT-6 = 6-item short-form Headache Impact Test.

With respect to the type of endpoints used for the headache/migraine-specific PROMs, the most common endpoint type was a fixed timepoint comparison (67.4%) with 30.6% of the publications using a change from baseline analysis.

Non-headache related PROMs (n=232)

Non-headache specific PROMs are scales/items that are not directly related to headache and often are used in a variety of disease areas. As seen in Table A-27, the most commonly used non-headache specific PROM measures/items seen in the examined acute migraine trials were related to treatment satisfaction (33.2%), treatment efficacy (33.6%), treatment preference (31.9%), and there were several 'other' scales/items that did not fit within existing categories (complete list in Table A-28). Treatment satisfaction was often measured using an ordinal scale with seven categories (35.1%) or five categories (22.1%). Treatment efficacy was often measured using an ordinal scale with four categories (37.2%) or five categories (35.9%). Treatment preference was most frequently measured using a binary scale (48.7%) or ordinal scale with three categories (28.4%).

Table A-27. Non-headache-specific PROMs Used in Acute Publications and Their Response Scales (n=232)

Rating scale features of non-headache specific PROMs	Percent	N
Patient Global Impression of Change	4.74	11
7 categories	81.82	9
Unknown	18.18	2
Treatment Satisfaction	33.19	77
2 categories	7.79	6
3 categories	0.00	0
4 categories	14.29	11
5 categories	22.08	17
6 categories	1.30	1
7 categories	35.06	27
Other	19.48	15
Treatment Efficacy	33.62	78
2 categories	5.13	4
3 categories	7.69	6
4 categories	37.18	29
5 categories	35.90	28
6 categories	1.28	1

7 categories	5.13	4
Other	7.69	6
Treatment Preference	31.90	74
2 categories	48.65	36
3 categories	28.38	21
4 categories	4.05	3
5 categories	12.16	9
6 categories	1.35	1
7 categories	1.35	1
Other	4.05	3
36-Item Short-Form Health Survey (e.g., Ware & Sherbourne, 1992)	2.15	5

Table A- 28. Breakdown of "Named" non-headache specific PROMs used in acute publications (n=232)

PROM	Percent	N
Treatment Satisfaction Questionnaire for Medication (Atkinson et al., 2004)	1.29	3
Short form McGill Pain Questionnaire (Melzack, 1987)	0.86	2
EuroQoL-5 Dimension (e.g., Herdman et al, 2011)	0.43	1
Hamilton Depression (Hamilton, 1960)	0.43	1
Hamilton Anxiety (Hamilton, 1959)	0.43	1
Headache Care Center-Automated Neuropsychological Assessment Metrics (Levenson & Reeve, 1994)	0.43	1
Minor Symptom Evaluation Profile (Dahlöf, 1990)	0.43	1
Profile of Mood States (McNair, Lorr, & Droppleman, 1981)	0.43	1
Stanford Sleepiness Scale (Hoddes et al., 1972, 1973	0.43	1
Symptom Checklist-90-R (Derogatis & Fitzpatrick, 2004)	0.43	1

Like the headache-related PROMs, the non-headache specific PROM endpoint type was primary based on fixed timepoints (93.5%) and change from baseline was less often observed (5.2%).

Table A-29. Endpoint Types Used for Non-headache specific PROMs (n=232)

Endpoint Type	Percent	N
Change from Baseline	5.17	12
Fixed Timepoint	93.53	217
Fixed and Change from Baseline	0.86	2
Unknown	0.43	1

APPENDIX B REFERENCES

Atkinson, M.J., Sinha, A., Hass, S.L. (2004). Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire for Medication (TSQM), using a national panel study of chronic disease. *Health Qual Life Outcomes*, 2, 12.

Dahlof, C. (1990). Minor Symptoms Evaluation (MSE) Profile – a questionnaire for assessment of subjective CNS-related symptoms. *Scandinavian Journal of Primary Health Care (supplement)*, 1, 19-25.

Davis, K.H., Black, L., & Sleath, B. (2002). Validation of the Patient Perception of Migraine Questionnaire. *Value in Health*, 5, 422-430.

Derogatis, L. R., & Fitzpatrick, M. (2004). The SCL-90-R, the Brief Symptom Inventory (BSI), and the BSI-18. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment: Instruments for adults* (p. 1–41). Lawrence Erlbaum Associates Publishers.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*, 56-62.

Hamilton M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology*, *32*, 50–55.

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, 20(10), 1727–1736.

Hoddes, E., Dement, W. & Zarcone, V. (1972). The development and use of the Stanford Sleepiness Scale (SSS). Psychophysiology, 9: 150.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R. & Dement, W. C. (1973). Quantification of sleepiness: A new approach. Psychophysiology, 10(4), 431-436.

Jhingran, P., Osterhaus, J. T., Miller, D. W., Lee, J. T., & Kirchdoerfer, L. (1998). Development and validation of the Migraine-Specific Quality of Life Questionnaire. *Headache*, *38*, 295-302.

Kosinski, M., Bayliss, M.S., Bjorner, J.B., Ware Jr, J.E., Garber, W.H., Baternhorst, A., … Tepper, S., (2003). A six-item short form survey for measuring headache impact: The HIT-6™. *Quality of Life Research*, 12, 963-974.

Levinson, D. M., & Reeves, D. L. (1994). Automated Neuropsychological Assessment Metrics (ANAM): ANAM V1.0 Normative Data. (Report No. NCRF-TR-W-01). San Diego, CA: National Cognitive Recovery Foundation.

Martin B. C., Pathak, D. S., Sharfman, M. I., Adelman, J. U., Taylor, F., Kwong, W. J., & Jhingran, P. (2000). Validity and reliability of the Migraine-Specific Quality of Life Questionnaire (MSQ Version 2.1). *Headache*, *40*, 204-216.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1981). Profile of mood states. San Diego: Educational and Industrial Testing Service.

Melzack, R. (1987). The short-form McGill pain questionnaire. *Pain*, 30(2), 191-197.

Richard, A., Henry, P., Chazot, G., Massiou, H. Tison, S., Marconnet, R., ... D'Allens, H., (1993). Quality of life and migraine: validation of the QVM questionnaire. *Thérapie*, 48(2), 89-96.

Stewart, W.F., Lipton, R.B., Kolodner, K., Liberman, J., & Sawyer, J. (1999). Reliability of the migraine disability assessment score in a population-based sample of headache sufferers. *Cephalalgia*, 19, 107-114.

Wagner, T.H., Patrick, D.L., Galer, B.S., & Berzon, R.A. (1996). A new instrument for the long-term quality of life effects from migraine: development and psychometric testing of the MSQOL. *Headache*, 36, 484–492.

Ware, J. E. Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, 30, 473–483.