



Intra- and interindividual attack frequency variability of chronic cluster headache

Cephalalgia
2023, Vol. 43(2) 1–8
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/03331024221139239
journals.sagepub.com/home/cep



Roemer B Brandt¹ , Wim Mulleners²,
Leopoldine A Wilbrink³, Paul Brandt⁴ , Erik W van Zwet⁵,
Frank JPM Huygen⁶, Michel D Ferrari¹ and Rolf Fronczek¹;
for the ICON study group*

Abstract

Background: The lack of knowledge about the intra- and interindividual attack frequency variability in chronic cluster headache complicates power and sample size calculations for baseline periods of trials, and consensus on their most optimal duration.

Methods: We analyzed the 12-week baseline of the ICON trial (occipital nerve stimulation in medically intractable chronic cluster headache) for: (i) weekly vs. instantaneous recording of attack frequency; (ii) intra-individual and seasonal variability of attack frequency; and (iii) the smallest number of weeks to obtain a reliable estimate of baseline attack frequency.

Results: Weekly median (14.4 [8.2–24.0]) and instantaneous (14.2 [8.0–24.5]) attack frequency recordings were similar ($p = 0.20$; Bland-Altman plot). Median weekly attack frequency was 15.3 (range 4.2–140) and highest during spring ($p = 0.001$) compared to the other seasons. Relative attack frequency variability decreased with increasing attack frequency ($p = 0.010$). We tabulated the weekly attack frequency estimation accuracies compared to, and the associated deviations from, the 12-week gold standard for different lengths of the observation period.

Conclusion: Weekly retrospective attack frequency recording is as good as instantaneous recording and more convenient. Attack frequency is highest in spring. Participants with ≥ 3 daily attacks show less attack frequency variability than those with < 3 daily attacks. An optimal balance between 90% accuracy and feasibility is achieved at a baseline period of seven weeks.

The ICON trial is registered in ClinicalTrials.gov under number NCT01151631.

Keywords

Cluster headache, clinical trials, headache

Date received: 27 June 2022; revised: 4 October 2022; accepted: 31 October 2022

Introduction

Cluster headache is a relatively rare primary headache disorder occurring in 0.1% of the general population and characterized by attacks of excruciating unilateral headache beside, behind, or in one eye. The headaches

⁴Department of Electronic Systems, Technical University Eindhoven, The Netherlands

⁵Department of Biomedical Data Sciences, Leiden University Medical Centre, Leiden, The Netherlands

⁶Department of Anaesthesiology, Erasmus MC, Rotterdam, The Netherlands

*ICON study group members are listed in the appendix.

¹Department of Neurology, Leiden University Medical Center, Leiden, The Netherlands

²Department of Neurology, Canisius-Wilhelmina Hospital, Nijmegen, The Netherlands

³Department of Neurology, Zuyderland Hospital, Heerlen, The Netherlands

Corresponding author:

Rolf Fronczek, Leiden University Medical Center, Department of Neurology, Albinusdreef 2, 2333 ZA, the Netherlands.
Email: r.fronczek@lumc.nl



last 15–180 minutes, can occur 8–10 times per day, and are typically accompanied by ipsilateral cranial autonomic symptoms, restlessness, or both (1,2). The most common form of cluster headache is episodic cluster headache, in which the attacks occur in clusters of several weeks to months, separated by remission periods of several months to years. Up to 20% of patients have the chronic form, in which remission periods are absent or last less than three months (2,3). Up to 15% of the patients with chronic cluster headache do not respond (adequately) to standard prophylactic medications, do not tolerate them, or have contraindications (medically intractable chronic cluster headache; MICCH) (4,5).

It is noteworthy that two thirds of patients with episodic cluster headache and one third of those with the chronic form have a seasonal pattern of attack frequency, with exacerbations in winter or spring (6–10). Often, there is also a diurnal pattern to the attack frequency, with many attacks occurring at night between 00:00 and 04:00 (9,11). Little is known about the intra- and interindividual variability of the attack frequency, making it difficult to reach consensus on the optimal length of baseline observation periods in clinical trials, especially for chronic cluster headache (12).

To obtain a reliable estimate of the variability of the attack frequency and intensity in patients with chronic cluster headache, we used data from the 12-week baseline observation period of the ICON trial (13,14). This is a randomized, double-blind, electrical dose-controlled clinical trial of the efficacy and safety of occipital nerve stimulation in patients with MICCH. Although these are the most difficult patients to treat, we consider the natural (untreated) variability of the attack frequency in this population to be representative of that in most patients with chronic cluster headache.

Prior to the ICON trial, we hypothesized that a short (two weeks) baseline observation period would not be accurate due to the fluctuating disease course and that weekly retrospective registrations would not be as accurate as instantaneous attack registration. In this study, we therefore aimed to (i) assess whether weekly retrospective registration of attack frequency are as reliable as daily instantaneous attack recordings; (ii) estimate intra-individual variability in diurnal and seasonal attack-frequency; and (iii) identify the optimal length of baseline observation periods for trials in chronic cluster headache that provides the best balance between feasibility and best estimate for baseline attack frequency.

Methods

Participants and questionnaires

Detailed data from the 12-week baseline observation period of the randomized, double-blind, dose-controlled

ICON trial of the efficacy and safety of occipital nerve stimulation in patients with MICCH were used (13,14). Participants diagnosed with chronic cluster headache according to the ICHD-II criteria, had an uneventful brain MRI since the onset of CH, a history of ≥ 4 weekly attacks, and failure of, intolerance to, or contraindications for verapamil *and* lithium in addition to at least one of methysergide, topiramate or gabapentin, were included (1,4,15–17). Participants did not change their current prophylactic therapy during the baseline period.

Using web-based diaries, participants completed (i) a screening questionnaire with baseline questions, (ii) a weekly retrospective diary throughout the 12-week baseline period with the total number and mean intensity (0–10) of attacks in the previous week and (iii) instantaneous recordings of each individual attack in the last six weeks of the baseline period.

Written informed consent was obtained from all participants and the study protocol was approved by the ethical committee of the LUMC (METC-LDD) and each participating center's local ethics committee. The ICON trial is registered in ClinicalTrials.gov under number NCT01151631.

Attack frequency variability

For each individual participant, the mean attack frequency over the entire 12-week baseline period was determined and regarded as the “true mean attack frequency” per participant.

Subsequently, all participants were placed in one of four subgroups:

- A. 0–14 weekly attacks (≤ 2 daily)
- B. 15–21 weekly attacks (2.1–3 daily)
- C. 22–28 weekly attacks (3.1–4 daily)
- D. >28 attacks (>4 daily)

We analyzed the variability of the intra-individual attack frequency in two ways:

1. The mean absolute weekly deviation, an absolute measure calculated as the average over the intra-individual absolute differences in each week from the 12-week mean, was determined and averaged per patient.
2. The coefficient of variation, a relative measure of intra-individual variation in attack-frequency calculated as the ratio of each participants' standard deviation to their mean.

Accuracy of different baseline periods

To determine the reliability of different lengths of baseline periods, the mean attack frequency per each

consecutive block of n weeks per participant was calculated. We thus obtained 11 averages for each consecutive block of two weeks, 10 averages for each consecutive block of three weeks, and so on. For each block of n weeks, the deviation from the participant's mean attack frequency was compared with its 12-week mean. The total percentage of blocks with a deviation of less than 10%, 20%, 30% or 50% was determined, thus providing an "accuracy measure" for each length. In addition, the total percentage of blocks with an absolute deviation of less than one, two, three, five or seven attacks from the 12-week mean was determined. For each of the subgroups, these calculations were performed only for participants who had completed all diaries during the 12-week baseline period.

Seasonal variability

For each week number, mean attack frequencies were obtained from the data provided by the patients for that specific week, creating a representation of the average attack frequency during the year. Subsequently, the mean weekly attack frequency per season was calculated using the following seasons: spring (weeks 10 to 22), summer (weeks 23 to 35), fall (weeks 36 to 48) and winter (weeks 49 to 9).

Statistics

Descriptive statistics were used to report median, mean, interquartile range and standard deviations. Absolute values for deviation from mean attack frequency were used. Separate analyses on subgroups were performed. A paired sample t-test was used to calculate the difference between instantaneous and retrospective weekly attack recordings and a Bland-Altman plot was used to visualize the data. Two-tailed p -values were deemed significant if $p < 0.05$ and were corrected using a Bonferroni correction for post-hoc testing. Statistical analysis was performed using SPSS version 25.

Results

Baseline characteristics

A total of 142 of the 150 patients who started the baseline period completed the final six weeks of both instantaneous and weekly registration and were included in this study. The 12-weekly attack journals were completed by 133/142 (94%) participants and only data of these participants were used for the baseline accuracy calculations. As is to be expected for studies in cluster headache, the majority of participants were male (63%). The side of the pain was equally divided

between left and right with 21 patients (15%) reporting attacks alternating between both sides. In total, 136 participants (96%) reported at least one autonomic symptom, with tearing eyes being the most frequent (80%) and eyelid edema the least frequent (37%) (Table 1). Only weeks in which participants filled out the diary were used for the analyses. In 62 out of 780 weeks, no attacks were registered (7.9%) and those weeks were regarded as missing.

Instantaneous versus weekly retrospective recording of attacks

In total, 15,889 attacks were recorded instantaneously (daily) in the last six weeks of the baseline period. Two participants did not reliably complete the registration and were excluded. The median attack frequency of the instantaneously (daily) recorded attacks was comparable with the median attack frequency based on weekly retrospective recordings (14.2 [8.0–24.5] vs. 14.4 [8.2–24.0], $p = 0.20$; Related-samples Wilcoxon Signed Rank Test). A Bland-Altman plot shows a normal distribution of the differences between the two measures without any evident regression (online Supplementary Material 1). The mean instantaneously recorded attack intensity was marginally, but significantly, lower than the mean weekly retrospectively recorded intensity (7.39 ± 1.64 vs. 7.53 ± 1.65 , $p < 0.01$; paired t-test).

Table 1. Baseline characteristics.

Patients (n)	142
Male, n (%)	90 (63%)
Age (years), mean \pm SD	45.5 \pm 12.9
BMI (kg/m^2), median (IQR)	25.1 (22.6–29.5)
Weekly attack frequency, median (IQR)	15.3 (9.3–23.4)
Attack intensity (0–10 scale), mean \pm SD	7.5 \pm 1.6
Absolute weekly attack variability, median (IQR)	3 (1.6–4.7)
Pain side, n (%) ^a	
Unilateral left	57 (40%)
Unilateral right	57 (40%)
Unilateral alternating	21 (15%)
Both sides simultaneously	2 (1%)
Autonomic symptoms (%)	
Lacrimation	80%
Rhinorrhea	68%
Ptosis	68%
Red eye	61%
Nasal congestion	62%
Hidrosis of the head	56%
Myosis	44%
Eyelid oedema	37%

BMI, body mass index; IQR, inter quartile range; SD, standard deviation.
^a4/142 missing.

Attack frequency variability during the entire 12-week baseline period

Interindividual variability in weekly mean attack frequencies ranged from 1.83 to 140. The distribution was skewed to the lower frequencies with a median of 15.3 weekly attacks (IQR: 9.3–23.4). Most participants (46%) reported a weekly attack frequency between 0 and 14 (0–2 daily attacks) (Table 2).

The median absolute intra-individual variability per group was 2.0 (IQR: 1.4–3.3) for group A (0–14 weekly attacks), 4.1 (IQR: 2.4–5.8) for group B (15–21 weekly attacks), 4.0 (IQR: 2.0–6.3) for group C (22–28 weekly attacks) and 4.6 (IQR: 2.2–8.0) for group D (>28 weekly attacks) (Table 2, online Supplementary Material 2a). Absolute attack variability differed between groups ($H(3) = 21.63$, $p = 0.000$). Post-hoc pairwise comparisons showed that the group with 0–14 weekly attacks (group A) exhibited a significantly lower absolute attack variability compared to the groups with 15–21 weekly attacks (group B, $p = 0.006$, $r = -0.327$); 22–28 weekly attacks (group C, $p = 0.023$, $r = -0.322$); and >28 weekly attacks (group D, $p = 0.001$, $r = -0.387$).

Although the absolute attack frequency variability increased with increasing attack frequency, relative variability decreased, as shown by a decrease in the coefficient of variation. The median coefficient of variation differed significantly between groups ($H(3) = 24.411$, $p = 0.000$). Post-hoc comparisons showed that the highest attack group exhibited a significantly lower coefficient of variation compared to that of the lowest attack groups (group A: $p = 0.000$, $r = 0.49$ and group B: $p = 0.040$, $r = -0.27$).

Baseline length accuracy

The percentage of blocks within various relative accuracy ranges of the 12-week mean per baseline block length is shown in Table 3. The table should be interpreted as follows: if one uses a 2-week baseline period, there is an 82% chance that the observed weekly mean attack frequency for this participant will be within 30%

of the 12-week mean attack frequency of the participant. However, if we want to have at least a 90% chance that the observed mean weekly attack frequency is within 20% accuracy range of the 12-week mean, the baseline should be at least seven weeks (Table 3A). The same logic applies to Table 3B that presents the absolute range from the 12-week mean. Due to the difference in attack variability between low and high attack frequencies, the tables differ for different baseline attack frequencies. These tables are provided in the supplemental materials (online Supplementary Material 3A and 3B).

Seasonal variability

Seasonal average weekly attack frequencies varied (One-way ANOVA; $F = 10.1$, $p < 0.001$). Post-hoc comparisons showed a higher mean attack frequency in the spring (mean \pm SD: 23.7 ± 3.5) compared to the summer (mean \pm SD: 19.7 ± 3.2 , $P = 0.001$), fall (mean \pm SD: 18.6 ± 1.8 , $P < 0.001$) and winter (mean \pm SD: 18.5 ± 2.0 , $P < 0.001$) (Figure 1).

Discussion

We had the unique opportunity to prospectively collect, analyze and compare daily and weekly retrospectively recorded attack frequency and intensity data from 142 patients with chronic cluster headache for 12 weeks. We made several observations that are important for future trial design in chronic cluster headache. Firstly, weekly retrospective recordings of the attack frequency are as good and reliable as instantaneous registrations of the attacks. Secondly, the interindividual distribution of the attack frequency was skewed, with most patients in the low attack frequency range. Thirdly, the higher the average intra-individual attack frequency, the more stable it was, with high absolute but lower than mathematically expected relative variability in participants with >3 daily attacks on average, and little absolute but higher than mathematically expected relative variability of the attack frequency in participants with a lower

Table 2. Characteristics per group based on number of weekly attacks.

	A (0–14 weekly attacks)	B (15–21 weekly attacks)	C (22–28 weekly attacks)	D (>28 weekly attacks)
Number (n, %)	66 (47%)	35 (25%)	15 (11%)	26 (18%)
Male (n, %)	46 (70%)	18 (51%)	12 (80%)	14 (54%)
Age (mean \pm SD)	48.9 \pm 13.0	44.7 \pm 12.2	38.1 \pm 11.4	42.1 \pm 12.2
Attack intensity (mean \pm SD)	7.2 \pm 1.7	7.4 \pm 1.6	7.3 \pm 1.6	8.2 \pm 1.5
Mean absolute deviation (median, IQR)	2.0 (1.4–3.3)	4.1 (2.0–5.8)	4.0 (2.0–6.3)	4.6 (2.2–8.0)
CoV (median, IQR)	0.34 (0.22–0.44)	0.27 (0.16–0.37)	0.20 (0.11–0.30)	0.14 (0.07–0.22)

CoV, coefficient of variation; IQR = interquartile range; SD = standard deviation.

Table 3. Percentage of week blocks that fall within different *relative* (A) and *absolute* (B) accuracy ranges of the 12-week mean, depicted against the week block lengths.

A. Percentage of week blocks that fall within different *relative* accuracy ranges of the 12-week mean, depicted against the week block lengths

Accuracy range of the 12-week mean	<10%	<20%	<30%	<50%
Week block length				
2 weeks	43	68	82	93
3 weeks	49	73	86	96
4 weeks	53	78	89	98
5 weeks	58	82	92	98
6 weeks	63	86	94	98
7 weeks	68	90	96	99
8 weeks	76	94	98	99
9 weeks	83	97	99	100
10 weeks	90	98	99	100
11 weeks	98	100	100	100

<50% of week blocks
 50%–75% of week blocks
 76%–90% of week blocks
 >90% of week blocks

B. Percentage of week blocks that fall within different *absolute* accuracy ranges of the 12-week mean, depicted against the week block lengths

Weekly attack range of the 12-week mean	≤1	≤2	≤3	≤5	≤7
Week block length					
2 weeks	28	47	61	76	86
3 weeks	34	54	65	81	90
4 weeks	34	55	69	85	92
5 weeks	39	61	74	88	94
6 weeks	45	66	80	91	96
7 weeks	49	71	83	94	97
8 weeks	56	76	87	96	98
9 weeks	64	81	93	97	99
10 weeks	73	90	96	99	99
11 weeks	85	96	99	100	100

<50% of week blocks
 50%–75% of week blocks
 76%–90% of week blocks
 >90% of week blocks

average attack frequency. Attack intensity did not vary much over time. Fourthly, although individual attack frequencies were highest during spring, mean and median group attack frequency varied little over time and there was no spontaneous decrease reflecting regression to the mean. Finally, a baseline observation period length of at least four weeks, most likely seven weeks, will probably provide an optimal balance between accuracy and feasibility. This length is dependent on the desired accuracy range of the baseline observation period the predefined outcome measure (relative or absolute change in attack frequency) and the expected baseline attack frequency levels of the included patients.

The finding that weekly recordings are as good as daily registration is of great convenience for participants in trials. Drop-out rate and number of missing data will likely be reduced, increasing the validity of the results.

Information on the interindividual variation and intra-individual variability of the attack frequency over time is important for power and sample size calculations, for deciding on whether the primary outcome measure is absolute or relative change in attack frequency, and for determining the optimal length of the baseline observation period. In the ICON study, intra-individual variability of the attack frequency differed between participants with a high and those with a

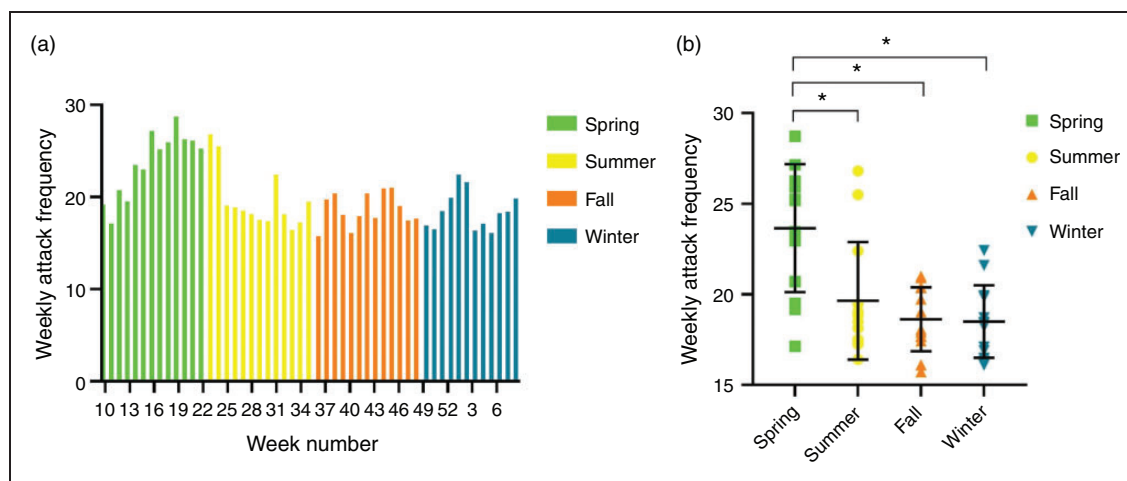


Figure 1. Mean weekly attack frequency. (a) Mean weekly attack frequency during the year and (b) mean weekly attack frequency is higher in spring (mean \pm SD: 23.7 ± 3.5) when compared to the summer (mean \pm SD: 19.7 ± 3.2 , $P = 0.001$), fall (mean \pm SD: 18.6 ± 1.8 , $P < 0.001$) and winter (mean \pm SD: 18.5 ± 2.0 , $P < 0.001$). * $p < 0.001$.

low pre-trial attack frequency. Participants with ≤ 3 attacks per day showed high relative and low absolute variability, whereas in participants with > 3 attacks per day the absolute variability was high and the relative variability was low.

By nature, absolute variability increases with increasing means since there is more ‘room’ to vary in. Relative variability however, is expected to remain constant if the data follows a “normal” Poisson distribution. However, in the ICON study baseline data, relative variability decreases, meaning that the attack variability does not increase as much with increasing means as would be mathematically expected if the data followed a normal Poisson distribution. This suggests a more stable attack frequency in participants with > 3 attacks per day than in those with ≤ 3 attacks per day. In practical terms, when calculating a sample size, this would equate to a constant standard deviation in patients with ≥ 3 daily attacks.

Thus, the required minimum lengths of the baseline and treatment periods for accurate estimates of the baseline attack frequency and treatment effects are strongly dependent on whether the primary outcome is an absolute or relative change and whether the predominant pre-trial attack frequency is greater or lower than three attacks per day. Baseline and treatment periods can be shorter in trials with participants with pre-trial on average > 3 attacks per day and relative (rather than absolute) change in attack frequency as primary outcome measure.

The shorter the baseline observation period, the more convenient this is for trial participants, which will likely translate into a lower risk of dropouts and

missing data, and thus an increased feasibility of the trial. On the other hand, the longer the length of the baseline period, the more accurate the estimate of the baseline attack frequency, but the greater the risk of dropouts and missing data. To derive the optimal length for a baseline period providing an optimal balance between accuracy and feasibility, we computed Tables 3A and B. These tables show the relation between different relative and absolute accuracy levels and associated required observation period lengths and can be used in trial design. Because of the relationship between attack frequency variability and attack rate, these tables differ between different pre-trial attack frequencies. If the pre-trial attack frequency can be estimated, one could choose to use a more specific table (Supplemental Tables), further improving the accuracy of the trial design. The challenge then lies in the estimation of the likely pre-trial attack frequency.

In this prospective study, the attack frequency was slightly higher in spring compared to that in other seasons. In retrospective studies, 41–56% of patients with episodic or chronic cluster headache reported seasonal variation in attack frequency (6–10), with most cluster bouts occurring in the spring (10). Retrospective studies, however have a high risk of bias. Circannual (8,9) and diurnal (6) rhythmicity appears to be less common in chronic compared to episodic cluster headache. In the ICON study, we observed only minimal rhythmicity, perhaps because the participants were medically intractable. The participants represented the extreme end of the severity spectrum and are possibly less susceptible to rhythmic influences. Due to the potentially increased attack frequency in spring, preferably

baseline observation periods should not be in spring or winter. In the first case there will be a high risk of a spontaneous reduction of the attack frequency during the treatment period causing a high placebo response (due to regression to the mean). In the second case, there will be a spontaneous increase in attack frequency, potentially complicating a true evaluation of the treatment effect. In the ICON study, there was little if any variability over time in the attack frequency for the whole participant group, presumably because patients were recruited during different periods over many years, thereby leveling out seasonal variations.

Limitations

Since the participants recorded attack occurrence both daily and weekly, we cannot exclude that this has increased the accuracy of the weekly retrospective registration. However, there were small but distinct differences in weekly and daily registrations of the attacks, making it less likely that patients just added up the daily attack recordings for a retrospective weekly attack frequency estimate.

In this study, we only describe the most severely affected patients with chronic cluster headache. Firstly because the diagnosis of chronic cluster headache was based on the ICHD-II criteria in which chronic cluster headache is defined as less than one month attack freedom instead of the less stringent

ICHD-III criterium of less than three months attack freedom, and secondly because only patients with MICCH were included. However, we would not expect a different study population if the ICHD-III criteria were used since sustained attack freedom (more than one month) is highly unlikely in patients with MICCH. Having said that, we can only speculate that the observed variability in this study is similar in patients with non-MI chronic cluster headache. Finally, we only had a 12-week registration period per patient limiting extrapolation to a true annual rhythmicity.

Conclusion

In conclusion, weekly retrospective recording of the attack frequency and intensity is as accurate as instantaneous daily recordings. Participants with ≥ 3 daily attacks show less attack frequency variability than those with < 3 daily attacks. As attack frequency is highest during spring due to natural fluctuations, it is recommended to avoid planning baseline observation periods in spring or winter. The optimal baseline length is strongly dependent on the preferred accuracy of the baseline attack frequency estimation, the desired outcome (relative or absolute attack frequency reduction) and the estimated baseline attack frequency. Most likely, a baseline length of at least seven weeks seems to offer the optimal balance between feasibility and accurate estimation of the baseline attack.

Article highlights

- Weekly retrospective recording of the AF is as good as instantaneous recording and more convenient.
- The attack frequency in MICCH patients with > 3 attacks per day is more stable than in those with ≤ 3 attacks per day.
- Attack frequency is highest in spring.
- An optimal balance between feasibility and accuracy is achieved at an observation period of seven weeks.



Declaration of conflicting interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: RBB and PB reports support for conference visits from Allergan; WM reports honoraria from Novartis, Teva, AbbVie, Lundbeck and Lilly, and consultancy and lecture fees from Lilly; RF reports consultancy and lecture fees from Novartis, Lundbeck, AbbVie, Lilly and TEVA, and independent support from the Dutch Brain Foundation, Leiden University Fund and Innovation Fund Dutch Healthcare Providers; FJPMH reports consultancy and lecture fees from ABBOTT, Saluda, Grunenthal and Pfizer; EWvZ, LAW, PB and MDF report no relevant conflict of interest.

Ethical approval and patient consent

Written informed consent was obtained from all participants and the study protocol was approved by the ethical committee of the LUMC (METC-LDD; protocol number P10.016) and each participating center's local ethics committee.

ORCID iDs

Roemer B Brandt  <https://orcid.org/0000-0002-2932-4872>
Paul Brandt  <https://orcid.org/0000-0002-2353-967X>

References

1. Headache Classification Committee of the International Headache Society (IHS) The International Classification

- of Headache Disorders, 3rd edition. *Cephalalgia* 2018; 38: 1–211.
2. Hoffmann J and May A. Diagnosis, pathophysiology, and management of cluster headache. *Lancet Neurol* 2018; 17: 75–83.
 3. Favier I, Haan J and Ferrari MD. Chronic cluster headache: a review. *J Headache Pain* 2005; 6: 3–9.
 4. Goadsby PJ, Schoenen J, Ferrari MD, et al. Towards a definition of intractable headache for use in clinical practice and trials. *Cephalalgia* 2006; 26: 1168–1170.
 5. Mitsikostas DD, Edvinsson L, Jensen RH, et al. Refractory chronic cluster headache: a consensus statement on clinical definition from the European Headache Federation. *J Headache Pain* 2014; 15: 79.
 6. Rozen TD and Fishman RS. Cluster headache in the United States of America: demographics, clinical characteristics, triggers, suicidality, and personal burden. *Headache* 2012; 52: 99–113.
 7. Bahra A, May A and Goadsby PJ. Cluster headache: a prospective clinical study with diagnostic implications. *Neurology* 2002; 58: 354–361.
 8. Barloese MCJ, Beske RP, Petersen AS, et al. Episodic and chronic cluster headache: differences in family history, traumatic head injury, and chronorisk. *Headache* 2020; 60: 515–525.
 9. Barloese MC, Lund NLT, Petersen AS, et al. Sleep and chronobiology in cluster headache. *Cephalalgia* 2015; 35: 969–978.
 10. Ofte HK, Berg DH, Bekkelund SI, et al. Insomnia and periodicity of headache in an arctic cluster headache population. *Headache* 2013; 53: 1602–1612.
 11. de Coo IF, van Oosterhout WPJ, Wilbrink LA, et al. Chronobiology and sleep in cluster headache. *Headache* 2019; 59: 1032–1041.
 12. Lipton RB, Micieli G, Russell D, et al. Guidelines for controlled trials of drugs in cluster headache. *Cephalalgia* 1995; 15: 452–462.
 13. Wilbrink LA, De Coo IF, Doesborg PGG, et al. Safety and efficacy of occipital nerve stimulation for attack prevention in medically intractable chronic cluster headache (ICON): a randomised, double-blind, multicentre, phase 3, electrical dose-controlled trial. *Lancet Neurol* 2021; 20: 515–525.
 14. Wilbrink LA, Teernstra OP, Haan J, et al. Occipital nerve stimulation in medically intractable, chronic cluster headache. The ICON study: Rationale and protocol of a randomised trial. *Cephalalgia* 2013; 33: 1238–1247.
 15. Favier I, van Vliet JA, Roon KI, et al. Trigeminal autonomic cephalgias due to structural lesions: a review of 31 cases. *Arch Neurol* 2007; 64: 25–31.
 16. de Coo IF, Wilbrink LA and Haan J. Symptomatic trigeminal autonomic cephalalgias. *Curr Pain Headache Rep* 2015; 19: 514.
 17. Wilbrink LA, Ferrari MD, Kruit MC, et al. Neuroimaging in trigeminal autonomic cephalgias: when, how, and of what? *Curr Opin Neurol* 2009; 22: 247–253.

Appendix I

ICON study group

Investigators are listed by country and center

1. The Netherlands: Leiden University Medical Center – MD Ferrari (Chair), LA Wilbrink, IF de Coo, PGG Doesborg, EC Bartels, EW van Zwet
2. Erasmus Medical Center – FJPM. Huygen (Vice Chair)
3. Canisius Wilhelmina Hospital – W Mulleners
4. Radboud Medical Center – RTM van Dongen
5. Zuyderland Hospital – OP Teernstra, PJJ Koehler, GH Spincemaille
6. Diaconessenhuis Zeist – F Wille
7. Alrijne Hospital – K Burger, J Haan
8. Boerhaave Medical Center – EGM Couturier
9. Rijnstate Hospital Arnhem – JW Kallewaard
10. University of Twente – PH Veltink
11. Medtronic BV – R Buschman
12. Belgium: CHR La Citadelle hospital – J Schoenen, D Magis
13. Hungary: National Institute of Neuroscience – LG Eross, J Afra
14. Germany: Schmerzlinik Kiel – H Göbel
15. Universitätsklinikum Schleswig-Holstein – V Tronnier