Published in final edited form as:

Headache. 2019 April ; 59(4): 495–508. doi:10.1111/head.13507.

Headache Triggers as Surprise

Dana P. Turner, MSPH PhD,

Massachusetts General Hospital, Harvard Medical School, Anesthesia, Critical Care and Pain Medicine (Boston, MA, USA)

Adriana D. Lebowitz, BA,

Massachusetts General Hospital, Harvard Medical School, Anesthesia, Critical Care and Pain Medicine (Boston, MA, USA)

Ivana Chtay, BS, and

Massachusetts General Hospital, Harvard Medical School, Anesthesia, Critical Care and Pain Medicine (Boston, MA, USA)

Timothy T. Houle, PhD

Massachusetts General Hospital, Harvard Medical School, Anesthesia, Critical Care and Pain Medicine (Boston, MA, USA)

Abstract

Objectives: To examine the hypothesis that surprising experiences of headache triggers are associated with daily headache activity.

Background: Little is known about the specific environmental or behavioral interactions that might trigger a headache attack in a prone individual. We propose that headache trigger exposures can be usefully characterized, not only by their mechanisms of action, but also on the degree of “surprise” they present to the individual. This hypothesis is based on elements of information theory: that unusual events and experiences carry more information than common events and experiences and that headache attacks are associated with reactions to uncommon or unexpected biopsychosocial exposures.

Methods: A secondary analysis of the Headache Prediction Study (HAPRED), this prospective cohort study followed N = 95 individuals with episodic migraine who contributed 4195 days of diary data. Information was collected on daily levels of several common headache triggers: number of caffeinated beverages, number of alcoholic beverages, stress (Daily Stress Inventory), and mood disturbance (Profile of Mood States). The probability of observing variations in each trigger was used to estimate the “surprisal” of experiencing each trigger, and this information, measured in bits, was used to predict headache attacks.

Results: Participants experienced a headache attack on 1613 of 4195 days (38.5%). Each of the triggers was modestly related to the probability of experiencing a future headache in a complex manner that involved interactions between current headache status, current levels of the trigger, and lagged (previous) levels of the trigger. However, when expressed as a surprisal, the

associations were simplified and strengthened. For example, each of the individual trigger surprisals exhibited a meaningful association with the development of a future headache attack (expressed as a 1 SD change in surprisal), with odds ratios ranging from a low of 1.11 (95%CI: 1.00 to 1.24) for alcohol, to a high of 1.30 (95%CI: 1.14 to 1.46) for stress. The individual surprisals could be summed for total trigger surprise and exhibited a reliable association with new onset headache, OR 1.35 (95%CI: 1.17 to 1.49). This score exhibited superior discrimination between headache and non-headache days from all of the individual triggers (ignoring base-rate, AUC: 0.61; AUC: 0.71 with base-rate).

Conclusions: Diverse headache triggers can be uniformly described using their probability distributions. Rare values of headache triggers, or surprising values, were found to have consistent associations with headache activity across a variety of triggers. This finding, if validated in external data using other triggers, has potential to be used to conceptualize the influence of a wide range of headache triggers.

Keywords

Migraine; Trigger; Information Theory; Stress; Forecasting

Introduction

Despite our general understanding of the pathophysiological and neurobiological cascades underlying the initiation of headache attacks,¹ little is known about the specific environmental or behavioral interactions that might trigger an attack in a prone individual.² Since the time of Hippocrates, a vast array of headache trigger candidates has been hypothesized ranging from foods, activities, behaviors, environmental conditions, and many others.³ Potency beliefs in these diverse factors vary considerably across individuals.^{4,5} Therefore, conceptualizing a common pathway to causing or initiating an attack is difficult. Nevertheless, several elegant hypotheses have been proposed to unify the causal influence of these triggers.

Martin⁶⁻⁹ and colleagues¹⁰⁻¹³ have proposed a fear-conditioning model that postulates that a trigger factor creates anxiety, and avoidance of this trigger reinforces this anxiety. Future (accidental) exposures to the feared trigger then elicit a fear-response, and the physiological responses to this fear trigger the attack. In this model, virtually anything can be conditioned to be a trigger, but careful exposure to the trigger extinguishes the fear response and corresponding potency of the trigger^{11,12}. Borkum has expounded a very different model based on oxidative stress.^{14,15} This model postulates that trigger factors all increase oxidative stress, either directly or indirectly, and the physiological response to this stress initiates a migraine attack. Borkum credits observations of Blau¹⁶ and Benemei¹⁷ in that migraine triggers can be uniformly characterized by their oxidative challenges to the brain and that migraine attacks are a defensive response to such challenges.¹⁵ A third model proposed by Burstein and Jakubowski organizes headache triggers by their inputs to the central nervous system.¹⁸ This model proposes that the diverse set of triggers activates a correspondingly diverse set of brain regions, but these regions all share projections to the superior salivatory nucleus. These projections, when activated, cascade to reduce activation of parasympathetic neurons in the meninges, and this leads to increased activation of the

meningeal nociceptors. Each of these hypotheses enjoys considerable face validity based on biological plausibility, and none are mutually exclusive or require the others to be untrue based on their own validity.

In this manuscript, we describe the evaluation of a new model to unify the conceptualization of headache triggers. We propose that headache trigger exposures can be usefully characterized, not only by their mechanism of action, but also on the degree of “surprise” they present to the individual. By “surprise” we do not imply a change in conscious awareness (i.e., shock or alarm) that is related to the exposure, but rather the rarity or unexpectedness in relation to past exposures. In this way, an exposure or trigger can be surprising even if it escapes conscious awareness. This hypothesis is entirely based on elements of information theory: that unusual events and experiences carry more information than common events and experiences and that headache attacks are triggered by reactions to uncommon or unexpected biopsychosocial challenges. Unlike existing theories, this model is agnostic about the exact biological mechanisms responsible for this proposed association. Rather than focusing on the proposed mechanisms that are initiated after exposure to a trigger, the model simply describes the amount of “information” being conveyed to an individual through this person’s interaction with the world. If supported, this trigger surprise model would have immediate implications for the assessment of daily headache risk based on the summation of trigger surprise.

Methods

The examination of the influence of trigger surprise on initiating new headache attacks is a secondary analysis of twice-daily headache diaries collected as part of the Headache Prediction Study (HAPRED). The primary analysis of the study was fully reported previously.¹⁹ HAPRED was a prospective cohort study conducted between 2009 and 2014 in Winston-Salem, NC. Participants recruited from the general community, neurology practices, and primary care clinics were eligible to participate in the study if they experienced episodic migraine with or without aura (International Classification of Headache Disorders 3rd edition, diagnosed by a study neurologist), experienced greater than two headaches per month of pain > 0 on a numerical scale, and experienced 4 to 14 headache days per month. Exclusion criteria included the presence of a secondary headache disorder, medication overuse, change in headache activity over the past 6 weeks, inability to read or speak English, psychiatric hospitalization within the past year, history of substance abuse, or pregnancy during observation period. The Wake Forest School of Medicine Institutional Review Board approved the data collection procedures (IRB00008743), and all participants provided written informed consent to participate in the study. The study protocol was not pre-registered in a clinical trial repository.

Summary of Diary Collection Procedures

As part of the study, 95 of 100 enrolled individuals contributed at least one diary entry for a total of 4626 days of electronic diaries. Individuals were followed between 11 and 96 days (mean: 49 days). The diaries were completed twice each day, once in the morning or on first awakening (AM) and once in the evening time or before going to bed (PM). Both AM and

PM diaries contained items related to headache activity and mood, while the PM diary also contained items related to stress and other commonly reported headache triggers like alcohol and caffeine intake that were encountered during the day. For the purposes of this analysis, the $n = 431$ (9.3%) of embedded missing diaries (i.e., diary entries that were missing during the observation period) were ignored.

Measures

Several well-validated daily assessment instruments and several single-item assessments were collected. In the original HAPRED study, these daily trigger assessments were specifically chosen because they represented popular headache triggers. A trigger candidate was defined as any factor that could be directly associated with the onset of a new headache attack or merely associated with other factors that are associated with a new headache attack. For this analysis, only several of the available measures were chosen based on sufficient between-person variability during the observation period (i.e., exposure to the triggers exhibited variability within and across individuals). Each of the daily measurements is described below.

Headache Activity Assessment (AM, PM diaries): The diaries assessed the presence or absence of headache activity of any kind, headache associated symptoms, and abortive medications since making the last diary entry. Headache intensity ratings were assessed using a 0 to 10 Numeric Rating Scale (NRS) corresponding to current level of pain and the worst pain experienced since the last entry.

Daily Stress Assessment (PM diary): Daily stress, or “daily hassles”, was assessed using the Daily Stress Inventory (DSI).²⁰ The DSI is a 58-item scale that allows an individual to identify stressful events experienced in the last 24 hours. For each endorsed item, a stressfulness impact rating is made ranging from 1 (“occurred but was not stressful”) to 7 (“caused me to panic”). Although several scales can be calculated from the measure, only the number of stressful items endorsed (frequency of stressful events, FREQ) was used in this analysis.

Mood Assessment (AM, PM diaries): The Profile of Mood States-Short Form (POMS-SF) was used to assess distinct mood states at each diary occasion.²¹ The POMS-SF consists of 37 Likert items that can be scored into six subscales and one total mood disturbance (TMD) scale. For the purposes of this analysis, only the TMD was used.

Alcohol Intake (PM diary): The number of alcoholic beverages consumed during the day was assessed using a 0 to 5 scale with 5 corresponding to “5 or more” alcoholic beverages.

Caffeine Intake (PM diary): The number of caffeinated beverages consumed during the day was assessed using a 0 to 5 scale with 5 corresponding to “5 or more” caffeinated beverages.

Definition of “Surprise” from Information Theory

When studying the influence of a potential headache trigger, a useful concept could be the amount of information contained in that trigger. Quantification of the amount of information contained in a variable, like a headache trigger, has been formalized in a set of methods known as information theory.²² Although quite mathematically complex, the core tenets of information theory arise from the astute observation that relates the amount of “information” in a variable to the amount of uncertainty in it. For variables that vary randomly, their variance determines the average amount of information they convey. For example, if two headache triggers such as red wine and cheese are observed over the course of one month, we could examine the number of days when either is encountered and estimate their variance (or distribution). In this example, if the probability of eating cheese = 0.35 (i.e., cheese is encountered in 35% of meals), and the probability of drinking wine = 0.00 (i.e., wine is never imbibed), then these triggers each have different amounts of information. Because red wine is never encountered, this trigger contains no information (i.e., there is no uncertainty), as the experience of it has no variance. On the contrary, cheese is sometimes encountered, so observing the day-to-day variation in cheese exposure does communicate something about the uncertainty in the environment, and the average amount of information expected from it can be estimated in a quantity known as information entropy.²³

In addition to the average amount of information contained in a headache trigger, the amount of information contained in a single exposure to a trigger can also be quantified. The information content of a single observation is referred to as the amount of surprise in observing it (i.e., “surprisal”). Although not necessarily intuitive, surprising events contain much more information content than expected ones. For example, if a trigger is only extremely rarely encountered (e.g., a probability of encounter = 0.01), actually encountering that trigger is extremely surprising, and there is substantial information content in the encounter. The quantification of surprisal for a binary (present or absent) trigger is:

$$\text{Surprisal} = -\log_2(p)$$

Where the log of base two is used with the probability of encountering the event, p , and the result is expressed in bits of information (taking the log using a different base results in different units).²⁴ For example, if a fair coin is flipped “heads”, the amount of surprisal is 1 bit ($-\log_2[.5]$). If two consecutive coins are flipped “heads”, the amount of surprisal would be 2 bits ($-\log_2[.5*.5]$), and this increases additively with each coin. This formulation allows surprising trigger information to be communicated on a scale akin to 1 to ∞ , with many random human events ranging between 1 (50:50 chance, like a coin flip) and 10 (1:1024 chance, like rolling “snake eyes” twice in a row). The definition of surprisal can be extended to any probability distribution and can certainly be applied to the surprisal experienced by an individual when observing a particular instance of a headache trigger.

For example, the left panel of Figure 1 displays two randomly generated normally distributed variables with a mean of 0 and a SD of either 1 (pink) or 2 (blue). These distributions could represent anything from the amount of daily sunlight exposure to mg of caffeine intake. The amount of surprisal contained in values that deviate from the mean

value can be observed in the right panel. As can be seen, the amount of surprisal in both distributions increases in either direction from the mean. The direction of the deviations does not impact surprisal, values can be either greater or less than expected, and the degree of surprise is related to the amount of variance in the distribution. For variables with less day-to-day variability (pink), there is less surprisal near the mean (as these values are observed more often) and greater surprisal for values farther from the mean.

In this analysis, the degree of surprisal contained in particular values of caffeine, alcohol, stress, and mood was calculated using methods summarized by Itti and Baldi²⁵ and McDaniel, Jordan, and Fleeman.²⁶ In brief, the empirical probability of observing a particular value of the trigger on any given day was estimated using the probability distributions for that variable for that person. The surprisal of this value, estimated in bits, was then plotted in relation to the probability of headache attacks, and this estimate then served as one predictor in the statistical models as described below.

Primary Outcome

The primary outcome for this analysis is identical to the one used in the primary analysis of the HAPRED study: the presence or absence of any headache attack (head pain > 0 on a NRS) occurring over the 24 hour period recorded following a PM diary entry. In essence, this relates the chances of experiencing a headache tomorrow to the levels of surprisal today.

Sample Size Considerations

This is a secondary analysis of previously collected data, and all available data were used for analysis. The original data collection was conducted for the purpose of evaluating the influence of stress on future headache attacks.¹⁹ No a priori statistical power calculation was conducted to guide this analysis.

Statistical Analyses

The statistical models used in this analysis are very similar to those employed in the primary analysis.¹⁹ All analyses utilized a generalized linear mixed model for binomial outcomes (headache present or absent) with a logit link function and random intercepts by subject to reflect the fact that individuals have varying base rates of headache activity. All models regressed the primary outcome, future headache attacks (Headache_{t+1}), on current and lagged versions of the predictors. Two sets of analyses were conducted separately for each predictor (caffeine, alcohol, stress, and mood). The first set utilized the current levels of the predictor, a lagged term of that predictor, and the current levels of headache activity (Headache_t). All main effects and interactions of these predictors (i.e., full factorial) were forced into the model to predict future headache. In the second set of models, the current surprisal values of the predictor, current headache activity, and current headache x surprisal were forced into the model to predict future headache activity. In a final model, the surprisals from each trigger were summed and used as a total predictor along with current headache activity to predict future headache.

The data are presented using frequency counts (%) for categorical data, means (SD) for normally distributed continuous data, and medians [25th, 75th] for ordinal or skewed data.

Additionally, Pearson correlations were used to express the linear relationship between variables and were applied to daily diary data assuming that the values were independent of one another (i.e., each individual contributed multiple days of surprisal data). Models are evaluated using model fit statistics (AIC) and discrimination values (AUC) and calibration plots. Embedded missing data were ignored in the analyses. All analyses were conducted using R version 3.4 statistical software (Vienna, Austria). All hypothesis testing was two-tailed with statistical significance interpreted at $P < .05$.

Results

The full results from the original primary analysis of these data, including descriptive characteristics of the participants, are reported elsewhere.¹⁹ Briefly, participants experienced a headache attack on 1613 of 4195 days (38.5%). Missing values of 396/4195 (9.4%) were embedded for caffeine, 422/4195 (10.1%) for alcohol, 431/4195 (9.3%) for frequency of stressful events, and 364/4195 (8.7%) for total mood disturbance. The participants were primarily female, 86/95 (90.5%), Caucasian, 83/95 (87.4%), and middle-aged, mean: 40 (SD = 13). Participants primarily had migraine without aura, 79/95 (83.2%) or migraine with aura, 10/95 (10.5%), but several had mixed migraine and tension headache, 6/95 (6.3%). Participants reported experiencing mild-moderate disability from their attacks on the MIDAS questionnaire, median [25th, 75th]: 10 [3.25, 18].¹⁹

Caffeine

Figure 2 displays the analyses related to daily caffeine intake. Across all individuals, the daily change in caffeine intake was normally distributed (histogram, Figure 2A). As the magnitude of changes in either direction increased (i.e., more or less caffeine than typical for that individual), the proportion of days with future headaches increased (scatterplot, Figure 2A). This relationship appeared to form a “U” distribution, consistent with surprisal as depicted in Figure 1. A factorial model using a lagged term for caffeine (to account for change) and current headache activity fit the data better than an intercepts-only model both in terms of AIC (4148.4 versus 4131.9) and likelihood ratio ($p < 0.0001$). This model exhibited statistically significant parameters for caffeine, caffeine (lag), current headache, and the interaction of caffeine x caffeine (lag).

Simulated predictions from this model, stratified by existing headache activity, are displayed in Figure 2B, and the modest calibration of this model to the data is displayed in Figure 2D. For days without an existing headache attack, predictions near the expected value of caffeine (i.e., 0 change from the day before) are quite variable but somewhat lower than the base rate of headache in the sample (38%). As greater changes in daily levels are observed, increased headache activity is predicted, again forming a “U” or “V” distribution. Different predictions are observed when an existing headache is present. However, when the level of caffeine is expressed as a surprisal (in bits of information), the relationships are linear for both strata (Figure 2C) with statistically significant effects ($p < 0.001$) for current headache, surprisal, and headache x surprisal.

Alcohol

Figure 3 displays the analyses related to daily alcohol intake. Across all individuals, the change in daily alcohol intake was leptokurtic, or exhibiting a tall peak around zero change (histogram, Figure 3A). As the magnitude of changes decreased, the proportion of days with future headaches increased (scatterplot, Figure 3A). This relationship appeared to form the left half of a “U” distribution, partly consistent with surprisal as depicted in Figure 1. A factorial model using a lagged term for alcohol (to account for change) and current headache activity fit the data better than an intercepts-only model both in terms of AIC (4096.4 versus 4076.3) and likelihood ratio ($p < 0.0001$). This model exhibited statistically significant parameters for alcohol, current headache, and the interaction of alcohol x current headache.

Simulated predictions from this model, stratified by existing headache activity, are displayed in Figure 3B, and the modest calibration of this model to the data is displayed in Figure 3D. For days without an existing headache attack, predictions near the expected value of alcohol (i.e., 0 change from the day before) are extremely variable. As greater changes in daily levels are observed, increased headache activity is predicted, again forming a “U” or “V” distribution for days with a headache attack. Different predictions are observed when an existing headache is present. However, when the level of alcohol is expressed as a surprisal (in bits of information), the relationships are linear for both strata (Figure 3C) with statistically significant effects ($p < 0.001$) for current headache, surprisal, and headache x surprisal.

Stress (Frequency of Stressful Events)

Figure 4 displays the analyses related to daily stress. Across all individuals, the change in daily stress was normally distributed but slightly leptokurtotic (histogram, Figure 4A). As the magnitude of changes in either direction increased (i.e., more or less stress than typical for that individual), the proportion of days with future headaches increased (scatterplot, Figure 4A). This relationship appeared to form a “U” distribution, consistent with surprisal as depicted in Figure 1. A factorial model using a lagged term for stress (to account for change) and current headache activity fit the data better than an intercepts-only model both in terms of AIC (4196.7 versus 4175.2) and likelihood ratio ($p < 0.0001$). This model exhibited statistically significant parameters for stress, stress (lag), current headache, the interaction of current headache x stress (lag), and current headache x stress x stress (lag).

Simulated predictions from this model, stratified by existing headache activity, are displayed in Figure 4B, and the acceptable calibration of this model to the data is displayed in Figure 4D. For days without an existing headache attack, predictions near the expected value of stress (i.e., 0 change from the day before) are quite variable but somewhat lower than the base rate of headache in the sample (38%). As greater changes in daily levels are observed, increased headache activity is predicted, again forming a “U” or “V” distribution. Different predictions are observed when an existing headache is present (i.e., an interaction between headache and stress). However, when the level of stress is expressed as a surprisal (in bits of information), the relationships are linear for both strata (Figure 4C) with statistically significant effects ($p < 0.001$) for current headache, surprisal, and headache x surprisal.

Mood (Total Mood Disturbance)

Figure 5 displays the analyses related to mood. Across all individuals, the change in daily mood was normally distributed (histogram, Figure 5A). As the magnitude of changes in either direction increased (i.e., more or less mood disturbance than typical for that individual), the proportion of days with future headaches increased (scatterplot, Figure 5A). This relationship appeared to form a “U” distribution, consistent with surprisal as depicted in Figure 1. A factorial model using a lagged term for mood (to account for change) and current headache activity fit the data better than an intercepts-only model both in terms of AIC (3966.5 versus 3939.3) and likelihood ratio ($p < 0.0001$). This model exhibited statistically significant parameters for mood and the interaction of mood x current headache.

Simulated predictions from this model, stratified by existing headache activity, are displayed in Figure 5B, and the modest calibration of this model to the data is displayed in Figure 5D. For days without an existing headache attack, greater mood change (i.e., worse mood) is highly associated with increased predictions for an attack. For days with an attack, as greater changes in daily levels are observed, increased headache activity is predicted, again forming a “U” or “V” distribution. However, when the level of mood is expressed as a surprisal (in bits of information), the relationships are linear for both strata (Figure 5C) with statistically significant effects ($p < 0.001$) for current headache, surprisal, and headache x surprisal.

Total Surprisal and Future Headache

Each of the surprisals from the four triggers alone exhibited modest associations with future headache attacks. The surprisals were also only modestly correlated with each other, $r = 0.17$ to $r = 0.42$ (see Supplemental Table 1). Stratifying by days with no current headache attack, each 1 SD change in bits of surprisal was associated with an increased risk of a future headache attack, with a low of OR 1.11 (95%CI: 1.0 to 1.24), $p = 0.050$ for alcohol (not statistically significant), to a high of OR 1.30 (95%CI: 1.14 to 1.46), $p < 0.0001$ for stress. However, a major advantage of using information theoretic measures for indexing trigger risk is that individual surprisals can be summed to estimate a total surprisal.²⁴ Therefore, surprisals from each of the four headache triggers can be summed to create a trigger surprise score that can be used to estimate daily headache risk from total trigger surprise encountered from that day. This trigger surprise score had a mean of 26.5 (SD 5.7) bits (see Supplemental Table 2, Supplemental Figure 1) and was modestly correlated with each individual's headache frequency, $r = 0.38$. Stratifying by days with no current headache attack, each 1 SD change in bits of total surprisal was associated with an increased risk of a future headache attack, OR 1.35 (95%CI: 1.17 to 1.49), $p < 0.0001$. This surprisal score exhibited superior performance compared to any individual trigger alone ($p < 0.0001$), including stress, with modest discrimination between headache and non-headache days from all of the individual triggers (ignoring base-rate, AUC: 0.61; AUC: 0.71 with base-rate).

Discussion

In this secondary analysis, small to modest associations between four common headache triggers (caffeine, alcohol, stress, and mood) and future headache activity were observed. However, these associations were often complex and entailed statistical interactions between

previous values of the triggers (i.e., yesterday's level), current values (i.e., today's level), and current headache status (i.e., the effects were different for the presence vs. absence of a current headache). When closely evaluating the pattern of these associations, the influences were well characterized by the degree of surprise posed by encountering the trigger. For example, the number of caffeinated beverages encountered today was not predictive of later headache, but encountering *surprising* levels of caffeinated beverages today was predictive. This implies that drinking one cup of coffee each day would not pose a risk for headache, but drinking four cups in a single day when only one cup is typically encountered (i.e., expected) would be predictive of increased headache risk. Thus, in this trigger surprise model, the actual nature of the event is not necessarily relevant, although it is intuitive that certain events would have greater relevance, but any surprises in encountering the event are relevant.

This trigger surprise model is based on a realization that all headache trigger candidates can be similarly expressed using the amount of surprising information they convey. This allows the presence or absence of consumptions of certain foods, "dosages" of certain chemicals (e.g., mg of caffeine) or events, weather patterns, mood changes, etc. all to be scaled using the same dimension, bits of information. This approach departs from a simpler model where mere exposure to a trigger (e.g., caffeine causes headache) is posited to underlie any association. In contrast, the trigger surprise model describes the amount of information being conveyed in any given exposure to the trigger. This model is based on elements of information theory: that unusual events and experiences carry more information about the environment than common events and experiences and that headache attacks are associated with reactions to uncommon or unexpected biopsychosocial exposures.

Why Would Surprise be Associated with Headache Attacks?

The current data demonstrated reliable relationships between individual elements of surprisal and the presence of a new-onset headache attack. Although this measurement simply describes the amount of information contained in the individual (or sum of individual) triggers, several underlying mechanisms are plausible. First, convergent findings from numerous domains support the important nature of surprising information for the nervous system. Surprising information has been found to be crucial for attention, learning, and sensation.^{25,26} In learning contexts, surprising information plays a key role and enhances association strength.²⁷ In biological terms, environmental surprise (e.g., hotter than usual, less food, infection) is often analogous with threat, and the primary purpose of the brain has been posited as one of reducing uncertainty or surprise to better ensure homeostasis and reduce future threat.^{28,29} Thus, trigger surprise might be measuring biopsychosocial signals to the individual about increased uncertainty or even threat. In these terms, any deviations from expected experiences (e.g., eating and drinking patterns, environmental conditions, etc.) are indicative of inputs to the nervous system that may require attention or action and that increase a reflexive "allostatic load"³⁰ and lead to an attack.

A second plausible mechanism is that an individual is avoiding certain trigger candidates because they are actually causing headache attacks. Because these triggers are mostly

avoided, actually experiencing them would be surprising. In this case, surprisal confirms what the patient suspected to be true. Furthermore, this rationale is concordant with the fear-avoidance hypothesis proposed by Martin⁹ in which avoidance of a trigger serves to maintain its potency to induce an attack (i.e., avoidance creates later surprise). However, this explanation makes sense for avoidable behaviors like drinking red wine, but it does not seem to apply to omnipresent influences such as the weather, mood, sleep, and stress. Nevertheless, quantifying these triggers using surprisal allows a direct test of the hypothesis that specific triggers are not what causes headache attacks, but surprising amounts of triggers are associated with attacks.

Surprise is Not Just Change in Triggers

While, historically, most literature has focused on the association of trigger presence with a migraine attack,^{31–34} several studies have highlighted the predictive utility of calculating *changes* in trigger experiences from yesterday to today.^{35–40} For example, Lipton et al.³⁷ examined the impact of change in stress by analyzing 771 sets of observations and presenting the change scores of stress from one night to the next in association with onset of migraine attack over the next 24 hours. This analysis supports the hypothesis of the “let-down headache,” which postulates that there is an increased probability of migraine onset following a decline in perceived stress. However, to our knowledge, the trigger surprise model is the first to establish that absolute changes in triggers are not as important as how unexpected these changes are in relation to their typical levels. Thus, as we have defined it, surprise is not merely change in a trigger but individual-specific contextual variation.

Limitations

This analysis has several important limitations that must be considered for interpretation. First, the analyses reported here are secondary, and the original data collection effort was not designed to examine this hypothesis. For example, the original study measurements were not designed to examine precise differences in daily caffeine intake and only assessed imprecise measurements of alcohol and caffeine (e.g., number of beverages). Second, the analyses presented here are entirely data driven, with the same data set used to describe a model form (surprise) and estimates of that model form (associations with headache). Such analyses tend to present overly optimistic associations between the predictors and outcomes given that the same data are used to specify and estimate the models. For this reason, we did not employ internal validation techniques given that the model forms were also data dependent, and it would have been difficult to account for the “phantom degrees of freedom” (i.e., the many model choices not accounted for in the estimation). Thus, no attempts were made to consider this issue, and replication of these associations is paramount.

Third, although surprisal could be calculated in many ways, we chose to estimate expectation based on empirical probability distributions that defined expectation as the probability of encountering a trigger based only on the frequency counts observed for that individual during the observation period. This method worked well for this retrospective analysis (i.e., all of the data were already collected) but would not be useful for prospective prediction studies, such as forecasting, given that the probabilities used to define surprise would not yet be fully observed for this individual. Stated simply, the degree of surprise of a

trigger will need to be defined using a different set of methods than the ones used here. To operationalize this model, a lengthy pre-observation period, such as the 49 days of data used here, would be needed to estimate the probability distributions for each trigger. Alternatively, a Bayesian form of surprise could be used that combined an a priori expectation of these probabilities with the actual experiences from an individual.²⁵

Fourth, we did not evaluate the timing of surprisals (e.g., lags of surprise), nor did we consider alternative definitions of surprise that are based on temporal considerations. It is plausible that more recent events should receive higher weight in developing expectations than events that are more temporally distant. For example, if an individual drinks a greater amount of caffeine in the past week compared to their historical levels, also drinking more tomorrow could be defined as less surprising because of recent exposures. There is ample evidence that lagged relationships are important for the development of headache risk (see: Feuerstein⁴¹). Finally, as with many trigger exposures, it is difficult to distinguish between surprising experiences as actual triggering events versus surprising events that simply mark prodromal symptoms.

Future Directions

The use of surprise to predict future headache activity must be externally validated with data not used to specify these models. The superiority of conceptualizing triggers using surprise as opposed to modeling the mere change in the predictors should be evaluated using predictive utility metrics. The participants who provided data for this study were demographically similar to typical individuals with migraine but were extremely motivated and willing to provide extensive amounts of daily data. Thus, the extent to which these results generalize to the overall population of individuals with migraine should be evaluated. Further, observing more trigger candidates using more precise measurements should add reliability to each surprisal and allow consideration of a more diverse range of biopsychosocial events. In the present analysis, each of the trigger candidates was given equal weight in creating a total risk score, but it is not clear if this choice was justified. For example, the independent association between stress and headache was substantially greater than for alcohol and headache. Alcohol appeared to only exhibit half of the “U” shaped pattern that is predicted by this model, so perhaps this trigger should be given a lower weight compared to stress. Future work should consider the relative importance of each domain, correlations among any domains, and the use of a weighting system to craft a reliable estimate of total risk. In this analysis, all headache attacks in individuals with migraine were included, but future works should consider if the model applies differentially to migraine attacks. Finally, the calibration of the trigger surprise model to the actual data was only modest. Individual triggers appeared to fit the data well for risks below a 50% chance of headache, but as risk predictions exceeded this value, the models systematically under-represented true risk. This could be due to a lack of relevant triggers in the model (i.e., unmeasured predictors), or it could be due to something fundamentally biased about this model form for predicting risk. Future research should closely monitor the calibration of predicted versus actual risk.

Conclusions

Using information theoretic methods, a trigger surprise model appears to hold great promise for conceptualizing the headache risks elicited from a diverse set of triggers. By utilizing the probability distributions of the triggers (i.e., the chances they are encountered), a model can be used to uniformly describe the risks from any conceivable trigger. Rare values of headache triggers, or surprising values, were found to have consistent associations with headache activity across a variety of triggers (caffeine, alcohol, stress, and mood). This finding, if validated in external data using other triggers, has potential to be used to conceptualize the influence of a wide range of headache triggers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support: Research reported in this publication was supported by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under award number R01NS065257.

References

1. Pietrobon D, Moskowitz MA. Pathophysiology of migraine. *Annu Rev Physiol.* 2013;75:365–391. [PubMed: 23190076]
2. Martin VT, Behbehani MM. Toward a rational understanding of migraine trigger factors. *Med Clin North Am.* 2001;85:911–941. [PubMed: 11480265]
3. Houle TT, Turner DP. Natural experimentation is a challenging method for identifying headache triggers. *Headache.* 2013;53:636–643. [PubMed: 23534852]
4. Turner DP, Houle TT. Influences on headache trigger beliefs and perceptions. *Cephalalgia.* 2018;39:1545–1553.
5. Walters Pellegrino AB, Davis-Martin RE, Houle TT, Turner DP, Smitherman TA. Perceived triggers of primary headache disorders: A meta-analysis. *Cephalalgia.* 2018;38:1188–1198. [PubMed: 28825314]
6. Martin PR. Headache triggers: To avoid or not to avoid, that is the question. *Psychol Health.* 2000;15:801–809.
7. Martin PR. How do trigger factors acquire the capacity to precipitate headaches?. *Behav Res Ther.* 2001;39:545–554. [PubMed: 11341251]
8. Martin PR. Behavioral management of migraine headache triggers: Learning to cope with triggers. *Curr Pain Headache Rep.* 2010;14:221–227. [PubMed: 20425190]
9. Martin PR. Managing headache triggers: Think ‘coping’ not ‘avoidance’. *Cephalalgia.* 2010;30:634–637. [PubMed: 19673895]
10. Martin PR, MacLeod C. Behavioral management of headache triggers: Avoidance of triggers is an inadequate strategy. *Clin Psychol Rev.* 2009;29:483–495. [PubMed: 19556046]
11. Martin PR, Reece J, Forsyth M. Noise as a trigger for headaches: Relationship between exposure and sensitivity. *Headache.* 2006;46:962–972. [PubMed: 16732842]
12. Martin PR, Lae L, Reece J. Stress as a trigger for headaches: Relationship between exposure and sensitivity. *Anxiety Stress Coping.* 2007;20:393–407. [PubMed: 17999239]
13. Martin PR, Reece J, Callan M, MacLeod C, Kaur A, Gregg K, Goadsby PJ. Behavioral management of the triggers of recurrent headache: A randomized controlled trial. *Behav Res Ther.* 2014;61:1–11. [PubMed: 25108482]
14. Borkum JM. Migraine triggers and oxidative stress: A narrative review and synthesis. *Headache.* 2016;56:12–35. [PubMed: 26639834]

15. Borkum JM. The migraine attack as a homeostatic, neuroprotective response to brain oxidative stress: Preliminary evidence for a theory. *Headache*. 2018;58:118–1135. [PubMed: 29034461]
16. Blau JN. Migraine: Theories of pathogenesis. *Lancet*. 1992;339:1202–1207. [PubMed: 1349944]
17. Benemei S, Fusi C, Trevisan G, Geppetti P. The TRPA1 channel in migraine mechanism and treatment. *Br J Pharmacol*. 2014;171:2552–2567. [PubMed: 24206166]
18. Burstein R, Jakubowski M. Unitary hypothesis for multiple triggers of the pain and strain of migraine. *J Comp Neurol*. 2005;493:9–14. [PubMed: 16258903]
19. Houle TT, Turner DP, Golding AN, Porter JA, Martin VT, Penzien DB, Tegeler CH. Forecasting individual headache attacks using perceived stress: Development of a multivariable prediction model for persons with episodic migraine. *Headache*. 2017;57:1041–1050. [PubMed: 28699328]
20. Brantley PJ, Waggoner CD, Jones GN, Rappaport NB. A daily stress inventory: Development, reliability, and validity. *J Behav Med*. 1987;10:61–73. [PubMed: 3586002]
21. Shacham S A shortened version of the Profile of Mood States. *J Pers Assess*. 1983;47:305–306. [PubMed: 6886962]
22. Shannon CE. A mathematical theory of communication. *Bell System Technical Journal*. 1948;27:379–423.
23. Rényi A On measures of entropy and information Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics. Berkeley: University of California Press, 1961.
24. Ince RA. Measuring multivariate redundant information with pointwise common change in surprisal. *Entropy*. 2017;19:318. doi:10.3390/e19070318.
25. Itti L, Baldi P. Bayesian surprise attracts human attention. *Vision Res*. 2009;49:1295–1306. [PubMed: 18834898]
26. McDaniel RR Jr, Jordan ME, Fleeman BF. Surprise, surprise, surprise! A complexity science view of the unexpected. *Health Care Manage Rev*. 2003;28:266–278. [PubMed: 12940348]
27. Barto A, Mirolli M, Baldassarre G. Novelty or surprise?. *Front Psych*. 2013;4:907. doi: 10.3389/fpsyg.2013.00907.
28. Friston K The free-energy principle: a rough guide to the brain?. *Trends Cogn Sci*. 2009;13(7): 293–301. [PubMed: 19559644]
29. Friston K Is the free-energy principle neurocentric?. *Nat Rev Neurosci*. 2010;11(8):605. doi: 10.1038/nrn2787-c2.
30. McEwen BS. Stress, adaptation, and disease: allostasis and allostatic load. *Ann N Y Acad Sci*. 1998;840(1):33–44. [PubMed: 9629234]
31. Prince PB, Rapoport AM, Sheftell FD, Tepper SJ, Bigal ME. The effect of weather on headache. *Headache*. 2004;44:596–602. [PubMed: 15186304]
32. Wöber C, Brannath W, Schmidt K, Kapitan M, Rudel E, Wessely P, Wöber-Bingöl Ç, PAMINA Study Group. Prospective analysis of factors related to migraine attacks: The PAMINA study. *Cephalalgia*. 2007;27:304–314. [PubMed: 17376107]
33. Schoonman GG, Evers DJ, Ballieux BE, De Geus EJ, De Kloet ER, Terwindt GM, van Dijk JG, Ferrari MD. Is stress a trigger factor for migraine?. *Psychoneuroendocrinology*. 2007;32:532–538. [PubMed: 17459597]
34. Zebenholzer K, Frantal S, Pablik E, Lieba-Samal D, Salhofer-Polanyi S, Wöber-Bingöl Ç, Wöber C. Reliability of assessing lifestyle and trigger factors in patients with migraine—findings from the PAMINA study. *Eur J Neurol*. 2016;23:120–126. [PubMed: 26228627]
35. Spierings EL, Sorbi M, Haimowitz BR, Tellegen B. Changes in daily hassles, mood, and sleep in the 2 days before a migraine headache. *Clin J Pain*. 1996;12:38–42. [PubMed: 8722733]
36. Sorbi MJ, Maassen GH, Spierings EL A time series analysis of daily hassles and mood changes in the 3 days before the migraine attack. *Behav Med*. 1996;22:103–113. [PubMed: 9116381]
37. Lipton RB, Buse DC, Hall CB, Tennen H, DeFreitas TA, Borkowski TM, Grosberg BM, Haut SR. Reduction in perceived stress as a migraine trigger: testing the “let-down headache” hypothesis. *Neurology*. 2014;82:1395–1401. [PubMed: 24670889]
38. Giffin NJ, Lipton RB, Silberstein SD, Olesen J, Goadsby PJ. The migraine postdrome: An electronic diary study. *Neurology*. 2016;87:309–313. [PubMed: 27335112]

39. Peris F, Donoghue S, Torres F, Mian A, Wöber C. Towards improved migraine management: Determining potential trigger factors in individual patients. *Cephalalgia*. 2017;37:452–463. [PubMed: 27179352]
40. Martin VT, Allen JR, Houle TT, Powers SW, Kabbouche MA, O'Brien HL, Kacperski J, LeCates SL, White S, Vaughan P, Mandell K. Ovarian hormones, age and pubertal development and their association with days of headache onset in girls with migraine: An observational cohort study. *Cephalalgia*. 2018;38:707–717. [PubMed: 28474986]
41. Feuerstein M, Bortolussi L, Houle M, Labbe E. Stress, temporal artery activity, and pain in migraine headache: a prospective analysis. *Headache*. 1983;23:296–304. [PubMed: 6643038]

Figure 1. Surprisal in 2 normal distributions.

Left panel: two normal distributions with a mean of 0 and a standard deviation of 1 (pink) or 2 (blue). The probability density (y-axis) is displayed as a function of z-score (x-axis). This plot illustrates the probability of observing a particular score for two distributions with different variances. Right panel: the surprisal in bits (y-axis) associated with observing a particular z-score (x-axis). This plot shows the differences in surprisal between the two distributions from the left panel with a standard deviation of 1 (pink) versus 2 (blue).

Figure 2. Associations between caffeine and future headache activity.

A. Change in number of caffeinated beverages from previous day (x-axis) and the mean proportion (error bars are SD) of future headache days (left y-axis). The probability distribution of observing these changes (right y-axis). B. Predicted probability (y-axis) of future headache from today's caffeine, yesterday's caffeine, and current headache status (full factorial model). C. The association between surprisal (x-axis) and future headache activity (y-axis) stratified by current headache status. D. Calibration between predicted probability of future headache (x-axis) and observed probability of headache (y-axis) from the full factorial model.

Figure 3. Associations between alcohol and future headache activity.

A. Change in number of alcoholic beverages from previous day (x-axis) and the mean proportion (error bars are SD) of future headache days (left y-axis). The probability distribution of observing these changes (right y-axis). B. Predicted probability (y-axis) of future headache from today's alcohol, yesterday's alcohol, and current headache status (full factorial model). C. The association between surprisal (x-axis) and future headache activity (y-axis) stratified by current headache status. D. Calibration between predicted probability of future headache (x-axis) and observed probability of headache (y-axis) from the full factorial model.

Figure 4. Associations between frequency of stressful events and future headache activity.

A. Change in frequency of stressful events from previous day (x-axis) and the mean proportion (error bars are SD) of future headache days (left y-axis). The probability distribution of observing these changes (right y-axis). B. Predicted probability (y-axis) of future headache from today's stress, yesterday's stress, and current headache status (full factorial model). C. The association between surprisal (x-axis) and future headache activity (y-axis) stratified by current headache status. D. Calibration between predicted probability of future headache (x-axis) and observed probability of headache (y-axis) from the full factorial model.

Figure 5. Associations between total mood disturbance and future headache activity.

A. Change in total mood disturbance from previous day (x-axis) and the mean proportion (error bars are SD) of future headache days (left y-axis). The probability distribution of observing these changes (right y-axis). B. Predicted probability (y-axis) of future headache from today's mood, yesterday's mood, and current headache status (full factorial model). C. The association between surprisal (x-axis) and future headache activity (y-axis) stratified by current headache status. D. Calibration between predicted probability of future headache (x-axis) and observed probability of headache (y-axis) from the full factorial model.