Modeling Migraine Triggers

Kevin Ma *
Department of Mechanical Engineering
University of California, Berkeley
Berkeley, CA

Christopher A. Strong
Department of EECS
University of California, Berkeley
Berkeley, CA

Jeffrey Wu
Department of EECS
University of California, Berkeley
Berkeley, CA

Abstract

Migraine is the second leading cause of disability worldwide and the first among young adults. Despite the impacts of migraine attacks, we still lack a fundamental understanding of their causes. Triggers of migraines vary from person to person and can be difficult to identify. However, a better understanding of a person's personal triggers may help them to navigate their condition. Unfortunately, the process of identifying migraine triggers can be ad-hoc and frustratingly uninformative. In this work we (i) introduce simple models of the dynamics of migraines and (ii) begin to explore how triggers can be identified under these models. We also present a preliminary investigation of closed-loop control with the goal of reducing pain from migraines. Although simplified and exploratory, these represent our initial attempts to bring analytical tools developed in system identification, control, and experimental design to the individualized treatment of migraines.

1 Introduction

In the United States, an estimated 23 million individuals have been diagnosed with migraines Cady [1999], Goadsby et al. [2002] and migraines rank as the second leading cause of disabilities worldwide. Avoiding known triggers that could lead to a migraine attack is one component of treating migraines. However, the challenge lies in the fact that these triggers are highly individualized, so what affects one person may not affect another [Peris et al., 2017].

This variability among patients makes migraine trigger identification a prime candidate for N=1 studies, which are research designs that focus on the individual rather than on a group Lillie et al. [2011], Guyatt et al. [1986]. In these studies, the patient is recognized to have unique genetic makeup and physiological characteristics that may lead to each patient responding differently to potential migraine triggers. Under this framework, a typical approach to identify these triggers would be based on trial and error. Patients would be advised to maintain a record of their migraine occurrences under normal conditions to establish a baseline. Then they would experiment by eliminating one potential trigger at a time to see if there is a reduction in migraine frequency. If the migraine frequency decreases after a trigger has been eliminated (and vice versa), the trigger may be identified as a cause.

While this method can be effective, it is also time-intensive and requires considerable self-discipline. As a result, it may not be a realistic option for many patients. Therefore, it is crucial to explore alternative methods for identifying migraine triggers. In our research paper, we delve into several

*Author orders are based on last name

Preprint. Under review.

methodologies for simulating and modeling the process of migraine trigger identification. Our work is exploratory in nature, examining how several simple system identification approaches could be applied in this context. We discuss potential avenues for future research that could build upon our findings, as well as speculate on the practical application of these algorithms in supporting doctors and patients in the identification and treatment of migraine triggers.

2 Related Works

Headaches are broadly categorized into two categories: primary headaches with no known underlying cause and secondary headaches that are often a warning sign of some other serious condition. Primary headaches are further divided into migraines, tension headaches, and cluster headaches [Gobel], each of which have unique diagnostic criteria. Broadly, tension headaches are more diffusely felt across the head while migraines are often felt more strongly on a particular side and accompanied by light and/or sound sensitivities. Cluster headaches appear for a period of time with a daily or near-daily frequency before receding. In this work we focus on migraines although we comment briefly on the modeling of cluster headaches as well in section 5.

Although understudied with respect to their prevalence, migraines have still received significant research attention [Tfelt-Hansen and Koehler, 2011, Goadsby et al., 2002, Robbins, 2021]. Despite this work, there remains a significant lack of understanding of the underlying physiology behind migraines and how this physiology relates to triggering events. In our case, we are particularly interested in environmental or behavioral factors that can trigger migraines. [Kesserwani] provide a detailed overview of a variety of triggers and their respective prevalence, including stress, hormones, missing meals, weather change, and sleep disturbances. Unfortunately, the study of triggers is complicated by the need to distinguish them from early symptoms of migraines [Lipton et al., 2014, Pavlovic et al., 2014,?]. For example, in some cases, neck pain might be an early symptom of a migraine rather than a trigger, or the craving to consume certain foods can be caused by the onset of a migraine rather than vice versa, complicating survey-based studies of migraine triggers. Triggers also vary widely between individuals as Peris et al. [2017] demonstrated through the application of $N = 1$ analysis to paper migraine diaries.

In general, migraines can be diagnosed as episodic, where they occur on less than half of someone's days, or as chronic, where they occur on more than half of their days. The causes of transitioning from experiencing episodic to chronic migraines are not well understood and are an active area of research [May and Schulte, 2016, Buse et al., 2019]. Roughly 3% of episodic migraine sufferers progress to experiencing chronic migraines each year. Better understanding these transitions may help us to prevent them as well. Additionally, these transitions are not permanent, providing hope that we can discover ways to reduce migraine frequency back to episodic levels for those who have crossed that threshold. Although not explored in detail here, we expect that the study of migraine chronification could be another fruitful area for control and system identification to contribute to the medical field.

There has been a variety of previous work combining dynamical system modeling with healthcare. To highlight a few papers, Lee et al. [2022] use MDPs to model the health state of patients and the effectiveness of interventions. Rieckmann et al. [2022] apply causal inference inspired ML to untangle complex interwoven causes of diseases. Lastly, [Barra et al., 2020] make use of data from headache diaries to estimate the probability of a migraine attack continuing from one day to the next. This approach differs from this work in that it doesn't explicitly consider migraine triggers. A more thorough literature review in this area may uncover other approaches that can directly inspire our approaches to trigger identification.

We also draw on the concept of system identification, which broadly is the process of estimating parameters that govern the functioning of a system. There a variety of classical surveys of the field [Ljung and GUNNARSSONt, Kerschen et al., 2006,?, Åström and Eykhoff, 1971] that can be referenced. For estimating transition models for MDPs and POMDPs we applied a simple counting approach mentioned in Kochenderfer [2015].

3 Methods

To begin exploring the intersection between system identification, control, and the study of migraines we (i) created two simple dynamic migraine models, (ii) attempted two approaches to system identification on the first of these models, and (iii) applied an integral controller to select triggers on the second model.

3.1 Migraine models

We propose two migraine models. The first is a markov decision process (MDP) with binary states corresponding to the presence of a migraine and binary actions corresponding to the presence of various triggers. The second is a discrete-time linear time invariant system model where the state is a continuous measure of discomfort and the control is a continuous measure of the amount of presence of each trigger.

Markov decision process. An MDP can be described by its state space, action space, transition function, and reward function. Our state is a single binary value with 1 corresponding to having a migraine at that point in time and 0 corresponding to not having a migraine at that moment in time. This will be denoted by y_k at time k. Our action is expressed as $n_{triggers}$ binary values, denoted by a binary vector u_k at time k, with 1 corresponding to the presence of that trigger and 0 corresponding to the lack of that trigger. We model the transition probabilities as

$$P(y_{k+1} \mid y_k, u_k) = \sigma(ay_k + b^\top u_k + c)$$

where $\sigma(\cdot)$ is the sigmoid function

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

and we let the reward function be 0 as for this model we will only investigate system identification on its transitions rather than designing a policy based on its rewards.

This model does not account for the fact that many triggers occur with varying levels of severity in practice. We also note that including only the most recent headache is a limitation. In comparison Barra et al. [2020] model a migraine's probability of continuing as dependent on how many days an individual has already been experiencing it. A final limitation is that we assume all triggers are able to be intervened on at will. In reality, some triggers will be impossible to change (the weather). Even those you can change will require differing levels of effort. We have left a more careful incorporation of the process of changing a trigger for future work.

Linear Time-Invariant (LTI) Model. When constructing the LTI model, we have drawn significant conceptual inspiration from the framework developed by Gradu and Recht [2023]. Within this framework, we proposed that an individual's pain score denoted as y_t is governed by a linear time-invariant system. The LTI assumption may be justified in some scenarios but it does mark a significant modeling assumption as previous work has suggested the importance of a time-varying threshold in whether someone experiences a migraine [May and Schulte, 2016]. The state at time $t + 1$ is given as

$$y_{t+1} = \left(\sum_{j=1}^{n} \sum_{k=0}^{t+1} h_j(k) u_{j,(t+1)-k} \right) + y_{t+1}^{nat}$$

where n is the number of migraine triggers the model is accounting for, $h_j(k)$ is the impulse response function that describes how trigger j at time $(t + 1) - k$ influences the well-being at time $t + 1$, and $u_{j,(t+1)-k}$ is the intensity of a specific trigger j at time $(t + 1) - k$. For instance, $u_{j=\text{caffeine},t-1}$ would be the amount of caffeine consumed one time step before time t, and $u_{j=\text{stress},t-1}$ would be the stress level one time step before time t. Finally, the natural progression of the patient's well-being not influenced by specific triggers is defined as y_{t+1}^{nat}.

To design the model with concrete values, we treated the action space $u_{j,t-k}$ as a continuous value from 0 to 1. Likewise, we treated the state space y_t as a continuous pain score from 0 to 7. We defined the impulse response $h_j(k)$ as an exponentially decaying function with a constant variable dependent on the effects of each migraine trigger on the patient as shown below:

$$h_j(k) = c_j e^{-\frac{t}{\tau_j}}$$

where c_j corresponds to the scale of impact of a unit of trigger j and k_j is a constant that dictates how fast the impulse response will decay. For the purpose of this project, we simplify the model by defining $c_j = 0.5$ and $\tau_j = 10$. In addition, for this paper, we defined the autonomous (in the absence of triggers) progression of the patient's well-being y_{t+1}^{nat} as a scalar value sampled from a normal distribution of $N(1, 1)$, which implies that the patient on average has a pain score of 1 with standard deviation of size 1. In practice, when using this, we recommend selecting these parameters based on the patient's history with migraine.

3.2 System identification

In this section we focus on the MDP model introduced in section 3.1. We take two approaches to system identification outlined below.

Simple transition counting. One of the simplest approaches to estimating the transition probabilities for an MDP consists of rolling out trajectories and counting how often you experience different transitions. If we let $N(y', y, u)$[2] be the number of times you've taken action u from state y and transitioned to state y' and $N(y, u)$ be the number of times you've taken action u in state y. Then our estimate of the transition probabilities is given by

$$P(y' \mid y, u) = N(y', y, u)/N(y, u)$$

We apply two strategies for choosing the triggers to apply while producing these rollouts. In the first strategy the controls are chosen uniformly at random. In the second, the action which has been chosen the least often from the current state is selected.

Least squares. Another approach is to fit a linear model mapping from the state y and action u to the next state y' using least-squares. Compared to simple transition counting, least-squares should be more sample efficient as it can interpolate between visited states and as a result does not require having visited every state-action pair many times to obtain an accurate estimate. However, we cannot guarantee the strength of the model's interpolations as the underlying migraine dynamics is likely highly non-linear. While producing the rollouts for the training dataset, we choose the controls uniformly at random.

3.3 Closed loop control

In this section, we focus on creating a closed loop controller for the LTI model described in section 3.1. We again draw inspiration from previous work done by Gradu and Recht [2023]. At all times k we set the desired pain tolerance to be $y_{min} = 0$ (no pain). To control the trigger intensity, we devised a similar integral controller as described in Gradu and Recht [2023] but simplified further. Here, we update $u_{j,t-k}$ based on the following integral control law:

$$uj, t = max[0, u_{j,t-1} - K(y_t - y_{min})]$$

The preceding integral control law is less general than the one described in Gradu and Recht [2023], but in section 4 we show that the intensity level for each migraine trigger converges to 0.

4 Results & Discussion

We implemented two migraine models, a markov decision process model with binary triggers and state and a linear time invariant model with continuous triggers and state. We then applied two system identification algorithms to the first model, and a closed loop control policy to the second. Although largely unsurprising, the results we obtain represent our first foray into exploring the simple models we have built.

[2]We maintain notation similar to that in linear systems theory rather than typical MDP/POMDP notation for consistency with our second model

Figure 1

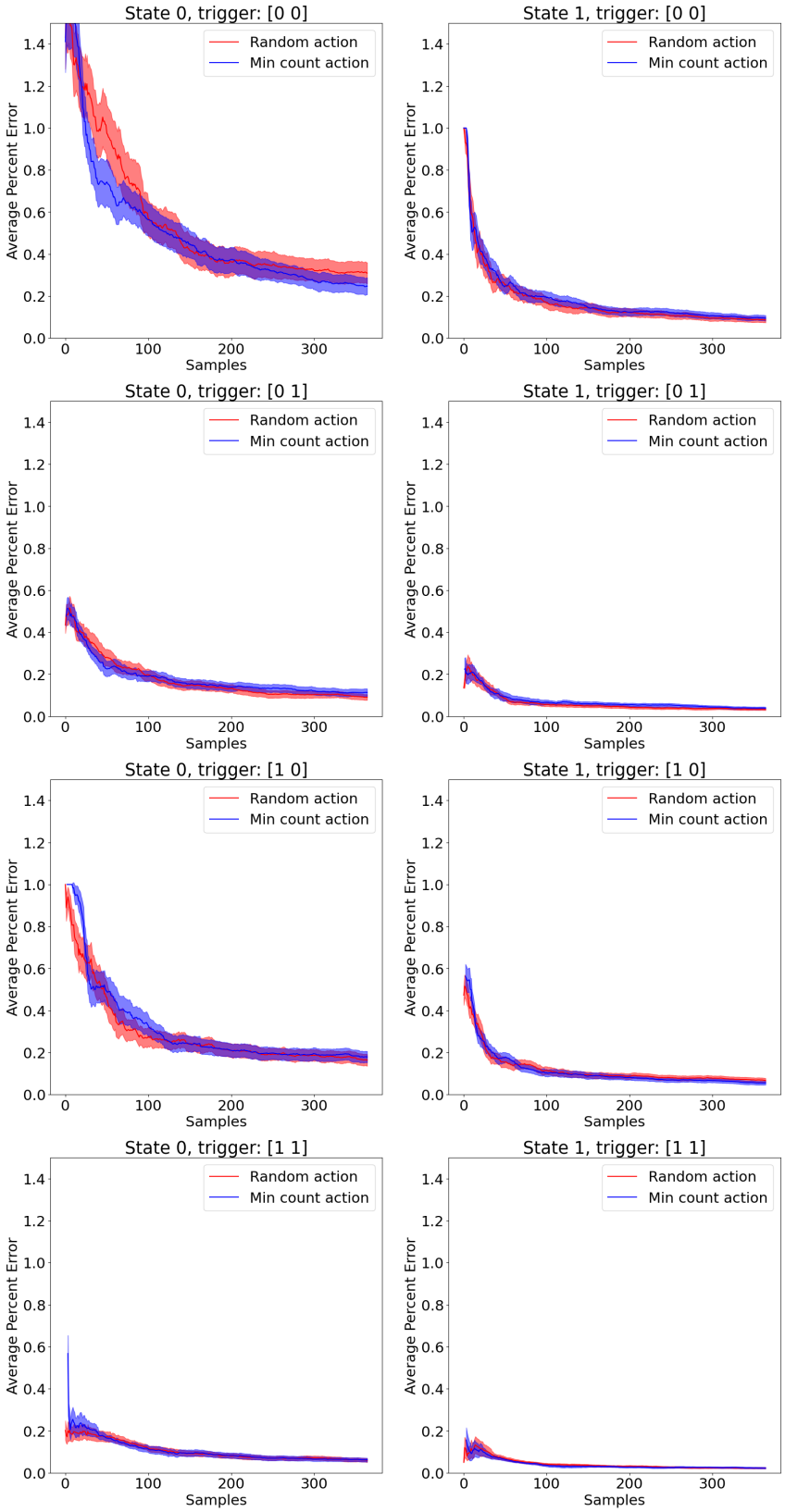
State 0, trigger: [0 0]
State 1, trigger: [0 0]
State 0, trigger: [0 1]
State 1, trigger: [0 1]
State 0, trigger: [1 0]
State 1, trigger: [1 0]
State 0, trigger: [1 1]
State 1, trigger: [1 1]
Random action
Min count action
Average Percent Error
Samples
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0
100
200
300

Figure 1: **Transition counting shows unacceptably slow convergence even with only two triggers**. In order to see how quickly a simple counting approach can estimate the transition function of our migraine model, we tested this approach on the MDP migraine model. The state had coefficient 1.0, the triggers had coefficients 1.0 and 2.0, and there was a shift c of -1 corresponding to a probability of roughly 27% of having a migraine when there are no triggers. We rolled out the model with two policies for choosing the triggers at each time step: (i) randomly choosing the action and (ii) choosing the action that had been observed the fewest times from the current state. We performed 100 rollouts under each strategy for 365 steps (one year) each. Each plot contains the average percent error across the 100 rollouts between the true probability and the estimated probability of having a migraine at the next time step given a particular state and action pair ($|(\text{true} - \text{predicted})/\text{true}$). The shaded region around the line gives two standard errors above and below the mean. The left column corresponds to a state of 0 (not having a migraine) while the right column corresponds to a state of 1 (having a migraine). The rows correspond to each combination of the two binary triggers. We observe each error converging towards 0 although at different rates. Convergence occurs on the order of several hundred steps. We don't observe a consistent difference between the two exploration strategies. Our main takeaway is that since this protocol would need to be applied with actual patients to be helpful this rate of convergence is unacceptably slow (despite including only two triggers), suggesting we should explore alternative methods for system identification.

4.1 System Identification

We applied two approaches to system identification to the MDP model of migraine dynamics as described in section 3.2. For the first, where we simply count transitions from each state-action pair, figure 1 shows the average convergence of the estimated transition model to the ground truth model. We note that even when limited to only two triggers convergence happens on the order of several hundred samples. In practice, we expect this is prohibitively long to require a migraineur to randomly vary their triggers. Even though this is a toy model it suggests that much more efficient approaches to trigger identification will need to be employed for it to be useful in practice. To demonstrate our second approach, section 4.1 shows the accuracy of a linear model fit with least squares on single-step prediction of migraines. We observe reasonable prediction accuracy despite the linear model's mismatch with the groundtruth sigmoid model, even with a limited number of samples. Although not a direct comparison to the metrics of evaluation for our first approach, this suggests that a functional predictive model may be produced in a much smaller number of steps.

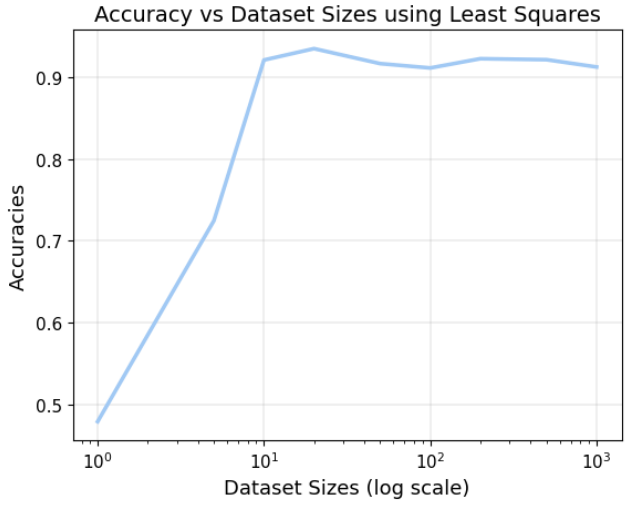

Figure 2: **Least-squares shows fast convergence with five triggers.** We use five triggers, with coefficients 0.0, 1.0, 2.0, 3.0, 4.0 with a shift c of 0. We rolled out the model with actions chosen randomly for a total of $10\,000$ transitions. We see that even with a small dataset of only 10 transitions, least-squares converges to a 90+$\%$ accuracy.

4.2 Integral Controller

In our implementation of the integral controller as described in section 3.1, we chose to include three triggering mechanisms $n = 3$. To initialize the computational model the initial migraine trigger intensities $u_{j,0}$ were randomly generated from a uniform distribution within the continuous interval of $[0, 1]$. Furthermore, the initial pain score y_0 was randomly sampled from a uniform distribution within the interval $[0, 7]$. This initialization protocol can be expressed by the following mathematical expressions:

$$y_0 \sim U(0, 7) \qquad u_{j,0} \sim U(0, 1)$$

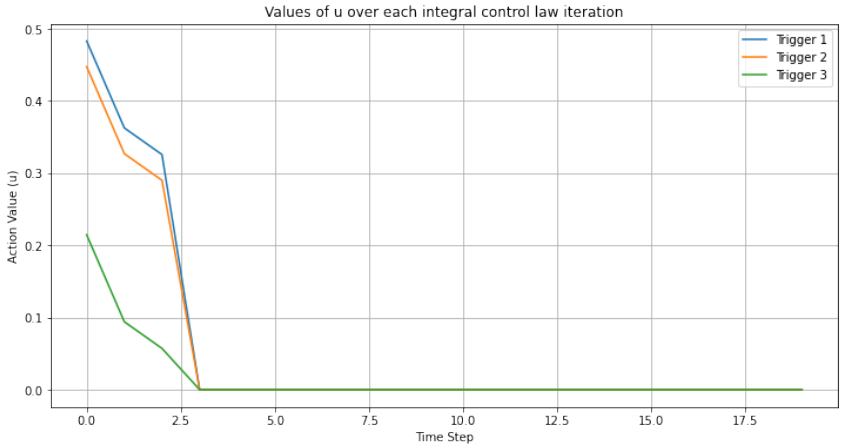

Figure 3: **Values of $u_{j,t-k}$ over each time iteration**. Results shown from the method done in section 3.3. We tested on three different triggers using the *linear time invariant model* as described in section 3.1. The initial action $u_{j,0}$ and state y_0 was set by uniformly sampling from a continuous space between $[0, 1]$ and $[0, 7]$, respectively.

Empirical outcomes from this study indicate that irrespective of the initial value from which the migraine intensity level was sampled, convergence to a value of 0 was observed for all migraine triggers under consideration. The outcomes of the study are consistent with expectations. The control law we outlined yields a policy that advocates for the complete avoidance of migraine triggers whenever possible. This policy is, however, impossible to implement because, realistically speaking, a patient can never avoid all triggers. The ideal objective, in reality, would be to have a policy based off a control law that would present a more nuanced dilemma between the effort to avoid triggers and the experience of migraines.

5 Future Works

Future research should focus on gaining a thorough understanding of the existing literature on migraines and incorporating solutions from various angles, even if they are not mathematically based and more conceptually grounded on clinical frameworks. By more deeply drawing inspiration from the medical literature, future works will be able to have better baseline models. On top of this, we also need to explore whether models from the literature are well-suited to be cast as optimal control problems, and if so, we need to carefully consider the specific objectives we aim to optimize.

Moreover, if there is an optimal controller that is effective in theory, it is, consequently, important to consider whether this can be applied in reality. As a result, we need to take into account the current standard practices of medical professionals to ensure the feasibility of adopting such a controller for both doctors and patients. Building on top of that, in the future, we would like to investigate how existing practices for managing migraines, such as migraine diaries, can be incorporated into mathematical frameworks. By collecting observational data of patients, we hope to design models that better reflect real migraine patterns. As a result, future work would need to determine the accessibility and appropriate form of data for migraines.

On top of those suggested future works, we also need to acknowledge the limitations of our system identification methods for both the LTI and markov decision process models. The assumptions made in these models are likely over-simplified because they do not account for interaction effects between migraine triggers and do not accurately capture the time-varying nature of an individual's internal state. Additionally, the equations proposed to model the existence of a migraine trigger may not fully represent the complexity of a patient's situation when it comes to their individual migraine triggers. Future research should carefully address these limitations and compare the LTI and markov models to determine which is a more useful representation of migraine trajectories.

In addition to the limitation of our system identification work, we also have limitations in our controller for migraine triggers. Our initial explorations did not produce informative results that could lead to better decision-making for a patient with respect to potential migraine triggers because the general policy from the framework was to always avoid encountering migraine triggers. However, this is not feasible, so future research should develop a more realistic and effective model to address those issues.

Beyond just migraines, it would be interesting to explore the modeling of other types of headaches, such as cluster headaches and to identify factors that influence the spacing of these clusters. This can be approached using a similar system identification framework as proposed in this study and previous research. Overall, by pushing for these "N of 1" studies, we can contribute to the development of individualized interventions and prevention mechanisms in future studies.

6 Conclusion

In summary, this study has investigated multiple approaches for the identification of migraine triggers within a variety of frameworks and assumptions. Furthermore, we have examined a strategy for controlling the intensity of these triggers. It is important to note that the findings presented in this paper are of a preliminary nature. Future research endeavors should aim to enhance the robustness of these results through a comprehensive review of current literature and a more thorough understanding of contemporary medical practice.

References

Mathias Barra, Fredrik A Dahl, Kjersti Grøtta Vetvik, and E Anne MacGregor. A markov chain method for counting and modelling migraine attacks. *Scientific Reports*, 10(1):3631, 2020.

Dawn C. Buse, Jacob D. Greisman, Khosrow Baigi, and Richard B. Lipton. Migraine Progression: A Systematic Review. *Headache: The Journal of Head and Face Pain*, 59(3):306–338, 2019. ISSN 1526-4610. doi: 10.1111/head.13459. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/head.13459. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/head.13459.

Roger K Cady. Diagnosis and treatment of migraine. *Clinical cornerstone*, 1(6):21–32, 1999.

Peter J Goadsby, Richard B Lipton, and Michel D Ferrari. Migraine—current understanding and treatment. *New England journal of medicine*, 346(4):257–270, 2002.

Hartmut Gobel. The International Classification of Headache Disorders. URL https://ichd-3.org/.

Paula Gradu and Benjamin Recht. Online Control for Adaptive Tapering of Medications, September 2023. URL http://arxiv.org/abs/2309.11629. arXiv:2309.11629 [math].

Gordon Guyatt, David Sackett, D Wayne Taylor, John Ghong, Robin Roberts, and Stewart Pugsley. Determining optimal therapy—randomized trials in individual patients. *New England Journal of Medicine*, 314(14):889–892, 1986.

Gaëtan Kerschen, Keith Worden, Alexander F. Vakakis, and Jean-Claude Golinval. Past, present and future of nonlinear system identification in structural dynamics. *Mechanical Systems and Signal Processing*, 20(3):505–592, April 2006. ISSN 0888-3270. doi: 10.1016/j.ymssp.2005.04.008. URL https://www.sciencedirect.com/science/article/pii/S0888327005000828.

Hassan Kesserwani. Migraine Triggers: An Overview of the Pharmacology, Biochemistry, Atmospherics, and Their Effects on Neural Networks. *Cureus*, 13(4):e14243. ISSN 2168-8184. doi: 10.7759/cureus.14243. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8088284/.

Mykel J. Kochenderfer. *Decision making under uncertainty: theory and application*. Lincoln Laboratory series. The MIT Press, Cambridge, Massachusetts, 2015. ISBN 978-0-262-02925-4.

Sujee Lee, Philip A Bain, Albert J Musa, Christine Baker, and Jingshan Li. A causal network-based markov decision process model for intervention planning. *IEEE Transactions on Automation Science and Engineering*, 2022.

Elizabeth O Lillie, Bradley Patay, Joel Diamant, Brian Issell, Eric J Topol, and Nicholas J Schork. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Personalized medicine*, 8(2):161–173, 2011.

Richard B. Lipton, Jelena M. Pavlovic, Sheryl R. Haut, Brian M. Grosberg, and Dawn C. Buse. Methodological Issues in Studying Trigger Factors and Premonitory Features of Migraine. *Headache: The Journal of Head and Face Pain*, 54(10):1661–1669, 2014. ISSN 1526-4610. doi: 10.1111/head.12464. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/head.12464. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/head.12464.

Lennart Ljung and SVANTE GUNNARSSONt. Adaptation and Tracking in System Identification A Surveyt.

Arne May and Laura H. Schulte. Chronic migraine: risk factors, mechanisms and treatment. *Nature Reviews Neurology*, 12(8):455–464, August 2016. ISSN 1759-4766. doi: 10.1038/nrneurol.2016.93. URL https://www.nature.com/articles/nrneurol.2016.93. Number: 8 Publisher: Nature Publishing Group.

Jelena M. Pavlovic, Dawn C. Buse, C. Mark Sollars, Sheryl Haut, and Richard B. Lipton. Trigger Factors and Premonitory Features of Migraine Attacks: Summary of Studies. *Headache: The Journal of Head and Face Pain*, 54(10):1670–1679, 2014. ISSN 1526-4610. doi: 10.1111/head.12468. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/head.12468. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/head.12468.

Francesc Peris, Stephen Donoghue, Ferran Torres, Alec Mian, and Christian Wöber. Towards improved migraine management: Determining potential trigger factors in individual patients. *Cephalalgia*, 37(5):452–463, 2017.

Andreas Rieckmann, Piotr Dworzynski, Leila Arras, Sebastian Lapuschkin, Wojciech Samek, Onyebuchi Aniweta Arah, Naja Hulvej Rod, and Claus Thorn Ekstrøm. Causes of outcome learning: a causal inference-inspired machine learning approach to disentangling common combinations of potential causes of a health outcome. *International journal of epidemiology*, 51(5):1622–1636, 2022.

Matthew S. Robbins. Diagnosis and Management of Headache: A Review. *JAMA*, 325(18):1874–1885, May 2021. ISSN 0098-7484. doi: 10.1001/jama.2021.1640. URL https://doi.org/10.1001/jama.2021.1640.

Peer C. Tfelt-Hansen and Peter J. Koehler. One Hundred Years of Migraine Research: Major Clinical and Scientific Observations From 1910 to 2010. *Headache: The Journal of Head and Face Pain*, 51(5):752–778, 2011. ISSN 1526-4610. doi: 10.1111/j.1526-4610.2011.01892.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1526-4610.2011.01892.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1526-4610.2011.01892.x.

K. J. Åström and P. Eykhoff. System identification—A survey. *Automatica*, 7(2):123–162, March 1971. ISSN 0005-1098. doi: 10.1016/0005-1098(71)90059-8. URL https://www.sciencedirect.com/science/article/pii/0005109871900598.