Bachelor's Thesis Computing Science

Radboud University Nijmegen

---

**Natural Language Processing - Migraine and Diet**

---

*Author:*
George Dan
s1046976

*First supervisor/assessor:*
dr. Ilona Wilmont

*Second assessor:*
dr. Faegheh Hasibi

January 15, 2025

# Contents

# Chapter 1

# Abstract

Dieting strategies have been shown to benefit migraine and other neurological conditions. The scientific consensus on diet and migraine shows promising benefits to dieting but has so far been considered inconclusive. In this paper we aim to test the current scientific knowledge on this topic against migraine sufferers lived experience. We provide a literature review on migraine and dieting, and we develop a natural language processing method using sentiment analysis strategies, to document consistencies and inconsistencies, between the scientific consensus and people's lived experience. We find a general positive sentiment regarding 4 frequently studied diets, which is predominantly consistent with the scientific literature, but we also spot some inconsistencies, such as positive sentiment surrounding caffeinated drinks and beer.

# Chapter 2

# Introduction

In this chapter we will begin by defining our problem and argue why solving this problem is relevant. We will explain how it can be mapped to a computing science problem and solved in the manner which we have chosen. We will then outline the intuition and strategy behind our solution and provide our hypothesis, research question, and contributions. We will integrate our work within the existing body of research by reviewing previous work and will finish by outlining the rest of this thesis.

## 2.1 Background

Migraine is a very common neurological affliction that can have serious negative impacts on a person's daily life. It manifests itself with episodes of unilateral headache (pain on one side of the head) as the main symptom. Other symptoms may include nausea, dizziness, and light and sound sensitivity. There are multiple forms of migraine, such as chronic or episodic (acute) migraine, with or without aura. Migraine is considered chronic if a headache occurs at least 15 days a month for more than 3 months, with migraine symptoms present at least 8 or more days a month. Both migraine with and without aura manifests unilaterally, but migraine with aura features fully reversible attacks lasting minutes, whereas migraine without aura presents itself with longer attacks lasting between 4 and 72 hours. [26].

### 2.1.1 Prevalence and impact of migraine

To highlight the importance of addressing this issue, we begin with some information on the consequences migraine has on individuals personally and professionally, as well as at a societal level. It is believed that between 12% and 28% of adults suffer from migraine at some point in their life. The financial impact of migraine is estimated at €27 billion per year in Europe and €17 billion per year in the United States. [21] This prevalence remains stable through the years. In the United States in 2018, migraine affected 15.9% of adults at a stable sex ratio (21% of women and 10.7% of men). [6]

A study by C.E. CLARKE published in 1996 looked at the impact of migraine in the workplace within a hospital. He found an average 2 years of work per year are lost for each employee due to migraine and an added more than 5 days work lost in terms of reduced effectiveness. The costs of these losses within the hospital over a year, rose to over £100,000. Having obtained such results through a questionnaire that required people to be open about their absenteeism, which is likely to render some unwilling to participate, further emphasizes the issue. The price of migraine in the workplace is not to be underestimated. [8]

Furthermore, migraine brings difficulties that span out into one's personal life. It accounted for 4.3 million office visits and 4 million ED (emergency department) visits in the US in 2016. It was also found that around 40% of adults with migraine in the United States were unemployed, and just as many are struggling financially. Roughly one in five has no health insurance and about a third have no higher studies than a high school education. [6] Migraine attacks also affect many common daily activities such as driving (55%), cooking or eating (42%), taking care of family/childcare (40%) and getting medicines at the pharmacy (40%). It is also a frequent occurrence that patients have to see multiple specialists to obtain a correct diagnosis of migraine and it often takes years for them to arrive to this point (more than 5 years in 40 percent of cases). [30]

### 2.1.2 History of Migraine & Diet

Ever since the beginning of the study of migraine, there have been experiments showcasing the impact of dietary choices on migraine sufferers. Earliest experiments date as far back as the 1920's - Schnabel's experiments with the ketogenic diet in 1928 already found promising results. The idea of the use of the ketogenic diet to mitigate migraines was born from the documented positive impact that it had in treating epilepsy, another condition of a neurological nature. Prescription of a high fat diet resulted in 9 out of 18 patients in the study experiencing some relief. [28] Barborka's 1930 experiment showed positive results of a similar nature with improvement in 25 out of 50 patients included in the study, as well as recurrence of migraine symptoms when the ketogenic diet was abandoned. It also cites multiple studies available at the time hypothesizing various metabolic causes of migraine such as protein sensitization, anaphylactic manifestation (allergy), inherited impaired metabolism with intolerance of nitrogenous foods, or dysfunction of the liver and duodenum. More promising than such theories however, seems to be the theory for alkalosis as a metabolic cause of migraine. Alkalosis is defined as an excess of alkali in the body and was observed prior to migraine attacks. [4] This also explains how a high fat diet such as keto may assist in migraine mitigation by correction a metabolic imbalance, given that high fat foods' are likely to increase acidity.

Since these early hypotheses, the study of migraine and diet has broadened to include many potential metabolic causes for migraine. Should the cause of migraine not be metabolic, dietary strategies may still be a useful tool to help mitigate symptoms. Commonly referred diets include keto, and comprehensive/elimination diets focused on removing certain foods or food categories suspected to cause migraine or aggravate migraine.

The promise of dietary strategies to mitigate migraine remains strong. To showcase the persistence of the issue of migraine to the present day, and the continued positive impact that dieting can have on migraineurs, we have also collected first-hand testimonies from current migraine patients :

*Patient 1: I am aware that it has not been scientifically proven that a low-carbohydrate diet helps to reduce migraines. For me, however, it is clear that it works. Sceptical as I was when starting out, I am now fully convinced to such an extent that I do not want to go back to my old ways of eating because I can enjoy everything around me. This is also a reason why I consider it very important that this will be subjected to scientific scrutiny.*

*Patient 2: I recently met fellow patients who had changed their nutrition resulting in fewer symptoms. As dietician, this appealed to me. For about half a year now, I have been eating a low-carbohydrate diet, around 30-50g of carbohydrates per day. [...] I hardly have any real migraine symptoms anymore. [...] To be able to give a large group of patients, who have been plagued for years by everything migraine bring about, nutritional advice [...] it appears to me to be highly desirable to possibly answer these questions scientifically.*

*Patient 3: I could hardly work anymore. Then I read about the ketogenic diet [...] and started it. [...] For a month, I have not had any migraine episodes. The episode that followed was a lot less severe. [...] My neurologist saw no reason to continue medication. [...] A year into it, I still eat a ketogenic diet. It is not easy, but it becomes doable. I have my life back. I work again for three days. I do household tasks. I play with our children again. I am happy again. [...] My neurologist has dismissed me as a patient. [...] I called the neurology department to ask if they could recommend a dietician who has experience with this diet. They said no, but* *maybe the migraine specialist nurse had ideas. [. . . ] The nurse was honest and said that nobody in the department believed in the diet, and that they did not want to make patients' lives more difficult with a diet. [. . . ] My neurologist has called me twice since. Both times I was positive about the diet, but he did not ask any question. Remarkable.*

*Patient 4: I started the ketogenic diet in April, that month I had 4 episodes. Since then I have had no episodes for months, and if one comes through it is less painful and I can function with it. [. . . ] I don't just have my life back, I have never had such a life before. [. . . ] Still, it is not only glory. [. . . ] I am looking for a good dietician to help me [. . . ] but I am not eligible for it. [. . . ] Given my success I cannot imagine that it will not also help other patients. That is why I am convinced more research is needed.*

## 2.2 Natural Language Processing

Recently, artificial intelligence has made its way into many fields in medicine, and the study of migraine has not escaped its grasp. Migraine diagnosis is facilitated by AI through image analysis [17], increased diagnostic accuracy, and classification . Management of migraine is also assisted through AI better identifying fitting treatment choices, or therapy response prediction. [29] AI can also aid in reaching a diagnosis or identifying appropriate treatment faster, thus reducing physician workload. However, such benefits come at the cost of necessity of technological literacy. [27]

### 2.2.1 Social Media Sentiment Analysis

A specific point of focus within the study of migraine with AI is the use of natural language processing (NLP) techniques. NLP is the process of using technology to interpret natural human language and extract insight from it.

Sentiment Analysis is a branch of NLP that focuses on the emotional tone of text. It sets out to quantify and measure sentiment value in written text. This strategy has proven successful in extracting meaningful insight from bodies of social media text.

### 2.2.2 Previous contributions

When looking at the current state of language processing strategies in the study of migraine, we see a budding yet promising field of research that has brought significant benefits. Back in 2018, language processing had already been employed in the analysis of social media discussions about dietary supplements, and successfully weeded out supplements associated with excessive effects. [25] Sentiment analysis has also been used to capture the expressive tone of tweets about migraine, showcasing a general negative tone and a tendency for expressions extreme sentiment and profanity. [9] Language Processing has additionally been used on patient reported narratives on headaches in order to classify cluster headache and migraine patients by detecting differences in lexical choices. [31] The overarching benefit of such studies is that they break out of the pattern established by clinical trials, findings therein not being limited to predetermined variables. Thus the accuracy of language processing mechanisms is of utmost importance, and it has been recently shown that supervised machine learning methods have achieved F1 scores of above 0.90 on tweets and Reddit posts. [18]

### 2.2.3 Knowledge Gap

We have seen that the study of dietary strategies to treat migraine is a rich field with a long history of experimenting and a consistent body of knowledge. We have also seen that in recent years, AI and machine learning strategies in various forms have been used to aid migraine research. Among these recent studies, sentiment analysis of migraine-related patient narratives has been proven both insightful and reliable. What we have not yet seen is the use of language processing extend into the study of dieting and migraine, and we aim to fill this gap in our research.

## 2.3 Research Question

We have noticed a discord between the great utility that dieting can have in mitigating migraine and the lack of easy access to this information by patients. Present research of dieting and migraine has been of great benefit, however, it is known that this research is by no means complete and sometimes inconsistent as different studies may reach different conclusions. In short, a consensus on dieting and migraine has not been achieved. [16] We hypothesize that there is additional insight on the topic to be found in social media discussion of migraineurs experiences with dieting. This information can highlight inconsistencies in the current scientific consensus on diet and migraine, as well as discover additional insight. We arrive at our research question:

**RQ: How does migraine patients' lived experience with diet compare to the scientific consensus on migraine and diet?**

## 2.4 Outline

In the basis of our research question, for the rest of our study we develop a sentiment analysis based NLP method, to measure migraineurs experience. We begin by performing a literature review on the connection between diet and migraine, extracting insight about 4 main diets commonly used to alleviate migraine. We create a corpus of text with migraine discussions from Reddit as well as a set of self-reported stories. We then find a way to access emotional value in the context of discussions surrounding diet and to measure the emotional tone of pieces of text. We then provide a lengthy analysis of our findings, examining consistencies and inconsistencies between our findings from our literature review and out language processing method, and we conduct 2 interviews with migraine sufferers to further verify our results.

# Chapter 3

# Related Work

In this section, we look at research related to migraine and Natural Language Processing, and we integrate our study within the present body of knowledge. We find that the use of language processing strategies to learn about migraine is no novelty. Studies processing language on migraine themed text do not look far into its connection with diet. However we still find related studies focusing on migraine text either from social media, self-reports or health records. [9]

- **Dietary Supplements:** Probably the most closely related study to our own is a 2018 study on dietary supplements [25]. The development of a search engine that could return supplements associated with extreme effects was developed. The study produced a list of supplements with the most risky ones at the top, and also found that previously banned products made it to the top of the list, thus aiding migraineurs in steering clear from potentially harmful supplements. This is the only study that deals with migraine and language processing we have found that (indirectly) deals with dieting.

- **Sentiment Analysis Studies:** Other studies that we have found do not attempt to tackle any dietary connection to migraine through language processing, but bear more resemblance to ours in terms of the method used, focusing on social media language processing with sentiment analysis. Analyzing migraine related tweets found that among those with expressive value, the emotional tendency was negative, with a high prevalence of extreme sentiment and profanity. [9] Sentiment analysis has also been used for classifying self-reported narratives by migraine patients during a clinical trial [31] finding predominantly negative sentiment, and evaluated the accuracy of machine learning algorithms in classifying such stories finding satisfying performance ratings (F1 scores between 0.82 and 0.86).

- **Models and Frameworks:** We also find studies that focus primarily on developing language processing strategies and fine-tuning them for specific types of text such as one project developing a language model based on 4 different NLP frameworks and designed to extract headache frequency from Electronic Health Records [7]. Language processing frameworks also escape the somewhat predictable format of Electronic Health Records and move into social media reporting in order to access information that isn't typically found in EHR's [18]

# Chapter 4

# Preliminaries

In this chapter we look at a piece of software that we will be using in our method, the main diets we will be evaluating, and some key concepts we will use in said evaluation

## 4.1 Lancsbox

Lancsbox is a free use language processing software that allows importing an entire corpus of text and performing various operations to extract insight. It has several tools that can be used for quick searching of words and linguistic structures as well as filtering, sorting, and performing statistical analysis on results. The tool that we will be using is *GraphColl*, which allows querying for collocates (words found in the vicinity of the search term) and returns a table of collocates packed with information we can use, such as statistics on the strength of association between the search term and collocated words, or frequency of collocated words.

## 4.2 Diets

In our research we will focus on 4 diets frequently used to mitigate migraine symptoms. We look at each of them in turn:

### 4.2.1 The Ketogenic Diet

The Ketogenic Diet (Keto) is a high-fat and low-carb diet associated with the management of neurologic conditions such as epilepsy and migraine. Common keto foods are dairy products, fresh meat and lean protein foods, seeds, nuts and low-carb fruits and vegetables. Keto is also (mostly) gluten free.

### 4.2.2 IgG Elimination Diet

IgG Elimination diet is a dietary strategy that aims to reduce the production of Immunoglobulin G (IgG) antibodies. The production of this antibody results as an immune response to certain foods and can end up leading to migraine attacks. While some foods are commonly known to trigger the production of the IgG antibody, this diet brings with it the challenge of dealing patients' subjective profiles. Different people have different immune responses and the antibody is not produced the same way in everyone.

### 4.2.3 DASH Diet

DASH stands for *Dietary Approaches to Stop Hypertension*, and it primarily focuses on restring high salt (high sodium) foods as well as sugars and saturated fats. Common DASH compatible foods are fruit, vegetables and whole grains.

### 4.2.4 Low Calorie Diet

A simple weight loss diet may also be helpful to mitigate migraine, either because of the nature of the diet, or because obesity is positively correlated with migraine.

## 4.3 Sentiment Lexicons and Classification Metrics

In order to measure sentiment around the diets above we will need to use a Sentiment Lexicon. A sentiment lexicon is a lexical database used to attribute a positive, negative or sometimes neutral sentiment value to a selection of words. Sentiment Lexicons can be built by interviewing people and having them score words to compile a collection of sentiment scores, it can be more general, focusing on common words, or more specialized focusing on words likely to contain emotional value. Later in our research in section 4.4.4. we will evaluate 4 such lexicons in order to evaluate which is most appropriate for our problem. In order to do so we will consult a comparative study on sentiment lexicons that uses 4 different metrics to evaluate performance. These metrics can be computed using the proportion of *true positives* ($TP$) and *true negatives* ($TN$) found by a classifier from a dataset with actual *positives* ($P$) and *negatives* ($N$). We explain these metrics here:

1. **Accuracy**
   Formula: $\frac{(TP+TN)}{(P+N)}$
   Accuracy is the percentage of correctly classified elements, more specifically it is a classifier's capacity to accurately classify positive / negative elements in a dataset. In our case these elements are units of text with positive or negative sentiment.

2. **Precision**
   Formula: $\frac{TP}{TP+FP}$
   Precision is a classifier's capacity to not classify negatives as positives.

3. **Recall**
   Formula: $\frac{TP}{P}$
   Recall is a classifier's capacity to correctly classify positives

4. **F1-Score**
   Formula: $\frac{2 \times Precision \times Recall}{Precision + Recall}$
   F1-Score is a combination of precision and recall measuring a classifier's capacity to classify both positives and negatives (similar to Accuracy)

# Chapter 5

# Method

In this chapter, we explain our research methodology. We discuss the different components of our research, why they are necessary and how they are conducive to answering our research question.

Social media platforms are abundant in raw unaltered text displaying people's thoughts, feelings and behavior. With the use of Natural Language Processing (NLP) strategies, text can be turned into data and analyzed to spot trends and uncover valuable information that other studies may miss. In this study we focus particularly on sentiment analysis. This very common NLP strategy involves measuring the emotional tone of text fragments and scoring it positively, negatively or neutral. In our case, we aim to search for trends in sentiment scores surrounding dietary choices in the context of migraine discussions online.

## 5.1 Data Collection

In order to gain access to people's stories and experiences with migraine, we need a body of text including such discussions. We obtain our data from 2 sources: a collection of Reddit threads from several migraine themed subreddits, and a collection of self-reports by migraineurs provided by provided by Storyconnect, commissioned by Je Leefstijl Als Medicijn. Storyconnect is a Dutch company that collects stories of any type of lived experience to help people connect, find commonalities and patterns in their experiences, and makes them available for discussion with other members of an organization or community. https://storyconnect.nl/en/. We look at each of these in turn:

1. Reddit: To access Reddit threads, we scrape need to scrape subreddits with Reddit's own official API named praw. This allows us to reach both posts and comments up to a limited maximum. We have obtained a total of 1558 reddits threads (posts with comments) from 9

different subreddits on the topic of migraine: r/migraine, r/migraine-science, r/headache, r/cgrpMigraine, r/NDPH, r/ChronicHeadaches, r/ClusterHeadaches, r/OcularMigraines, r/VestibularMigraines.

2. Je Leefstijl Als Medicijn: An additional set of 52 self-reported stories from migraine patients who have had experience with dieting has been provided to us by 'Je Leefstijl als Medicijn', a Dutch foundation advocating and informing on dieting and lifestyle changes as a way to manage various diseases.

In total we have obtained a corpus of language of 888602 tokens.

## 5.2 Literature Review

The reason we are doing a literature review on migraine and dieting is to assess what the current knowledge is on the topic, so that we can verify how accurately research findings reflect people's lived experience with migraine. We begin our literature review in a narrative way, by looking up the general view on migraine and dieting. This allows us to find what diets are being studied, what particular products seem to be points of focus, and how studies on migraine and diet are being conducted overall. Once we know the general scope of the field, we begin to operate in a more systematic manner. We seek to find what diets are popular, why they are helpful, how effective they are, and also what foods are migraine attack triggers or just commonly considered safe or unsafe for migraineurs.

### 5.2.1 The Ketogenic Diet

The Ketogenic Diet is by far the most studied with a body of research dating back since the 1920's [28] and 30's [4] finding a metabolic side to migraine. After these studies, pharmacologic treatments have taken over, but have the study of keto has been reviewed and brought back with hopeful skepticism in recent years [20]. Since then multiple studies have shown that keto is effective in ameliorating migraine even in short-term studies with visible effects within 1 month [13], improving frequency, severity and duration of attacks, as well as resulting in reduced drug intake. Its effectiveness can be further improved by fatty acid and amino acid supplements [13]. Comparative studies [11] [12] also show keto to be more effective than a simple weight loss diet, attributing positive effect to ketosis rather than weight loss, which may naturally occur during keto [10]. Despite potential non-negligible side effects, keto remains a promising avenue for managing migraine.

### 5.2.2 Elimination Diets

Another common dietary strategy for migraine is to follow elimination diets, which seek to eliminate migraine triggers (foods that may cause an attack). It is not quite as frequently studied as the ketogenic diet, but promising results were found nonetheless. Different elimination strategies may be used [23] but the most common is removal from one's diet foods that produce the Immunoglobulin G (IgG) antibody, since it can cause an inflammatory or immune reaction that can trigger an attack. This strategy has proven promising in studies as well, reducing attack frequency and ameliorating migraine. Common IgG positive triggers have also been identified such as nuts, seeds, spices, seafood, gluten, grains, dairy, vegetables, sugary products, eggs etc. [14] [33]

### 5.2.3 DASH and Low Calorie Diet

The DASH diet has also shown some promising results leading to a reduced risk of migraine attacks [2] although no improvement was found in terms of attack duration and intensity.
A simple low calorie diet has been suggested in studies as well, tested in comparison to the ketogenic diet. It is treated in 3 studies that we have included in our literature review, and in 2 of them, it was found to be ineffective. In the third it was found to be less effective than keto.

## 5.3 Language Analysis

To be able to measure people's sentiment regarding migraine, we need to be able to:

1. Identify discussions about food and dietary choices in our corpus of text.
2. Quantify the emotional tone of the relevant text
3. Identify the relevant sentiment charged text within these discussions
4. Verify that our sentiment measurements accurately reflect people's experience within the text

We will deal with each of these problems in turn in the following subsections.

### 5.3.1 Diet Lexicon

Identifying the parts of our corpus of text that discusses migraine in relation to diet can be done by searching for diet related words (foods, drinks, diet names etc.) within our corpus. To reliably access a consistent amount of such conversations, we compile a lexicon of diet related terms with the use of ChatGPT. We ask ChatGPT to provide a complete list of all foods, one with all drinks one with all diets and one with all nutrients (e.g. vitamin and minerals) that it can think of. We gather all instances of these lists into a single collection of words. We obtain a collection of 599 diet terms. We then iteratively search for each of these terms in our corpus of text and filter out the ones that are absent. We remove duplicates from the remainder and end up with a diet lexicon of 204 words that are present in our corpus of text to be analyzed.

### 5.3.2 Sentiment Lexicons

Quantifying the emotional tone of text requires being able to find emotionally charged words within a piece of text and be able to assign a sentiment value to such words. This is where sentiment lexicons come into play. We consider at 4 such lexicons:

1. XANEW - ANEW stands for "Affective Norms for English Words" and it is a sentiment score lexicon developed by Margaret M. Bradley & Peter J. Lang in 1999 [5] providing scores for English words over 3 dimensions: pleasure, arousal and dominance. This initial lexicon has been extended in 2013 to the extended ANEW (XANEW) version including 13915 English Words. [32] The word list is not curated to include only sentiment words, but also any words that may have emotional value (e.g. pizza, medicine, football). XANEW uses a scoring scale from 1 to 9 with the neutral point at 5.

2. AFINN - otherwise known as "A new ANEW" [24] is a sentiment analysis wordlist especially built for microblogs (Twitter posts). It is much shorter than XANEW, with only 2477 words in the AFINN-en-165 version that we will be using, and it includes affective words as well as Internet slang and obscene words. AFINN works on a sentiment scale of -5 to +5 with the neutral point at 0. The lexicon we have made use of can be found at https://github.com/fnielsen/afinn/tree/master/afinn/data.

3. Vader Lexicon - VADER is a pre-trained sentiment analysis model that can evaluate the emotional tone of an entire piece of text. It is based on a lexicon that contains 7521 terms that includes sentiment words, slang, emojis and even acronyms that may contain sentiment value. It works on a sentiment scale from -4 to +4 with the neutral point at 0. Vader Lexicon is considered by its authors a gold standard list of words for microblogs sentiment analysis. [19]

4. SentiWordNet 3.0 - SentiWordNet [15] is the most vast sentiment lexicon available, relying on the lexical network WordNet [22] and includes over 100,000 synsets (sets of synonym words that are equivalent in sentiment score). It assigns three scores (a positivity score, a negativity score, and an objectivity score) to every synset, measuring sentiment over 3 dimensions. SentiWordNet 3.0 [3]is the latest version of the lexical network, with an approximate 20% higher accuracy of sentiment value.

### 5.3.3 Lancsbox: Collocates

Once we have our diet lexicon and our selection of sentiment lexicons, we need to be able to find these scored sentiment words in our corpus of text, to be able to measure emotional tone. But we cannot just look for them anywhere. If we want to use emotional tone as a metric for how beneficial or detrimental to one's migraine a specific food may be, we need to asses whether a sentiment charged word is actually connected to a diet word. One way to increase the likelihood that this is the case, is to only look for sentiment words in the vicinity of diet terms.

Lancsbox can help us find collocated words in such manner with the GraphColl menu that allows us to receive a list of all words found within a given distance of any given word we query. Thus by querying the words in our diet lexicon, we receive a list of collocations for each one of them. We can then cross reference the sentiment lexicons with the lists of collocates to find common words and assign score to the context surrounding diet terms. By further averaging out the scores of all collocates we found for a diet term, we receive a score for the diet term itself that can tell us how people feel about it.

### 5.3.4 Test Corpus

We have so far seen how to locate diet terms in our corpus of text, how to find relevant sentiment words in their vicinity, and how to use these collocations and sentiment lexicons to evaluate the emotional tone around diet terms. However, if we want this method to work, we need test whether the results that it gives accurately reflect what people are trying to express. In order to do this, we have selected a mini-corpus of 8 Reddit threads rich in discussion of diet from our main database. This very reduced sample is small enough to be manually evaluated, allowing us to run our method repeatedly on the mini-corpus to find the right parameters. We need to figure out the following parameters:

1. **Window Size:** how many words to the left and right of a word do we search for collocates?
2. **Statistic:** what statistics do we want Lancsbox to record about the collocations we find
3. **Thresholds:** what Thresholds should we input in Lancsbox for collocation frequency and statistics
4. **Lexicon:** Which of the 4 sentiment lexicons we have available gives the best results?

We address each of these in turn:

1. **Window Size:** When querying collocations to find sentiment words around a diet term, we need to make sure we find as much of the relevant sentiment words as possible to accurately measure context. However, social media text is messy and unpredictable. One post may be lengthy and detailed with relevant sentiment words appearing far away from the target term. This would justify selecting a large window, in order to not miss out on any meaningful sentiment. On the flip side, a different post, or a comments sections may be full of short one-liners, changing topic from one comment to another, meaning that a larger window may end up capturing emotional tone from outside of the scope of conversation surrounding a queried diet term. Additionally, repetition of a diet term would mean a large window size is even likely to end up measuring the same context multiple times, providing a skewed final result. Take for example the text below:

   *Pizza and ice cream are the only ones I love without a migraine haha. Frozen pizza products. Not pizza in general, frozen pizza specifically.*

   We see that *"pizza"* appears 4 times in just 2 short lines. If we then query the collocations of *"pizza"* with a window size of 30 words (15 to the left and 15 to the right) we would capture the word *"love"* 4 times, skewing the average sentiment score positively.A window of 14 (7 to the left and 7 to the right) would only record *"love"* once and a window of 18 (9 to the left and 9 to the right) would record it twice. After sifting through the entire test-corpus and purposely looking up some of the longer format stories, we find that short ideas and quickly changing topics are frequent on social media and a shorter window serves us better. We also find that if a situation with longer stories and wider context appears, the food/drink/diet that it is about is often mentioned repeatedly through the story, meaning that a short window still has a decent chance of capturing a lot of the relevant context.

After several attempts we settle on window size of 20 words (10 to the left and 10 to the right).

2. **Statistic:** We have mentioned in section 4.4.3. that seeking the emotional tone surrounding a diet term means we need to increase the likelihood that a collocate with sentiment value is actually correlated to the queried diet term. We have argued that ensuring that a sentiment charged collocate thus needs to be in the vicinity of the queried diet term in order to be correlated. However this may not be sufficient. It may be the case at times that a sentiment charged word is found near by a diet term, but that the occurrence is purely accidental. To account for this, Lancsbox allows us to query collocates with an associated statistic called a MI Score (Mutual Association Score) which measures how tightly linked two words are.

$$MI = \log_2 \frac{O}{E}$$

If it is observed (O) that 2 words are found together far more often than expected (E), then the MI score will be higher. While we cannot filter out words that are collocated purely by accident, we can use the MI score to reduce their ability to skew our results. Since we compute the sentiment score of a diet term by averaging out the sentiment scores of its collocates, we can instead perform a weighted mean, where the MI scores are the weights, ensuring that random collocates influence the mean less and meaningful collocates influence the mean more.

$$score = \frac{\sum s \times MI}{\sum MI}$$

3. **Thresholds:** We have seen that Lancsbox provides us with statistical insight into our collocations, but it also allows us to set thresholds that determine what is or isn't considered a collocation. We can decide how many times minimum a word should be found within our window size for it to be considered a collocate (frequency - how many times a word was found within **window size** of the queried word) and also what the minimum value for the chosen statistic (in our case MI score) should be. We set both the the MI score and frequency threshold to a minimum of 0, because we want to filter out words with negative MI scores (which would suggest that they are negatively correlated), and because we want to catch all words that may include relevant emotional value, even if they are only found once (frequency = 1). However, accounting for collocations being potentially found more than once, means that frequency too should be included in the weight of the average, so our formula for the sentiment score of a diet term becomes:

$$score = \frac{\sum s \times MI \times freq}{\sum MI \times freq}$$

4. **Lexicon:** Finally, we need to test which of our 4 candidates for sentiment lexicons to provide the actual sentiment scores for our collocates gives the results that most closely resemble our manual evaluation of the text in our aforementioned mini-corpus. With the chosen **window size**, **Statistic**, and **Thresholds** that we have selected above, we compute the sentiment value of the surrounding context for each diet term in our diet lexicon that is also present in the 8-thread mini-corpus. We first adapt all 4 lexicons to use the same scoring scale (-4,4) so that we may easily compare results. For the total 49 diet terms present within the mini-corpus, we find the following results:

| Lexicon | Diet Terms found | Negative Scores | Positive Scores | Range |
|---|---|---|---|---|
| XANEW | 49 | 48 | 1 | (-0.06, 1.97) |
| AFINN | 39 | 15 | 23 | (-1.6, 2.5) |
| VADER | 43 | 10 | 33 | (-1.7, 3.1) |
| SentiWordNet 3.0 | 49 | 25 | 24 | (-1.25, 1.87)) |

This already tells us that XANEW skews too far into positive scores to be reliable. We see that with such a small corpus of text, many of the diet terms we spot actually occur only once in the entire body of text, so we can simply check whether the sentiment delivered by the lexicons fits. By manually evaluating the available text in our mini corpus we find that we have both positive and negative sentiment. Much of the negative sentiment revolving around a few diet terms such as gluten, cheese and dairy, coke, wine and alcohol, and it fails to be correctly classified by scoring with XANEW.

We still need to evaluate how AFINN, VADER and SentiWordNet. We consult Al Shabi's 2020 comparative study [1] on sentiment lexicons, which includes all 3 lexicons that we are using here. This study however does use a different (older) version of AFINN (AFINN-111). Al Shabi compares lexicons over 4 dimensions, Accuracy, Precision, Recall and F1-Score while classifying tweets as positive, neutral or negative. He finds that VADER scores the best accuracy, precision and F1-Score. Vader also scores the best recall for positive and negative tweets, but AFINN outperforms VADER exclusively with regard to recall of neutral tweets. SentiWordNet on the other hand scores the poorest accuracy and precision indicating it is less reliable at identifying whether sentiment is positive/negative. In our case, we don't only need to classify correctly, but we also have to use a scoring scale to discriminate between less or more positive/negative sentiment. We see that our scores for SentiWordNet 3.0 are restricted to a smaller range, sticking much closer to neutral than AFINN and VADER do, failing to properly grasp the polarity of sentiment. We thus decide that SentiWordNet is not a good choice and will proceed to use both AFINN and VADER to run our method on the main corpus of text we intend to analyze. Due to VADER's stellar performance on almost all dimensions but poor recall of neutral tweets, we also keep AFINN results to fall back on since VADER seems a bit more likely to push neutral scores towards more positive and negative scores.

## 5.4 Verification Interviews

Once we have completed our literature review and language analysis, we have additionally conducted 2 verification interviews with migraine sufferers to ask them about their experience. One of the respondents had been dieting gluten-free for 5 years and had switched to keto for 4 years since. The second had only been on keto for 4 years. We asked both about their experience with each diet that they had tried, as well as about their identified personal triggers, and how their dietary strategy had shifted from plain keto to include / exclude certain foods.

# Chapter 6

# Results

In this section, we first look at the results that we obtain by applying our language processing method, and in our literature review. Then we compare our findings from language processing with what we have seen in the literature review and further check them against the responses from our verification interviews. We categorize the foods we have present in our Diet Lexicon to look in turn at those that fit each diet. We then look further within different 10 food categories: alcoholic beverages, caffeinated beverages, carbohydrates, dairy, fruit, meats, sweets, vegetables and vitamins.

## 6.1 Per Diet:

1. **Ketogenic Diet**
   We find 25 foods in our lexicon that are compatible with a ketogenic diet. We examine the computed sentiment scores and we find that we have obtained 13 positive and 12 negative scores from VADER and 12 positive and 13 negative scores from AFINN. However, when we look at a handful words with richer context (7 or more collocates used to compute scores) we find for both lexicons the words: *"keto", "ketones", "ketogenic", "meats", "beef", "cheese", "omega-3", "turmeric"*. 7 out of 9 of them score positively for VADER and 6 out of 9 for AFINN. We also see that for the entire set of keto compatible foods, positive scores tend to be further away from neutral than negative scores. 7 positive scores are greater than 0.5 and only 3 negative scores are under -0.5 for VADER, while for AFINN also 7 words score positively above 0.5, whereas there are only 4 negative scores under -0.5. Summing up, these results show us a general slightly positive tilt in the measured emotional tone surrounding mentions of keto compatible foods.

2. **IgG Elimination Diet**
   For the IgG Elimination diet, we find 41 words in total, all of which get scored by VADER and 40 of them which gets scored also by AFINN.

From VADER we get 27 positive scores and 14 negative and from AFINN we get 25 positive scores and 15 negatives. If we once again focus separately on words with a higher amount of collocates used to compute the score, we find the same 8 words for both lexicons: *"turmeric", "omega-3", "herbs", "zinc", "antioxidant", "coffee", "water", "meats"*. All of these score positively from both lexicons. Thus we can see that for this selection, discussion around foods compatible with the IgG Elimination diet also tilts positively in emotional tone. It is worth noting also that while the selected foods are typically accepted within an IgG Elimination diet, it is a subjective matter from one person to another whether a food will or will not trigger production of immunoglubulin G antibody, and a potential migraine attack in consequence.

3. **DASH Diet**
   For the DASH diet we find 36 words in total that give us 24 positive and 12 negative scores from VADER and out of 36 positive scores from VADER. AFINN has only managed to find collocates for 34 words out of the 36, resulting in 24 positive and 10 negative scores. When filtering for sentiment scores computed using 7 or more collocations we find 7 words *"omega-3", "fruits", "turmeric", "herbs", "beans", "rice", "lentil"* out of which 6 score positively from VADER and the same 7 words, 5 of which score positively for AFINN scores. We conclude that our analysis shows a positive emotional tone for the DASH diet.

4. **Low Calorie Diet**
   Looking at the results for the low calorie diet we find 32 words that give us 21 positive and 11 negative scores for VADER, and 31 of those words also have scores from AFINN out of which 20 are positive and 11 negative. When we adapt for words with richer context again we find that the same 2 words *("herbs", "coffee", "antioxidant", "water")* use 7 or more collocates for both lexicons. We find that all 9 of these words all of which score positively from both sentiment lexicons. Thus we find that there is a positive emotional tone around a low calorie diet as well.

## 6.2 Per category:

1. **Alcoholic Beverages:**
   We find 10 words describe alcoholic beverages in our Diet Lexicon that have also yielded sentiment scores. 8 out of the 10 score negatively and 2 score positively from VADER and 7 out of 10 for AFINN. We find that the only alcoholic drinks that score positively are beer and

vodka.

2. **Caffeinated Beverages:**
We find words for 14 caffeinated beverages in our Diet Lexicon that have yielded sentiment score, and we find that all of them score positive scores from both AFINN and VADER lexicons, with some *"red bull", "cola"* scoring closer to neutral and others *latte, americano* scoring higher than 1.

3. **Carbohydrates:**
We find 17 words that describe carbohydrates. From VADER, 10 receive positive scores and 7 receive negative scores while from AFINN 12 receive positive scores and 5 receive negative scores.

4. **Dairy:**
We find only 5 mentions of dairy products giving us 3 positive and 2 negative scores from VADER and 4 positive and 1 negative from AFINN.

5. **Fruit:**
We find 13 words describing fruit, all of which are scored by VADER and 12 by AFINN. We get 9 positive scores and 4 negatives from VADER and 9 positives and 3 negatives from AFINN.

6. **Meat:**
We find 10 words describing meat, yielding 2 positive and 8 negative scores from VADER, and 3 positive and 7 negative scores from AFINN. For VADER, 5 of the 8 negative scores are close to neutral and for AFINN 4 of them are close to neutral.

7. **Sweets:**
We find 11 words describing sweets and sugary products out of which all 11 were scored by VADER and only 10 by AFINN. VADER yields 5 positive and 6 negative scores and AFINN yields 7 positive and 3 negative scores. We find a high prevalence of scores coming close to neutral from both the positive and negative side.

8. **Vegetables:**
We find 13 words describing vegetables, all of which are score by VADER and only 12 by AFINN. We find 8 positive and 5 negative scores from VADER and 8 positive and 4 negative scores from AFINN.

9. **Vitamins:**
We find 17 words describing vitamins, all of which are scored by both lexicons, revealing a very high prevalence of positive scores with only 2 negative scores from VADER and 1 from AFINN.

## 6.3 Literature Review Findings

Our review included a total of 40 academic articles. Out of these, 15 papers study the ketogenic diet, 2 study IgG elimination diets, 2 study the DASH diet and 3 are looking at low calorie and weight-loss diets. Another 3 papers focus specifically on alcohol, 10 focus on other related topics such as less commonly used diets, obesity or identifying migraine triggers and 5 more are looking at the general state of migraine and dieting. Most of these studies take the form of clinical trials, recruiting patients for the studies and monitoring their evolution over a pre-determined span of time. The beneficial effect of dieting becomes apparent in most studies, most commonly the conclusion being that dieting is undeniably helpful, but that more research is needed. We decide to focus our study on 4 diets: Ketogenic Diet, IgG Elimination Diet, DASH diet, Low Calorie Diet.

## 6.4 Verification Interviews Findings

We interviewed 2 respondents Common findings were the high risk of an attack due to sugar, alcohol, and lactose and dairy products. Both respondents have seen improvements in their energy levels and frequency, duration and severity of attacks. Both were following keto at the time of the interview but had adapted the diet around their subjective triggers and preferences. The first respondent had restricted the diet further, cutting out all sugars (even in fruit) and all alcohol, and had noticed a preference for fresh products to processed foods, even within keto compatible choices. The second was at times allowing foods that don't fit the ketogenic diet with moderation. Alcohol, caffeine, fruits and vegetables were all safe. The second responder had also noticed that keto had created a sort of "reset" where going on the diet for 2 years meant that afterwards, foods that were triggering attacks beforehand had become safe.

One of the respondents had additional experience with a forum of migraine sufferers and could confirm that it is common practice for migraineurs to seek out personalized dietary choices and avoid triggers as opposed to just following a pre-determined diet. Additionally a difference was spotted between American and European migraine sufferers where Americans were more open to lifestyle changes whereas the Europeans were less aware of the curative potential outside of pharmacologic solutions.

# Chapter 7

# Discussion

## 7.1 Discussion

We have reviewed the scientific literature on the connection between diet and migraine, we have performed a language processing method on a body of self-reported stories and Reddit threads about migraine, and we have conducted 2 interviews to verify our findings. We will now discuss our results we have obtained through language processing and compare it with the literature review and the findings from the interviews. We will then treat the limitations of our study and point out future avenues of research that result from our findings.

### 7.1.1 Comparing Findings

Our literature review has found promise in the treatment of migraine by dietary means through several dieting options. We have chosen to focus on four dietary strategies: the Ketogenic Diet, the IgG Elimination Diet, the DASH Diet and Low Calorie Dieting. We have seen that the ketogenic diet is the most studied and seems to yield better improvements for patients than other diets such as low calorie diets or DASH. Results from our language processing analysis show that sentiment value of online discussions revolving foods compatible with these dietary strategies results in a positive sentiment score, which seems in accordance with the general direction of the literature. We have seen an even split in positive and negative scores for the ketogenic diet, but with positive scores tilting further away from neutral than the negative scores. We have also seen for all four diets that words that incorporate a higher number of collocations in the computation of the sentiment scores tend to lean positive. This is relevant because words that provide a sentiment score from few collocations are doing so based on lesser context and are more likely to be circumstantial. In our analysis, the DASH Diet, IgG Elimination Diet and Low Calorie diet have a higher preponderance of positive scores than negative scores, higher than keto which resulted in an even split. This may suggest that while keto is an effective dietary strategy, some keto compatible foods may be migraine triggers.

We have also categorized foods to spot trends within food classes that may be diet triggers or not. We have seen mostly negative scores for alcoholic beverages which is in line with the scientific literature. The odd outlier with regard to alcohol is a positive sentiment score for beer. We have also seen that while caffeine is known to be a trigger, but in our analysis caffeinated drinks score positively. It may be that this result is circumstantial since sentiment measures for caffeinated drinks have relied on few collocations. One of the most meaningful findings is that we have also seen predominantly positive scores with regard to vitamins. For carbs, meats, sweets, and dairy we end up with mostly inconclusive numbers, as whether the score tilts positive or negative for more words in a given category seems mostly accidental.

It then seems we can confirm that dieting for migraine is a promising avenue as a general direction, and could find a few surprising trends with regard to alcohol, caffeine and vitamins, but our results currently don't allow us to draw any conclusions that are more specific than that. The subjective nature of migraine makes it further more likely that this is the case, as migraine triggers differ from one person to another.

We have also conducted 2 interviews where respondents have reported positive experiences with keto, but have also reported preferring to add or remove foods from the diet to adapt to their own triggers. The interview responses confirm our positive impression of the ketogenic diet. Both respondents also can safely drink caffeine and avoid lactose which is in alignment with our findings. However, the opposing responses from interviewees about triggers seem to reflect the ambiguous nature of our sentiment scores for carbs, meat and sweets.

### 7.1.2 Limitations

Our language analysis method suffers from a few limitations. We list them and discuss their potential implications below:

1. **Circumstantial Scores**
   We have already pointed out that sentiment scores computed from a higher number of collocations are more likely to be accurate as the context is less likely to be circumstantial. It is also the case that words that appear more frequently in our corpus are more likely to give accurate measurements since they expose us to more context. However, a lot of diet words for which we have computed sentiment scores only ap-

pear a few times in the corpus and rely on few collocations to compute a sentiment score which in those cases leads to unreliable results.

2. **Language Nuance**
   Our method for sentiment analysis doesn't account for nuances of language such as sarcasm, irony, humor or negations. This may also result in erroneous sentiment measurements, as the sentiment score of a word becomes warped by its purpose.

### 7.1.3 Future Work

Our work shows that sentiment analysis, used in parallel with a literature review can be used to find trends that either confirm or contradict the scientific consensus on a given topic. We have been able to confirm the positive impact of dieting strategies on migraine through sentiment analysis. However, more research is needed to make such methods more precise. We have found ourselves unable to rely on some of our results, indicating that better sampling of text from social media and online forums can provide more accurate measurements.

Responses from our interviewees combined with the ambiguous nature of our sentiment scores for some food categories also highlights the importance for migraineurs to identify individual triggers. At the moment, most studies of migraine and dieting set out to test pre-determined dietary patterns in studies. More research into finding triggers and dealing with the subjective nature of migraine is needed. The reported difference in mindset and level of information between European and American patients by one of our interviewees also raises an interesting possibility. Should a comparative language analysis study be done on forums from different parts of the world, we could asses the importance for people to be informed about diet and lifestyle choices as a way to mitigate their health issues.

# Chapter 8

# Conclusion

We set out to verify whether the scientific consensus on dieting and migraine reflects patients' experience. As part of our research we have conducted a literature review, we have developed a natural language processing method for computing sentiment score, and we have conducted 2 verification interviews to check our results. We have measured sentiment for 4 diets and found a general trend of positive sentiment correlated to the optimistic findings in the literature review and in our interview respondents' answers. Further attempting to extract more insight from our results has shown some consistencies such as an overall negative sentiment around alcohol, but also contradictions such as an overall positive sentiment revolving around talks on caffeinated drinks. The ambiguity in our results highlights the subjective nature of migraine and the importance of studying individual food triggers for patients, but also reveals limitations in our study. Such shortcomings may be an insufficient body of text that would provide a reliable sample of sentiment charged context, and failure to account for linguistic nuance such as irony, sarcasm, humor or negation.

# Bibliography


[1] Mohammed Al Shabi. Evaluating the performance of the most important lexicons used to sentiment analysis and opinions mining. 2020. , https://osf.io/exqr9/download.

[2] Muhammad Amer. Effects of dietary sodium and the dash diet on the occurrence of headaches: results from randomised multicentre dash-sodium clinical trial. 2014. , https://pubmed.ncbi.nlm.nih.gov/25500372/.

[3] Stefano Baccianella. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. 2010. , https://aclanthology.org/L10-1531/.

[4] Barborka. Migraine - results of treatment by ketogenic diet in 50 cases, 1930. , https://jamanetwork.com/journals/jama/article-abstract/245128.

[5] Margaret M. Bradley. Affective norms for english words (anew). 1999.

[6] Rebecca Burch. The prevalence and impact of migraine and severe headache in the united states: Updated age, sex, and socioeconomic-specific estimates from government health surveys, 2020. Department of Neurology, Graham Headache Center, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA https://pubmed.ncbi.nlm.nih.gov/33349955/.

[7] Chia-Chun Chiang. A large language model-based generative natural language processing framework finetuned on clinical notes accurately extracts headache frequency from electronic health records. 2023. , https://pubmed.ncbi.nlm.nih.gov/37873417/.

[8] C.E. Clarke. Economic and social impact of migraine, 1995. From the Department of Neurology, Hull Royal Infirmary, Hull, and Glaxo Pharmaceuticals UK Ltd, Uxbridge, UK, https://www.researchgate.net/publication/14472805_Economic_and_social_impact_of_migraine.

[9] Hao Deng. Sentiment analysis of real-world migraine tweets for population research. 2019. , https://www.researchgate.net/publication/338600696_Sentiment_analysis_of_real-world_migraine_tweets_for_population_research.

[10] Cherubino Di Lorenzo. Diet transiently improves migraine in two twin sisters: possible role of ketogenesis? 2013. , https://pmc.ncbi.nlm.nih.gov/articles/PMC3951260/.

[11] Cherubino Di Lorenzo. Short term improvement of migraine headaches during ketogenic diet: a prospective observational study in a dietician clinical setting. 2013. , https://www.researchgate.net/publication/344101629_Short_term_improvement_of_migraine_headaches_during_ketogenic_diet_a_prospective_observational_study_in_a_dietician_clinical_setting.

[12] Cherubino Di Lorenzo. Migraine improvement during short lasting ketogenesis: a proof-ofconcept study. 2014. . https://pubmed.ncbi.nlm.nih.gov/25156013/.

[13] Cherubino Di Lorenzo. A ketogenic diet normalizes interictal cortical but not subcortical responsivity in migraineurs. 2019. , https://pubmed.ncbi.nlm.nih.gov/31228957/.

[14] Mustafa Ertas. Diet restriction in migraine, based on igg against foods: A clinical doubleblind, randomised, cross-over trial. 2010. , https://pubmed.ncbi.nlm.nih.gov/20647174/.

[15] Andrea Esuli. Sentiwordnet: A publicly available lexical resource for opinion mining. 2006. , https://www.researchgate.net/publication/200044289_SentiWordNet_A_Publicly_Available_Lexical_Resource_for_Opinion_Mining.

[16] Parisa Gazerani. Migraine and diet, 2020. Laboratory of Molecular Pharmacology, Department of Health Science and Technology, Faculty of Medicine, Aalborg University, 9220 Aalborg East, Denmark; gazerani@hst.aau.dk, url{https://pmc.ncbi.nlm.nih.gov/articles/PMC7986733/.

[17] Chen Gou. Functional connectivity of the language area in migraine: a preliminary classification model. 2023. , url{https://bmcneurol.biomedcentral.com/articles/10.1186/s12883-023-03183-w.

[18] Yuting Guo. Generalizable natural language processing framework for migraine reporting from social media. 2022. , https://pubmed.ncbi.nlm.nih.gov/37350878/.

[19] C.J Hutto. Vader: A parsimonious rule-based model for sentiment analysis of social media text. 2014. , https://github.com/cjhutto/vaderSentiment.

[20] Ferdinando Maggioni. Ketogenic diet in migraine treatment: A brief but ancient history. 2011. ,https://pubmed.ncbi.nlm.nih.gov/21727144/.

[21] Ananya Mandal. Economic impact of migraines.

[22] George A. Miller. Introduction to wordnet: An on-line lexical database. 1993. , https://academic.oup.com/ijl/article-abstract/3/4/235/923280.

[23] Kattia Valentine Nguyen. The evidence for diet as a treatment in migraine—a review. 2024. , https://pubmed.ncbi.nlm.nih.gov/39408380/.

[24] Finn Arup Nielsen. A new anew: Evaluation of a word list for sentiment analysis in microblogs. 2011. , https://www.semanticscholar.org/paper/A-New-ANEW%3A-Evaluation-of-a-Word-List-for-Sentiment-Nielsen/d38763b4b0bbf1ed91b55107c658fcc96f8ef82d.

[25] Nelleke Oostdijk. Fora fuelling the discovery of fortified dietary supplements – an exploratory study directed at monitoring the internet for contaminated food supplements based on the reported effects of their users. 2017. Radboud University Nijmegen, Department of Language and Speech, Nijmegen, the Netherlands, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0215858.

[26] Marco Pescador Ruschel. Migraine headache, 2024. , https://www.ncbi.nlm.nih.gov/books/NBK560787/.

[27] Igor Petrušić. Influence of next-generation artificial intelligence on headache research, diagnosis and treatment: the junior editorial board members' vision – part 1. 2024. Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations, url{https://thejournalofheadacheandpain.biomedcentral.com/articles/10.1186/s10194-024-01847-7.

[28] Schnabel. An experience with a ketogenic dietary in migraine, 1928. , https://www.semanticscholar.org/paper/An-Experience-with-a-Ketogenic-Dietary-in-Migraine-Schnabel/a839b1daf1c12890d0d0456318f7b912500965dd.

[29] Angelo Torrente. The clinical relevance of artificial intelligence in migraine. 2024. , url{https://pmc.ncbi.nlm.nih.gov/articles/PMC10813497/.

[30] Gloria Vaghi. Who cares about migraine? pathways and hurdles in the european region - access to care iii, 2023. , https://thejournalofheadacheandpain.biomedcentral.com/articles/10.1186/s10194-023-01652-8.

[31] Nicolas Vandenbussche. Using natural language processing to automatically classify written self-reported narratives by patients with migraine or cluster headache. 2022. , https://pubmed.ncbi.nlm.nih.gov/36180844/.

[32] Amy Beth Warriner. Norms of valence, arousal, and dominance for 13,915 english lemmas. 2013. , https://pubmed.ncbi.nlm.nih.gov/23404613/.

[33] Yangzhi Xie. Effects of diet based on igg elimination combined with probiotics on migraine plus irritable bowel syndrome. 2019. , https://pubmed.ncbi.nlm.nih.gov/31531150/.