Stojchevska et al.
BMC Medical Informatics and Decision Making (2025) 25:88
https://doi.org/10.1186/s12911-025-02916-w

BMC Medical Informatics and Decision Making

RESEARCH

Uncovering the potential of smartphones for behavior monitoring during migraine follow-up

Marija Stojchevska[1*], Jonas Van Der Donckt[1], Nicolas Vandenbussche[2,3], Mathias De Brouwer[1], Koen Paemeleire[2], Femke Ongenae[1] and Sofie Van Hoecke[1]

Abstract

Background Migraine is a neurological disorder that affects millions of people worldwide. It is one of the most debilitating disorders which leads to many disability-adjusted life years. Conventional methods for investigating migraines, like patient interviews and diaries, suffer from self-reporting biases and intermittent tracking.

Methods This study aims to leverage smartphone-derived data as an objective tool for examining the relationship between migraines and various human behavior aspects. By utilizing built-in sensors and monitoring phone interactions, we gather data from which we derive metrics such as keyboard usage, application interaction, physical activity levels, ambient light conditions, and sleep patterns. We perform statistical analysis testing to investigate whether there is a difference in user behavioral aspects during headache and non-headache periods.

Results Our analysis of 362 headaches reveals differences in behavioral aspects such as ambient light, use of leisure apps, and number of keystrokes during headache periods and non-headache periods.

Conclusions This exploratory study shows on the one hand that it is possible to monitor various human behavioral aspects using the smartphone sensors and interaction data only. On the other hand it shows that we can observe difference in human behavior between headache and non-headache periods. Our work is a step towards objectively measure the effects that migraine has on people's lives.

Keywords Migraine, Real-world data, Smartphones, Human behavior

Background

Migraine is a highly prevalent and debilitating neurological disorder. It affects millions of individuals worldwide, imposing a significant social and economic burden on both patients and healthcare systems. It has a diverse symptomatology, which besides headache, includes symptoms such as photophobia, nausea, cognitive decline and vomiting [1, 2]. There is little understanding of what triggers migraine attacks [3, 4], posing a challenge for doctors and patients in developing effective treatments and disease management methods. Over the years, conventional methods, such as patient interviews, diaries, and clinical assessments have been utilized to investigate migraine characteristics, symptoms, and triggers, and to decide on treatment and management [5–7]. However, these approaches often rely on self-reporting, i.e. with a diary, and retrospective data,

*Correspondence:
Marija Stojchevska
Marija.Stojchevska@ugent.be
[1] IDLab, Ghent University - imec, Technologiepark-Zwijnaarde, Ghent 9052, Belgium
[2] Department of Neurology, Ghent University Hospital, Ghent 9000, Belgium
[3] 4BRAIN, Institute for Neuroscience, Department of Head and Skin, Ghent University, Ghent 9000, Belgium

leading to potential inaccuracies and limitations in capturing the real effects of migraine on people's lives. These inaccuracies arise from recall bias (error in recollecting information regarding past events or experiences), confirmation bias (error stemming from the tendency to look for or interpret information that confirms person's existing beliefs) and intermittent logging. Additionally, migraine patients have expressed frustrations and negative emotions about tracking, including concerns that tracking could worsen their experience with the condition [4]. Current treatment and management mostly rely on medication and lifestyle adaptation [8]. However, due to all the shortcomings of the information used in treatment decision making, treatments are often ineffective [8]. Additionally, patients do not always use the drugs they have been prescribed correctly. They discontinue their treatments early [9], and they often self-medicate and/or over-medicate [10]. This can have adverse effects and worsen the patient's situation, highlighting the importance of well-informed and patient-centered treatment decision making.

In recent years, the widespread adoption of smartphones has revolutionized the way we collect and analyze data in various fields, including healthcare. Much of the information gathering, such as diaries, questionnaires and assessments, now have a digital form, primarily incorporated into smartphone applications. Examples of such applications include Migraine Buddy [11], Migraine Coach, and N1 Headache [12]. These applications offer easy headache tracking, logging various types of information, such as activities, weather, food and drink intake, and allow taking additional notes. Many of them provide statistical overviews of the logged information, which should serve as tools for people to better analyze and understand their symptoms and triggers. All of these apps share a common shortcoming: they mostly rely on subjective information and self-logging of symptoms and triggers. However, symptoms and triggers may present themselves long before the headache occurs, which makes self-reporting more difficult. Additionally, a trigger may be a combination of factors or it can be dose dependent. Both symptoms and triggers may also change over time, introducing additional challenge for self-tracking [3, 4]. Therefore, current solutions do not address the issue of inaccuracies regarding logging the context prior to the headache or in which the headache occurred, nor the frustrations and concerns patients have regarding headache tracking. Hence, there is a need for obtaining objective data in a continuous, yet non-intrusive manner. Besides digitizing self-tracking, smartphones can be used to collect objective data, as they are equipped with various sensors, such as motion sensors (accelerometer, magnetometer, and gyroscope), a GPS location sensor, a proximity sensor, an ambient light sensor, and a step detector. Additionally, phone usage patterns, like a screen status detector, keyboard interaction and foreground applications, can be tracked. As smartphones have become an integral part of our daily lives, and we use them frequently throughout the day, they enable the continuous monitoring of various behavioral parameters, such as movement, sleep, phone usage and additional environmental context, such as location and ambient light. Using such objectively collected real-world data might bring us closer to understanding migraine mechanisms and their relation to lifestyle and behavior. Real-world monitoring of behavioral aspects for migraine patients also finds its relevance in clinical settings by supporting decision making for treatment and following-up after treatment changes. It also facilitates more detailed and thorough information flow between patients and physicians [13, 14].

In this work we explore the possibility of monitoring various human behavioral aspects using real-world smartphone data, specifically focusing on keyboard interaction, app usage, activity index, ambient light and sleep duration, as these aspects have been frequently associated with migraine attacks [3, 5, 8, 15]. We propose a methodology for processing this raw smartphone data and translating it into meaningful behavioral aspects. Using these modalities, we look for difference in these behavioral aspects during headache periods and non-headache periods. This work aims to address the following research questions:

1. Can smartphone data be used for real-world monitoring of behavioral aspects in the migraine context, and if so, which ones?
2. What data processing steps are needed to obtain usable and interpretable information from this smartphone data linked to migraine symptoms and triggers?
3. What behavioral changes do migraine patients show between headache periods (ictal) and headache-free periods (interictal) attacks?
4. What are the limitations of this approach and how can they be accounted for?

This paper is organized as follows: In Related work section we discuss related migraine studies and studies using smartphone data for monitoring behavioral aspects. Next, in Methods section we present our methodology of data collection, processing the raw smartphone data, and data analysis. Afterwards, the obtained results are outlined in Results section. The discussion of our findings and limitations can be found in Discussion section. Finally, we formulate our conclusion and discuss future work possibilities in Conclusions section.

Related work

In this section we discuss the related research. We first focus on migraine symptomatology and which behavior changes can be noted in migraine attacks, both in-between attacks as well as during attacks. Afterwards we discuss studies that have used smartphone data for behavioral monitoring.

Related migraine studies

A lot of research is done on how migraine impacts people's lives and how lifestyle impacts migraine attacks and headaches. Some of the symptoms that patients report, and that are most pertinent to this research, include cognitive decline, photophobia and disturbed sleep.

Cognitive decline has been extensively researched [16] in both interictal (i.e., period between headaches) and ictal (i.e., period during headache) studies. Interictal studies, focusing on cognitive differences between migraine patients during migraine-free periods and subjects with no headaches, show conflicting results. Studies focusing on cognitive performance during a migraine attack consistently find impairments across various cognitive aspects [16]. Gil-Gouveia et al. found in their study that cognitive performance decreases during migraine attacks compared to headache-free periods [17]. Their study included tests on executive functions (e.g., attention, processing speed, working memory), long-term memory (e.g., visual memory, verbal memory and learning), perception and motor control (e.g., spatial and visual perception, motor function and speed) and language (e.g., naming, verbal initiative). The results indicated a nominal performance decline during migraine attacks for the majority of tests. A significant decrease in performance was observed in two tests: one measuring processing speed and reading, and the other measuring learning and memory. Edwards et al. found a statistically significant decline in overall cognitive performance at the onset of the headache compared to the migraine-free baseline period, in a study including 30 migraine patients[18]. A significant drop was observed in cognitive tasks including simple reaction time (e.g., visual-motor reaction time, immediate attention), procedural reaction time (e.g., sustained attention, concentration), matching to sample (e.g., working memory) and pursuit tracking (e.g., fine motor skills). The limitation of these studies lies in their cross-sectional nature and the use of rigorously designed cognitive tests. Such studies fail to provide a comprehensive understanding of how this cognitive impairment manifests and impacts patients in their day-to-day lives.

Photophobia, or light sensitivity, is one of the most prevalent symptoms among migraine patients both during ictal and interictal periods [15, 19, 20]. It is one of the diagnostic criteria for migraine defined by the International Classification of Headache Disorders, 3rd Edition (ICHD-3) [1]. Bright light has also often been reported as a triggering factor of headache attacks by migraine patients [3, 21]. As a result, management strategies often include seeking a dark room [8] and avoiding bright light exposure. Studies give conflicting results on triggering migraine with light [15]. Some show results in which migraine can be triggered with specific color lights [20] regardless of the intensity, and others present results of experiments in which light did not induce migraine [22]. Moreover, researchers argue that early manifestation of photophobia as a symptom may be the basis for believing that light is a trigger [15].

Sleep and migraine have a complex and incompletely understood bidirectional relationship [23, 24]. Poor sleep quality and various other sleep disturbances, such as insomnia and sleep-related breathing disorders, have often been associated with migraine [7, 24, 25]. On the other hand, sleep deprivation and poor sleep quality are commonly reported as migraine triggers among migraine patients [3, 8, 21]. Whereas the association between frequency of migraine attacks and insufficient sleep, insomnia and lower sleep quality has been repeatedly found in studies [7, 24, 26], there is insufficient evidence on the association between sleep quality, pain intensity and burden [24, 27]. There is also limited evidence that supports the subjective perception of poor sleep quality and lack of sleep as migraine triggers [23, 27]. The Pittsburgh Sleep Quality Index (PSQI) is a self-reported sleep quality questionnaire used in many studies researching the relationship between sleep and migraine [24, 27]. Studies have shown that people with insomnia overestimate the time it took them to fall asleep and underestimate the total sleep time compared to objective sleep estimates, such as polysomnography [27]. Polysomnography is however expensive and impractical for longitudinal studies.

Further, migraines have been linked to disruptions in circadian rhythms, with studies suggesting that migraine attacks often follow a predictable daily pattern influenced by hormonal fluctuations and sleep-wake cycles [28]. Studies have pointed melatonin, often referred to as the sleep hormone, to be an important regulator of sleep, circadian rhythm and headache disorder [29]. Excessive artificial light at night or insufficient daylight exposure during the day can disrupt melatonin production and increase migraine susceptibility. To conclude, within the research community, there is an interest in understanding the relationship between cognitive decline, photophobia, sleep behavior, and stress, with migraine, and headache frequency and intensity. Most of the studies analyzing such associations are however cross-sectional or questionnaire-based and thus suffer from several limitations.

Smartphone data for behavioral monitoring

Recently, due to their accessibility and ubiquity, smartphones have been considered as good device alternatives for objective estimation of different behavioral parameters. They are equipped with sensors, such as accelerometer, screen status (on/off) detector, ambient light, microphone and GPS whose data can be used for inferring behavioral aspects such as sleep behavior [30, 31], activities [32], phone usage [33] and keystrokes dynamics [34]. Methodologies for analyzing data obtained from smartphones and inferring these behavioral aspects have been employed for monitoring and gaining insights in different health conditions such as mental well-being [31, 33, 35, 36], and cognitive decline in elderly [37] and Multiple Sclerosis (MS) patients [34]. The results of these studies show that it is possible to infer different behavioral patterns, such as sleep duration and physical activity, using objectively gathered smartphone data that can be used to gain insights into different health-related conditions. Smartphone usage aspects, such as app use duration and keystroke dynamics, have been linked to stress in college students and cognitive decline in older adults and MS patients. We conclude that passively collected data from smartphone has great potential in inferring different behavioral and environmental aspects previously linked to migraine. To the best of our knowledge, no research has yet analyzed such data in migraine patients.

Methods

In this section we explain the methodology of our study. We start by providing information on our study group and the details of the data collection process in Study group and data collection section. In Headache and non-headache periods section we explain the label cleaning process, and in Data processing section we describe the data processing steps. Finally, in Data analysis section we describe the data analysis methods we have adopted for this study.

Study group and data collection

All the data analyzed comes from migraine patients that participated in the mBrain study [14]. The general goal of the mBrain study was to support both the doctor and the patient in the follow-up of the patient's headache disorder. Patients were monitored for about 90 days in the real-world (i.e. patients led their normal lives) using wrist-worn wearable devices and smartphones. The participants installed two applications on their smartphone. The first application collected smartphone data from motion sensors (accelerometer, gyroscope, magnetometer, gravity, rotation, and linear accelerometer), GPS, ambient light sensor, proximity sensor, as well as data about screen state, application usage, and keystrokes. Data was buffered locally on the smartphone for two minutes before being uploaded to our servers via Wi-Fi. The second application, named mBrain, allowed patients to log daily life events, headaches and medication use. When reporting headache attacks, patients were required to provide additional contextual information about these events, such as medication intake, headache location, symptoms, and triggers. They also indicated their pain intensity on a scale from 1 to 5, with 1 being a headache with no pain and 5 indicating a very severe headache. For more details about the data collection and the smartphone applications we refer the reader to the work of De Brouwer et al. [14].

On the collected data we performed several data and headache record quality checks, and filtered the data accordingly. First, we define valid study days. These are days in which enough data has been streamed, which we consider to be at least 13h of data streamed between 7 a.m. and midnight (so we tolerate 4h, or about 25% of missing data during the day). Second, we check the headache reports from the patients, and define valid headache and non-headache periods. This process is explained in details in Headache and non-headache periods section. Third, we form corresponding pairs of headache and non-headache periods, in order to perform appropriate statistical testing. The process of forming pairs is explained in Headache and non-headache periods section in greater detail. Finally, after all the filtering, we considered only participants that had at least two pairs of headache - non-headache periods. This is to avoid including data sampled by chance (as only one pair may not be representative for the general behavior during headache and non-headache periods). We keep it to at least two pairs, even though more pairs would lead to more representative values, as a trade off against having enough participants in the tests. Our initial study cohort consisted of 29 migraine patients, 22 female and 7 male patients. The average age was 36.9 years (std +/- 12.7 years). Table 1 shows information on how many headaches each participant recorded, how many headaches we considered to be valid after filtering, and how many times a symptom was reported. Finally, this table also reports the number of valid headache - non-headache pairs formed. After applying all the requirements to get valid days and periods, p-002, p-006, p-009–p-014, p-016, p-017, p-024, and p-026 were excluded from the analysis, which is a total of 12 excluded participants. Our final dataset consisted thus of 17 participants, aged 38.9 years on average (std +/- 14.2).

Table 1 Demographics info and additional statistics on recorded headaches and symptoms for all participants (mean) and (std), and for only the included participants (mean final participants) and (std final participants)

	Age	Sex	#Reported Headaches	#Considered Headaches	Symptom Cognition	Symptom Light	Symptom Movement	Other Symptoms	Days considered in study	#Pairs
p-001	29	F	18	16	13	12	12	2	72	10
p-002	35	F	3	2	0	0	0	1	21	1
p-003	29	F	19	14	0	0	0	4	74	4
p-004	59	F	31	28	4	12	0	16	65	26
p-005	35	F	6	5	4	0	0	1	41	3
p-006	21	F	1	1	0	0	0	0	1	0
p-007	26	M	6	6	1	0	0	3	15	3
p-008	29	F	40	36	10	24	5	0	103	9
p-009	38	M	4	3	0	0	0	2	6	1
p-010	41	F	1	1	0	0	0	0	31	0
p-011	38	M	0	0	0	0	0	0	0	0
p-012	56	F	1	1	1	1	0	0	4	0
p-013	36	F	0	0	0	0	0	0	0	0
p-014	24	F	2	2	1	1	0	1	3	0
p-015	21	F	10	10	1	2	0	4	52	5
p-016	22	F	6	5	2	0	1	0	81	1
p-017	42	F	0	0	0	0	0	0	0	0
p-018	53	F	15	15	16	15	13	0	86	15
p-019	40	F	38	38	14	25	6	6	88	29
p-020	27	F	43	39	26	31	23	0	90	14
p-021	44	F	22	22	8	5	3	5	89	15
p-022	20	F	16	13	7	3	4	2	82	5
p-023	38	F	17	15	7	0	4	5	88	9
p-024	23	F	0	0	0	0	0	0	0	0
p-025	34	M	52	50	10	9	1	9	90	5
p-026	33	M	1	1	0	0	0	0	2	0
p-027	59	M	41	39	3	6	1	26	76	12
p-028	61	M	9	9	1	0	0	6	26	2
p-029	58	F	7	7	1	0	0	3	45	4
(mean)	36.9		14.1	13.0	4.5	5.0	2.6	3.3	45.9	6.0
(std)	12.7		15.6	14.7	6.3	8.7	5.2	5.6	37.4	7.7
(mean final participants)	38.9		22.9	21.3	7.4	8.5	4.3	5.4	69.5	10.0
(std final participants)	14.2		14.9	14.1	6.9	10.0	6.3	6.6	25.2	7.9

Participants that were excluded from final analysis are indicated in gray color. Cognition symptom: decline in concentration, memory, and ability to form and/or pronounce words and/or sentences; Light symptom: sensitivity to light; Movement symptom: sensitivity to movement

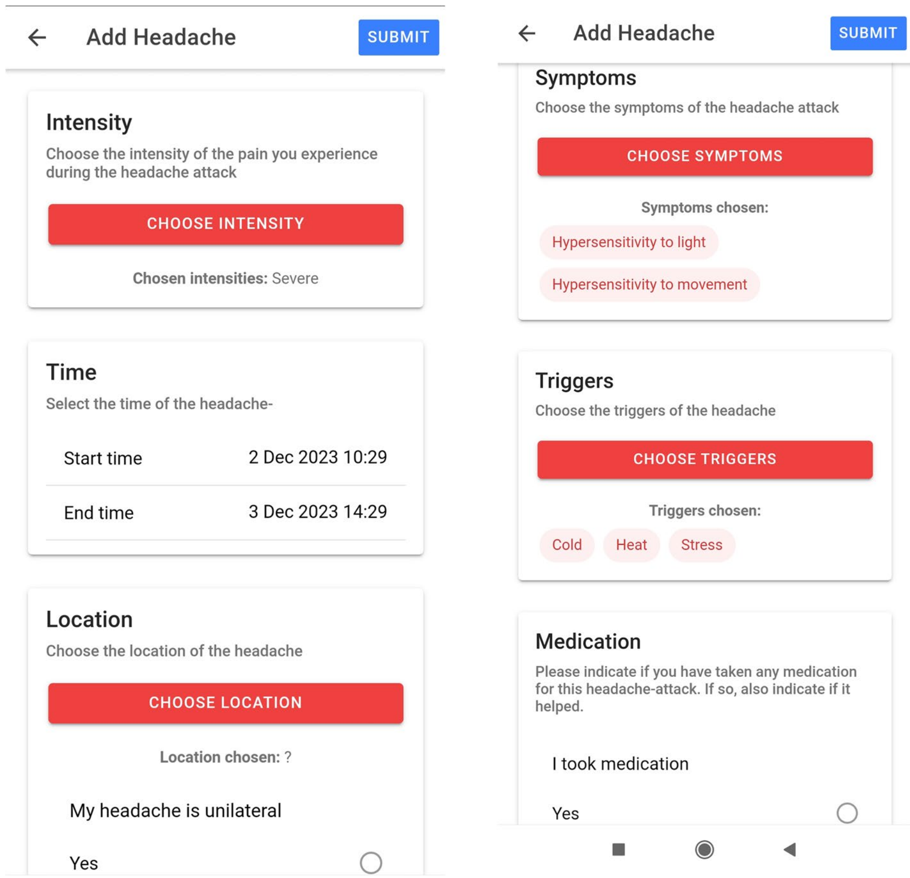

Fig. 1 User interface for adding a headache event

Headache and non-headache periods

Headache events were self-reported by patients through the mBrain (second) application. Besides the start and end time of their headache periods, patients were instructed to provide additional information including location of the pain in the head, intensity, symptoms and possible triggers. Figure 1 shows the user interface for adding a headache event. We considered all reported headaches in this study, as long as there was enough data available for the reported period (at least 75% of the indicated period). Additionally, the 24 hours before the start and after the end of a valid headache period were also considered as uncertain periods, in order to account for prodromal and postdromal periods. Even though prodromes are defined as up to 48 hours before the onset and after the cessation of pain[38], some patients can report symptoms up to 3 days before and after headache, and thus these phases have no strict duration[38]. As a trade-off between not using data in which symptoms possibly already present themselves, but also not throw away too much data, we decided to set this window to 24 hours before onset and after cessation of the headaches. The data originating from these uncertain periods was excluded from the data analysis. All of the other periods, i.e., moments that do not fall into a valid headache period nor into an uncertain period were considered to be non-headache periods.

Finally, we formed headache - non-headache pairs. As we have more non-headache periods than headache periods, we searched, for each valid headache period, one or more corresponding non-headache periods. A corresponding non-headache period has the same start and end moment (hour and minute of the day) as its headache period counterpart. Additionally it is on the same type of day (weekday or weekend) and it is not more than a week before or after the headache period. We also made sure that no non-headache period appears twice (nor an intersection of two non-headache periods) in the list of paired headache - non-headache periods.

For certain behavioral aspects, we group the pairs based on relevant symptoms reported by the patients during their headaches. The symptoms we considered are: decline in concentration, decline in ability to properly pronounce words/sentences, decline in ability to properly form words/sentences, decline in memory (we group these 4 symptoms in a "cognitive decline" group), sensitivity to light, and sensitivity to movement.

Data processing

In this section we describe the data modalities considered in this research. We explain in detail the processing of the data and the derivation of behavioral aspects from these data.

Accelerometer data, activity index, and On-Table detection

The first data source we consider is the accelerometer data from the smartphone. Every smartphone is equipped with a 3-axial accelerometer that measures the acceleration in the three axes, expressed in meters per second squared (m/s^2). Though the range of measurement may differ from phone to phone, the axes of the accelerometer are aligned in the same way in all the smartphones and can be seen in Fig. 2. The accelerometer signal was sampled at 32Hz during the study. From this modality we calculate an activity index [39] and we determine periods during which the phone is laying still on a flat surface, e.g. a table.

The activity index is defined by Bai et al. as the square root of the mean of the variance along the three axes, and it is calculated over windows of 1s [39]. Equation (1) shows the equation used to calculate the activity index (AI).

$$\mathrm{AI} = \sqrt{\frac{1}{3}\left(\sigma_x^2 + \sigma_y^2 + \sigma_z^2\right)} \quad (1)$$

It is predominantly used with on-body accelerometers as an estimate for activity energy expenditure. Using a smartphone's accelerometer, we cannot make strict statements on the energy expenditure of the person, as it can happen that the phone is laying still while the person has gone running. We can, however, estimate how much the phone is being moved around or interacted with, which is also a characteristic of human behavior.

The phone-on-table detection is a threshold based algorithm. For the phone to be considered lying still the acceleration measured in the x-axis should be larger than $9m/s^2$ or smaller than $-9m/s^2$. Additionally, the standard deviation in each axis should be less than $0.5m/s^2$.

The complete processing of the accelerometer data, based on these two algorithms, is as follows. First, we chunk the data in continuous (uninterrupted) chunks and each chunk is windowed in windows of 1s with a stride of 1s. Figure 3a shows a raw accelerometer signal and Figure 3b shows the split of this raw signal in two uninterrupted chunks. Afterwards, for each window we calculate the activity index and whether or not the phone is on the table. Figure 3c shows the windowing in 1s long windows, and the decision of the phone-on-table detection, for 10s of data. All visualisations have been made using the plotly-resampler [40] library. For the phone-on-table detection we have an additional smoothing step, by performing a majority voting on 1min windows with 1min stride.

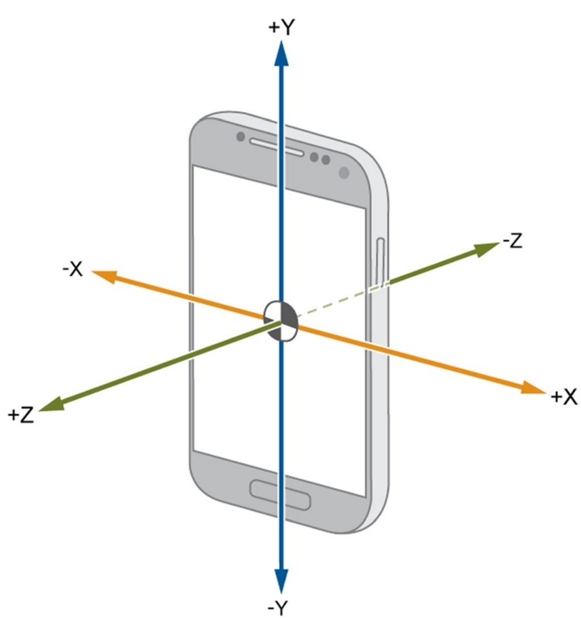

Fig. 2 Accelerometer axes in a smartphone

Keyboard interaction

Another modality we investigate in our research is the keyboard interaction. The data collection application registered the timestamp of each keystroke the person made, more precisely the timestamp every time the content of a typing field changed. The first step was to detect typing sessions, which we defined to be all the keystrokes until there was 5s of inactivity, as we consider 5s to be reasonable time for looking for a character or an emoji during typing within the same typing session. After 5s of inactivity we assume the patients finished the current typing session. This can be due to finishing the chat, waiting for a reply, or other reasons (e.g. taking a photo within the chat). For each session we calculated the mean rate of typing. The typing rate is calculated as the number of keystrokes in a session, divided by the duration of the session. In other words, the typing rate is the number of keystrokes per second.

The aspects regarding keyboard interaction that were tested included: mean typing rate (expressed as keystrokes per second), number of sessions per second (i.e., the total number of sessions within the period of interest, divided by the duration of the period) and total number of keystrokes, normalized over the duration of the period.

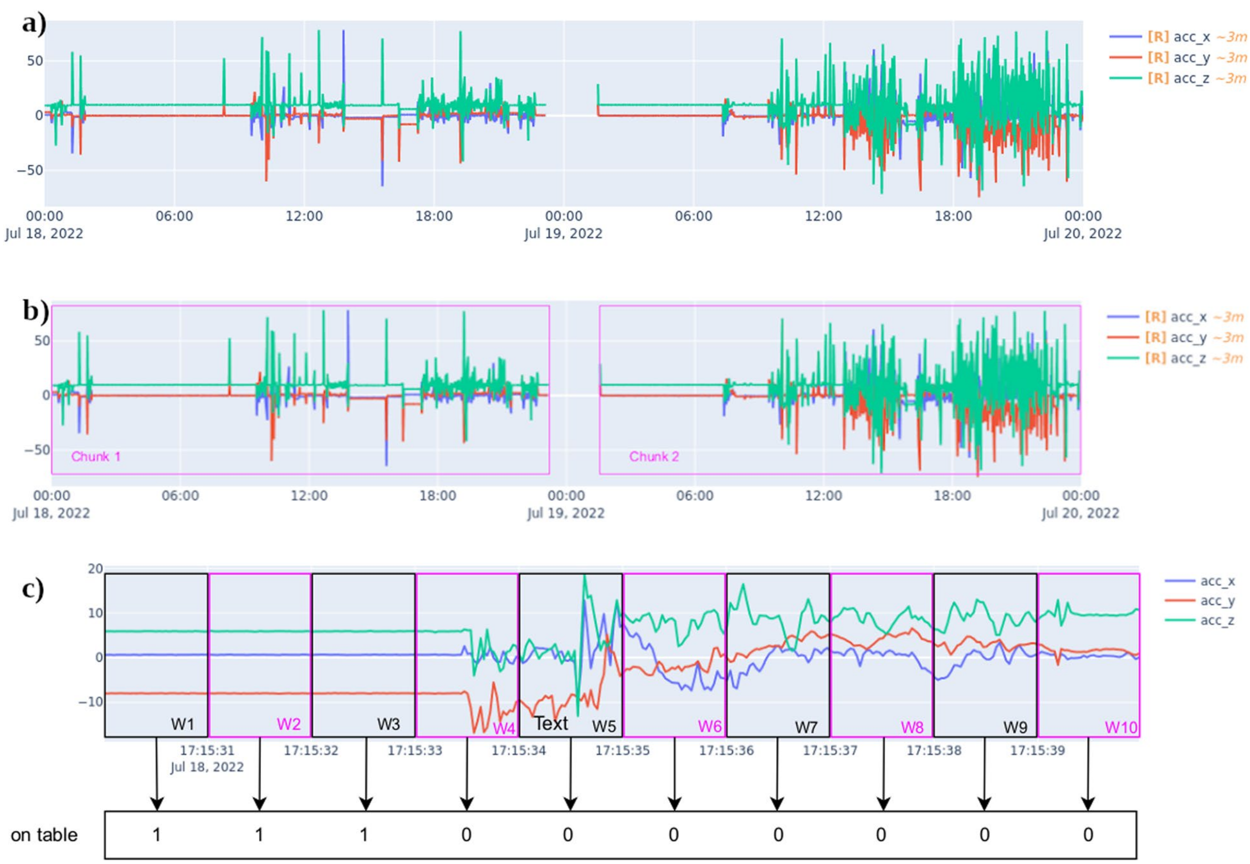

Fig. 3 Top **a** Raw accelerometer data, middle **b** Accelerometer data split in two uninterrupted chunks, and bottom **c** windowed 10s of accelerometer data, and corresponding outcome of the phone-on-table detection

Time spent using smartphone applications

The next behavioral aspect of interest was the time spent using smartphone apps. The data collection application registered every moment an application came to the foreground, accompanied by the package name of this application. Additionally, it registered the screen states, namely Locked, Unlocked, Off and On. The duration of a single application is thus the period between the moment the application came to the foreground and the moment another application came to the foreground or the screen state was changed to Off or Locked.

In order to obtain the effective application usage accurately, we had to address several issues and edge cases. First, we had to consider periods of missing data due to data not being streamed properly from the data collection application to the back-end. We need to ensure that we do not mistakenly assume that during these periods the person did not use their phone. We tackled this issue by leveraging the continuity aspect of the accelerometer data, i.e. accelerometer data is always streamed irregardless of how the phone is being used. Specifically, whenever we had a gap of at least one second in this accelerometer data modality, we considered that during this period no data was streamed at all. As such, it might have been that the person was using an application, but this data was not captured. Moreover, we noticed that the package names of many system apps would be registered even though these processes do not represent significant user interactions, such as displaying a clock on the lock-screen (e.g., *com.google.android.deskclock*) or switching between apps in the launcher (*net.oneplus.launcher*). Additionally, the keyboard is an application on its own, and thus would come to the foreground every time the person is typing within another app. Examples of keyboard package names include *com.android.inputmethod.latin*, and *com.sec.android.inputmethod*. Figure 4 schematically shows the procedure for estimating app usage duration. To do this as precisely as possible, we took the following steps. First, from the original list of application usage entries as logged by the data collection application, we removed the records that include keyboard package names (removing second row from *App opening* list in Figure 4). We then inserted records indicating when the screen state changed to Off or Locked, and the starting

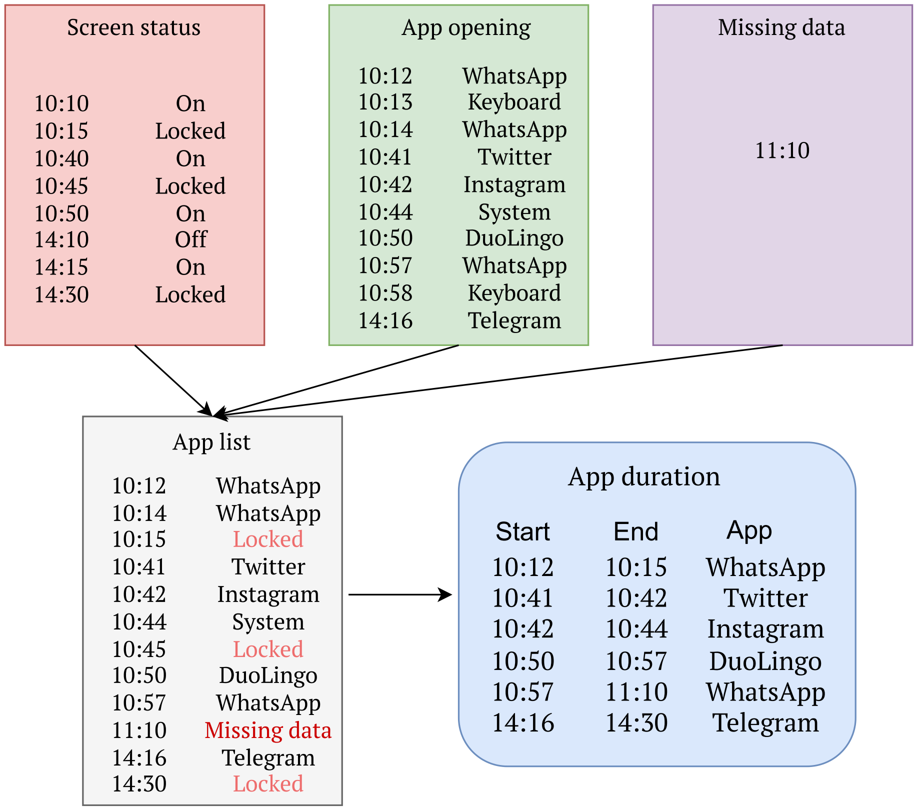

Fig. 4 Procedure for estimating app duration

moments of the periods with missing data (this results in *App list* in Figure 4). From this list we calculate the app usage period as follows: the start time is the timestamp recorded at each entry and the end timestamp is the start of the successive entry, i.e. a successive application, a screen Locked or Off state or a period of missing data. We then remove all Off or Locked screen state entries, system process entries, and missing data periods, as these are all the periods in which the person was not using the smartphone. The step of removing the keyboard entries may result in successive entries with the same application (this situation is shown in the first three rows of *App opening* list in Figure 4). We merge these two entries into one as can be seen in Figure 4. This way, we get the final start and end period for each application usage session.

People tend to use a wide variety of applications that have the same purpose, e.g. playing different types of games, using different mail clients, etc. To have a better understanding of what types of apps patients were using, we categorized all apps in different categories. The list of categories and example apps can be seen in Table 2.

Regarding the app usage behavior we considered the percentage of time spent using an app during a given period (the total duration on the phone divided by the duration of the period of interest). We additionally looked at usage duration of subsets of apps. More specifically, we divided the app categories in two: one category, *leisure* that we consider to contain apps that are used in free time (leisure), such as games, social media, and e-commerce, and another category, *mandatory* that we consider to contain apps that mandate (immediate) interaction, such as banking, chat, email, transport. Which categories are considered leisure or mandatory can be seen in Table 2.

Ambient light

Almost every smartphone is equipped with an ambient light sensor, which measures the amount of ambient light. This illuminance is measured in lux, whose scale shows a logarithmic relationship with human perceived light [41]. Therefore, in order to have interpretable values, we took the logarithm of the raw lux values. When the phone was on the table facing down, or in a pocket, no light reached the sensor, resulting in capturing invalid ambient light estimates.

Initially we incorporated the proximity sensor to filter out these periods. The proximity sensor is a binary sensor that measures whether there is an object close

Table 2 Chosen app categories with app examples, and category type: leisure (L) or mandatory (M)

Category	Type	Example 1	Example 2
System	X	com.motorola.msimsettings	com.samsung.android.net.wifi.wifiguider
Game	L	com.playrix.township	com.ripostegames.shopr
Ecommerce	L	com.mcdonalds.mobileapp	be.bluestores.lolaliza.highstreet.app
Social	L	com.instagram.android	com.zhiliaoapp.musically
Organize	M	com.google.android.apps.walletnfcrel	com.sunnyportal.ui
Health	M	com.polarsteps	io.yuka.android
Listen	L	be.vrt.radioplus.radio1	com.audible.application
Banking	M	com.droid4you.application.wallet	com.saxobank.investor
Domotica	M	com.philips.lighting.hue2	com.tao.wiz
Transport	M	de.hafas.android.sncbnmbs	hr.infoart.epk
Media	L	com.huawei.videoeditor	com.qeexo.smartshot
Watch	L	com.google.android.apps.chromecast.app	com.disney.disneyplus
Camera	L	com.motorola.camera2	com.hantor.CozyMag
Productivity	M	com.google.android.apps.docs	com.microsoft.office.officehubrow
Browser	M	com.android.chrome	com.google.android.googlequicksearchbox
Sports	L	com.tayu.tau.pedometer	com.strava
Call	M	com.android.phone	com.samsung.android.app.telephonyui
Chat	M	com.whatsapp	com.google.android.apps.messaging
Keyboard	X	com.samsung.android.aremoji	com.google.android.inputmethod.latin
Learn	L	com.duolingo	es.aroundpixels.hsk5lite
Email	M	com.android.email	com.samsung.android.email.provider
Read	L	com.innologica.inoreader	com.google.android.apps.books
Crypto	L	com.coinmarketcap.android	com.coinbase.android
Calendar	M	com.samsung.android.calendar	com.google.android.calendar

System and Keyboard apps were not considered (X)

to the smartphone. Both the proximity sensor and the ambient light sensor are placed near the front camera of a regular smartphone. Hence, in case of the proximity sensor indicating that there is an object close to it, we would discard this period as we consider it unreliable for estimating ambient light. However, we empirically observed that very often, even if the proximity sensor would indicate an object close to the smartphone, the ambient light sensor would still capture light. As such we decided to exclude periods for which the light sensor indicates "pitch black" (i.e., lux values lower than 10), assuming these were measured when the sensor was covered. We additionally excluded periods in which the smartphone was lying still on a surface (possibly facing up) as detected from our on-table detection algorithm. This is to account for situations in which the person is not in close proximity to the phone (e.g., different room) or has left the smartphone on the table but covered their eyes to protect themselves from the light.

Estimated sleep

As a proxy for sleep periods we used our phone-on-table detection module. In accordance with the nighttime definition of Böttcher et al. [42], we considered periods from 8 p.m. to 8 a.m., and looked for the longest uninterrupted phone-on-table period within that time period. We considered this period to be the sleeping period of the person.

Data analysis

In this section we describe the data analysis methodology. For each behavioral aspect discussed in the previous sections, i.e. activity index, keyboard interaction, application usage, ambient light & estimated sleep, we performed statistical significance testing of the differences between matched Headache period (HP) and Non-headache period (NHP). The pair making process, for a single participant, is visually shown in Fig. 5. We first explain the general approach and in the following subsections we give the specific steps behavioral aspect

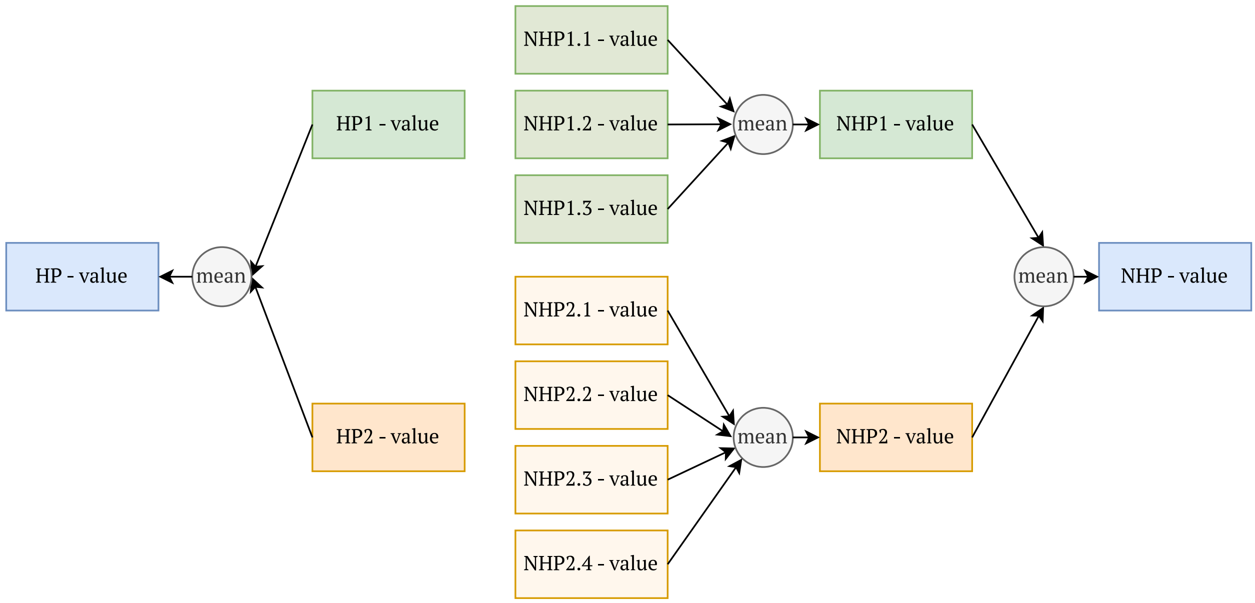

Fig. 5 Calculating paired values approach. Example for one participant with two headache periods (HP1 and HP2), and three and four corresponding non-headache periods (NHP1.1-NHP1.3 and NHP2.1-NHP2.4) respectively. Final headache non-headache value pair from this participant is HP and NHP

we have taken, if any. For each headache period, we took the behavioral aspect values that were observed during that period and calculated the mean. Next, we took the behavioral aspect values that were observed during all the corresponding non-headache periods and we calculated the mean. This resulted in one headache - non-headache behavioral aspect value pair. We did this for all headache periods per person. We then took the mean of all headache values, and the mean of all non-headache values. This way, we obtained one headache - non-headache value pair per participant. We did this for all participants, and then used the Wilcoxon signed-rank test to test whether the difference of the behavioral aspect values observed during headache and non-headache values was symmetric around zero [43].

Keyboard interaction When analyzing the keyboard interaction behavior we considered only headache periods which were accompanied by a cognitive decline symptom, as explained in Headache and non-headache periods section.

App usage duration For this modality we did not filter on any symptom and considered all of the headache - non-headache pairs.

Ambient light We filtered the headache - non-headache pairs based on a self-reported sensitivity to light symptom.

Estimated sleep duration For this modality we took a slightly different approach than for all other modalities. To test whether there is difference in the sleep on the night preceding a headache/non-headache period, we considered all the nights preceding a valid headache period, and all the nights preceding a valid non-headache period. In other words, we did not take the corresponding pairs, since we are interested in the nights rather than the periods themselves.

Activity index For the Activity index, we considered only headache - non-headache pairs for which there was a self-reported sensitivity to movement symptom for the headache period.

Results

In this section we report the results from the statistical tests from moderate to severe headaches (pain level 3 or higher) for each behavioral aspect. We provide visualisations on the pairs included in the tests. A summary of the results and additional info, such as number of participants, number of pairs, and considered symptom can be

Table 3 Overview of the performed tests

Behavior aspect tested	Relevant symptom	Weekday	Weekend	#Subj.	#Pairs	*p*-value
Typing rate	Cogn. Decl.	✓	✓	12	68	0.8833
		✓		11	48	0.7934
			✓	7	16	0.7109
Number of typing sessions	Cogn. Decl.	✓	✓	12	68	0.0646
		✓		11	48	0.0610
			✓	7	16	0.0390
Number of keystrokes per hour	Cogn. Decl.	✓	✓	12	68	0.0461
		✓		11	48	0.1030
			✓	7	16	0.0156
Time spent on leisure apps	None	✓	✓	16	164	0.1372
		✓		16	128	0.0106
			✓	10	36	0.0322
Time spent on productivity apps	None	✓	✓	17	168	0.3559
		✓		16	126	0.5100
			✓	10	36	0.1376
Ambient Light	Light sensitivity	✓	✓	8	59	0.0273
Estimated sleep duration	None	✓	✓	16	201	0.3528
Activity Index	Movement sensitivity	✓	✓	8	42	0.0546
		✓		6	28	0.2187
			✓	2	8	/

Relevant symptom column states which symptoms were reported in the headaches included in the test. The check marks in Weekday and Weekend columns indicate whether the headaches included in the test took place on a weekday, weekend, or all headaches (both checkmarked). #Subj. indicates the number of subjects (and number of final sample points) in the test. #Pairs indicates the total number of pairs that were used in the test

found in Table 3. The results of the tests from also mild headaches can be found in Results for different headache intensity section.

Keyboard interaction After applying the symptom filter on cognitive decline and splitting the observations in weekday or weekend, we had 48 paired observations from 11 patients during weekdays and 16 paired observations from 7 patients during weekends. From the 362 headaches, in 129 this symptom was reported indicating that people experience cognitive decline in about one third of their headache attacks on average. We did not observe a difference during headache - non-headache periods that took place on a weekday for any of the keyboard interaction aspects we tested: typing rate, number of sessions per second, and total number of keystrokes per hour. The top subfigure of Figure 6 shows the difference between the mean values we have obtained during

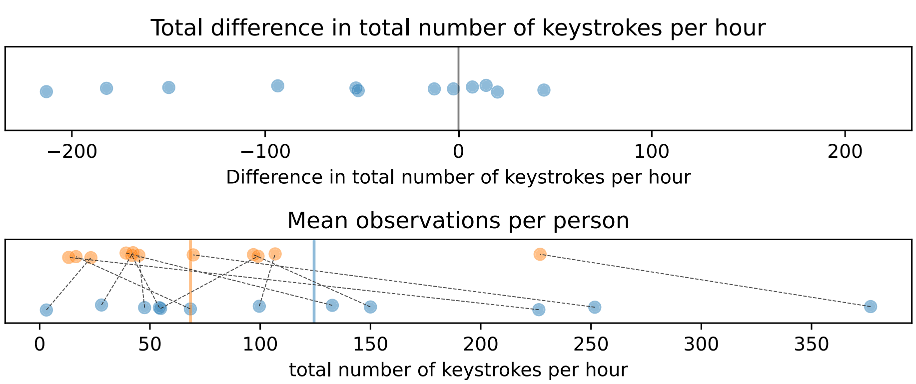

Fig. 6 Top: difference between paired headache vs non-headache observed values per participant for the keystrokes per hour, regardless of the type of day. (small jitter has been added in the vertical axis to better visualize clusters of (overlapping) values); Bottom: observed mean values of the keystrokes per hour during headache and non-headache periods during weekends per participant. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

Table 4 Mean, median and standard deviation of headache and non-headache paired total number of keystrokes per hour values

	mean	median	std
headache	68.37	43.65	59.49
non-headache	124.44	83.96	110.55

Table 5 Mean, median and standard deviation of headache and non-headache paired percentage of time spent on leisure apps values

	mean	median	std
headache	0.043	0.027	0.04
non-headache	0.048	0.04	0.03

headache periods and non-headache periods. The bottom subfigure of Figure 6 shows the mean values per specific period for each participant. As can be derived from these figures, for periods that took place in the weekend we observed some difference for the number of typing sessions and the total number of keystrokes. The mean, median and standard deviation of the headache and non-headache paired values are presented in Table 4. In the first three rows of Table 3, information such as number of subjects, pairs and *p*-value regarding these tests can be found. The visualizations of the other tests can be found in Visualization of additional tested aspects section.

App usage duration We did not observe a difference in the use of "mandatory" apps, regardless of the type of day. We observed however, a difference in the use of "leisure" apps, in weekdays and weekends, separately. In these tests we had 164 paired observations from 16 participants. The mean, median and standard deviation of the headache and non-headache paired values are presented in Table 5. Figure 7 shows visually the difference and the mean observations of the pairs in the test for the "leisure" apps, from which we can see that the headache values are smaller than the non-headache values for the majority of the participants. The fourth row of Table 3 provides the number of subjects, pairs and *p*-value of the performed tests. The visualizations of the other tests can be found in Visualization of additional tested aspects section.

Ambient light After applying the symptom filter "sensitivity to light", we had 59 paired observations from 8 patients. The mean, median and standard deviation of the headache and non-headache paired values are presented in Table 6. We observed a difference in the ambient light observed during headache and non-headache periods, regardless the day of the week. Figure 8 shows visually the difference between the illuminance values, and the mean observed values for each period respectively. As can be seen, the difference between the headache and non-headache values is negative for all participants but two. The fifth row of Table 3 provides the number of subjects, pairs and *p*-value of the performed tests.

Estimated sleep duration For this test we had observations from 16 patients. We did not observe difference in the estimated sleep duration regardless of the type of day. The mean, median and standard deviation of the headache and non-headache paired values are presented in Table 7. The top subfigure of Figure 9 shows the difference between the mean sleep duration estimates we have obtained during headache periods and non-headache periods. The bottom subfigure of Figure 9 shows the mean sleep duration per specific period for each

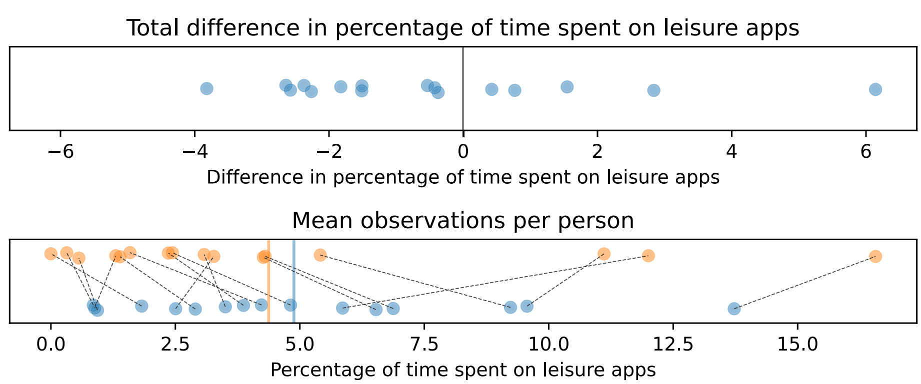

Fig. 7 Top: difference between paired headache vs non-headache observed values per participant for the time spent on leisure apps regardless of the type of day; (small jitter has been added in the vertical axis to better visualize clusters of (overlapping) values); Bottom: observed mean values of time spent in leisure apps during headache and non-headache periods per participant. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

Table 6 Mean, median and standard deviation of headache and non-headache paired light values

	mean	median	std
headache	3.99	3.98	0.86
non-headache	4.22	4.22	0.85

participant. The sixth row of Table 3 provides the number of subjects, pairs and *p*-value of the performed tests.

Activity index After applying the symptom filter "sensitivity to movement" we had 42 paired observations from 8 patients. We did not observe difference in the phone's activity index, regardless of the type of day. The mean, median and standard deviation of the headache and non-headache paired values are presented in Table 8. The top subfigure of Fig. 10 shows the difference between the mean activity index we have obtained during headache periods and non-headache periods. The bottom subfigure of Fig. 10 shows the mean values per specific period for each participant. In the last row of Table 3 the number of subjects, pairs and *p*-value of the performed tests can be found.

Discussion

The main goal of our research was to look for behavioral changes in migraine patients during headache periods. In Related work section, we discussed that symptoms, such as light sensitivity, lower sleep quality, and cognitive decline are often reported by migraine patients. Additionally, light and lack of sleep, among others, have repeatedly been reported as headache triggers. There is however little understanding on these, often, bidirectional relationships. Much research has focused on bringing insights, but the studies are mostly cross-sectional and rely on self-reports or questionnaires, laboratory data collection, or rigorously defined tests. The findings of these studies can, as a result, not easily be projected to the real-world situations. There is thus a need for research that is based on real-world and objective data.

Smartphones have been employed for collecting real-world data in longitudinal studies. They are unobtrusive, affordable, and people use them all the time in their daily lives. These devices can collect diverse data: movement (e.g., accelerometer), contextual information (e.g., GPS), as well as phone interaction. This data offers the possibility to monitor people's behavior. Several studies have already used particularly smartphone data to monitor people's behavior and detect changes and/or differences thereof. These studies focused on people with stress, anxiety and cognitive decline in elderly people and MS patients. In our work we explored the possibility of using real-world smartphone data to observe behavioral changes in migraine patients. We processed and analyzed data collected in the real-world from migraine patients for a period of up to 90 days. The insights we gained allow us to answer the questions formulated in Background section.

Discussion of the research questions

Can smartphone data be used for real-world monitoring of behavioral aspects in the migraine context, and if so which ones? In this work we developed a methodology for processing raw smartphone data to obtain and monitor behavrior aspects, such as keyboard interaction, app usage, environmental context (i.e., ambient light), activity index, and sleep duration. All of these translate to different aspects regarding migraine: symptom and trigger detection or behavior change. For example, monitoring keyboard interaction could help detect changes in

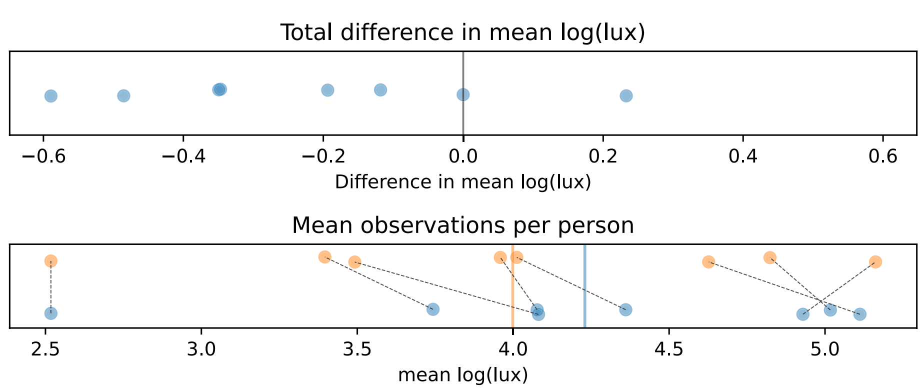

Fig. 8 Top: difference in illuminance between paired headache vs non-headache observed values per participant regardless of the type of day. (small jitter has been added in the vertical axis to better visualize clusters of (overlapping) values); Bottom: observed mean values of ambient light during headache and non-headache periods per participant. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

Table 7 Mean, median and standard deviation of headache and non-headache paired estimated hours of sleep values

	mean	median	std
headache	6.52	7.13	1.82
non-headache	6.70	7.11	1.58

typing speed, number of typing sessions, and total number of keystrokes, which may be an indication of change in cognition capabilities which has ofteen been reported as a migraine symptom. App usage duration could give insights on what type of apps are used less during headache attacks, and whether or not patients limit their phone usage. Ambient light monitoring may be a way of detecting photophobia or light sensitivity as a symptom. Activity index can translate to amount of movement and help monitor for increase or decline of physical activity. Finally, sleep duration estimation is one step towards modelling sleep behaviors, changes in sleeping trend, detecting sleep interruption or lack of sleep, all of which are often related to migraine attacks.

What data processing steps are needed to obtain usable and interpretable information from this smartphone data linked to migraine symptoms and triggers? To infer each of the behavioral aspects, one or more raw data modalities should be first adequately preprocessed. Afterwards they can be further processed to give them meaning. For example, from the raw stream of app coming to foreground data and the screen status, we could find periods of each app usage and calculate the duration. Additionally, categorizing the apps leads to additional insight on the person's behavior. The keyboard strokes data was first split into typing sessions in order to obtain reliable typing speed. To obtain reliable ambient light values, we incorporated the accelerometer data to exclude periods in which the smartphone was on the table, as we consider those moments to be unreliable for determining the ambient light surrounding the person. In a similar way we used the accelerometer data to estimate the sleep period during the night. We found that translating the raw data into meaningful behavioral aspects is not straightforward. It requires detailed insights of each modality, correct and well-thought through processing, and visual validation of processing steps.

What behavioral changes do migraine patients show between headache periods (ictal) and headache-free periods (interictal) attacks? In our research we observed several behavioral changes. Regarding the keyboard interaction, we observed a lower number of keystrokes during headache periods compared to non-headache periods, but only in the weekends. We did not observe a difference in the typing speed nor the number of sessions. This indicates that the typing sessions contain less keystrokes during weekend headache periods.

We also observed that patients, during headaches, lower the time of using apps that we consider to be free-time apps, regardless of the type of day (week or weekend). A similar change is however not observed in the usage of apps that we consider to be mandatory. This may indicate that patients try to decrease the phone usage time during headache, if possible. However, it seems that, when necessary, they maintain their non-leisure smartphone usage as usual. This may reflect outside of the digital world as well, that patients reduce their free-time activities, but maintain performing as normally as possible in other situations, such as work, or family obligations.

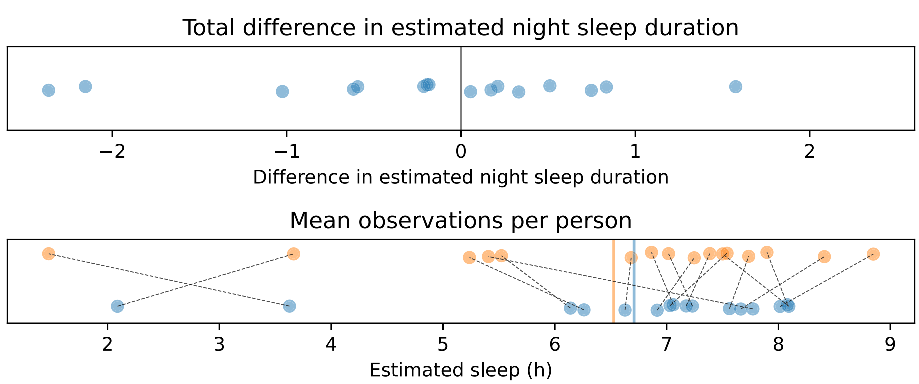

Fig. 9 Top: difference in sleep duration between paired headache vs non-headache observed values per participant regardless of the type of day. (small jitter has been added in the vertical axis to better visualize clusters of (overlapping) values); Bottom: observed sleep duration values during headache and non-headache periods per participant. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

Table 8 Mean, median and standard deviation of headache and non-headache paired activity index values

	mean	median	std
headache	0.19	0.16	0.13
non-headache	0.23	0.16	0.15

Our results show further that patients with light sensitivity stay in darker environments during their headache periods in the weekend, but this difference is not to be observed during the weekdays. Similarly to the app usage, this may indicate that patients are often in situations in which they cannot adapt their environment according to their needs, e.g., when they are at work.

We further observed no difference in the physical activity measured through the smartphone, nor the estimated sleep duration, regardless of the type of day.

What are the limitations of this approach and how can they be accounted for? There are several limitations in our approach that we would like to discuss. First, we recorded only the timestamps of the keystrokes, which allowed us to calculate the typing speed. There are however other aspects of keyboard interaction that may be useful and worth researching. For example, the type of stroke (alpha-numeric, emoji, auto-correct/auto-complete or backspace) may reveal additional insights, such as how often people make mistakes typing, whether they type full words, or how often they use auto-complete. Additionally, we did not consider the app in which they were typing. With that additional info, we could detect whether the person is chatting, sending email, taking notes, or something else.

Another limitation is the sensitivity of the ambient light sensor. The values that this sensor produces depend a lot on how directly the light falls on the sensor. Given the same environment, if the phone sensor is turned towards the light source, it will yield higher values compared to if the light comes from more indirect angles relative to the ambient light sensor. There is, however, no difference in the brightness of the room in general. Additionally, the phone may not be in direct vicinity of the patient and therefore produce values that are not representative for the light in which the patient is. To account for this we excluded periods on which the phone is lying on the table.

Using the activity index calculated from the smartphone accelerometer does not necessarily correctly reflect the amount of physical activity the person is performing. The phone may be in the person's backpack or on the table while they could be out for a run or laying down in the sofa. Wearable devices, such as activity trackers, are more suitable for reliably estimating the person's movement. In this study we focused on only smartphone data as this device is ubiquitous and most people already own one.

Finally, our approach for estimating sleep is limited to the longest uninterrupted phone on table period. We do not account for interruptions (checking the time or using the phone in the middle of the night) and going back to sleep. We additionally do not detect awake periods in which the person does not interact with the phone. Additional smartphone sensors, such as microphone, may be helpful for detecting such situations. Alternatively, similarly to the activitiy index, wearable devices may be

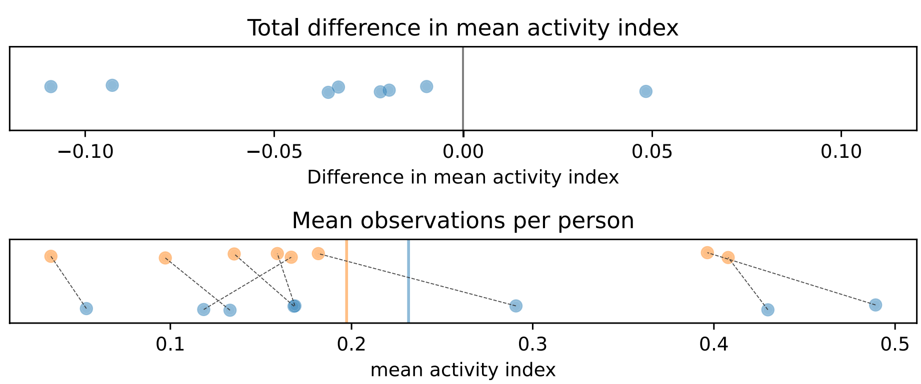

Fig. 10 Top: difference in activity index between paired headache vs non-headache observed values per participant regardless the type of day. (small jitter has been added in the vertical axis to better visualize clusters of (overlapping) values); Bottom: observed activity index values during headache and non-headache periods per participant. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

useful to more precisely detect sleep periods and sleep interruptions.

General limitations and additional discussion

Working with real-world data comes with certain challenges. At the end of the study, patients had registered 558 headaches in total. After accounting for possible errors in reporting (i.e., “recall" and “predictive" bias as explained in Headache and non-headache periods section) and missing data (due to technical errors or decision of the patient to not stream data) there were 376 valid headaches left. Of those, 243 were paired with corresponding non-headache periods. This shows that a real-world study as this requires stable and reliable streaming infrastructure. Additionally, mechanisms should be implemented to support and motivate the patients to comply to the study and continuously collect data and correctly report their headache periods. Morning and evening questionnaires provide an opportunity to query the patient and ask them about their recent entries. There should be a good balance though, as querying too often may be cumbersome for the patient and work counter-productively.

In our study we considered all headache periods that we classified as valid, regardless of whether the patients took medication and whether the medication had an effect. Headache periods that were not treated or when the medication did not work might be more disabling and hence have greater impact on the person’s behavior.

In this work we tested different aspects of behavioral change during headache and non-headache periods. This increases the chance of making at least one type-I error, and thus finding a significant difference for at least one aspect. Therefore, if the significance level is corrected for the repeated testing, using the Bonferroni correction method, we cannot reject any null-hypothesis. Additionally, the sample size of our (sub)populations is small, which further limits the statistical power of our test results. However, this is an exploratory study and the results should be regarded as directions for further research rather than conclusive findings.

Conclusions

In this work we explored the possibility for longitudinal, real-world and objective behavior monitoring of migraine patients through smartphones. We identified behavioral aspects of interest in the migraine context and we proposed a methodology for monitoring them. These behavioral aspects include: keyboard interaction, app usage, ambient light, activity index, and sleep duration. We described how the raw smartphone data sources can be processed to obtain meaningful behavioral aspects. Our results indicate that there is a behavioral change in aspects such as app usage, keyboard interaction, and ambient light.

Future work

Future work should consider working with a larger sample size by recruiting more patients. As we have learned in this study, some patients will drop-out of the study, others will provide less reliable labels, so having a greater sample size will increase the power of the statistical analysis and conclusions made from this study.

Future work could incorporate more specific data from certain sources. For example, as discussed in Discussion section, instead of only looking at the timestamp at which a keystroke happened, one could look to the type of keystroke. We see potential improvement in estimating patients’ physical movement and sleep duration by incorporating wearable devices since they are attached to the person’s body. Additional behavioral aspects can be analyzed if more data modalities are included. For example, eating behavior can be monitored by asking people to take pictures of their meals. This should of course be considered keeping in mind privacy concerns and additional user burden. Furthermore, GPS data holds great potential for providing relevant context information. The type of location can influence people’s behavior and thus it can be interesting to look for changes in behavior given the location they are at.

Our study focused on finding difference of behavior during headache periods versus non-headache periods. Since our results show that there are certain changes, future work may focus on methodologies for detecting migraine periods based on smartphone data. Additionally, looking for behavioral changes preceding headache periods might be the next step towards headache prediction. Our study included primarily white-collar (desk job) workers. Future studies should consider including a more diverse population including blue-collar, night-shift workers, and students, since their daily activities and surroundings greatly differ from the white-collar workers. Conclusions may differ based on different sub-populations.

Ethics and privacy

Monitoring human behavior raises certain ethical and privacy concerns. The data collected in these studies can be sensitive, and the granularity at which it is collected should be kept to a minimum. For example, even though as previously discussed, the type of they keystroke can be more informative than just the timestamp, it is also much more sensitive and intrusive. Studies should explore methodologies that use the minimum required and

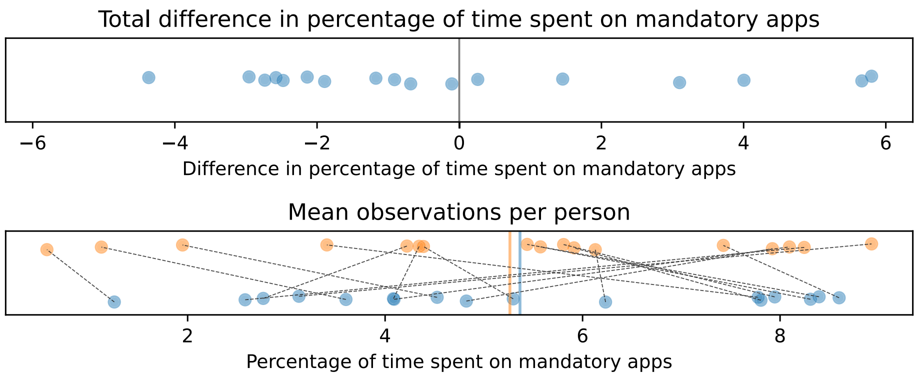

Fig. 11 Top: difference between paired headache vs non-headache observed values per participant for the time spent on mandatory apps regardless of the type of day, and bottom: observed mean values of time spent in mandatory apps during headache and non-headache periods per participant. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

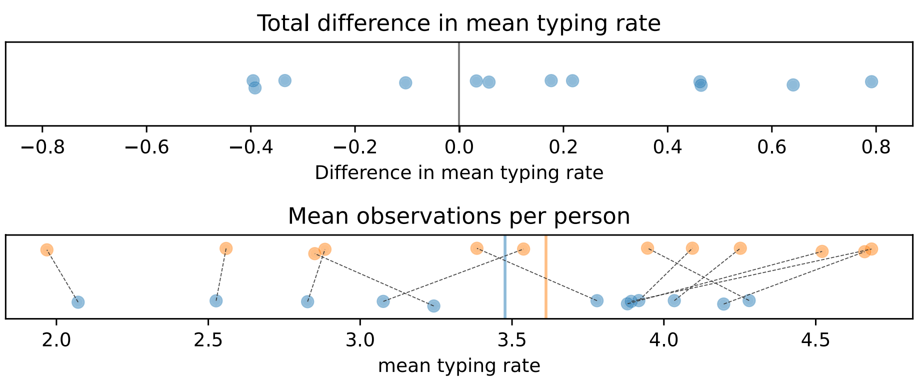

Fig. 12 Top: difference between paired headache vs non-headache observed values per participant for the typing speed regardless of the type of day; Bottom: observed mean values of typing speed during headache and non-headache periods per participant. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

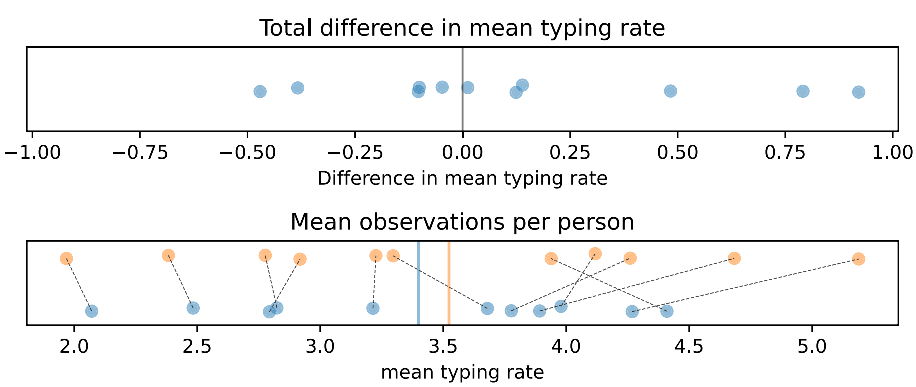

Fig. 13 Top: difference between paired headache vs non-headache observed values per participant for the typing speed during weekdays; Bottom: observed mean values of typing speed during headache and non-headache periods per participant. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

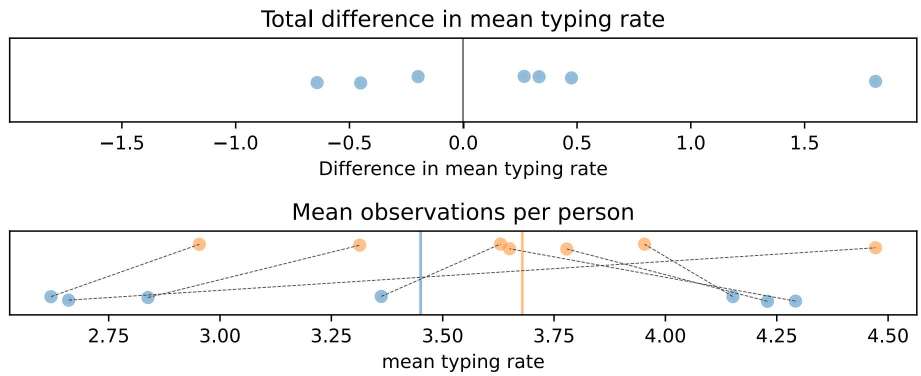

Fig. 14 Top: difference between paired headache vs non-headache observed values per participant for the typing speed during weekend; Bottom: observed mean values of typing speed during headache and non-headache periods per participant. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

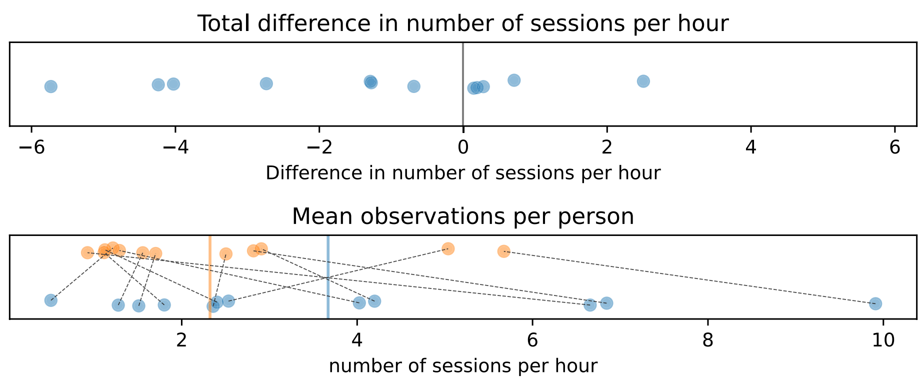

Fig. 15 Top: Difference between paired headache vs non-headache observed values per participant for number of sessions per hour regardless of the type of day; Bottom: observed mean values of number of sessions per hour during headache and non-headache periods per participant, both regardless of the type of day. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

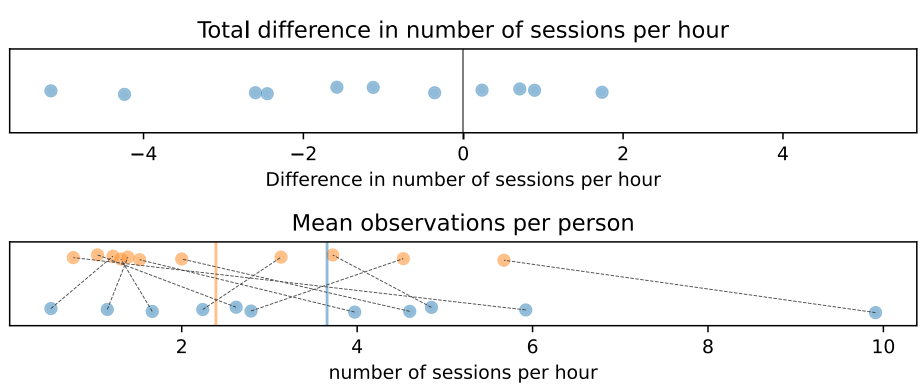

Fig. 16 Top: difference between paired headache vs non-headache observed values per participant for number of sessions per hour during weekday; Bottom: observed mean values of number of sessions per hour during headache and non-headache periods per participant, both during weekdays. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

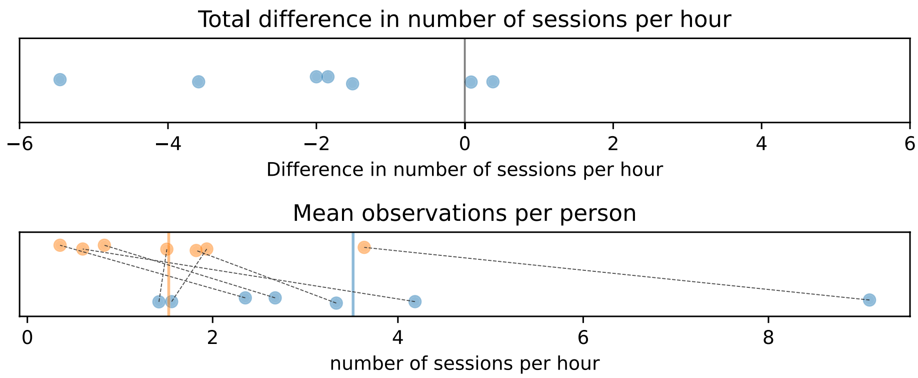

Fig. 17 Top: difference between paired headache vs non-headache observed values per participant for number of sessions per hour during weekends; Bottom: observed mean values of number of sessions per hour during headache and non-headache periods per participant, both during weekends. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

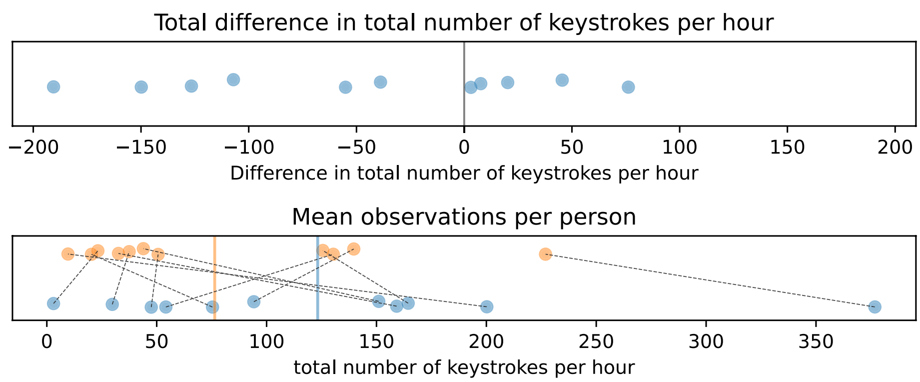

Fig. 18 Top: difference between paired headache vs non-headache observed values per participant for number of keystrokes per hour during weekdays; Bottom: observed mean values of number of keystrokes per hour during headache and non-headache periods per participant, both regardless of the type of day. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

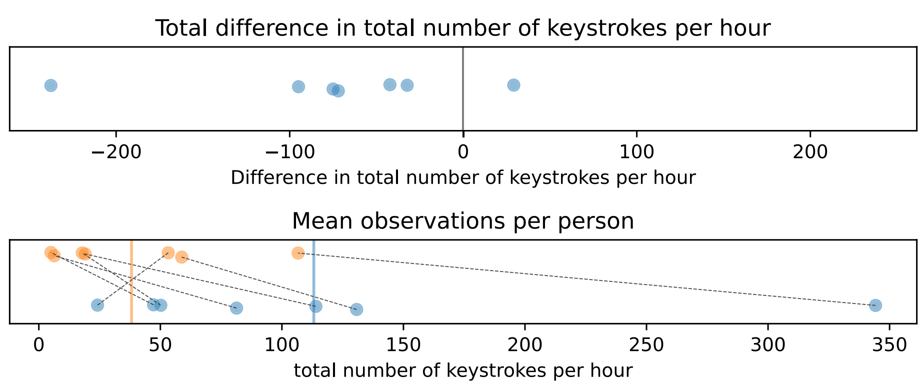

Fig. 19 Top: difference between paired headache vs non-headache observed values per participant for number of keystrokes per hour during weekends; Bottom: observed mean values of number of keystrokes per hour during headache and non-headache periods per participant, both during weekdays. Corresponding pairs are connected with dashed line. Vertical lines indicate the median of the values for each period

Table 9 Keyboard interaction

Behavior aspect tested	Weekday	Weekend	Minimum headache intensity	# Subj.	# Pairs	*p*-value
Typing rate	✓	✓	1	13	80	0.952881
Number of typing sessions	✓	✓	1	13	80	0.227417
Number of keystrokes per hour	✓	✓	1	13	80	0.136719
Typing rate	✓	✓	2	13	80	0.952881
Number of typing sessions	✓	✓	2	13	80	0.227417
Number of keystrokes per hour	✓	✓	2	13	80	0.136719
Typing rate	✓	✓	3	12	68	0.883301
Number of typing sessions	✓	✓	3	12	68	0.064697
Number of keystrokes per hour	✓	✓	3	12	68	0.046143
Typing rate	✓	✓	4	8	30	0.472656
Number of typing sessions	✓	✓	4	8	30	0.125000
Number of keystrokes per hour	✓	✓	4	8	30	0.097656
Typing rate	✓		1	13	60	0.945068
Number of typing sessions	✓		1	13	60	0.207153
Number of keystrokes per hour	✓		1	13	60	0.248657
Typing rate	✓		2	13	60	0.945068
Number of typing sessions	✓		2	13	60	0.207153
Number of keystrokes per hour	✓		2	13	60	0.248657
Typing rate	✓		3	11	48	0.793457
Number of typing sessions	✓		3	11	48	0.061523
Number of keystrokes per hour	✓		3	11	48	0.103027
Typing rate	✓		4	7	22	0.343750
Number of typing sessions	✓		4	7	22	0.343750
Number of keystrokes per hour	✓		4	7	22	0.289062
Typing rate		✓	1	8	18	0.769531
Number of typing sessions		✓	1	8	18	0.019531
Number of keystrokes per hour		✓	1	8	18	0.007812
Typing rate		✓	2	8	18	0.769531
Number of typing sessions		✓	2	8	18	0.019531
Number of keystrokes per hour		✓	2	8	18	0.007812
Typing rate		✓	3	7	16	0.710938
Number of typing sessions		✓	3	7	16	0.039062
Number of keystrokes per hour		✓	3	7	16	0.015625
Typing rate		✓	4	2	5	1.000000
Number of typing sessions		✓	4	2	5	0.500000
Number of keystrokes per hour		✓	4	2	5	0.500000

Overview of the test results for all behavior aspects regarding keyboard interaction

least sensitive data. If sensitive data, such as GPS, audio recordings, or type of keystrokes, is absolutely required, edge computing can be employed. Even for less sensitive data edge computing can help reduce the granularity of data. For example, the time spent using an application can be calculated on the smartphone itself, and only the duration together with the category of the application can be collected. This way the (raw) data never leaves the user's smartphone which resolves privacy risks.

List of keyboard applications that we filter out when calculating the using phone feature

- com.riffsy.FBMGIFApp
- com.google.android.inputmethod.latin
- com.samsung.android.honeyboard
- com.touchtype.swiftkey
- com.sec.android.inputmethod
- com.android.inputmethod.latin
- com.samsung.android.aremoji

Table 10 Time spent using smartphone apps

Behavior aspect tested	Weekday	Weekend	Minimum headache intensity	# Subj.	# Pairs	*p*-value
Time spent using leisure apps	✓	✓	1	19	221	0.146749
Time spent using productivity apps	✓	✓	1	19	226	0.399126
Time spent using leisure apps	✓	✓	2	19	221	0.146749
Time spent using productivity apps	✓	✓	2	19	226	0.399126
Time spent using leisure apps	✓	✓	3	16	164	0.137222
Time spent using productivity apps	✓	✓	3	17	168	0.355949
Time spent using leisure apps	✓	✓	4	10	56	0.116211
Time spent using productivity apps	✓	✓	4	10	58	0.312500
Time spent using leisure apps	✓		1	17	165	0.054443
Time spent using productivity apps	✓		1	17	170	0.555016
Time spent using leisure apps	✓		2	17	165	0.054443
Time spent using productivity apps	✓		2	17	170	0.555016
Time spent using leisure apps	✓		3	16	123	0.010696
Time spent using productivity apps	✓		3	16	126	0.510025
Time spent using leisure apps	✓		4	10	43	0.347656
Time spent using productivity apps	✓		4	10	45	0.577148
Time spent using leisure apps		✓	1	13	49	0.034058
Time spent using productivity apps		✓	1	13	49	0.317749
Time spent using leisure apps		✓	2	13	49	0.034058
Time spent using productivity apps		✓	2	13	49	0.317749
Time spent using leisure apps		✓	3	10	36	0.032227
Time spent using productivity apps		✓	3	10	36	0.137695
Time spent using leisure apps		✓	4	4	11	0.125000
Time spent using productivity apps		✓	4	4	11	0.312500

Overview of the test results for app usage duration

Table 11 Ambient light

Weekday	Weekend	Minimum headache intensity	# Subj.	# Pairs	*p*-value
✓	✓	1	9	70	0.326172
✓	✓	2	9	70	0.326172
✓	✓	3	8	59	0.027344
✓	✓	4	7	30	0.078125
✓		1	9	56	0.410156
✓		2	9	56	0.410156
✓		3	8	47	0.054688
✓		4	6	22	0.156250
	✓	1	5	13	0.593750
	✓	2	5	13	0.593750
	✓	3	5	12	0.593750
	✓	4	3	5	0.750000

Overview of the test results for ambient light

Table 12 Estimated sleep duration

Weekday	Weekend	Minimum headache intensity	# Subj.	# Pairs	*p*-value
✓	✓	1	16	256	0.469940
✓	✓	2	16	256	0.469940
✓	✓	3	16	201	0.352859
✓	✓	4	12	79	0.259277
✓		1	16	184	0.449966
✓		2	16	184	0.449966
✓		3	16	144	0.352859
✓		4	12	57	0.809814
	✓	1	12	72	0.633301
	✓	2	12	72	0.633301
	✓	3	12	57	0.574951
	✓	4	8	22	0.125000

Overview of the test results for estimated sleep duration

Table 13 Activity index

Weekday	Weekend	Minimum headache intensity	# Subj.	# Pairs	*p*-value
✓	✓	1	9	47	0.037109
✓	✓	2	9	47	0.037109
✓	✓	3	8	42	0.054688
✓	✓	4	5	21	0.156250
✓		1	6	31	0.218750
✓		2	6	31	0.218750
✓		3	6	28	0.218750
✓		4	4	12	0.187500
	✓	1	2	8	0.500000
	✓	2	2	8	0.500000
	✓	3	2	8	0.500000
	✓	4	2	6	0.750000

Overview of the test results for activity index

- com.samsung.android.aremojieditor
- com.samsung.android.samsungpassautofill
- com.samsung.android.mdx.quickboard

Visualization of additional tested aspects

App usage Figure 11a shows the difference between the mean time spent on mandatory apps during headache periods and non-headache periods. Figure 11b shows the mean time spent on mandatory apps per specific period for each participant.

Keyboard interaction Figures 12, 13, and 14 show visually the observed differences in typing speed regardless of the type of day, during the weekend and during the weekdays respectively.

Figures 15, 16, and 17 show visually the observed differences in number of typing sessions regardless of the type of day, during the weekdays and during the weekend respectively.

Figures 18, and 19, show visually the observed differences in number of keystrokes per hour during the weekdays and weekends respectively.

Results for different headache intensity

In this section we present the results of the tests using headaches with different minimum headache intensity (Tables 9, 10, 11, 12, and 13).

Abbreviations

HP	Headache period
ICHD-3	International classification of headache disorders, 3rd edition
MS	Multiple Sclerosis
NHP	Non-headache period
PSQI	Pittsburgh sleep quality index

Acknowledgements

The authors would like to thank all study subjects for their collaboration and efforts in the mBrain study.

Authors' contributions

All authors were involved in the methodology design for the data collection. MDB, JVDD & MS performed the technical implementation and deployment of the data collection set-up. MS implemented the data collection from smartphone behavior. NV performed the patient pre-evaluation, follow-up and patient post-evaluation. MS performed the data analysis, in close collaboration with the medical experts NV & KP. FO & SVH supervised the data collection & analysis. MS prepared the first draft of the manuscript, which was thoroughly reviewed by all authors. All authors read and approved the final manuscript.

Funding

This work was partially funded by imec via the AAA Context-aware Health Monitoring project. Nicolas Vandenbussche received funding from Ghent University Hospital for his research (Fonds voor Innovatie en Klinisch Onderzoek, 2019). Jonas Van Der Donckt (1S56322N) is funded by a strategic base research grant of Fonds Wetenschappelijk Onderzoek (FWO), Belgium. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Data availability

The datasets generated and/or analysed during the current study are not publicly available due to protection of participants' privacy but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This study was approved by the Committee for Medical Ethics of the Ghent University Hospital (internal ID BC-10031, named mBrain-21, approved June 7th 2021). Patients were fully informed on all the aspects of the study (duration, procedures, study visit, etc.) and gave written informed consent at the beginning of the study. Participants received a pseudonymized code throughout the study. Only physician-researchers had the key to decode the participant if required. The methods in the protocol are in accordance with all relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 14 December 2023 Accepted: 3 February 2025
Published online: 18 February 2025

References

1. Olesen J. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders. 3rd ed. 2018. https://doi.org/10.1177/0333102417738202.
2. World Health Organization. WHO - Headache disorders. 2016. https://www.who.int/news-room/fact-sheets/detail/headache-disorders. Accessed 27 Nov 2023.
3. Kelman L. The triggers or precipitants of the acute migraine attack. Cephalalgia. 2007;27. https://doi.org/10.1111/j.1468-2982.2007.01303.x.

4. Schroeder J, Chung CF, Epstein DA, Karkar R, Parsons A, Murinova N, et al. Examining Self-Tracking by People with Migraine. 2018. https://doi.org/10.1145/3196709.3196738.
5. Demir YP, Sümer MM. Effects of smartphone overuse on headache, sleep and quality of life in migraine patients. Neurosciences. 2019;24. https://doi.org/10.17712/nsj.2019.2.20180037.
6. Dodick DW, Tepper SJ, Lipton RB, Buse DC, Stewart WF, Bayliss M, et al. Improving Medical Communication in Migraine Management: A Modified Delphi Study to Develop a Digital Migraine Tracker. Headache. 2018;58. https://doi.org/10.1111/head.13426.
7. Lin YK, Lin GY, Lee JT, Lee MS, Tsai CK, Hsu YW, et al. Associations between Sleep Quality and Migraine Frequency. Medicine (United States). 2016;95. https://doi.org/10.1097/MD.0000000000003554.
8. Malone CD, Bhowmick A, Wachholtz AB. Migraine: Treatments, comorbidities, and quality of life, in the USA. J Pain Res. 2015;8. https://doi.org/10.2147/JPR.S88207.
9. Hepp Z, Dodick DW, Varon SF, Chia J, Matthew N, Gillard P, et al. Persistence and switching patterns of oral migraine prophylactic medications among patients with chronic migraine: A retrospective claims analysis. Cephalalgia. 2017;37. https://doi.org/10.1177/0333102416678382.
10. Vandenbussche N, Laterza D, Lisicki M, Lloyd J, Lupi C, Tischler H, et al. Medication-overuse headache: a widely recognized entity amidst ongoing debate. J Headache Pain. 2018;19. https://doi.org/10.1186/s10194-018-0875-x.
11. Healint. Migraine Buddy. 2023. https://migrainebuddy.com/. Accessed 27 Nov 2023.
12. Curelator I. N1-Headache. 2023. https://n1-headache.com/. Accessed 27 Nov 2023.
13. Van Der Donckt J, De Brouwer M, Moens P, Stojchevska M, Steenwinckel B, Pletinck S, et al. From selfreporting to monitoring for improved migraine management. In: Engineer meets Physician (EmP). 2022.
14. De Brouwer M, Vandenbussche N, Steenwinckel B, Stojchevska M, Van Der Donckt J, Degraeve V, et al. mBrain: towards the continuous follow-up and headache classification of primary headache disorder patients. BMC Med Inform Decis Mak. 2022;22. https://doi.org/10.1186/s12911-022-01813-w.
15. Artemenko AR, Filatova E, Vorobyeva YD, Do TP, Ashina M, Danilov AB. Migraine and light: A narrative review. 2022. https://doi.org/10.1111/head.14250.
16. Vuralli D, Ayata C, Bolay H. Cognitive dysfunction and migraine. 2018. https://doi.org/10.1186/s10194-018-0933-4.
17. Gil-Gouveia R, Oliveira AG, Martins IP. Cognitive dysfunction during migraine attacks: A study on migraine without aura. Cephalalgia. 2015;35. https://doi.org/10.1177/0333102414553823.
18. Edwards KR, Rosenthal BL, Farmer KU, Cady RK, Browning R. Evaluation of sumatriptan-naproxen in the treatment of acute migraine: A placebo-controlled, double-blind, cross-over study assessing cognitive function. Headache. 2013;53. https://doi.org/10.1111/head.12052.
19. Wilkins AJ, Haigh SM, Mahroo OA, Plant GT. Photophobia in migraine: A symptom cluster? 2021. https://doi.org/10.1177/03331024211014633.
20. Nir RR, Lee AJ, Huntington S, Nosed R, Bernstein CA, Fulton AB, et al. Color-selective photophobia in ictal vs interictal migraineurs and in healthy controls. Pain. 2018;159. https://doi.org/10.1097/j.pain.0000000000001303.
21. van Casteren DS, Verhagen IE, Onderwater GLJ, MaassenVanDenBrink A, Terwindt GM. Sex differences in prevalence of migraine trigger factors: A cross-sectional study. Cephalalgia. 2021;41. https://doi.org/10.1177/0333102420974362.
22. Hougaard A, Amin F, Hauge AW, Ashina M, Olesen J. Provocation of migraine with aura using natural trigger factors. Neurology. 2013;80. https://doi.org/10.1212/WNL.0b013e31827f0f10.
23. Vgontzas A, Pavlović JM. Sleep Disorders and Migraine: Review of Literature and Potential Pathophysiology Mechanisms. 2018. https://doi.org/10.1111/head.13358.
24. Duan S, Ren Z, Xia H, Wang Z, Zheng T, Liu Z. Association between sleep quality, migraine and migraine burden. Front Neurol. 2022;13. https://doi.org/10.3389/fneur.2022.955298.
25. Tiseo C, Vacca A, Felbush A, Filimonova T, Gai A, Glazyrina T, et al. Migraine and sleep disorders: a systematic review. 2020. https://doi.org/10.1186/s10194-020-01192-5.
26. Song TJ, Yun CH, Cho SJ, Kim WJ, Yang KI, Chu MK. Short sleep duration and poor sleep quality among migraineurs: A population-based study. Cephalalgia. 2018;38. https://doi.org/10.1177/0333102417716936.
27. Buse DC, Rains JC, Pavlovic JM, Fanning KM, Reed ML, Adams AM, et al. Sleep Disorders Among People With Migraine: Results From the Chronic Migraine Epidemiology and Outcomes (CaMEO) Study. Headache. 2019;59. https://doi.org/10.1111/head.13435.
28. Burish MJ, Chen Z, Yoo SH. Emerging relevance of circadian rhythms in headaches and neuropathic pain. Acta Physiol. 2019;225(1):e13161. https://doi.org/10.1111/apha.13161.
29. Masruha MR, de Souza Vieira DS, Minett TSC, Cipolla-Neto J, Zukerman E, Vilanova LCP, et al. Low urinary 6-sulphatoxymelatonin concentrations in acute migraine. J Headache Pain. 2008;9(4):221–4. https://doi.org/10.1007/s10194-008-0047-5.
30. Fino E, Mazzetti M. Monitoring healthy and disturbed sleep through smartphone applications: a review of experimental evidence. 2019. https://doi.org/10.1007/s11325-018-1661-3.
31. Alamoudi D, Breeze E, Crawley E, Nabney I. The Feasibility of Using Smartphone Sensors to Track Insomnia, Depression, and Anxiety in Adults and Young Adults: Narrative Review. 2023. https://doi.org/10.2196/44123.
32. Straczkiewicz M, James P, Onnela JP. A systematic review of smartphone-based human activity recognition methods for health research. 2021. https://doi.org/10.1038/s41746-021-00514-4.
33. Aalbers G, Hendrickson AT, Abeele MMV, Keijsers L. Smartphone-Tracked Digital Markers of Momentary Subjective Stress in College Students: Idiographic Machine Learning Analysis. JMIR mHealth uHealth. 2023;11. https://doi.org/10.2196/37469.
34. Chen MH, Leow A, Ross MK, DeLuca J, Chiaravalloti N, Costa SL, et al. Associations between smartphone keystroke dynamics and cognition in MS. Digit Health. 2022;8. https://doi.org/10.1177/20552076221143234.
35. Demasi O, Feygin S, Dembo A, Aguilera A, Recht B. Well-being tracking via smartphone-measured activity and sleep: Cohort study. JMIR mHealth uHealth. 2017;5. https://doi.org/10.2196/mhealth.7820.
36. Ben-Zeev D, Scherer EA, Wang R, Xie H. Next-Generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. Psychiatr Rehabil J. 2015;38. https://doi.org/10.1037/prj0000130.
37. Gordon ML, Garys L, Guestrin C, Bigham JP, Trister A, Patel K. App usage predicts cognitive ability in older adults. 2019. https://doi.org/10.1145/3290605.3300398.
38. Peng KP, May A. Redefining migraine phases - a suggestion based on clinical, physiological, and functional imaging evidence. Cephalalgia. 2020;40. https://doi.org/10.1177/0333102419898868.
39. Bai J, Di C, Xiao L, Evenson KR, LaCroix AZ, Crainiceanu CM, et al. An Activity Index for Raw Accelerometry Data and Its Comparison with Other Activity Metrics. PLoS ONE. 2016;11(8):1–14. https://doi.org/10.1371/journal.pone.0160644.
40. Donckt JVD, Donckt JVD, Deprost E, Hoecke SV. Plotly-Resampler: Effective Visual Analytics for Large Time Series. 2022. https://doi.org/10.1109/VIS54862.2022.00013.
41. Bridge K, Sharkey K, Satran M. Understanding and Interpreting Lux Values. 2021. https://learn.microsoft.com/en-us/windows/win32/sensorsapi/understanding-and-interpreting-lux-values. Accessed 27 Nov 2023.
42. Böttcher S, Vieluf S, Bruno E, Joseph B, Epitashvili N, Biondi A, et al. Data quality evaluation in wearable monitoring. Sci Rep. 2022;12(1):21412.
43. Wilcoxon F. Individual Comparisons by Ranking Methods. Biom Bull. 1945;1. https://doi.org/10.2307/3001968.

Publisher's Note