ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

DIGITAL MEASUREMENT OF EMPLOYEE STRESS AND FATIGUE IN ARTIFICIAL INTELLIGENCE-DRIVEN WORKFORCE MANAGEMENT ENVIRONMENTS: AN EMPIRICAL STUDY USING AICOACH

Adam Dalnoki[1], Md Iftekhar Islam[2]

Article DOI: https://doi.org/10.36713/epra15161
DOI No: 10.36713/epra15161

ABSTRACT

Effective fatigue management in the workplace is associated with employee well-being, safety, and productivity. Fatigue, often resulting from inadequate rest and high workloads, can lead to decreased cognitive function, diminished job performance, and increased risk of accidents, thus affecting both individual health and organizational efficiency. For this, we introduced an innovative approach to understanding and addressing workplace fatigue. By analyzing the interrelationships between employees' self-reported fatigue levels, their perceptions of managing fatigue, and their engagement with an AI-based coaching tool (AiCoach), we sought to uncover patterns that could inform more effective fatigue management strategies. Using advanced sentiment analysis adapted for the context of fatigue and Granger-causality tests, we examined these dynamics over time. Our findings highlight the importance of immediate perceptions of fatigue in predicting engagement with management interventions, underscoring the need for real-time monitoring and adaptive strategies in managing workplace fatigue.

KEYWORDS: *Workplace Fatigue, Stress Management, AiCoach, Sentiment Analysis.*

1. INTRODUCTION

Fatigue is recognized as a significant occupational hazard that has profound implications on the safety and health of employees and their co-workers. Fatigue poses a considerable challenge in today's fast-paced society, primarily due to intense work demands, extended working hours, disrupted sleep-wake cycles, various social and surrounding pressures, and often inadequate sleep [1,2]. This condition is a multifaced issue that arises from a combination of factors including the duration of wakefulness, specific times of day, extremes in workload, personal health status, and the balance between professional and personal responsibilities and lifestyle choices. Modern industrial society inherently brings about fatigue for several reasons. Operations that run 24/7, irregular work schedules, and frequent travel across time zones can significantly disrupt natural circadian rhythms. Additionally, brief and inconsistent periods of rest, long travel times to and from work, and suboptimal sleeping conditions often comprise both the amount and quality of sleep. Furthermore, people vary significantly in their sleep needs and in how they tolerate fatigue, which means some are more susceptible to its effects than others. Fatigue, along with excessive sleepiness during the day, can also stem from disorders affecting the central or peripheral nervous systems, as well as from various other health conditions, including common ailments like infections, asthma, gastrointestinal issues, and metabolic disturbances [3].

Even with advancements in technology and industry, the issue of work-related fatigue remains persistent. According to a report by the National Safety Council (NSC) in 2018, a significant portion of the United States workforce, about 107 million of the total 160 million workers, experienced work-related fatigue. A study conducted in 2007 indicated that fatigue among workers leads to productivity loss and other problems, costing the U.S. economy around 101 billion dollars annually [4]. Furthermore, it has been found that 13% of injuries in the workplace are linked to fatigue. In Spain, fatigue affects 30.8% of workers. In Europe, 3.2% of people aged between 15 to 64 experienced at least one work-related accident in the past year. About 70% of these non-deadly accidents happened because of losing control or failing, often due to stress or fatigue from work [5].

Given the aforementioned factors and to circumvent them, it becomes crucial to identify and quantify work-related fatigue. This step is essential to prevent injuries, accidents, or illness. Additionally, accurate detection and measurement of fatigue enable the provision of tailored recommendations aimed at reducing stress in the workplace.

In a previous study [6], we proposed a platform to analyze the fatigue level in patients with Multiple sclerosis (MS). Now in this study, we are trying to understand and manage workplace fatigue, leveraging the advanced data analysis techniques used in our previous research to explore the dynamic relationships between employees' self-reported fatigue levels, their perceptions of managing fatigue, and their engagement with an AI-assisted coaching tool (AiCoach). By employing different data analysis techniques adapted to the specific context of workplace fatigue, our research aims to unravel the complex interplay between these factors over time. Furthermore, we employed the Valence Aware Dictionary for sEntiment Reasoning (VADER) methodology and conducted the Granger-Causality Test to the open-ended question to estimate the fatigue level of the participants. This approach not only provides a deeper insight into the immediate and evolving impacts of fatigue on employee engagement with fatigue management interventions but also offers a data-driven foundation for developing more effective, responsive, and personalized strategies to address fatigue in the workplace. Through this study, we seek to contribute to the growing field of occupational health psychology by offering empirical evidence and analytical perspectives on fatigue management, a critical aspect of employee well-being and productivity.

2. LITERATURE REVIEW

To identify and quantify fatigue, various methods are employed. These include objective techniques, which analyze the body parts exerting force during a task, and subjective methods which gauge fatigue through the use of rating scales and questionnaires. These questionnaires assess an individual's perceived stress [11]. Perceived stress, a common metric for quantifying fatigue, is described by Borg as an individual's sense of how hard their body is working during an activity. This sense is a holistic interpretation based on various sensory inputs and perceptions [12]. In studying work-related fatigue, it's important to assess symptoms and discomfort subjectivity. Therefore, three psychophysical measurement techniques for perceived stress have been developed over recent decades: ratio scaling, category scaling, and acceptability scaling [13]. Ratio scaling aims to achieve the same metric qualities as those in physics and physiology, with absolute zero and equal distances between scale values [14]. A notable example is the Magnitude Estimation scale, introduced by Stevens in 1975. This perceptual scaling method asks participants to assign numbers proportional to their perceptual intensities [12]. Borg later introduced the Rating of Perceived Exertion (RPE) scale, which allows subjects to rate their effort and stress during physical tasks on a scale from 6 (no stress) to 20 (maximum stress) [15]. Borg also developed the CR10 scale, a category scale ranging from 0 to 10 with verbal anchors, where 10 represents extremely strong stress, categorized as 'maximal' [16]. The RPE scale is often preferred in simple applied studies of perceived stress and for predicting physical intensities, while the CR10 scale is more suitable for assessing subjective symptoms [15]. Other category scales include the CR100 (centiMax), ranging from 0 to 100, and the OMNI-RPE scale, which goes from O to 10 and includes mode-specific pictures [17].

In addition to the methods previously discussed, questionnaires are another effective tool for evaluating perceived stress. One such example is the Fatigue Severity Scale (FSS), a self-report questionnaire consisting of nine items, specifically designed for monitoring fatigue related to various diseases [18]. Another is the Multidimensional Fatigue Inventory (MFI-20), which is a more extensive 20-item questionnaire. It is divided into five subcategories: general fatigue, physical fatigue, reduced activity, reduced motivation, and mental fatigue, allowing for a comprehensive assessment of fatigue [19]. Additionally, there's the Chadler Fatigue Scale (CFQ), which focuses on both the physical and psychological aspects of fatigue. It comprises an 11-item questionnaire where responses are rated on a scale with four options: 0 (better than usual), 1 (not worse than usual), 2 (worse than usual), and 3 (much worse than usual), providing a nuanced view of fatigue levels [20].

Recently, there has been a surge in the utilization of wearable technology, which offers real-time monitoring, recording, and communication of an individual's physical activities and environmental conditions. These technological innovations come in various forms, including smartwatches, wristbands, eyeglasses, jewelry, skin patches, and even textiles embedded with smart technology [21]. Sensors are a key component in these wearable devices, and their application is predominantly seen in the sports sector. With the advances in semiconductor technology, these devices are now capable of monitoring a comprehensive range of parameters. This technological progress is bringing the use of wearable devices closer to practical applications in the field of medicine [22]. Various health monitoring wearables, such as ECG monitors, blood pressure monitors, and biosensors are some examples [23]. The integration of these devices into the healthcare sector has been gradual, mainly due to the necessity for their validation in the context of various medical conditions. However, numerous studies have shown the potential of these devices in research, demonstrating their feasibility for predicting, monitoring, or assessing a range of diseases and disabilities [24,25].

Beyond the previously discussed ratings of perceived fatigue for evaluating work-related stress, there is ongoing research into the potential use of wearable devices for monitoring bodily functions. Such devices could quantify physical exposures in the workplace by tracking brain activity with electroencephalography (EEG) or by observing changes in muscle activity using electromyography (EMG) [26]. However, methods based on EEG and EMG for measuring physical fatigue are considered intrusive, as they require the attachment of multiple electrodes. These methods are also not ideally suited for dynamic work environments, as they are better suited for stationary tasks, and they represent costly technologies [27,28]. Due to these limitations, researchers are exploring the use of non-intrusive wearable sensors in the workplace to monitor physical activities and movements. These devices are not only more affordable but also simpler to operate [29]. Among the various wearable sensors, one of the most frequently used is the inertial measurement unit (IMU), which provides valuable data on angular velocity and acceleration, aiding in the detection of subject movement.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

In recent years, a growing number of workplace-related studies have incorporated Inertial Measurement Units (IMUs) for various purposes. These include ergonomic evaluations [30], analyzing postures [31], assessing musculoskeletal disorders [32], examining body motion and the risks associated with lifting during manual handling tasks [33], and evaluating the risk of falls during everyday activities [34,35]. Additionally, more recent studies have started using IMUs to detect physical activity levels and fatigue [36,37,38].

To create comprehensive systems capable of identifying fatigue states by processing vast amounts of data, the application of Artificial Intelligence (AI) has been increasingly utilized. Several past research projects have employed Machine Learning (ML) or Deep Learning (DL) techniques on physiological signals. These techniques help in developing systems that more effectively extract pertinent features from the collected data sets, simplifying the process of data analysis [40].

In the realm of Deep Learning (DL) for fatigue classification, Maman et al. (2017) [11], utilized IMUs to gather data on acceleration and jerk from participants engaged in manufacturing tasks. This data, combined with Heart Rate (HR) readings and Rated Perceived Exertion (RPE) values, was used to implement a Least Absolute Shrinkage and Selection Operator (LASSO) model. This model helped in selecting significant features from the data for applying regression and logistic models to estimate levels of physical fatigue. Later, Maman et al., in 2020 [39], proposed a comprehensive framework focused on the detection, identification, diagnosis, and recovery from fatigue, aiming to quantify and predict shifts in worker's performance.

In a different approach, Karvekar et al., (2019) [36], employed accelerometers integrated into smartphones to measure the motion and gait parameters of participants, alongside RPE values for data labeling. They then applied a Machine Learning (ML) algorithm, specifically a Support Vector Machine (SVM) model, to classify the fatigue levels of subjects. This methodology was similarly adopted in studies by Zhang et al., (2013), Baghdadi et al., (2018), and Kuschan and Krüger (2021), who also used SVM models for physical fatigue detection [41,42,43].

Another research focus in this area involved measuring the cycle acceleration of workers performing various tasks with IMUs for fatigue detection, employing Statistical Process Control (SPC) techniques [37]. More recently, Lambay et al. (2021) leveraged the dataset from Sedighi Maman et al. (2017) to conduct fatigue prediction for manual material handling tasks [44]. These studies collectively contribute to the development of a proactive approach to the continuous monitoring of operator's fatigue levels, with the potential to enhance work performance and mitigate the earlier-mentioned risks [45,46].

3. METHODOLOGY

3.1 Study Population

This study was conducted with employees from multiple industrial plants across Hungary. The focus was on understanding the dynamics of workplace fatigue and the effectiveness of an AI-assisted coaching tool (AiCoach) in managing it. Our study particularly targeted the industrial workforce, a demographic where fatigue management is crucial yet often under-researched. The participant pool consisted of employees from various industrial sectors, representing a broad spectrum of occupational roles and environments. The inclusion criteria were: (1) being currently employed in one of the participating industrial plants, (2) willingness to engage with the AiCoach tool over the study period, and (3) providing informed consent for participation in the study. Employees with any known medical conditions that could significantly impact fatigue levels were excluded to ensure the reliability of the data regarding occupational fatigue

3.2 Study Design

The study was designed as a longitudinal observational study, where participants were asked to respond to a series of questions including both open and closed-ended, over five epochs, in two different phases with a duration of approximately three months. This design allowed for the monitoring of changes in fatigue levels and engagement with the AiCoach tool over time. Participants were introduced to the AiCoach tool and trained on its use at the onset of the study. They were asked to report their perceived fatigue levels and their experiences with managing fatigue using the tool. The data collected included not only their responses to the survey but also their engagement metrics with the developed software.

To ensure the robustness of the data, the study employed advanced analytical techniques, including an adapted version of the VADER sentiment analysis to interpret the textual responses and Granger-causality tests to explore the predictive relationships between reported fatigue levels, perceptions of fatigue management, and engagement with AiCoach. The study collected data in two different phases. The first phase lasted from 2021.05.10 to 2021.08.31, while the second phase lasted from 2021.09.01 to 2022.04.30. A total of 168 employees were recruited for the study. All participants provided written informed consent in accordance with the study protocol. Participants who completed all survey epochs and met the engagement criteria with the proposed tool were offered a nominal remuneration for their time and contribution to the research.

Table 1. Balance of Enrollment During the Study Period

Participants	Number	Percentage
Total enrolled	168	100%
Fully completed the study	105	62.5%
Partially completed the study	63	37.5%

3.3 Development of the mobile application

Mobilengine, an interdisciplinary team of physicians and scientists, developed the AiCoach mobile application. The organization has a long track record of developing mobile apps for monitoring and recording healthcare data [53]. The application development process involved collaboration across teams, including physicians and engineers, achieved through a series of virtual workshops. During the study, participants were asked to install the AiCoach app on their own Android smartphones. The study was conducted over approximately 3 months, and divided into two phases. During this time, participants integrated the use of AiCoach into their daily routines, providing valuable data for analysis by the Mobilengine research team.

The mobile app is structured into two distinct modules, which are elaborated upon below. (1) the first module consists of a series of one-time questions that participants were required to complete within the first three days post-enrollment. (2) The second module includes Visual Analog Scales (VASs) that participants used to self-report their level of fatigue, depression, anxiety, and pain twice a day.

It is important to note that the data collection through VAS was consistently conducted over around 200±30 days, commencing once all one-time questionnaires were completed. Consequently, the total duration for data collection varied between 10 to 13 days, depending on the time taken by each participant to answer all one-time questionnaires. Additionally, participants were given the flexibility to pause their entries in the app and return at a later time to complete the remaining questions.

Table 2. Overview of standardized survey questionnaires integrated into our one-time questionnaire module.

Questionnaire	Domain Assessed	Study
The Fatigue Severity Scale	Fatigue	Krupp et al (1989) [18]
The Neuro-QoL[a] fatigue questionnaire	Fatigue	Cella et al (2012) [47]
The Neuro-QoL depression questionnaire	Depression	Cella et al (2012) [47]
The Neuro-QoL anxiety questionnaire	Anxiety	Cella et al (2012) [47]
The Neuro-QoL sleep questionnaire	Sleep Quality	Cella et al (2012) [47]
The Modified Fatigue Impact Scale	Fatigue	Amtmann et al (2012) [48]
Symptoms of depression questionnaire (7 questions only)	Vegetative symptoms of depression	Pedrelli et al (2014) [49]
The Epworth Sleepiness Scale	Sleepiness	Johns (1991) [50]
The Godin Leisure-Time Exercise Questionnaire	Physical activity	Godin (1985) [51]
The Behavioral Approach System and Behavioral Avoidance System scale	Drive; fun seeking; reward responsiveness	Carver and White (1994) [52]

To mitigate the occurrence of slip errors in our study, the following protocols were established: (1) User-initiated response submission: answers were recorded only after the participant selected the navigational arrows to move forward or backward; (2) Navigational flexibility: participants could seamlessly traverse through questions within the same section, permitting on-the-fly amendments; (3) Uniform interface design: all visual scales were standardized in size with questions, options, and navigational cues consistently positioned on the interface. Additionally, the application supported a landscape orientation, offering expanded spacing on visual analog scales; and (4) Progress transparency: a counter indicating the number of questions completed relative to the total was displayed prominently at the interface's lower segment.

3.4 One-time questionnaire module

In the study, we incorporated a one-time questionnaire module comprising a total of 168 questions which encompassed a mix of validated questionnaires (Table 2) and newly developed inquiries tailored to investigate various facets of fatigue relevant to our research objectives. These additional questions aimed to explore specific aspects of fatigue such as (1) variations in fatigue severity over the day, (2) the potential influence of caffeine and nicotine consumption on perceived fatigue levels, and (3) open-ended queries to uncover other potential factors contributing to or alleviating fatigue among the participants. The questionnaire module retained the original wording and response formats, including single-choice, multiple-choice, free-text, analog scale, and date selection, which were consistent with the formats used in the source questionnaires. Questions were logically grouped into distinct sections that mirrored the structure of the original questionnaires, allowing participants to navigate through the questions one at a time. Importantly participants were alerted that they could not revisit previous sections after completing them, emphasizing the importance of careful consideration before final submission of their responses.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

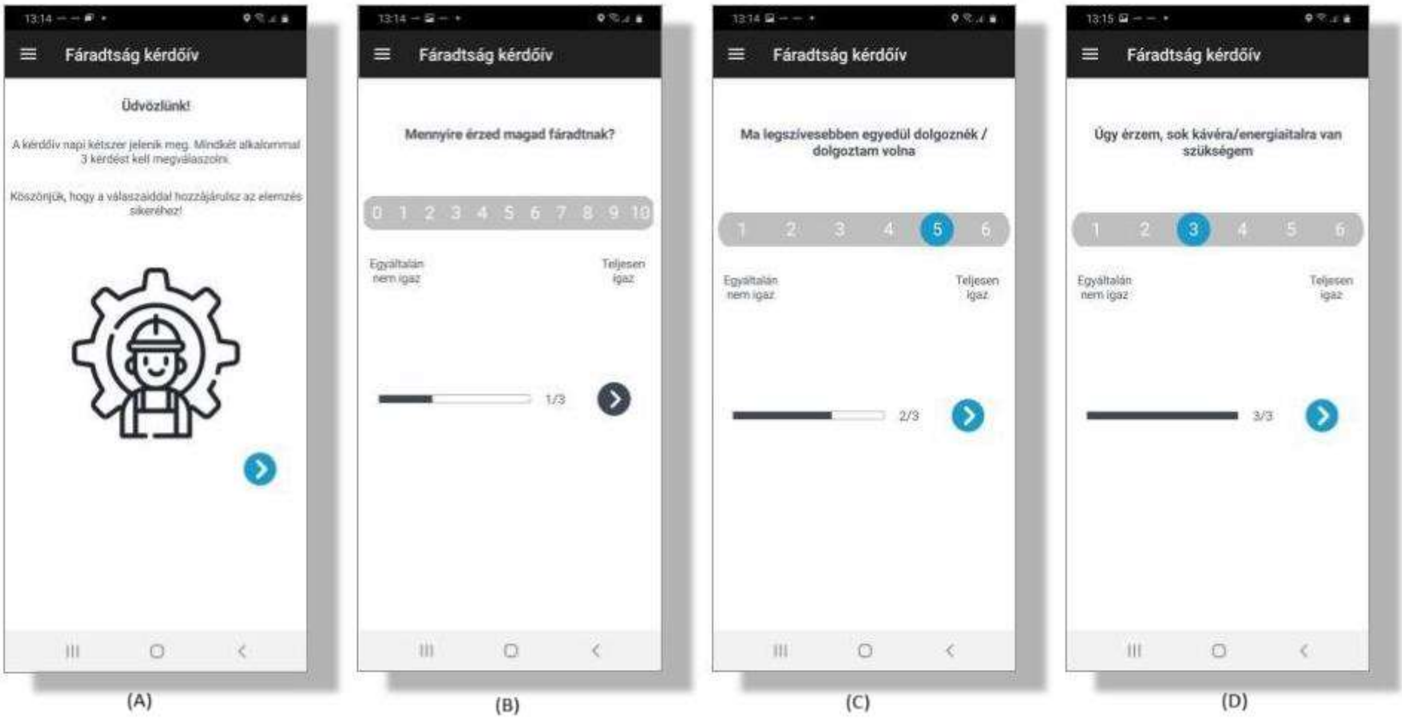

Figure 1: Sample interfaces for different types of questions within the one-time questionnaire module, including single-answer are depicted.

The developed questionnaire is in Hungary as local participants were selected and to facilitate their involvement in the study. In image (A) the participants are being greeted to take part in the study. The image (B) asks the participants how tired they are feeling right now. The image (C) translated that "I would prefer to work alone today". And the translation of image (D) says that "I feel like I need a lot of coffee/energy drinks".

3.5 VAS Module

The VAS (Visual Analog Scale) module was utilized to evaluate an individual's existing level of fatigue, anxiety, depression, and pain. Each of these four VASs is rated on a scale ranging from 0 to 10, where 0 signifies the absence (none) of the symptom (e.g., no fatigue), and 10 indicates the extreme presence of the symptom (as depicted in Figure 2). Prior research has demonstrated a substantial correlation between VAS scores and a series of visual depictions of facial expressions depicting increasing distress [7,8].

Our primary objective was to investigate circadian variations in fatigue, anxiety, depression, and pain by measuring these symptoms once in the morning (2 hour after starting work) and once in the afternoon (2 hour before finishing work). The participants were given the flexibility to choose preferred times for completing their VAS (Visual Analog Scale) assessments, once before going to work and once in the afternoon. To facilitate this, the AiCoach app was programmed to send reminders. These reminders took the form of notifications, tailored to each participant's chosen times for the survey.

Figure 2: Workflow of the data collection process using survey questionnaire

In our study, participants were required to respond to all questions within a questionnaire before they could submit it. A progress bar at the bottom of the screen visually displayed the participant's advancement through the one-time questionnaires or VAS assessments (as shown in Figure 1). The app recorded the exact date and time when each question was opened and submitted. Every participant was assigned a unique subject ID to maintain anonymity, and no additional personally identifiable information was collected or stored within the app.

3.6 Extracting Fatigue level from Open-ended texts

Adapting the Valence Aware Dictionary for sEntiment Reasoning (VADER) methodology [9], we propose an engine specifically tuned for assessing fatigue levels from free text. This adaptation maintains the original system's advantages, such as no requirement for a training phase, suitability for short texts, and rapid processing for near real-time applications. Its domain-agnostic nature and white-box model approach ensure interpretability and adaptability across various languages. The foundation of this system is a modified sentiment lexicon, now focused on fatigue-related terms. This lexicon is an extension of the original VADER lexicon, supplemented with terms and expressions commonly associated with fatigue and tiredness. These terms are annotated to reflect varying degrees of fatigue on a scale (e.g., [0, 4], where 0 is no fatigue and 4 is extreme fatigue), using a crowd-sourcing platform similar to Amazon Mechanical Turk [10].

The key step in this fatigue assessment engine is the identification of linguistic patterns and cues that are indicative of fatigue. To this end, we adapt the VADER's sentiment polarity shifters to fatigue-specific shifters, identifying terms and constructs that either amplify or diminish the expression of fatigue. For instance:

- Punctuation: Repetitive punctuation (e.g., ellipsis ...) may indicate a trailing off of thought, potentially a sign of fatigue.
- Capitalization: Random or inconsistent capitalization could reflect a lack of attention or focus, often associated with fatigue.
- Degree modifiers: Certain adverbs and adjectives (e.g., 'extremely tired', 'barely awake') will be key in determining the level of fatigue expressed.

- Contrastive particles: Phrases like 'but still tired' can indicate a persistent state of fatigue, despite changing circumstances.
- Negation: The use of negation in the context of rest or sleep (e.g., 'not rested', 'couldn't sleep') can be a strong indicator of fatigue.

The extension to languages other than English, such as Hungarian, involves translating and adapting these fatigue-specific terms and rules. Just as in VADER, certain universal aspects, like the use of punctuation and capitalization, remain consistent across languages, while others, like specific idioms or culturally specific expressions of tiredness, require localization. This adapted method aims to provide a nuanced analysis of fatigue levels in text, offering valuable insights in domains such as health monitoring, workplace wellbeing assessments, and psychological studies.

3.7 Granger-Causality Testing

In the context of analyzing fatigue levels from textual data, Granger-causality is utilized as a statistical hypothesis testing model to determine if there is a directed relationship between two time series in terms of fatigue expression [7]. Specifically, a time series X, representing a measure of fatigue indicators in text, is said to be Granger-cause time series Y, which could be a series of outcomes or states related to fatigue (e.g., performance metrics, error rates, health indicators), if it can be shown that including past values of X (i.e., lagged values of fatigue indicators) alongside Y significantly improves the prediction of future values of Y.

For this adaptation, the Granger-causality test was applied to the lagged values of the fatigue indicator time series (X). All lags ranging from one to four were tested, aligning with the considered periods of data collection or observation epochs minus one. The alternative hypothesis in this context is that the time series of fatigue levels derived from textual analysis Granger causes the time series representing the related outcomes or states. The level of significance for these tests was set at 5%, i.e., a p-value < 0.05 was considered statistically significant. It's important to note that the Granger-causality test presupposes that the time series being investigated are stationary. Therefore, to ensure the validity of the test, the augmented Dickey–Fuller method was employed to verify the stationarity conditions of both the fatigue indicator series and the outcome series [8]. This adaptation of Granger-causality testing is aimed at understanding the impact of textual expressions of fatigue over time on various outcome measures, providing a novel approach to assessing and predicting the implications of fatigue in different contexts, such as workplace productivity, academic performance, or health-related quality of life."

3.8 Monitoring the Database

The application functioned entirely offline, storing all responses from participants on the smartphone's internal memory. The survey data from the participant's phones were collected by the Mobilengine server only when they were connected to Wi-Fi. This setup enabled real-time, on-demand monitoring of data collection. The developed database also categorized participants based on their activity. Participants who actively completed the surveys were marked green and participants lagging in their daily tasks were marked in red.

Once a participant completed the VASs, their monitoring period ended, and they were no longer able to respond to further questions. Subsequently, their code was removed from the daily update emails. This monitoring capability was crucial, as it allowed investigators to quickly follow up with participants not adhering to the study timeline, rather than waiting until the end of the period. The database was designed for easy data extraction, with each response stored in a table format that included the subject ID, individual question ID, response, and time of response, with each response generating a new row in the table, following a long data format.

4. RESULT

4.1 Exploratory Data Analysis

The research focused on identifying fatigue levels among employees, and exploratory data analysis was conducted on responses gathered from various industrial plant workers. A total of 105 fully completed and 63 partially completed responses were collected and analyzed in line with the system architecture outlined in the previous section. This analysis is critical for gleaning insights into data collection and integration processes related to workplace fatigue.

Responses were submitted over five subsequent periods, approximately 3 months in two different phases, to track changes and trends in fatigue over time. To visualize the distribution of fatigue levels across these different time points, a violin plot was generated (Figure 5). This plot illustrates the variation in fatigue scores (expressed as a percentage) corresponding to each submission epoch. Concurrently, Table 3 presents this data in a tabular format for a more detailed examination. While the mean fatigue scores were found to range between [38.67%, and 51.45%] across different epochs, the standard deviation, as well as the minimum and maximum values, highlighted significant variations in fatigue levels among the employees. These variations were evident from the extremes of some employees reporting minimal fatigue to others indicating levels considerably higher than average. Such disparities underscore the complex nature of workplace fatigue and its varied manifestations among employees in different roles and working conditions.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

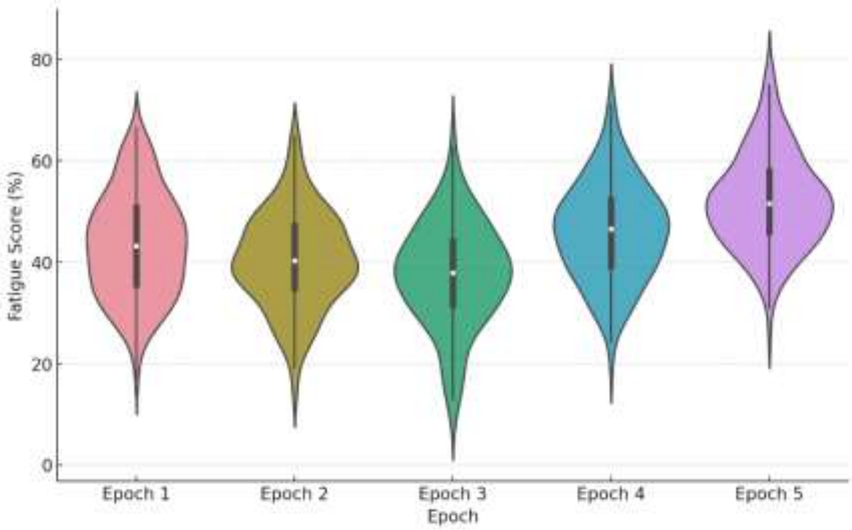

Figure 3: Distribution of fatigue percentage across five subsequent submission epochs.

Table 3. Fatigue Score (%): Descriptive Statistics Across Five Subsequent Epochs

	Epoch 1	**Epoch 2**	**Epoch 3**	**Epoch 4**	**Epoch 5**
Mean	42.67	41.77	38.67	47.39	51.45
Std	57.25	57.70	59.64	59.54	51.82
Min	0.00	0.00	0.00	0.00	0.00
Max	100%	100%	100%	100%	100%

The data presented in Table 3 reveals an intriguing pattern in the fatigue scores of employees over five different epochs. Initially, the average fatigue score starts at 42.67% in Epoch 1, slightly decreases to 41.77% in Epoch 2, and reaches its lowest at 38.67% in Epoch 3. This trend could suggest an initial adaptation or improvement in managing fatigue. However, this trend reverses in the later epochs, with a notable increase to 47.39% in Epoch 4 and further up to 51.45% in Epoch 5. This upward trend could be indicative of accumulating fatigue over time or changes in workplace dynamics or stressors. The standard deviation remains consistently high across all epochs, hovering around the high 50s, which indicates a substantial variation in individual fatigue levels within the workforce. This high variability suggests that while some employees might be coping well, others are experiencing significantly higher levels of fatigue. The minimum values across all epochs stand at 0.00%, indicating that there are individuals who report no fatigue. However, the maximum values show an extreme range, peaking at 100% in all Epochs. These maxima are notably higher and reflect instances of extreme fatigue or possibly the way fatigue is being reported or calculated, suggesting a need for closer examination of these outlier responses.

4.2 Testing Granger-Causality

A comprehensive analysis of the responses to the fatigue identification survey is beyond the scope of this paper. Instead, only answers to two open-ended questions collected from the survey are discussed for testing Granger-Causality.

Q1: “Explain how fatigued you are now?”

Q2: “How successful are you in managing fatigue since using AiCoach?”

Three-time series were evaluated: the fatigue intensity scores corresponding to Q1 ("Explain how fatigued you are now?") and Q2 ("How successful are you in managing fatigue since using AiCoach?"), as well as the time series of engagement levels with the AiCoach program. The Augmented Dickey–Fuller Test confirmed that each of the three-time series met the stationarity condition (p-value = 4.6124×10^{-18}, p-value = 3.2185×10^{-7}, and p-value = 0.0035, respectively). Two Granger-causality tests were carried out to examine responses to Q1 Granger-cause engagement scores and responses to Q2 Granger-cause engagement scores. Additionally, given the concurrent consideration of all three series, it was also necessary to assess whether engagement scores Granger-cause responses to Q1 and Q2. Three different test statistics—F-test, chi-square, and likelihood-ratio—were utilized, with the number of lags tested ranging from one to four. The outcomes, presented in terms of p-values in Tables 4 and 5, indicate that both Q1 and Q2 have a Granger-causal relationship with engagement scores at lag = 1. Conversely, engagement scores do not appear to Granger-cause the responses to Q1 or Q2.

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

Table 4: Q1 and Engagement Score: p-value of Granger-causality test performed with three different statistics and four different lags.

Lag	Q1 → Engagement Score F-test	Q1 → Engagement Score Chi2 test	Q1 → Engagement Score Likelihood-ratio	Engagement Score → Q1 F-test	Engagement Score → Q1 Chi2 test	Engagement Score → Q1 Likelihood-ratio
1	0.0368	0.0339	0.0350	0.4076	0.4027	0.4031
2	0.0998	0.0908	0.0936	0.6667	0.6585	0.6592
3	0.1063	0.0917	0.0963	0.6628	0.6479	0.6496
4	0.2032	0.1760	0.1833	0.8200	0.8061	0.8064

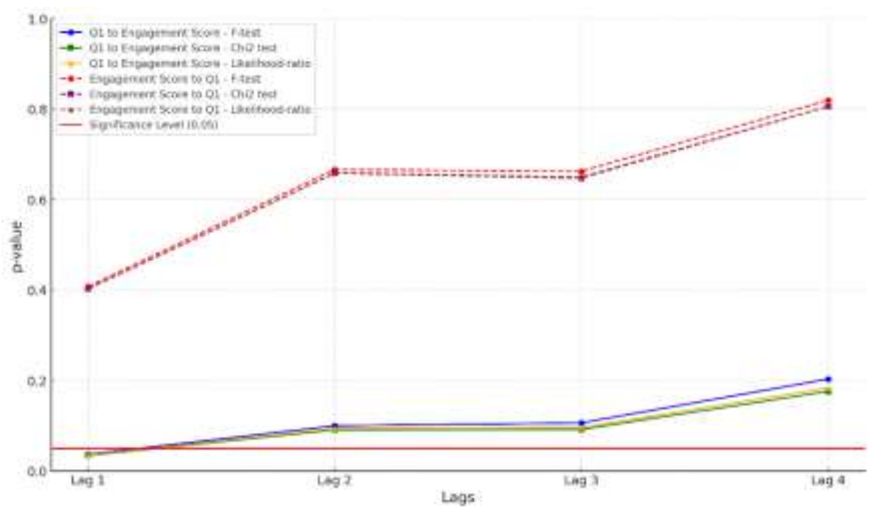

Figure 4: Granger-Causality Analysis of Fatigue Levels and Engagement Scores for Q1

Figure 6 presents the Granger-causality test p-values for the relationship between employee-reported fatigue levels (Q1) and their Engagement Scores with the AiCoach tool over four sequential time lags. A striking feature of the graph is the initial set of p-values for Q1 influencing Engagement Scores at Lag 1, where all three statistical tests—the F-test, Chi2 test, and Likelihood-ratio—yield p-values below the 0.05 significance threshold. This suggests a statistically significant relationship at this initial lag, indicating that earlier reported fatigue levels can predict subsequent engagement with the AiCoach tool. As the lags increase, however, the p-values rise above the 0.05 threshold, indicating that the predictive power of Q1 decreases, implying that fatigue levels reported further in the past are less indicative of future engagement behaviors. Conversely, the Engagement Scores do not present a significant Granger-causal influence on Q1 at any lag, as evidenced by the consistently high p-values across all tests. This lack of statistical significance suggests that how employees engage with the AiCoach tool does not serve as a reliable indicator of their future self-reported fatigue levels.

Table 5: Q2 and Engagement Score: p-value of Granger-causality test performed with three different statistics and four different lags.

Lag	Q1 → Engagement Score F-test	Q1 → Engagement Score Chi2 test	Q1 → Engagement Score Likelihood-ratio	Engagement Score → Q1 F-test	Engagement Score → Q1 Chi2 test	Engagement Score → Q1 Likelihood-ratio
1	0.0020	0.0339	0.0350	0.4076	0.4027	0.4031
2	0.0174	0.0908	0.0936	0.6667	0.6585	0.6592
3	0.0547	0.0917	0.0963	0.6628	0.6479	0.6496
4	0.0865	0.1760	0.1833	0.8200	0.8061	0.8064

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

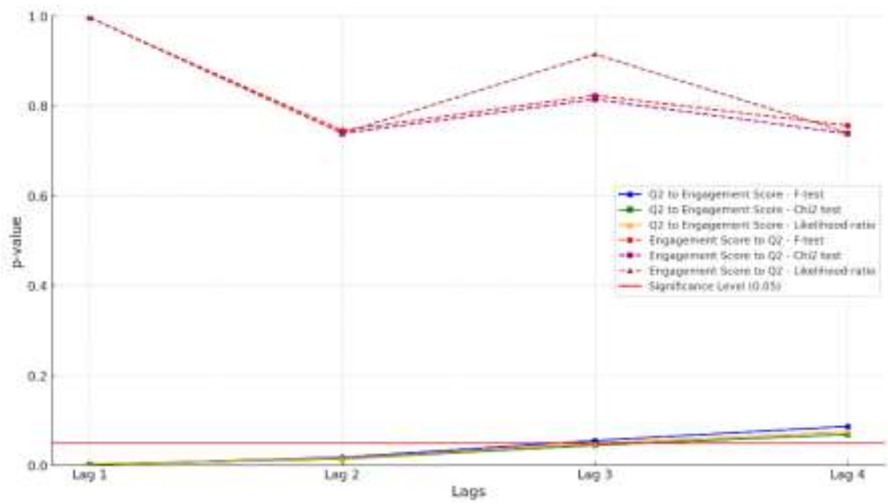

Figure 5: Granger-Causality Analysis of Fatigue Levels and Engagement Scores for Q2

Figure 5 illustrates the Granger-causality test p-values between responses to Q2 and Engagement Scores. A key observation from this graph is the marked significance of the causal relationship from Q2 responses to Engagement Scores, especially at Lag 1, where the p-values for all tests (F-test, Chi2 test, and Likelihood-ratio) fall significantly below the 0.05 threshold. This indicates a strong predictive relationship, suggesting that responses to Q2 have a substantial impact on predicting subsequent Engagement Scores. As the lags increase, the p-values for the relationship from Q2 to Engagement Scores rise, but they remain below the threshold in some tests up to Lag 4, pointing to a sustained, albeit diminishing predictive influence. In contrast, the relationship from Engagement Scores back to Q2 does not exhibit statistical significance at any lag. The p-values in this direction remain consistently high, well above the 0.05 significance threshold for all lags and statistical tests. This lack of significance suggests that the Engagement Scores do not have a predictive effect on the subsequent responses to Q2.

4.3 Fatigue level analysis

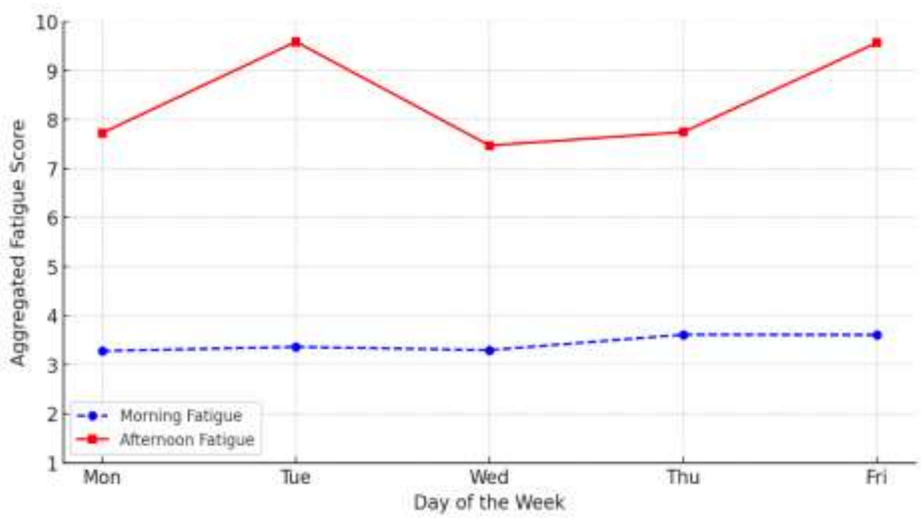

Figure 6: Average fatigue level of participant

Figure 6 demonstrates the statistical analysis conducted on a single randomly selected participant's aggregated fatigue scores over a five-day workweek. The statistic reveals a distinct pattern in the variance of fatigue levels between morning and afternoon. In the morning, the participant exhibited a moderate level of fatigue with a mean score of 3.27 and a closely aligned median of 3.20, indicating a symmetrical and consistent distribution of fatigue levels. The standard deviation of 0.38 for these morning scores suggests minimal variability, with fatigue levels generally confined to a moderate range, as evidenced by the narrow range of scores (minimum of 2.69 and maximum of 3.86).

In stark contrast, the afternoon fatigue scores were significantly higher, with a mean of 8.66, reflecting a pronounced increase in fatigue as the day progressed. This increase is further underscored by the median score of 8.58, which, like the morning, suggests a symmetrical distribution of scores in the afternoon. However, the afternoon scores exhibited a higher standard deviation of 0.93, indicating a greater fluctuation in fatigue levels. The broader range of scores in the afternoon, from a minimum of 7.53 to a maximum of 9.85, highlights a more varied experience of fatigue, encompassing moderately high to very high levels.

The analysis underscores a clear escalation in fatigue levels from the morning to the afternoon for the participant. The consistency in morning fatigue levels contrasts sharply with the more varied and elevated levels of afternoon fatigue. This pattern could have significant implications for understanding individual energy cycles, optimizing work schedules, and tailoring personal health and wellness strategies.

5. DISCUSSION

This research aimed to understand the dynamic relationship between employees' self-reported fatigue levels, their perceptions of managing fatigue, and their engagement with the AiCoach program. Utilizing Granger-causality analysis over multiple time lags, our study revealed several key insights that can inform future workplace fatigue management strategies. The results from the Granger-causality tests for Q1 ("Explain how fatigued you are now?") indicated a significant predictive relationship between employees' reported fatigue levels and their subsequent engagement with the AiCoach program, particularly at the immediate lag (Lag 1). This finding suggests that employees' current perception of their fatigue levels is a strong predictor of how actively they engage with interventions designed to manage fatigue. The significance of this relationship diminishes with time, which could be attributed to the evolving nature of fatigue and its management over longer periods.

The dynamics in Figures 4 and 5 for the two causal directions diverge as the number of lags increases, further reinforcing the one-way predictive relationship from fatigue reports to engagement levels, specifically in the short term. This unidirectional causality highlights the potential impact of immediate fatigue perceptions on the engagement with interventions but does not support the reverse; that is, engagement with the AiCoach does not appear to influence how employees will report fatigue levels thereafter.

6. CONCLUSION

In this research, we investigated workplace fatigue through the lens of self-reported metrics and engagement with an AI-based coaching tool that has yielded significant insights. The research found a clear and immediate relationship between employees' reported fatigue levels and their interaction with the AiCoach program. This underscores the potential of real-time monitoring and intervention in the management of workplace fatigue. The absence of a reciprocal predictive relationship suggests that while the AiCoach tool is a valuable resource in responding to fatigue, its impact on subsequent self-reported fatigue levels is not immediate or direct. This may indicate the need for a sustained and adaptive engagement strategy to see a measurable change in fatigue over time. Furthermore, the research emphasizes the importance of user-friendly interfaces in digital health tools. By allowing participants to navigate freely, make corrections, and track their progress, we minimized user errors and enhanced the accuracy of the collected data. The study's findings advocate for the integration of such design considerations in the development of digital health interventions.

In conclusion, the implementation of AI-assisted tools in the workplace for fatigue management holds promise. It encourages active employee participation and offers organizations a proactive approach to addressing fatigue. However, it also highlights the complexity of fatigue as a multifaceted issue that requires comprehensive strategies, including but not limited to digital intervention. Future research should aim to longitudinally assess the impact of such tools on fatigue and explore the integration of these tools with broader organizational health initiatives for a holistic approach to employee well-being and productivity.

Funding

Mobile Engine Kft (Mobilengine) (R&D project code GINOP-2.1.2.-8-1-4-16-2017-00032 received funding from the EU and the Hungarian government to develop the software solution and conduct the research.

Conflict of interest

The authors of this paper have no conflict of interest

REFERENCES

1. *Z. Yazdi and K. Sadeghniiat-Haghighi, "Fatigue management in the workplace," Industrial Psychiatry Journal, vol. 24, no. 1, p. 12, 2015. doi:10.4103/0972-6748.160915*
2. *Centers for Disease Control and Prevention (CDC), "Short sleep duration among workers--United States, 2010," MMWR. Morbidity and Mortality Weekly Report, vol. 61, no. 16, pp. 281–285, 2012.*
3. *C. Guilleminault, "Excessive daytime sleepiness: A challenge for the practising neurologist," Brain, vol. 124, no. 8, pp. 1482–1491, 2001. doi:10.1093/brain/124.8.1482*
4. *J. A. Ricci, E. Chee, A. L. Lorandeau, and J. Berger, "Fatigue in the U.S. Workforce: Prevalence and Implications for Lost Productive Work Time," Journal of Occupational and Environmental Medicine, vol. 49, no. 1, pp. 1–10, 2007. [Online]. Available: http://www.jstor.org/stable/44997095*
5. *I. Steinbuka, A. Clemenceau, A. Venema, S. van den Heuvel, and G. Geuskens, "Health and safety at work in Europe (1999-2007): A statistical portrait," 2010.*
6. *M. Palotai, M. Wallack, G. Kujbus, A. Dalnoki, and C. Guttmann, "Usability of a mobile app for real-time assessment of fatigue and related symptoms in patients with multiple sclerosis: Observational study," JMIR mHealth and uHealth, vol. 9, no. 4, 2021. doi:10.2196/19564*
7. *C. W. Granger, "Testing for causality: A personal viewpoint," Journal of Economic Dynamics and Control, vol. 2, pp. 329–352, 1980.*

ISSN (Online): 2455-3662
EPRA International Journal of Multidisciplinary Research (IJMR) - Peer Reviewed Journal
Volume: 9| Issue: 12| December 2023|| Journal DOI: 10.36713/epra2013 || SJIF Impact Factor 2023: 8.224 || ISI Value: 1.188

8. D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," Journal of the American Statistical Association, vol. 74, pp. 427–431, 1979.
9. O. Perski, A. Blandford, R. West, and S. Michie, "Conceptualising engagement with Digital Behaviour Change Interventions: A systematic review using principles from Critical Interpretive Synthesis," Translational Behavioral Medicine, vol. 7, no. 2, pp. 254–267, 2016. doi:10.1007/s13142-016-0453-1
10. M. Bradley and P. J. Lang, "Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings," 1999.
11. Z. Sedighi Maman, M. A. Alamdar Yazdi, L. A. Cavuoto, and F. M. Megahed, "A data-driven approach to modeling physical fatigue in the workplace using wearable sensors," Applied Ergonomics, vol. 65, pp. 515–529, 2017. doi:10.1016/j.apergo.2017.02.001
12. E. Borg, "On perceived exertion and its measurement," Doctoral dissertation, Psykologiska institutionen, 2007.
13. G. Borg, "Psychophysical scaling with applications in physical work and the perception of exertion," Scandinavian Journal of Work, Environment & Health, pp. 55-58, 1990.
14. F. Gamberale, "The perception of exertion," Ergonomics, vol. 28, no. 1, pp. 299-308, 1985.
15. G. A. Borg, "Psychophysical bases of perceived exertion," Medicine and Science in Sports and Exercise, vol. 14, no. 5, pp. 377-381, 1982.
16. B. Watt and R. Grove, "Perceived exertion: antecedents and applications," Sports Medicine, vol. 15, pp. 225-241, 1993.
17. C. Ritchie, "Rating of perceived exertion (RPE)," Journal of Physiotherapy, vol. 58, no. 1, p. 62, 2012.
18. L. B. Krupp, N. G. LaRocca, J. Muir-Nash, and A. D. Steinberg, "The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus," Archives of Neurology, vol. 46, no. 10, pp. 1121-1123, 1989.
19. A. Bunevicius, A. Stankus, J. Brozaitiene, S. S. Girdler, and R. Bunevicius, "Relationship of fatigue and exercise capacity with emotional and physical state in patients with coronary artery disease admitted for rehabilitation program," American Heart Journal, vol. 162, no. 2, pp. 310–316, 2011. doi:10.1016/j.ahj.2011.05.003
20. S. R. Aghdam, S. S. Alizadeh, Y. Rasoulzadeh, and A. Safaiyan, "Fatigue Assessment Scales: A Comprehensive Literature Review," Archives of Hygiene Sciences, vol. 8, no. 3, 2019.
21. S. Seneviratne, Y. Hu, T. Nguyen, G. Lan, S. Khalifa, K. Thilakarathna, et al., "A survey of wearable devices and challenges," IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2573-2620, 2017.
22. M. Haghi, K. Thurow, and R. Stoll, "Wearable devices in medical internet of things: scientific research and commercially available devices," Healthcare Informatics Research, vol. 23, no. 1, pp. 4-15, 2017.
23. S. M. Lee and D. Lee, "Healthcare wearable devices: an analysis of key factors for continuous use intention," Service Business, vol. 14, no. 4, pp. 503-531, 2020.
24. S. Beniczky, P. Karoly, E. Nurse, P. Ryvlin, and M. Cook, "Machine learning and wearable devices of the future," Epilepsia, vol. 62, pp. S116-S124, 2021.
25. M. J. Domínguez-Morales, F. Luna-Perejón, L. Miró-Amarante, M. Hernández-Velázquez, and J. L. Sevillano-Ramos, "Smart footwear insole for recognition of foot pronation and supination using neural networks," Applied Sciences, vol. 9, no. 19, 3970, 2019.
26. H. Dong, I. Ugaldey, and A. El Saddik, "Development of a fatigue-tracking system for monitoring human body movement," in 2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, May 2014, pp. 786-791.
27. A. Baghdadi, F. M. Megahed, E. T. Esfahani, and L. A. Cavuoto, "A machine learning approach to detect changes in gait parameters following a fatiguing occupational task," Ergonomics, vol. 61, no. 8, pp. 1116-1129, 2018.
28. T. J. Balkin, W. J. Horrey, R. C. Graeber, C. A. Czeisler, and D. F. Dinges, "The challenges and opportunities of technological approaches to fatigue management," Accident Analysis & Prevention, vol. 43, no. 2, pp. 565-572, 2011.
29. J. Fu, L. Ma, L. Tsao, and Z. Zhang, "Continuous measurement of muscle fatigue using wearable sensors during light manual operations," in Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management. Human Body and Motion: 10th International Conference, DHM 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part I 21, Springer International Publishing, 2019, pp. 266-277.
30. N. Vignais, M. Miezal, G. Bleser, K. Mura, D. Gorecky, and F. Marin, "Innovative system for real-time ergonomic feedback in industrial manufacturing," Applied Ergonomics, vol. 44, no. 4, pp. 566-574, 2013.
31. D. Battini, A. Persona, and F. Sgarbossa, "Innovative real-time system to integrate ergonomic evaluations into warehouse design and management," Computers & Industrial Engineering, vol. 77, pp. 1-10, 2014.
32. K. S. Tee, E. Low, H. Saim, W. N. W. Zakaria, S. B. M. Khialdin, H. Isa, and C. F. Soon, "A study on the ergonomic assessment in the workplace," in AIP Conference Proceedings, vol. 1883, no. 1, AIP Publishing, Sep. 2017.
33. M. S. Barim, M. L. Lu, S. Feng, G. Hughes, M. Hayden, and D. Werren, "Accuracy of an algorithm using motion data of five wearable IMU sensors for estimating lifting duration and lifting risk factors," in Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 63, no. 1, pp. 1105-1111, Nov. 2019.
34. F. Luna-Perejón, M. J. Domínguez-Morales, and A. Civit-Balcells, "Wearable fall detector using recurrent neural networks," Sensors, vol. 19, no. 22, 4885, 2019.
35. F. Luna-Perejón, L. Muñoz-Saavedra, J. Civit-Masot, A. Civit, and M. Domínguez-Morales, "AnkFall—Falls, falling risks and daily-life activities dataset with an ankle-placed accelerometer and training using recurrent neural networks," Sensors, vol. 21, no. 5, 1889, 2021.
36. S. Karvekar, M. Abdollahi, and E. Rashedi, "Smartphone-based human fatigue level detection using machine learning approaches," Ergonomics, vol. 64, no. 5, pp. 600-612, 2021.
37. S. R. Lamooki, J. Kang, L. A. Cavuoto, F. M. Megahed, and L. A. Jones-Farmer, "Challenges and opportunities for statistical monitoring of gait cycle acceleration observed from IMU data

for fatigue detection," in 2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob), pp. 593-598, Nov. 2020.

38. *M. Schmidt, C. C. Rheinländer, S. Wille, N. Wehn, and T. Jaitner, "IMU-based determination of fatigue during long sprint," in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, pp. 899-903, Sep. 2016.*
39. *Z. S. Maman, Y. J. Chen, A. Baghdadi, S. Lombardo, L. A. Cavuoto, and F. M. Megahed, "A data analytic framework for physical fatigue management using wearable sensors," Expert Systems with Applications, vol. 155, 113405, 2020.*
40. *A. Al-Saegh, S. A. Dawwd, and J. M. Abdul-Jabbar, "Deep learning for motor imagery EEG-based classification: A review," Biomedical Signal Processing and Control, vol. 63, 102172, 2021.*
41. *J. Zhang, T. E. Lockhart, and R. Soangra, "Classifying lower extremity muscle fatigue during walking using machine learning and inertial sensors," Annals of Biomedical Engineering, vol. 42, pp. 600-612, 2014.*
42. *A. Baghdadi, L. A. Cavuoto, A. Jones-Farmer, S. E. Rigdon, E. T. Esfahani, and F. M. Megahed, "Monitoring worker fatigue using wearable devices: A case study to detect changes in gait parameters," Journal of Quality Technology, vol. 53, no. 1, pp. 47-71, 2021.*
43. *J. Kuschan and J. Krüger, "Fatigue recognition in overhead assembly based on a soft robotic exosuit for worker assistance," CIRP Annals, vol. 70, no. 1, pp. 9-12, 2021.*
44. *A. Lambay, Y. Liu, P. Morgan, and Z. Ji, "A data-driven fatigue prediction using recurrent neural networks," in 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), IEEE, June 2021, pp. 1-6.*
45. *M. T. Darbandy, M. Rostamnezhad, S. Hussain, A. Khosravi, S. Nahavandi, and Z. A. Sani, "A new approach to detect the physical fatigue utilizing heart rate signals," 2020.*
46. *F. Nasirzadeh, M. Mir, S. Hussain, M. Tayarani Darbandy, A. Khosravi, S. Nahavandi, and B. Aisbett, "Physical fatigue detection using entropy analysis of heart rate signals," Sustainability, vol. 12, no. 7, 2714, 2020.*
47. *D. Cella et al., "Neuro-Qol: Brief measures of health-related quality of life for clinical research in neurology," Neurology, vol. 78, no. 23, pp. 1860–1867, 2012. doi:10.1212/wnl.0b013e318258f744*
48. *D. Amtmann et al., "Comparison of the psychometric properties of two fatigue scales in multiple sclerosis.," Rehabilitation Psychology, vol. 57, no. 2, pp. 159–166, 2012. doi:10.1037/a0027890*
49. *P. Pedrelli et al., "Reliability and validity of the symptoms of depression questionnaire (SDQ)," CNS Spectrums, vol. 19, no. 6, pp. 535–546, 2014. doi:10.1017/s1092852914000406*
50. *M. W. Johns, "A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale," Sleep, vol. 14, no. 6, pp. 540–545, 1991. doi:10.1093/sleep/14.6.540*
51. *G. Godin and R. J. Shephard, "A simple method to assess exercise behavior in the community," Canadian Journal of Applied Sport Sciences, vol. 10, no. 3, pp. 141-146, Sep. 1985. [Medline: 4053261]*
52. *C. S. Carver and T. L. White, "Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/Bas Scales.," Journal of Personality and Social Psychology, vol. 67, no. 2, pp. 319–333, 1994. doi:10.1037/0022-3514.67.2.319*
53. *Mobilengine, https://mobilengine.com/*