Occupational Fatigue Prediction for Entry-Level Construction Workers in Material Handling Activities Using Wearable Sensors

Wonil Lee

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2018

Committee:

Edmund Seto (Chair)

Peter W. Johnson

Ken-Yu Lin

Program Authorized to Offer Degree:

School of Public Health

University of Washington

Abstract

Occupational Fatigue Prediction for Entry-Level Construction Workers in Material Handling Activities Using Wearable Sensors

Wonil Lee

Chair of the Supervisory Committee:

Edmund Seto, PhD, MS

Department of Environmental and Occupational Health Sciences

Research on the measurement and prediction of occupational fatigue of construction workers using wearable sensors has been carried out using different types of sensor technologies and measurement variables. Previous studies have demonstrated promising results using wearable sensors for fatigue prediction, suggesting that they may have practical use as tools for the prevention of fatigue management on worksites. However, there are no clear guidelines on the type of sensors to use in fatigue management or on the relevant sensed variables. Moreover, the collection and processing of wearable sensor data should be sufficiently simple for safety professionals to use in practice. The current study aimed to address these challenges by using several of the most active wearable sensor technologies in occupational fatigue research—

actigraphy and electrocardiogram (ECG) sensors—to obtain different variables from participants performing simulated construction tasks in laboratory conditions.

A total of 22 participants participated in the experiment. Of these, 19 were exposed to different task workloads and completed four repetition measurements, while three participants only completed one or two experiment session(s). A total of 80 observations obtained by the experiment were used for the analysis. Stepwise logistic regression was used to identify the most appropriate and parsimonious fatigue prediction model. Among the different variables collected, heart rate variability (HRV) measurements, standard deviation of NN intervals (SDNN) and power in the low frequency range (LF) were found to be useful in predicting fatigue. Both the fast Fourier transform (FFT) and the autoregressive (AR) analysis in the frequency domain analysis methods were employed. Log transformed LF obtained by AR analysis method was found to be more suitable for daily management of worker's fatigue, while the SDNN was useful in weekly fatigue management. This study contributes to the body of knowledge on the use of wearable technology for the management of fatigue among construction workers.

Table of Contents

Table of Figures

Table of Tables

Acknowledgment

I would like to express my sincere gratitude to Dr. Edmund Seto for providing me insights from the beginning to the end of the Master's degree. I will try to continue my journey, resembling your genuine passion for exploring and scientifically inquiring what interests me as a researcher. I would like to thank Dr. Peter W. Johnson for your teaching. You have taught me the lessons from your experience, research methodology, ergonomics analysis technologies and occupational biomechanics research. Dr. Ken-Yu Lin, thank you for your generous support, encouragement for the research and your wisdom. I also appreciate Dr. Giovanni Migliaccio for approving me to use the data that initiated from one of our projects.

Lastly, I would like to thank my parents and parents-in-law, who encouraged me and showed everlasting love. Without your endless support, I could not have pursued and earned my Master's degree in Public Health program. I appreciate my loving wife who gave me an infinite love and always had faith in me. I have been able to study with courage and to complete this thesis with gratitude due to your support.

1. Background and Significance

In 2005, a hydrocarbon vapor cloud of combustible gases was released and ignited by a running diesel truck in a British Petroleum (BP) facility in Texas City, which was the America's third-largest refinery. The explosion killed 15 workers and caused more than 170 injuries (U.S. Chemical Safety and Hazard Investigation Board, 2007). The hydrocarbon vapor leaked from the isomerization unit. The alarms warning of the overpressure were not operated properly in the isomerization process. While this was not the direct cause of the disaster, the United States Chemical Safety and Hazard Investigation Board found that staff fatigue due to working overtime to cover for reduced staff also contributed to the Texas City refinery accident. It recommended the development of guidelines for managing staff fatigue during the work shift. This case study illustrates how fatigue is a key antecedent to occupational fatalities and injuries in hazardous working conditions.

Many occupational injuries and incidents that have occurred, and potential precursors to these events, have been associated with occupational fatigue. Zhang et al. (2015) found a statistically significant correlation between fatigue and physical and cognitive function in a study of 606 U.S. construction workers. They also found that construction workers who often feel fatigue had difficulties in physical and cognitive functioning, compared to groups that did report any fatigue. Fatigue reduces concentration and results in poor decision making, which ultimately has adverse effects on safety performance (Haslam et al., 2005). Nardone, Tarantola, Giordano, and Schieppati (1997) found that fatigue caused loss of balance, measured by the frequency of body oscillation in experiments with subjects aged 18 to 39 years. Loss of balance has been reported as the initial cause of slips, trips, and fall fatalities among roofing workers (Hsiao & Simeonov, 2001). Fatigue has a negative effect on the gait stability performance of firefighters, leading to accidents due to slips, trips, and falls (Park et al., 2011). The causes of safety accidents include workers' accumulated fatigue due to continuous work activity (Nag & Patel, 1998). Fatigue leads to impaired performance capabilities, such as slowed or incorrect responses, eventually leading to accidents (Williamson et al., 2011). Impaired cognitive performance due to inadequate sleep can increase the occurrence of fatigue-related errors in the workplace (Caruso & Hitchcock, 2010). Fatigue causes deterioration of overall human performance, which increases human error and reduces hand-eye coordination and memory (Murray & Thimgan, 2016). Murray and Thimgan also stated that fatigue causes inattentional blindness by slowing down reaction time and reducing the capacity to manage stimuli.

As fatigue is a significant cause of incidents and injuries on construction worksites, research has focused on the etiological factors of construction worker fatigue and its effects. Fatigue increases the likelihood of human error, such as slips and lapses of caution on construction jobsites, and causes long-term adverse health effects, such as hypertension and cardiovascular disease (Hallowell, 2010). Safety professionals and industrial hygienists on construction jobsites need to optimize workers' schedules to prevent excessive workloads. High-rise building construction workers are more likely to exhibit fatigue symptoms than ground-level construction workers (Hsu, Sun, Chuang, Juang, & Chang, 2008).

In the hierarchy of controls, the elimination of potential hazards is known to be more effective than other bottom levels of control methods, such as substitution, engineering controls, administrative controls, and personal protective equipment (Centers for Disease Control and Prevention, 2016). The proactive prediction and detection of worker fatigue status is the preferred approach to prevent injuries and accidents due to human errors caused by fatigue.

Therefore, monitoring the physiological status of workers in real time may help to prevent workers from experiencing excessive physical exertion, cardiovascular load, and potential accidents.

Fatigue prediction research has been conducted for a variety of occupations that would benefit from off-the-shelf sensor tools, including actigraphy, heart rate monitors, electroencephalography (EEG) (Zhu, Jankay, Pieratt, & Mehta, 2017). Researchers have developed a sleep, activity, fatigue, and task effectiveness model and have validated the model with flight crews, railroad crews, and military personnel (Hursh et al., 2004; Hursh, Raslear, Kaye, & Fanzone, 2006; Roma, Hursh, Mead, & Nesthus, 2012). However, these fatigue prediction studies have only considered individual factors, such as sleep duration, sleep efficiency, and circadian processes and have not fully considered job and task conditions, such as workload.

Additionally, a number of fatigue detection studies have been conducted using computer vision–based facial detection or wavelet analysis of heart rate variability (HRV) for automobile drivers (Li & Chung, 2013; Patel, Lal, Kavanagh, & Rossiter, 2011) and the effect of intervention (e.g., acupuncture) on the HRV and fatigue status was conducted with the driving simulation task (Li, Wang, Mak, & Chow, 2005). While these studies predict task-induced fatigue level, they would not be applicable to construction workers who perform more physically demanding tasks. Moreover, only factors relating to the types of tasks or jobs were considered in predicting fatigue. However, according to the National Institute for Occupational Safety and Health (NIOSH) approach (Hammer & Sauter, 2013; Sauter, 1999) to job stress, individual life factors can intervene in the relationship between job conditions and a worker’s safety and health and, therefore, parameters related to both work tasks and life factors should be considered.

Caterpillar Inc., which is the world's largest construction equipment manufacturer, has commercialized a wrist-worn CAT® Smartband capable of fatigue assessment that is part of a system for a scheduling scheme to avoid worker fatigue in the workplace (Caterpillar Inc., n.d.). Caterpillar Inc. has a partnership with the distribution of the smart band that was originally developed by Fatigue Science, the same company that introduced the Readiband™ Actigraph that has been validated for its use on sleep and wake detection (Russell et al., 2000). This wristband quantifies sleep factors, such as sleep interruption, cumulative sleep debt, sleep onset, and wake times as predictive model input variables to predict the fatigue state of workers based on the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) fatigue model developed by fatigue scientists (Hursh et al., 2004). Sleep deprivation and circadian rhythm disruption are also known as major contributing factors of fatigue syndrome (Lerman et al., 2012).

Zhu et al. (2017) found that heart rate (HR)/electrocardiogram (ECG) and ActiGraph sensors were most suitable for occupational fatigue research among different wearable sensor metrics. A review of the literature found that 57% of studies used these two types of sensors to measure physical fatigue. In the case of ActiGraph, it was frequently used for fatigue research to monitor sleep length, sleep efficiency, activity counts, sleep latency (Zhu et al., 2017). HR trends, high frequency power variability, RR interval variability, and mean % HR max were frequently used as the metrics measured by the ECG wearable sensors (Zhu et al., 2017). Through investigation of historical incident and injury data, fatigue was identified as one of the main causes of accidents in the construction industry. In a survey conducted in oil and gas construction companies, stakeholders in construction stated that they perceived fatigue to be the most critical accident risk (Chan, 2011). It is recognized that the intensity of the workload of workers in the construction industry is heavy (Hartmann & Fleischer, 2005). Construction workers' level of exposure to physical fatigue can also be high, depending on the elements of the work environment. Therefore, the early prediction and detection of worker fatigue levels are crucial in preventing fatal incidents caused by fatigue. Many studies of transportation workers (drivers and railroad workers), pilots, warfighters, firefighters, and hospital employees have predicted other workers' levels of fatigue by utilizing wearable technologies measuring their physiological state (Tran, Wijesuriya, Tarvainen, Karjalainen, & Craig, 2009; Patel et al., 2011; Hursh et al., 2004). However, the types of fatigue that can be detected among construction workers are still rarely studied even though they are exposed to a high risk of fatigue.

2. Specific Aims and Research Scope

There are no clear guidelines for the type of wearable sensor to use and which sensor variables to collect for the practical management of fatigue in construction workers. Data collection and processing should be simple, so that the methods can be adopted in practice by safety professionals. Moreover, it is not recommended to manage fatigue based solely on the sensors and their measurement data; it is necessary to combine these with traditional management methods, such as surveys and observations. This study is aimed at addressing these challenges by applying several of the most commonly used wearable-sensor technologies from the occupational fatigue research.

The term fatigue is usually defined as either muscular fatigue or general fatigue (Grandjean & Kroemer, 1997). The present study focuses on general bodily fatigue when the entire organism is overloaded. Acute fatigue due to the physical or mental labor of construction workers is the primary concern, rather than chronic fatigue in occupational settings (Techera, Hallowell, Stambaugh, & Littlejohn, 2016). Thus, the current study focuses on the prediction of acute fatigue based on subjective and objective fatigue measurement methods. Several studies have been conducted to measure fatigue using wearable sensors (Aryal et al., 2017). However, no study has been conducted to determine which variables are the most convenient and easiest for the prediction of fatigue. For practical reasons, types of sensors with the current off-the-shelf versions that are impractical to give to construction site workers (e.g., EEG and electrooculography) are excluded from this study. Therefore, physiological status and activity monitors were used. The respiratory rate was also not considered as a predictor of fatigue as there are no studies reported in the occupational fatigue research that employed respiratory rate (RR) as the metric to assess the fatigue status (Zhu, et al., 2017).

This study will investigate the following research questions:

- Do the measured variables derived from wearable sensors yield useful information that is associated with fatigue?
- Which factors among heart rate, heart rate variability, physical activity, and sleep measurements are more useful for predicting occupational fatigue?
- Can a wearable sensor be used to predict fatigue, yet still be effectively worn by a worker with minimal discomfort and interference with work?

In this study, the word "predicting" is used loosely, as the study is most interested in providing understanding about which survey and sensor variables are most associated with fatigue. The study uses multivariate statistical models (i.e., predictors or independent variables in the statistical model).

Although environmental factors, such as noise, vibration, and heat are also important in the etiology of fatigue (Yi & Chan, 2015), this research limits the scope to the prediction of sleep impairment and task-induced fatigue through ActiGraph and ECG. Therefore, this research controlled for the potential effects of noise, vibration, and heat by conducting the experiment in a laboratory environment. The experiment was designed to measure fatigue induced by repetitive lifting and carrying materials. Thus, the outcome of the research will contribute to understanding of how safety professionals/managers can mitigate fatigue resulting from material handling, which is a major cause of overexertion among construction, as well as a leading cause of nonfatal injuries and work-related musculoskeletal disorders (The Center for Construction Research and Training, 2013).

3. Methods

3.1. Data

This study used data collected for an experimental study conducted as part of a Doctor of Philosophy dissertation in the Built Environment Ph.D. program at the University of Washington, with the preliminary title "*Job Demands-Resources, Burnout and Performance of Construction Workers*". The Ph.D. work differs in that it is focused on understanding worker performance, whereas the current research extracts different research variables from the data and applies different analysis methods to understand worker fatigue.

The original data set was collected to examine how the unbalance between task demands and resources influences the exhaustion and engagement level of construction workers involved in repeated lifting and lowering of materials, and how these factors affect workers' productivity and ergonomic safety behaviors. To achieve this aim, a lab-based study was conducted using experiments with lifting and lowering materials during construction installation tasks.

3.2. Participants

The target population consisted of individuals at the entry level of construction or with no construction experience. Participants were recruited from a local pre-apprenticeship construction education program and were university students. Martin, Chalder, Rief, & Braehler (2007) reported that somatoform symptoms, such as back pain, pain in the arms/legs, joint pain, and headaches correlated closely with the fatigue state of the general population. To confirm their admissibility to the experimental study, the Medical History Questionnaire and a Physical Activity Readiness Questionnaire (PAR-Q) were administered to participants. Participants were excluded if there was any other physical reason preventing them from taking part in the experiment, such as asthma or orthopedic problems. Twenty-two healthy participants were enrolled in the experiments.

3.3. Design

Participants were asked to perform a task: installing pavers and pedestals of raised decks repeatedly for one hour. The working space was divided into three zones: installation area, traveling area, and material inventory area. To allow the subject to keep installing pavers and pedestals over one hour with a limited number of materials, the working space zone was subdivided into two installation areas and traveling areas, based on the material inventory area (Figure 1).

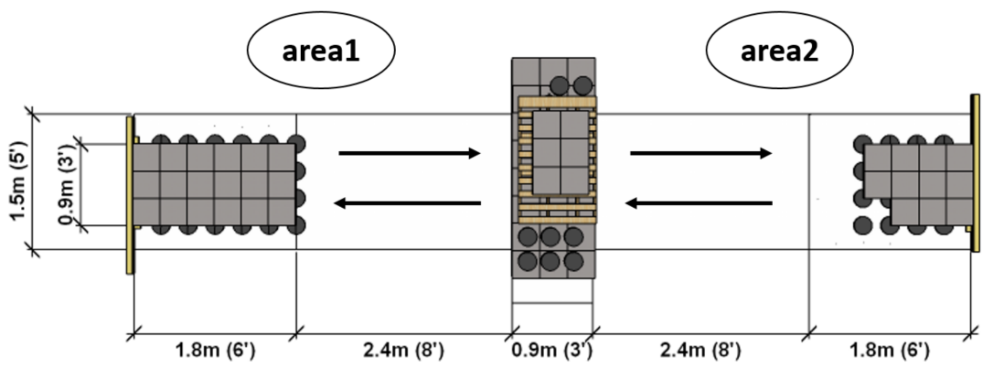

Figure 1. Scheme of experiment working space

Traveling distance between the installation area and the material area was 2.4 meters (8 feet). Participants moved a unit comprising a set of pedestals and a paver from the material area in the middle of working space and made a raised deck with a paver combination of six rows and three columns in the installation area, as shown on the left side of Figure 1. A unit of paver is supported by four pedestals. After completing the tasks in the installation area on one side (area 1, Figure 1), participants immediately went to the installation area on the other side (area 2, Figure 1) and installed a deck of the same size. During this time, an assistant disassembled the materials installed in the first phase and returned them to the material area between area1 and area 2. This experimental design enabled the participants to conduct material installation continuously for one hour, reaching a certain level of workload and physiological strain that eventually resulted in fatigue. Four Internet protocol cameras recorded the experimental working space and participants for one hour (Figure 2). This allowed the participants' production rate to be measured.

Figure 2. Example of views from the recorded videos

In order to reduce the cost and time required for the recruitment of participants, the participants were encouraged to engage in four experimental sessions on different days with differently designed workloads (with 1- or 2-week intervals) so that the measurements were repeated. In each of the four experimental sessions, the participants performed tasks of different workload levels by changing the weight of pavers and the pallet height in the material area. Figure 3 shows that the different heights of the pallets and weights of the pavers and pedestals exposed the same participants to different workload levels. Assigned task loads were also varied, with different weights of pavers [2 lbs (≈1kg) vs. 18 lbs (≈8kg)].

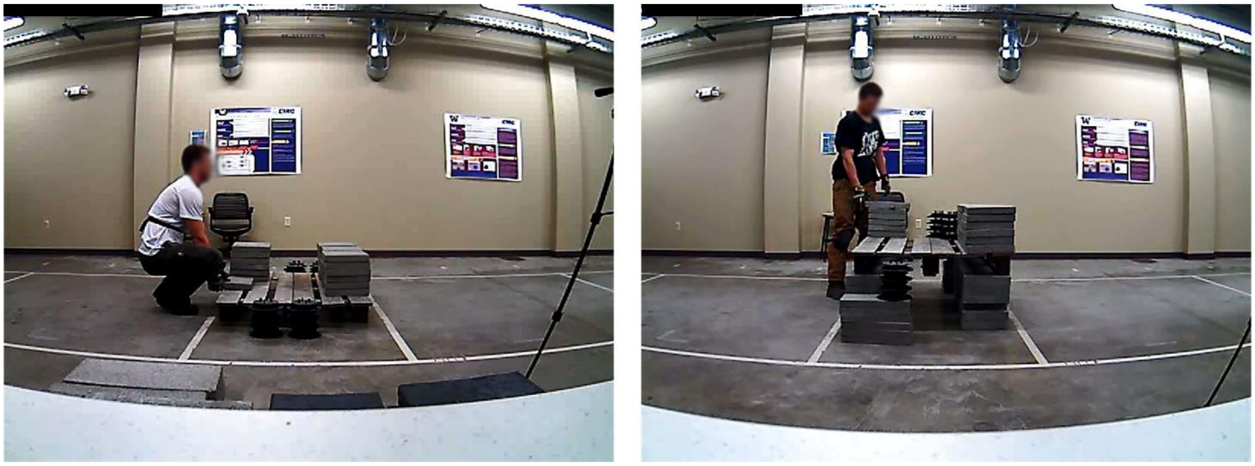

Figure 3. The different height of pallet placement to vary the workload

There are known factors that influence fatigue status, measured by both subjective and direct methods. Participants were asked to refrain from having a meal, consuming caffeine, taking medication or smoking for at least two hours prior to participating in the experiment, based on guidelines by Medicore (n.d.). Before and after the assigned task, the variables of fatigue and fatigue prediction indicators were obtained using a survey and wearable sensors. The data collection process is presented in Figure 4. The variables collected are described in detail in the following sections.

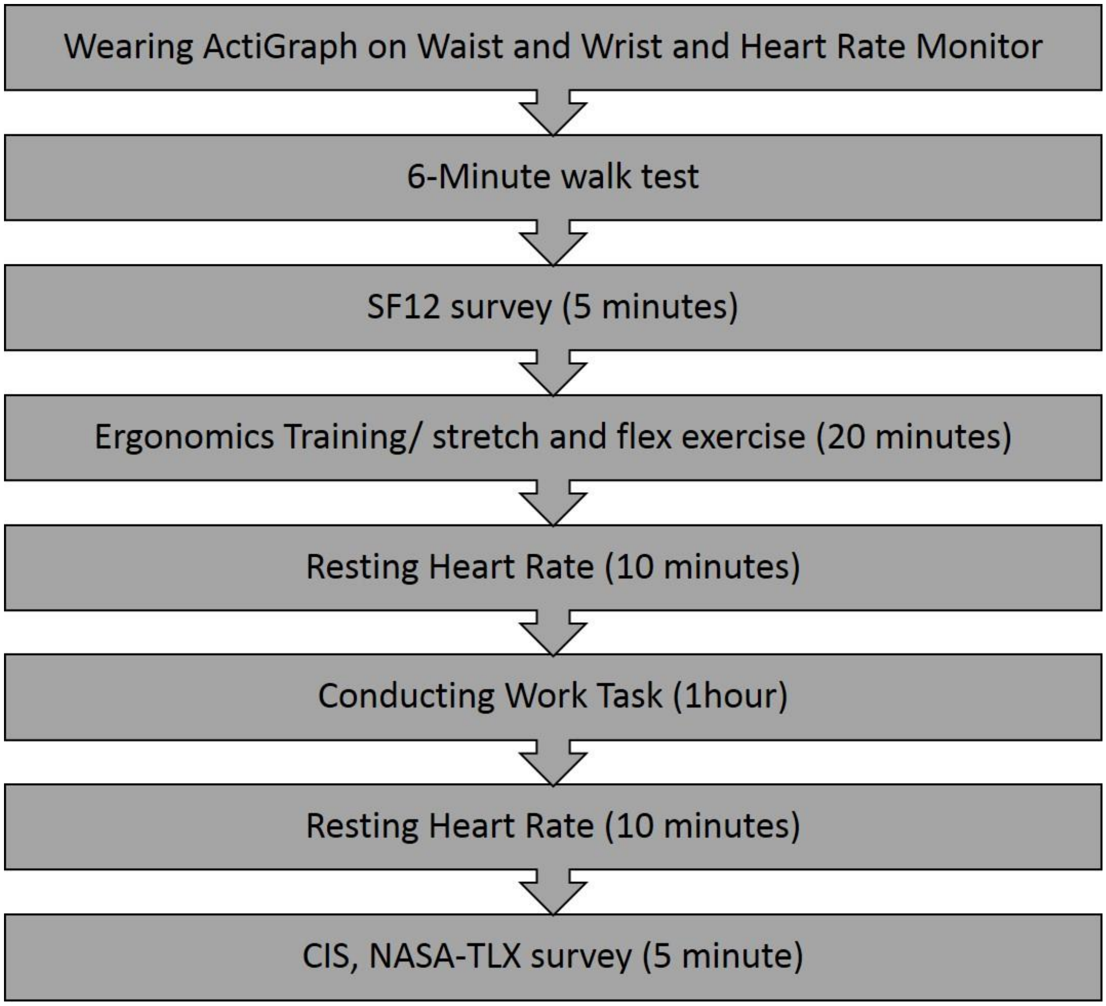

Figure 4. Data collection process

3.4. Fatigue measurement

Different subjective measurements of fatigue have been applied to occupational fatigue research. They are the Checklist Individual Strength (CIS), Fatigue Severity Scale (FSS), Fatigue subscale of the Visual Analogue Scale (VAS-F), the Need for Recovery Scale (NRS), the Swedish Occupational Fatigue Inventory (SOFI), the Fatigue Assessment Scale (FAS), and the Multidimensional Fatigue Inventory (MFI). These survey tools are used to measure the fatigue of the working population (Techera, Hallowell, Marks, & Stambaugh, 2016). Yamazaki et al. (2007) measured the fatigue of Japanese manufacturing workers using the SF-36 Health Survey vitality domain scale. Aryal et al. (2017) used Borg's Rating of Perceived Exertion (RPE) to measure physical fatigue. Zhang et al. (2015) developed the Fatigue Assessment Scale for Construction Workers (FASCW), a fatigue scale specifically applicable to construction workers.

Among the multitude of survey tools for measuring subjective levels of fatigue, a short version of the CIS (Beurskens et al., 2000) was selected for this study as the main outcome measure of fatigue because its subscale, physical fatigue, was the main focus of this research. Specifically, eight subscales of the CIS questionnaires (a.k.a., CIS8R) directly related to physical fatigue level were administered. The assessed statements included: "I feel tired," "Physically, I feel exhausted," "I feel fit," "I feel weak," "I feel rested," "Physically, I feel I am in a bad condition," "I get tired very quickly," and "Physically, I feel in good shape." Eight items were selected and used to assess subject fatigue status among the total 20 items (a.k.a., CIS20R) in the full CIS survey. In each item, participants chose the level of their current status from a seven-point Likert scale, and the sum of the selected scale values was calculated to determine each participant's fatigue score. The statements "I feel fit," "I feel rested," and "Physically, I feel in good shape" were scored using a reverse Likert scale. Therefore, a higher score on the CIS8R indicates a higher level of fatigue. Participants completed the CIS8R after task execution and resting heart measurement (Figure 4).

The CIS8R has been validated in other contexts. In a study of patients with rheumatoid arthritis (van Hoogmoed, Fransen, Bleijenberg, & van Riel, 2008), the Cronbach's alpha of CIS8R was 0.92, and the reliability coefficient was 0.81. Subscale's reliabilities estimated that while using Cronbach alpha, physical fatigue items (i.e., CIS8R) were 0.88, reduced concentration was 0.92, reduced motivation was 0.93 and the reduced physical activity level was 0.87 (Vercoulen, Alberts, & Bleijenberg, 1999).

3.5. Research variables to predict fatigue

3.5.1. Survey measurements

Age, gender, height, and weight to estimate body mass index were obtained as predictors of fatigue. Male gender was coded as 0 and female was coded as 1 in the dataset. Participant perceived health status was also considered a predictor of acute fatigue based on the 12-Item Short Form Health Survey (SF12) survey conducted before performance of the task. The SF12 survey (the short version of SF36) is specifically focused on the responder's subjective health status. Ware, Kosinski, & Keller (1996) validated the short version of SF12, and found that it can successfully measure the level of health and quality of life of participants.

A significant association between workload and fatigue has been reported in previous studies (MacDonald, 2003). To measure subjective workload, participants were administered the NASA Task Load Index (NASA-TLX), which consists of 6 subscales of task load: physical, temporal, mental, effort, frustration, and performance level (Hart & Staveland, 1988).

3.5.2. Wearable sensors measurements

ECG monitoring (e.g., heart rate monitoring) was conducted in this study, as it has been found in other studies to be useful in understanding fatigue (Zhu et al., 2017). Beat-to-beat RR-interval was collected using a Zephyr™ BioHarness 3 (Medtronic, Minneapolis, MN) with 1,000 Hz sampling frequency. The BioHarness system consists of a compact physiological monitoring module and a side strap that connects to the module. The strap includes a fabric sensor to measure ECG and breathing rate, and it transmits the acquired data to the sensor module. The sensor module contains a 3-axis accelerometer as well as memory for data storage. The BioHarness system provides measures such as heart rate, breathing rate, temperature, and body posture. Each participant selected a small or large strap, depending on chest size, and wore a BioHarness inside a garment so that the ECG sensor would have direct contact with the skin. According to the manufacturer's recommendation, the sensor modules were placed in the ideal position, which is in the apex of the rib curvature below the left arm of each participant.

Energy expenditure measured by direct or indirect calorimetry can also be used to evaluate the physical fatigue of construction workers (Abdelhamid & Everett, 2002). The activity level of the workers was assessed by having them wear the ActiGraph GT9X Link unit (ActiGraph, LLC., Pensacola, Florida). Because the participants were asked to wear the ActiGraph GT9X after hours, and while at home and during sleep, the device measured sleep quality. The ActiGraph includes an inertial measurement unit (IMU) that consists of an accelerometer, a gyroscope, a magnetometer, a primary accelerometer which is a separate sensor with the IMU, a wear-time sensor and a liquid-crystal display window (which displays the date, time and feedback information). To measure energy expenditure, each participant wore an ActiGraph with a wristband on the nondominant wrist and an additional ActiGraph on the waist (attached with a waist belt and pouch). The sample rate for the IMU data was set at 100 Hz. The 3-axis accelerometer data were used for data analysis.

The participants received an ActiGraph the night before the experiment began, after which they performed the simulated material-handling activities. The participants wore the ActiGraph on their nondominant wrists from when they received it until just before falling asleep. The participants were not required to wear the ActiGraph during showers or exercise; however, they were instructed to wear it before going to bed. The participants were advised to wear the ActiGraph on the wrist until the next day, when they would participate in the experimental session in the laboratory and return the ActiGraph to the researcher. Then, participants received new ActiGraph sensors for a data collection during the experiment session.

3.6. Other factors influencing individual fatigue level

Participant circadian rhythm, such as the time of day, affects fatigue levels (Murray & Thimgan, 2016). Therefore, the time of conducting the experiment was recorded and included as a predictor in the fatigue prediction model. There were three time slots when subjects performed assigned tasks. If subjects performed tasks between 6 am and 12 pm, the data were recoded as 'Morn'. If subjects performed tasks between 12 pm and 6 pm, the data were recoded as 'Aft'. If subjects performed tasks between 6 pm and 10 pm, the data were recoded as 'Even'.

A six-minute walk test, which is one of the common physical performance and endurance tests (Bennell, Dobson, & Hinman, 2011), was conducted to measure each participant's physical capacity affecting the occurrence of fatigue after conducting the simulated construction activity. Following the approach of Seynnes et al. (2004), the 15-meter track protocol of the six-minute test was designed and conducted for this study (Figure 5).

Figure 5. Six-minute walk test setup (15-meter track protocol)

3.7. Data Analysis

3.7.1. Missing data

At each of the 80 data points, three missing data were identified for sleep quality and sleep time variables, respectively. This occurs when the sleep detection function of the ActiGraph software is used and the life of the battery is exhausted earlier than expected or when the subject does not wear the ActiGraph while sleeping and the data cannot be collected. Missing data were found to be completely random, and imputation was done using multivariate imputation by the chained equations (MICE) R package (van Buuren et al., 2014) because missing data were under 5% of all data points. This research used predictive mean matching among MICE options. Missing values were imputed to approximate the predicted mean through the random rendering of observed values. A total of five imputed datasets were created and used for the main data analysis of the dataset.

3.7.2. HR and HRV measurements

HR measurements: Relative heart rate (RHR), which is a normalized HR estimate reflecting each participant's age and resting HR, was used an indicator of occupational fatigue. Janssen, Van Oers, Van der Woude, and Hollander (1994) used RHR as the relative level of physical strain of wheelchair users in daily life. The RHR was estimated as an indicator for assessing physical strain based on the equation that was introduced by Rodahl (1989): RHR (%) = [(HR-HRrest)/(HRmax-HRrest)]×100, where HRmax is the predicted maximum HR, and HRrest is the measured average HR with the participant sitting on a chair for 10 minutes. To estimate the participant's maximum HR, the method developed by Tanaka, Monahan, & Seals (2001) method was used: HRmax= 208-0.7×age. Dehydration during work is known to be a factor that increases fatigue (Cheuvront, Carter, & Sawka, 2003), and this experiment was controlled by allowing participants to consume sufficient water as needed during work. As an explanatory variable of occupational fatigue, heart rate recovery (HRR) was selected as one of the heart rate measurements. HRR was calculated as the absolute heart rate from peak levels upon completion of the task minus the heart rate two minutes after task completion.

HRV measurements: Camm et al. (1996) provided a guideline for the standard measurement of HRV, physiological interpretation, and clinical applications. In the time domain analysis, the mean HR, the standard deviation of NN (or RR) intervals (SDNN), and the square root of the mean squared differences of successive NN (or RR) intervals (RMSSD) were estimated as the statistical processing of the NN (or RR) interval change during the recording time. Kang et al.

(2004) used the SDNN and RMSSD as the parameters of HRV measurement to investigate their association with job stress among the shipyard male workers. In the analysis of the frequency domain, the method of analyzing the change of the RR interval into the waveforms of each frequency band was summarized as total power, very low frequency (VLF), low frequency (LF), high frequency (HF), normalized HF, normalized LF, and LF/HF ratio. According to Camm et al., the HRV frequency bands were calculated as follows. LF was the power spectrum in the frequency range 0.04 to 0.15 Hz. HF was the power spectrum with a frequency range of 0.15 to 0.4 Hz. This was the default range values of Kubios HRV analysis software (Premium version 3.0.2). The normalized units of LF and HF were calculated using the following formulas, respectively: LF(n.u) = LF(ms^2)/[Total Power(ms^2)-VLF(ms^2)]×100, and HF(n.u) = HF(ms^2)/[Total Power(ms^2)-VLF(ms^2)]×100 where the Total Power includes a power spectrum of 0-0.4 Hz, and VLF includes 0-0.04 Hz (Camm et al., 1996).

Tarvainen, Lipponen, Niskanen, & Ranta-aho (2014) stated that the LF component is known that reflects both the activities of the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS), but it could be used as an index of sympathetic nervous activity with the normalized value of the LF component. Parasympathetic influences of the LF component are present when a participant's respiration rates fall lower than seven breaths per minute (Medicore, n.d.), which is a very rare case in the material-handling activities. Thus, in this current study, the absolute value of the LF component can be used as an index of SNS activity as well. Usually, as Medicore (n.d.) informed the reduced LF indicates loss of energy, fatigue and insufficient sleep and Dishman et al. (2000) reported that the healthy people who have high stress levels and fatigue showed the reduced level of LF power. HF is an indicator of the activity of the PNS, and reduced PNS is related to the electrical stability of the heart under stress, anxiety, and panic

(Medicore, n.d.). The LF/HF ratio, which reflects the balance of the autonomic nervous system as a whole (Camm et al., 1996), and the ratio close to 1 indicate that the autonomic nerves were relatively well-balanced. If the LF/HF is high, it represents the autonomic overactivity with increased sympathetic nervous activity (Kang et al.,2004).

The heart anatomically consists of right and left ventricles and right and left atriums and linked by the aorta and pulmonary veins. Through the activities of the heart, including atrial and ventricular depolarization and repolarization, blood is pumped to the lungs and systemic arteries, and finally all tissues of body are supplied with oxygen and nutrients (Klaassen & Watkins, 2010) which the body can eventually utilize for necessary processes. As illustrated in Klaassen and Watkins, electrical currents generated during atrial depolarization, ventricular depolarization and deflection corresponding to ventricular repolarization were recorded on the ECG. The electrical activity by atrial depolarization, ventricular depolarization, and repolarization of the heart produces a QRS complex (Figure 6), and the HRV implies a variation in the time interval between the two successive R points, where the R point is the peak point of the QRS complex. The pattern of a cycle of the electrocardiogram is composed in the order of PQRST waves as illustrated in Figure 6, and the interval between R peaks is the RR (or NN) interval.

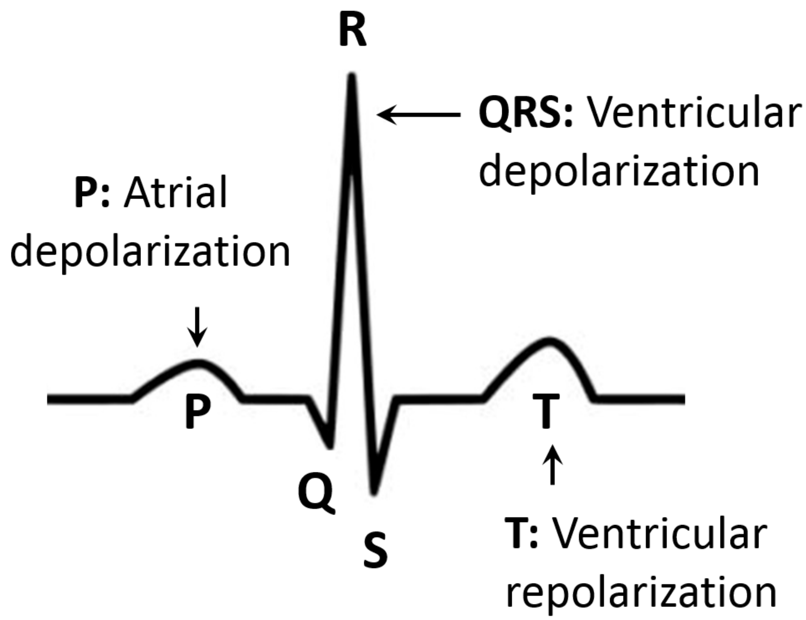

Figure 6. Electrocardiogram records during atrial depolarization, ventricular depolarization and ventricular repolarization (adapted from Klaassen & Watkins, 2010)

The changes in heartbeat interval is the most common indicator of the activity of the autonomic nervous system, reflecting sympathetic and parasympathetic activity. Therefore, because the changes in heartbeat interval is closely related to the activity of the autonomic nervous system, the autonomic nervous system activity can be estimated by analyzing the heartbeat interval. The degree of change in the heartbeat interval is referred to as HRV. HRV is a technique that numerically quantifies the characteristics of measured heartbeat intervals. It enables the evaluation of autonomic nervous system activities by analyzing the statistical characteristics of heartbeat intervals, HRV, and frequency characteristics. In the frequency domain analysis method, information on the biological activity of a certain period of coordinating the heartbeat is provided, so it is possible to intuitively obtain information on the human body’s activity. Fast Fourier transform (FFT) and the autoregressive (AR) analysis are analysis methods in the frequency domain. It is known that there is a discrepancy between the results of the two methods among healthy subjects (Pichon, Roulaud, Antoine-Jonville, de Bisschop, & Denjean, 2006) and non-healthy subjects who are suffering from diabetes (Chemla et al., 2005). Thus, for the current research, occupational fatigue prediction models were investigated by adding and removing (i.e., stepwise regression) both FFT and AR measurements into multivariate models.

FFT is a mathematical method for reconstructing and representing time-domain signals using the basic function of a myriad of frequencies. Introduced by the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, it is the most commonly used frequency analysis technique (Camm et al., 1996). FFT is widely used because it can intuitively provide information on frequency components, but spectrum leakage occurs when a signal of a limited length is handled (Acharya, Joseph, Kannathal, Lim, & Suri, 2006).

The AR model not only provides a way to predict frequency components with short time-domain signals, but it can also reduce spectral leakage using windowing (Jaffe & Fung, 1994). The AR model order value of 16 was selected as the default setting for the autoregressive model of Kubios HRV analysis software (using the premium version 3.0.2) (Tarvainen, Lipponen, Niskanen, & Ranta-aho, 2017). Boardman, Schlindwein, & Rocha (2002) recommended not using a model order value that is less than 16 for the HRV frequency domain analysis if short segments of tachograms are used in the analysis.

Data from the construction experiment were analyzed by dividing measurements into 5-minute segments which is a default setting of the Kubios HRV software. The time-domain and frequency-domain parameters of each 5-minute segment were obtained from the 60-minute RR intervals. The final yield included averages for each of the 12 data points (i.e., 60 minutes divided by 5 minutes) and each of the variables used in the logistic regression analysis. Outliers were managed using a threshold based artefact correction algorithm available in the Kubios HRV analysis software. The option of the strong level was applied for an artifact detection and removal, as Garza et al. (2015) used. If the RR intervals were above or below 0.15 seconds when compared to the local average, and these values were detected as an artifact, they were replaced using cubic spline interpolation (Tarvainen et al., 2017).

The change rate of the NN interval is continuously changed within a certain standard deviation range and is represented by the RR interval tachogram. The analysis of this is HRV. The standard deviation of the total NN intervals (SDNN) is an indicator of how much the HR changes during recording time. When the SDNN is large, it indicates that the HRV signal is irregular, alternatively, the HRV signal is monotonous. Indicators constituting HRV were classified into the time domain and frequency domain methods. The time domain analysis used SDNN and RMSSD, the mean square root of the square of the difference between SDNN of the total RR interval (i.e., time between two consecutive R waves) and the adjacent RR interval. SDNN is an indicator of how much the HR changes, and RMSSD is an indicator of the activities of the cardiac parasympathetic nervous system.

Frequency range analysis was performed by analyzing the change of RR interval between the same heartbeats by waveform and measuring HF and LF. HF is an indicator of the activities of the parasympathetic nervous system and LF reflects the activities of the sympathetic nervous system. LF/HF represents the overall sympathetic and parasympathetic balance, high LF/HF being the dominant sympathetic activity, and low LF/HF being the dominant parasympathetic activity.

In summary, SDNN (ms) and RMSSD (ms) are included in the logistic regression analysis, along with time-domain HRV. The frequency-domain HRV measurements that used the FFT method are LFFFT (ms^2), LFFFT/HFFFT (ratio), LFFFTNU (n.u.), HFFFT (ms^2), and HFFFTNU (n.u.). The frequency-domain HRV measurements that used AR include LFAR (ms^2), LFAR/HFAR (ratio), LFARNU (n.u.), HFAR (ms^2), and HFARNU (n.u.); these were used for the logistic regression analysis. Detailed descriptions of each variable are provided in the Results section of this thesis.

3.7.3. Activity and sleep measurements

ActiLife software (version 6.13.1) was used to estimate the energy consumption by the units of Kcal and metabolic equivalent of task (MET) and sleep measurements, including sleep efficiency, total sleep time, and wake after sleep onset (WASO). Energy consumption obtained from both waist-worn and wrist-worn ActiGraphs was estimated. Energy consumption measurement from the waist-worn ActiGraph is considered the validated method. The wrist-worn ActiGraph's energy consumption was also estimated using the internal algorithm of ActiLife (ActiGraph, 2012) because the participants were more involved in the carrying and lowering materials activities, involving wrist movement.

Energy expenditure (in kilocalories per an hour) was estimated over 1 hour of a simulated material-handling task using two variables: ENERKAL (kcal/hour), which was measured using the ActiGraph waist sensor, and WENERKCAL (kcal/hour), which was measured using the ActiGraph wrist sensor. ActiLife provides various algorithms for energy-expenditure calculations. In this study, the Freedson combination algorithm (Freedson, Melanson, & Sirard, 1998) was used to estimate energy expenditure. In addition, the variables ENERMET and

WENERMET were calculated as the MET at the waist and the wrist, respectively. For these MET [kcal/(kg·hour)] calculations, the Swartz adult overture and lifestyle algorithm (Swartz et al., 2000) was selected.

The sleep measurements were TOTALSLEEP (in minutes), which represented the total sleep time; WASO (in minutes), which represented the total number of minutes that participants were awake during sleep time; and SLEEPQUAL (%), which was the proportion of time spent asleep out of the total time in bed. These measurements were used for the logistic regression analysis. In the ActiLife software, the Cole-Kripke algorithm (Cole, Kripke, Gruen, Mullaney, & Gillin, 1992) was selected for sleep-period scoring, and the Tudor-Locke algorithm (Tudor-Locke, Barreira, Schuna, Mire, & Katzmarzyk, 2013) was chosen for sleep-period detection.

In this study, one-second epoch data exported from the raw data were reintegrated using 60-second epoch cycles. These reintegrated data were then used for the calculation of ActiLife's scoring algorithm, as recommended for adults. The raw data (.gt3x format) from the ActiGraph were exported to an ActiGraph data file (.agd format) with a 60-second epoch length. ActiLife was used to calculate the count level based on the vector magnitude of the three axes' combined accelerations in each epoch. This count level was used as an input parameter along with body mass index (BMI) to calculate the energy expenditure. Measuring the activity count from a waist ActiGraph is a validated method. The scaled-down count was measured from the wrist and the waist each minute to measure their correspondence.

ActiLife’s sleep-scoring tool provides measures of sleep onset, total sleep time, wake after sleep onset (WASO), number of awakenings, average length of awakening, and efficiency; it also provides sleep scores for each sleep period. The Results section of this thesis describes the type of variables selected for the logistic regression analysis, explain why these variables were selected, and provide detailed information about the measurements.

Before the activity and sleep measurements were calculated, the biometric information about each participant’s age, gender, height, weight, race, and both limb and side dominance were stored in ActiLife. This information was necessary for setting the input parameters for some of the selected algorithms and calculations.

3.7.4. Subjective fatigue survey cut-off point for classifying no fatigue and fatigue status

According to research conducted by van Hoogmoed et al. (2008), severe fatigue was reported with scores above 35 and heightened fatigue was reported with scores between 27 and 35. The fatigue status of a patient suffering from rheumatoid arthritis was classified as a binary variable, based on evidence that severe physical fatigue corresponds to a CIS8R score greater than or equal to 35 and heightened fatigue corresponds to a CIS score greater than or equal to 27 (van Hoogmoed, Fransen, Bleijenberg, & Van Riel, 2010).

Any information that is not known on the cutoff point of the CIS8R to divide fatigue or no fatigue status criteria for the healthy construction worker’s population. Only the full version of the CIS survey’s total cutoff point of the working population was reported, in which fatigue state is defined as a score above 76 (Beurskens et al., 2000). Therefore, the minimum value that divides the reference of the highest group in the tertile of the CIS of the sample is the cutoff point dividing fatigue and no fatigue. If there is no clear cutoff point for dichotomizing work stress or fatigue groups of the survey instrument measurement, the method using the tertile value is usually used. Vrijkotte, Van Doornen, & De Geus (2000) classified the upper tertile of over commitment into a high overcommitment group and the remaining 2 tertile groups into a low commitment group and determined that the high overcommitment group was exposed to work stress. Swaen, Van Amelsvoort, Bültmann, & Kant (2003) used CIS fatigue cut points to identify the state of fatigue as low, medium, or high fatigue. In this study, only the subjects belonging to the highest tertile were classified as being in the fatigue state, and all others were considered non-fatigued.

3.7.5. Logistic regression, model fit testing and variable selection

Multiple logistic regression was used to determine the association between independent variables and fatigue. The significance level was $p = 0.05$. The dependent variable in the logistic regression is the probability of fatigue being present after conducting a simulated construction task. Taking the natural log-transformation of the odds of fatigue present, the equation of the logistic regression is as follows (McDonald, 2009):

$$ln\frac{Y}{1-Y} = a + b_1 x_i + b_2 x_2 + b_3 x_3 \ldots$$

where a is the intercept; b_1, b_2, b_3… are the slopes of the multiple logistic regression; Y is the probability of the fatigue state classified by the CIS survey measurement; and x_1, x_2, x_3.. are the predictor variables.

Logistic regression does not require assumptions about the normal distribution of measurement variables (McDonald, 2009). However, linear relationship is assumed between the log of the odds and the measurement variables as stated by McDonald. Although violation of this assumption is difficult to determine, the log transformations of HRV measurements have also been applied to the current study, on the basis that log-transformed values of HRV measurements have been previously applied in logistic regression (Schmitt et al. 2013; Vrijkotte et al., 2000).

To find the fitted model and ultimately to determine which variables collected from wearable sensors are useful for predicting construction workers' fatigue, the most popular model selection criteria, the Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used in this study. A stepwise-variable selection method using AIC and BIC to determine variable inclusion was used in research conducted by Li and Nyholt (2001) to select markers as the input of the discriminant analysis to classify the sib-pair type and ultimately find a model that best fit the linkage data set from 97 German families.

The SPost user package for STATA13 (College Station, TX, USA: StataCorp LP) was used for AIC and BIC calculations developed by Long and Freese (2006). Among the different pseudo R-squares, McFadden's adjusted R-squared (McFadden, 1973) for each model was also estimated using STATA13. The stepwise procedure of comparing logistic regression models with or without specific predictors using the AIC and BIC criteria and sequentially adding (forward selection) or deleting (backward selection) explanatory variables can be performed based on criteria that reach stopping points (Murtaugh, 2009). In the AIC method, the model with the smallest information loss is selected as the most suitable model for the data (Akaike, 1974). The model with the smallest information loss has the lowest AIC value, so the model with the smallest AIC value is selected as the optimal model. Accordingly,

$$AIC = \frac{-2 \times LL + 2 \times k}{n}$$

where LL is the model log-likelihood, k is the number of predictors, and n is the number of observations for the model. The STATA statistical software package output yields the AIC dropping division by n, using the equation:

$$AIC = -2 \times LL + 2 \times k$$

which is used to compared to models to decide a preferred model (Hilbe, 2017). If the difference in the AIC between two models is greater than 0.0 and equal to or greater than 2.5, it is considered that there is no difference between the two models. When the sample size is greater than 256, the small AIC model is preferred if the difference between the two models is greater than 2.5 and less than or equal to 6.0. When the sample size is greater than 64 but less than or equal to 256, the model with smaller AIC is preferred if the difference of AICs is greater than 6.0 and less than or equal to 9.9. Regardless of the sample size, the smaller AIC model is preferred if the difference of AIC between the two models is bigger than or equal to 10 (Hilbe, 2017).

The BIC designed by Raftery (1986) is another popular metric for evaluating information captured within multivariate models. The BIC method is a statistical test derived by approximating the posterior probability distribution calculated using the likelihood function and the prior probability distribution in Bayesian theory. The model with the smallest BIC value is selected as the optimal model, like the AIC.

$$BIC = D - df \times \ln(n)$$

where D is the model deviance statistic, df is the degree of freedom, and n is the number of observations in the model (Hilbe, 2017). The Stata software output used the following equation that used the log-likelihood:

$$BIC = -2 \times LL + ln(n) \times k$$

where LL is the model log-likelihood, n is the number of observations, and k is the number of predictors of the model (Hilbe, 2017). The selection of the preferred model can be made based on the difference between the BIC in the models compared. The criteria to assess the degree of preference is weak when the difference is between 0 and 2, positive when the difference is between 2 and 8, and strong when the difference is between 6 and 10. If the difference of BIC is larger than 10, the degree of preference is very strong (Hilbe, 2017).

In the model selection procedure, the maximum number of variables chosen to include in the potential best fit model was limited to eight. This is because, as a general rule, at least 10 cases per predictor are necessary for the logistic regression analysis (Hilbe, 2017) and the total number of sample size is 80 in this current research. For the final model, meeting the model selection criteria of the pseudo R-squared, AIC and BIC, the simplest model was chosen for the pragmatic use of fatigue management. In the case of surveys and control variables, the backward variable deletion method removed variables that were not suitable for the model. Then, forward selections were performed to determine whether the AIC and BIC criteria could be improved by adding the heart rate, HRV, sleep, and energy expenditure variables to the model in turn, as these are the easiest variables to obtain from the wearable sensor and minimally affect the worker's work.

Ideally, the simplest model for fatigue should include only the minimum number of variables, and the variables should be obtained with minimum effort from both the safety professionals and construction workers. For example, from the workers' perspective, collecting 10-minute resting HR might be preferred rather than monitoring and collecting HR during working hours. Also, collecting HR variables during working hours would be easily manageable by the safety professionals rather than providing wearable devices to workers during off-duty to collect their sleep variables. Therefore, to select a final model, this research relied on the information criteria and the experiences of the author of this thesis in applying the wearable technology to several field and laboratory studies (Lee & Migliaccio, 2016; Lee, Lin, Seto, & Migliaccio, 2017; Lee, Seto, Lin, & Migliaccio, 2017).

4. Results

4.1. Dependent variable for fatigue classification

The 80 samples were divided by tertile distribution according to scores from the checklist of individual strengths (CIS8R). Twenty-eight participants were included in the lowest third, 27 participants were included in the medium third, and the remaining 25 participants were included in the highest third (Table 1). The participants included in the highest third were coded as having fatigue status after the task completion, and the participants included in the medium and lowest thirds were coded as not fatigued.

Table 1. Cumulated score of CIS in tertiles

Tertile	Number of Observations	Mean	Standard Deviation	Min	Max	Fatigue Status
1 (Low)	28	16.1	3.10	8	20	No Fatigue
2 (medium)	27	24.2	2.17	21	27	No Fatigue
3 (high)	25	31	3.44	28	40	Fatigue

The highest tertile group presented CIS8R scores of 28–40, the participants who reported this range of perceived fatigue level were classified as having fatigue status. Therefore, CIS8R scores of 28 or greater were considered to indicate the fatigue condition.

4.2. Description of research variables used in the data analyses

Table 2 summarizes the description of research variables and their units referred to in the data analyses. The acronyms of the variables are used in the results and discussion sections.

Table 2. Summary of research variables used in data analysis

Acronym	Description of variable	Unit
CIS	Total score of the eight physical fatigue subscales of checklist individual strength survey items	none
FATIGUE	Binary variable classified fatigue status based on the CIS tertile distribution	No Fatigue = 0; Fatigue = 1
AGE	Participant’s age	Year
GENDER	Participant’s gender	Male = 0; Female = 1

TLX	Perceived Workload surveyed by NASA task load index (scored using weights ratings)	None
SF12PCS	SF12 physical component summary including measurement items: general health (GH), physical functioning (PF), role physical (RP), bodily pain (BP)	None
SF12MCS	SF12 mental component summary including measurement items: mental health (MH), role emotional (RE), vitality (VT), social functioning (SF)	None
SMWT	Result of the six-minute walk test conducted on 15-meter track	centimeters
TIMEC	Daily time to conduct the simulated construction task that coded as 'Mon' if participants worked between 6am and 12pm; as 'Aft' if participants worked between 12pm and 6pm; as 'Even' if participants worked between 6pm and 10pm.	None (Dummy variable)
HRBPM	Average heart rate during the task	Beats per minute(bpm)
RESTHR	Resting heart rate calculated by the average heart rate measured over 10 minutes before task	Beats per minute (bpm)
RHR	Relative heart rate during the task	%
HRR	Heart rate recovery calculated the absolute heart rate from peak levels upon completion of the task minus the heart rate after 2 minutes from task completion	Beats per minute (bpm)

ENERKCAL	Energy Expenditure estimated from the ActiGraph worn on waist	Kcal/hour
ENERMET	Metabolic equivalent of task estimated from the ActiGraph worn on waist	None
WENERKCAL	Energy Expenditure estimated from the ActiGraph worn on wrist	Kcal/hour
WENERMET	Metabolic equivalent of task estimated from the ActiGraph worn on wrist	None
TOTALSLEEP	Total sleep time recorded as asleep	Minute
WASO	Total minutes during which participants were awake after sleep onset occurred.	Minute
SLEEPQUAL	Sleep efficiency estimating the total sleep divided by the total minutes spent in bed by the ActiGraph.	%
SDNN	Standard deviation of normal-to-normal RR intervals	ms
RMSSD	Root mean square of successive RR interval differences	ms
LFFFT	Heart rate variability LF measurement using FFT method	ms^2
HFFFT	Heart Rate Variability HF measurement using FFT method	ms^2
LFFFT/HFFFT	Ratio of heart rate variability LF measurement to HF measurement using FFT method	none
LFFFTNU	Normalized heart rate variability LF measurement using FFT method	Normalized unit (n.u.)

HFFFTNU	Normalized heart rate variability HF measurement using FFT method	n.u.
LFAR	Heart rate variability LF measurement using AR method	ms^2
HFAR	Heart Rate Variability HF measurement using AR method	ms^2
LFAR/HFAR	Ratio of heart rate variability LF measurement to HF measurement using AR method	none
LFARNU	Normalized heart rate variability LF measurement using AR method	n.u.
HFARNU	Normalized heart rate variability HF measurement using AR method	n.u.

4.3. Demographic information

Nineteen participants completed all four experiment sessions with different intended workload levels, one participant conducted two sessions, and two participants completed only one session. In total, 80 observations for data analysis were collected. Although measurements of participants' demographic information (e.g., age) were the same, each observation was considered an independent measurement in this study. The demographic information summarized by no fatigue and fatigue groups and the t-tests for a mean difference of the predictor variables are reported in Table 3.

Table 3. Summary of descriptive statistics (Continuous variables)

Variable	No Fatigue (n=55)	Fatigue (n=25)	p-value [a]	All group (n=80)

	Mean	SD	Mean	SD		Mean	SD
Demographic information							
AGE (years)	25.1	3.18	24.7	2.72	0.56	25.0	3.03
BMI (kg/meter^2)	23.2	3.50	22.9	2.94	0.72	23.1	3.32
Survey measurements							
TLX	33.3	13.37	49.2	12.57	<0.05	38.2	15.01
SMWT (meters)	526.7	39.82	552.6	42.42	<0.05	534.8	42.15
SF12PCS	56.9	4.06	53.6	5.37	<0.05	55.8	4.73
SF12MCS	52.2	7.64	49.6	7.42	0.15	51.4	7.62
Heart rate measurements							
HRBPM (bpm)	112.2	19.27	117.3	18.31	0.27	113.8	19.01
RHR (%)	35.6	14.50	37.7	13.79	0.54	36.2	14.23
RESTHR (bpm)	68.7	11.44	73.1	11.10	0.11	70.1	11.45
HRR (bpm)	17.7	12.42	17.5	11.40	0.92	17.7	12.04
Sleep measurements							
SLEEPQUAL (%)	83.0	7.09	85.2	4.79	0.10	83.7	6.51
TOTALSLEEP (minutes)	340.9	104.98	345.8	78.84	0.82	342.4	97.09
WASO (minutes)	67.7	29.72	60.6	23.89	0.25	65.5	28.08
Energy expenditure							
ENERKCAL (kcal/hour)	189.3	84.79	176.6	88.91	0.55	185.3	85.74
ENERMET	3.6	0.60	3.6	0.76	0.87	3.6	0.65
WENERKCAL (kcal/hour)	163.7	73.83	151.3	55.53	0.41	159.8	68.53
WENERMET	3.5	0.18	3.6	0.31	0.51	3.5	0.23

Note. a: Two-sided t-test

In the no fatigue group, 67% of participants were male and 33% were female; in the fatigue group, 64% were male and 36% were female. The gender distribution was relatively uniform between the no fatigue and fatigue groups. Both groups participated in the study most frequently in the afternoon, with 58% in the no fatigue group and 56% in the fatigue group.

Table 4. Summary of descriptive statistics (categorical variables)

Variable	No Fatigue (n=55)		Fatigue (n=25)		p-value [a]	All group (n=80)	
	n	%	n	%		n	%
Demographic information (GENDER)							
Male	37	67	16	64	0.77	53	66
Female	18	33	9	36	0.77	27	34
Day of task (TIMEC)							
Morn	17	31	5	20	0.31	22	28
Aft	32	58	14	56	0.85	46	57
Even	6	11	6	24	0.13	12	15

Note. a: Two-sample test of proportion

The mean and SD of the HRV for both groups, the difference between the mean values of the two groups, and the t-test results are summarized in Table 5. In the time domain measurements, if the HRV variables SDNN and RMSSD both showed low mean values in the fatigue group, the difference between the two groups was statistically significant.

Table 5. Summary of descriptive statistics for HRV measurements

Variable	No Fatigue (n=55)		Fatigue (n=25)		p-value [a]	All group (n=80)	
	Mean	SD	Mean	SD		Mean	SD
Time Domain HRV							
SDNN(ms)	50.2	20.57	36.4	19.78	<0.05	45.9	21.21
RMSSD(ms)	23.3	10.43	17.5	10.77	<0.05	21.5	10.81
Frequency domain HRV (FFT)							
LFFFT(ms^2)	2582.2	2306.33	1432.0	1332.27	<0.05	2222.7	2112.56
HFFFT(ms^2)	519.0	761.78	215.4	259.38	<0.05	424.1	661.18
LFFFTNU(n.u.)	85.5	9.71	87.8	7.40	0.24	86.2	9.07
HFFFTNU(n.u.)	14.5	9.66	12.1	7.37	0.24	13.7	9.03
LFFFT/HFFFT	2.3	0.78	2.4	0.67	0.74	12.9	9.88
Log (LFFFT)	7.5	0.96	6.7	1.29	<0.05	7.2	1.13
Log (HFFFT)	5.5	1.29	4.7	1.33	<0.05	5.2	1.35
Frequency domain HRV (AR)							
LFAR(ms^2)	2329.9	2122.66	1275.5	1166.59	<0.05	2000.4	1932.66
HFAR(ms^2)	470.6	667.74	198.6	234.61	<0.05	385.6	581.03
LFARNU(n.u.)	85.1	10.02	87.8	7.21	<0.05	85.9	9.28
HFARNU(n.u.)	14.9	9.97	12.1	7.17	<0.05	14.0	9.24
LFAR/HFAR	11.2	7.69	12.5	10.27	0.58	11.6	8.53

Log (LFAR)	7.4	0.93	6.6	1.24	<0.05	7.1	1.09
Log (HFAR)	5.4	1.25	4.6	1.30	<0.05	5.2	1.31

Note. a: Two-sided t-test

4.4. Multicollinearity among the HRV variables

To check for multicollinearity among the HRV variables obtained from both time and frequency domain methods, bivariate correlation analyses were performed. The Pearson correlations and significant levels are summarized in Table 6 (FFT frequency domain method) and Table 7 (AR frequency domain method), respectively. In both the FFT and AR methods, the SDNN measurement of the time domain analysis was correlated with LF and HF (as well as natural log-transformed LF and HF) in the frequency domain. RMSSD was also observed to have significant correlations with LF and HF (including natural log-transformed LF and HF) in both the FFT and AR methods. The RMSSD was negatively correlated with the LF/HF ratio in both the FFT and AR methods. Thus, it was highly expected that there would be multicollinearity problems in the logistic regression model, including both issues with time domain HRV variables and with frequency domain HRV variables concurrently.

Table 6. Correlation coefficients of the heart rate variability measurement (FFT)

	SDNN	RMSSD	LFFFT	HFFFT	LFFFTNU	HFFFTNU	LFFFT/ HFFFT	Log (LFFFT)	Log (HFFFT)
SDNN	1.000								
RMSSD	0.897**	1.000							
LFFFT	0.881**	0.740**	1.000						
HFFFT	0.656**	0.689**	0.735**	1.000					
LFFFTNU	-0.219	-0.507**	-0.164	-0.598**	1.000				
HFFFTNU	0.220	0.508**	0.165	0.599**	-1.000**	1.000			
LFFFT/HFFFT	-0.041	-0.307**	0.041	-0.322**	0.668**	-0.668**	1.000		
Log (LFFFT)	0.938**	0.791**	0.824**	0.548**	-0.079	0.080	0.088	1.000	
Log (HFFFT)	0.834**	0.883**	0.711**	0.764**	-0.587**	0.588**	-0.287**	0.815**	1.000

**. Correlation is significant at the 0.01 level (2-tailed)

Table 7. Correlation coefficients of the heart rate variability measurement (AR)

	SDNN	RMSSD	LFAR	HFAR	LFARNU	HFARNU	LFAR/ HFAR	Log (LFAR)	Log (HFAR)
SDNN	1.000								
RMSSD	0.897**	1.000							
LFAR	0.875**	0.743**	1.000						
HFAR	0.664**	0.707**	0.792**	1.000					
LFARNU	-0.220*	-0.510**	-0.179	-0.614**	1.000				
HFARNU	0.222*	0.510**	0.181	0.615**	-1.000**	1.000			
LFAR/HFAR	-0.088	-0.353**	-0.022	-0.361**	0.665**	-0.665**	1.000		
Log (LFAR)	0.946**	0.804**	0.827**	0.579**	-0.096	0.097	0.030	1.000	
Log (HFAR)	0.837**	0.894**	0.717**	0.780**	-0.610**	0.611**	-0.338**	0.814**	1.000

**. Correlation is significant at the 0.01 level (2-tailed)

*. Correlation is significant at the 0.05 level (2-tailed)

4.5. Stepwise logistic regression for variable selection

Stepwise regression analyses were performed with the survey measurements first, because the subjective task workload and health status have traditionally been regarded as important information for fatigue management, particularly from the perspective of project processes and total worker health interventions based workers' fatigue symptoms. Table 8 presents the backward selection approach used to find the best fitting model that included survey measurement variables. Therefore, the first model in the top row of the table included all three survey measurements, and continuous logistic regression analyses were conducted removing one variable at a time to generate every possible combination of the variables.

Table 8. Logistic regression results with AIC, BIC and Pseudo R-squared for survey measurements

Variables	AIC	BIC	Pseudo R^2	Note
TLX, SF12PCS, SF12MCS	77.5	87.1	0.300	
TLX, SF12PCS	77.1	84.3	0.284	Model 1 (Best fit)
TLX, SF12MCS	82.5	89.6	0.230	
SF12PCS, SF12MCS	93.6	100.7	0.119	
TLX	81.1	85.8	0.224	
SF12PCS	95.0	99.8	0.084	
SF12MCS	101.4	106.1	0.020	

The best fitting model for fatigue was identified with the backward selection using survey information only (Table 8, Model 1 with lowest AIC and BIC and the second highest R^2). After choosing the best fit model (Model 1, Table 8) with the survey variables, the backward selection method was used for selecting a best fit model that also included demographic variables in the model (Table 9, Model 2), which was found to include gender, BMI, and SMWT variables.

Table 9. Logistic regression results with AIC, BIC and Pseudo R-squared for demographic information

Variables confirmed to be added in the previous step	Variables	AIC	BIC	Pseudo R^2	Note
TLX, SF12PCS (Model 1, Table 8)	AGE, GENDER, BMI, SMWT	71.5	86.2	0.441	
	AGE, GENDER, BMI	82.5	94.8	0.311	
	AGE, GENDER, SMWT	74.5	86.8	0.391	
	AGE, BMI, SMWT	72.9	87.1	0.388	
	GENDER, BMI, SMWT	69.8	84.1	0.419	Model 2 (Best fit)
	AGE, GENDER	80.0	91.9	0.296	
	AGE, BMI	79.7	91.6	0.298	
	AGE, SMWT	73.0	85.0	0.366	
	GENDER, BMI	78.9	90.8	0.307	
	GENDER, SMWT	74.9	84.9	0.367	
	BMI, SMWT	71.4	83.3	0.382	
	AGE	78.7	88.3	0.288	

	GENDER	78.7	88.3	0.288	
	BMI	77.8	87.4	0.297	
	SMWT	72.0	81.5	0.356	

Information on the time of the task can be more easily obtained than the collecting data via wearable sensors. Therefore, stepwise logistic regression analysis was conducted with the TIMEC variable before adding the wearable sensor measurements. In the results of the analysis, Model 3 without TIMEC presented lower AIC and BIC values, indicating addition the TIMEC variable had minimal information value beyond the information from the other variables already included in the model (Table 10).

Table 10. Logistic regression results with AIC, BIC and Pseudo R-squared for time to conduct

Variables confirmed to be added in the previous step	Variables	AIC	BIC	Pseudo R^2	Note
TLX, SF12PCS GENDER, BMI, SMWT (Model 2, Table 9)	without TIMEC	69.8	84.1	0.419	Model 3 (Best fit)
	TIMEC[a]	72.5	91.6	0.431	

Note. a: For the variable TIMEC, three dummy variables, “Morn,” “Aft,” and “Even” were created, and two of the three dummy variables were included in the logistic regression model.

In subsequent stepwise logistic regression analyses for variable selections, heart rate, energy expenditure, and sleep measurements from the wearable sensors were included to Model 3 from Table 10.

By applying the forward model selection, AIC, BIC, and pseudo R-squared values were compared. Although AIC was not the lowest, Model 4 was selected as the best fit because BIC was the smallest and the difference of AIC between Model 4 and Model 5 is very minimal. This study selected Model 4 also because the purpose of the study was to identify a simple prediction model with the smallest number of variables.

Table 11. Logistic regression results with AIC, BIC and Pseudo R-squared for variables measured wearable sensors

Variables confirmed to be added in the previous step	Variables	AIC	BIC	Pseudo R^2	Note
TLX, SF12PCS, GENDER, BMI, SMWT (Model 3, Table 10)	**Resting heart rate before task**				
	RESTHR	67.8	84.5	0.459	Model 4 (Best Fit)
	Resting heart rate after task- Heart rate recovery				
	RESTHR, HRR	69.2	88.3	0.465	
	Heart rate				
	RESTHR, HRBPM	67.5	86.6	0.482	
	RESTHR, RHR	67.2	86.2	0.485	Model 5
	Energy expenditure				
	RESTHR, RHR, ENERKCAL	67.8	89.2	0.499	

	RESTHR, RHR, ENERMET	69.1	90.6	0.485	
	RESTHR, RHR, WENERKCAL	68.0	89.5	0.497	
	RESTHR, RHR, WENERMET	69.1	90.5	0.486	
	Sleep measurement				
	RESTHR, RHR, TOTALSLEEP	68.9	90.3	0.488	
	RESTHR, RHR, SLEEQUAL	69.1	90.6	0.485	
	RESTHR, RHR, WASO	68.7	90.1	0.564	

Because the HRV analysis process is more complex for obtaining fatigue prediction variables than for other variables added or removed from the previous model selection procedure, HRV variable selections were performed as the last step, as presented in Table 12. Due to the high possibility of multicollinearity issues in the model when more than two HRV measurement variables are included simultaneously, only one HRV measurement variable was added in Model 4 to compare AIC and BIC (along with Pseudo R^2).

Table 12. Logistic regression results with AIC, BIC and Pseudo R-squared for two different HRV measurements methods

Variables confirmed to be added in the previous step	Variables	AIC	BIC	Pseudo R^2	Note
TLX, SF12PCS GENDER, BMI, SMWT, RESTHR (Model 4, Table 11)	**HRV Time Domain**				
	SDNN(ms)	63.7	82.8	0.520	Model 6 (Best fit)
	RMSSD(ms)	63.8	82.9	0.519	
	HRV Frequency Domain- FFT				
	LFFFT(ms^2)	66.9	86.0	0.488	
	LFFFTNU(n.u.)	69.8	88.8	0.459	
	HFFFT(ms^2)	69.4	88.5	0.463	
	HFFFTNU(n.u.)	69.8	88.8	0.459	
	LFFFT/HFFFT	69.0	88.0	0.467	
	HRV Frequency Domain- AR				
	LFAR (ms^2)	66.6	85.7	0.491	Model 7
	LFARNU (n.u.)	69.8	88.8	0.459	
	HFAR (ms^2)	69.1	88.2	0.465	
	HFARNU (n.u.)	69.8	88.8	0.459	
	LFAR/HFAR	68.0	87.1	0.477	

The final results of variable selection for fatigue prediction using stepwise logistic regression are summarized in Table 13. According to the guidelines of the task force (Camm et al., 1996), the frequency domain requires five minutes of data collection, and 24 hours of data collection is recommended for the time domain method. Therefore, frequency domain analysis of HRV should be applied to daily fatigue management of the construction site. Model 6, which showed the smallest value of the calculated AIC and BIC, was analyzed as the best fit. However, Model 6 requires SDNN, which typically would involve 24-hour data collection, which may not be appropriate as a predictor of fatigue using short-term data collection on construction workers. Therefore, Model 7, which did not rely on SDNN, but instead used LFAR, which is an HRV variable based on spectral analysis, was found to be more suitable for short-term data collection.

Table 13. Logistic regression results with AIC, BIC and Pseudo R-squared for final model

Variables confirmed to be added in the previous step	Variables	AIC	BIC	Pseudo R^2	Note
TLX, SF12PCS, GENDER, BMI, SMWT, RESTHR	**HRV Time Domain and Frequency Domain**				
	SDNN(ms)	63.7	82.8	0.520	Model 6: Best fit
	LFAR(ms^2)	66.6	85.7	0.491	Model 7: For short term data collection
	SDNN(ms), LFAR(ms^2)	63.858	85.296	0.539	Model 8: Multicollinearity issue between SDNN and LFAR

4.6. Fatigue prediction model

In stepwise logistic regression, TLX, SF12PCS, Gender, BMI, SMWT, RESTHR, and SDNN were selected as the predictors for the lowest AIC and BIC values. The odds ratio for fatigue of the TLX, BMI, SMWT and RESTHR were greater than 1: TLX (OR = 1.15, 95% confidence interval = 1.053–1.258), BMI (OR = 1.18, 95% confidence interval = 0.838–1.674), SMWT (OR = 1.04, 95% confidence interval = 1.011–1.069), RESTHR (OR = 1.05, 95% confidence interval = 0.944–1.1161). However, there was no statistically significant association between BMI and the odds of fatigue. In addition, it was failed to reject the null hypothesis that the odds of fatigue is not associated with resting heart rate (i.e., RESTHR) because the two-sided P-value was 0.387 (Table 14). Lower SF12PCS and SDNN were associated with lower odds ratio for fatigue. Regarding the predictor for gender, the odds ratio was calculated to be 5.47 for female, but the association between gender and the odds of fatigue was not statistically significant.

Table 14. Factors associated with fatigue by logistic regression analysis (Odds ratio reported)

	Odds ratio	Standard Error	95% confidence interval		P-value
			Lower	Upper	
TLX	1.15	0.052	1.053	1.258	<0.01
SF12PCS	0.71	0.076	0.571	0.872	<0.01
GENDER (Female)	5.47	5.383	0.795	37.645	0.08
BMI	1.18	0.209	0.838	1.674	0.34
SMWT	1.04	0.015	1.011	1.069	<0.01
RESTHR	1.05	0.055	0.944	1.161	0.39

SDNN	0.94	0.029	0.880	0.995	0.03
Log likelihood = -23.862709					
Pseudo R^2 = 0.5197					

Table 15. Factors associated with fatigue by logistic regression analysis (Coefficient reported)

	Coefficient	Standard Error	95% confidence interval		P-value
			Lower	Upper	
TLX	0.14	0.045	0.052	0.229	<0.01
SF12PCS	-0.35	0.108	-0.560	-0.137	<0.01
GENDER (Female)	1.70	0.984	-0.230	3.628	0.08
BMI	0.17	0.176	-0.177	0.515	0.34
SMWT	0.04	0.014	0.011	0.067	<0.01
RESTHR	0.05	0.053	-0.058	0.149	0.39
SDNN	-0.07	0.031	-0.128	-0.005	0.03
Log likelihood = -23.862709					
Pseudo R^2 = 0.5197					

The reduced model was analyzed after eliminating predictors whose association with fatigue was not statistically significant.

Table 16. Reduced model with predictors that p-value is less than 0.05 (Odds ratio reported)

	Odds Ratio	Standard Error	95% confidence interval		P-value
			Lower	Upper	
TLX	1.11	0.035	1.041	1.179	<0.01
SF12PCS	0.75	0.071	0.620	0.898	<0.01
SMWT	1.02	0.011	1.004	1.046	<0.05
SDNN	0.94	0.023	0.895	0.985	<0.05
Log likelihood = -27.303766					
Pseudo R^2 = 0.4505					

Table 17. Reduced model with predictors that p-value is less than 0.05 (Coefficient reported)

	Coefficient	Standard Error	95% confidence interval		P-value
			Lower	Upper	
TLX	0.10	0.032	0.040	0.165	<0.01
SF12PCS	-0.29	0.095	-0.478	-0.107	<0.01
SMWT	0.02	0.010	0.004	0.045	<0.05
SDNN	-0.06	0.024	-0.110	-0.015	<0.05
Log likelihood = -27.303766					
Pseudo R^2 = 0.4505					

The LFAR data were skewed so that the linearity between log odds and predictor in the logistic regression could not be affirmed. Furthermore, the coefficient value of LFAR was estimated at a small decimal number (i.e., - 0.0008), which made it difficult to interpret the probability of fatigue and no fatigue. The log-transformed LFAR value was used for ease of interpretation.

Table 18. Model with HRV predictors frequency domain (log-transformed)

	Odds Ratio	Standard Error	95% confidence interval		P-value
			Lower	Upper	
TLX	1.10	0.035	1.037	1.173	<0.01
SF12PCS	0.76	0.069	0.635	0.906	<0.01
SMWT	1.02	0.011	1.004	1.046	<0.05
Log (LFAR)	0.33	0.145	0.144	0.781	<0.05
Log likelihood = -27.45624					
Pseudo R^2 = 0.4474					

From the multivariate logistic regression, fatigue status was significantly associated with the TLX, SF12PCS, SMWT, and Log (LFAR) predictors. While other predictors were constant, a 10% increase in the odds of fatigue status was associated with every one-unit increase in the TLX score. A 24% decrease in the odds of being in fatigue status was associated with a one-unit increase in the SF12PCS score when other predictors were constant. Withholding other factors, there was a 2% increase in the odds of being in fatigue status with a one-unit increase in SMWT test performance.

Table 19. Model with HRV predictors frequency domain (log-transformed)

	Coefficient	Standard Error	95% confidence interval		P-value
			Lower	Upper	
TLX	0.10	0.032	0.036	0.160	<0.01
SF12PCS	-0.28	0.090	-0.453	-0.099	<0.01
SMWT	0.02	0.010	0.004	0.045	<0.05
Log (LFAR)	-1.09	0.432	-1.940	-0.248	<0.05
Log likelihood = -27.45624					
Pseudo R^2 = 0.4474					

Regarding the log-transformed LFAR predictor, the odds ratio for two populations that differ by 1% in the predictor was $1.01^{-1.09} = 0.989$. The odds ratio for two populations that differ by 10% in the predictor was $1.1^{-1.09} = 0.901$. Estimated linear relationships for the log odds fatigue by four selected predictors were calculated as follows:

$$\log Odds\ FATIGUE_i = 4.66 + 0.10 \times TLX_i - 0.28 \times SF12PCS_i + 0.02 \times SMWT_i - 1.09 \times \log(LFAR_i)$$

where 4.66 is the estimated intercept and each estimated slope was stated with each predictor. All coefficients and intercept values were rounded from the third decimal points.

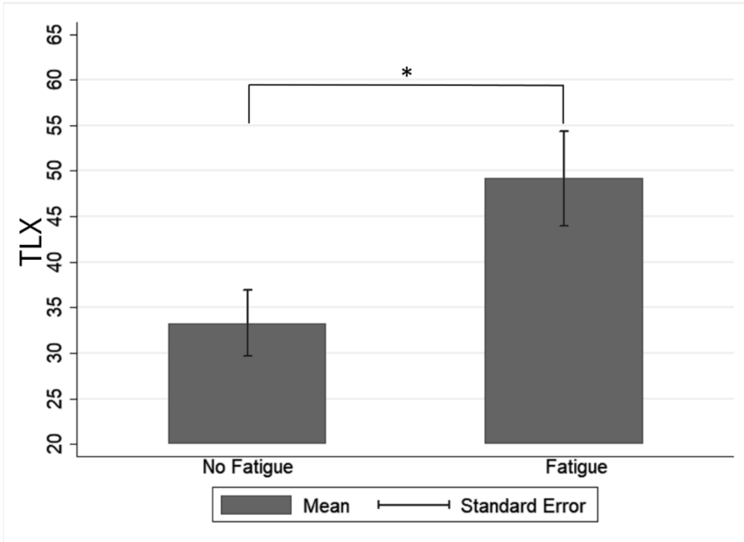

Figure 7. Difference in TLX between no fatigue and fatigue groups (* p<0.05)

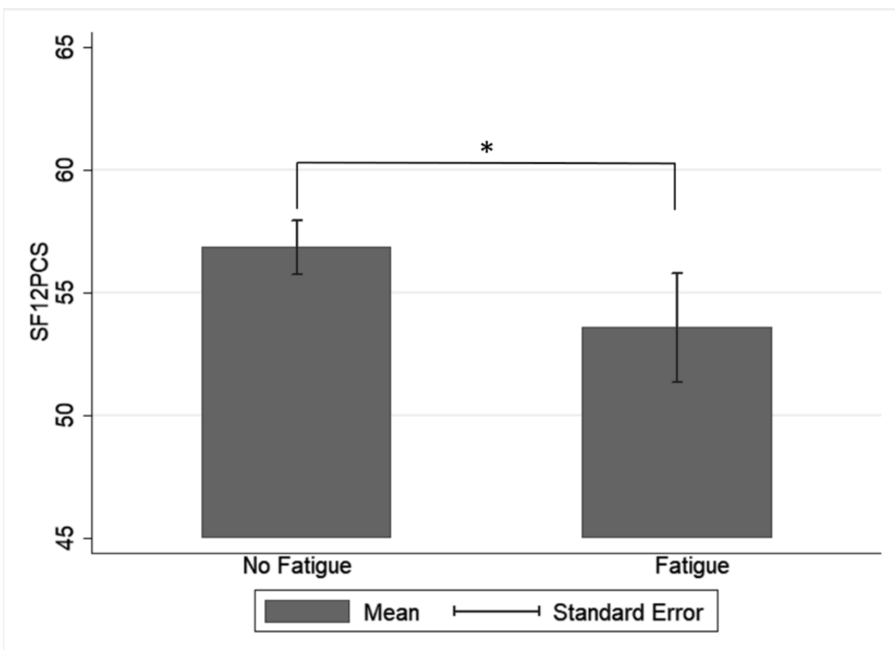

Figure 8. Difference in SF12PCS between no fatigue and fatigue groups (* p<0.05)

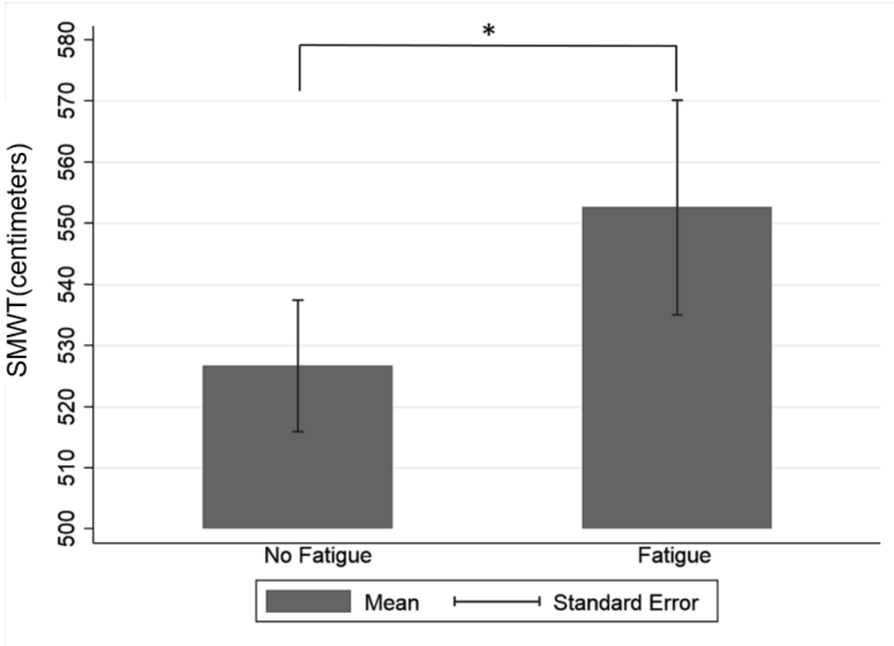

Figure 9. Difference in SMWT between no fatigue and fatigue groups (* p<0.05)

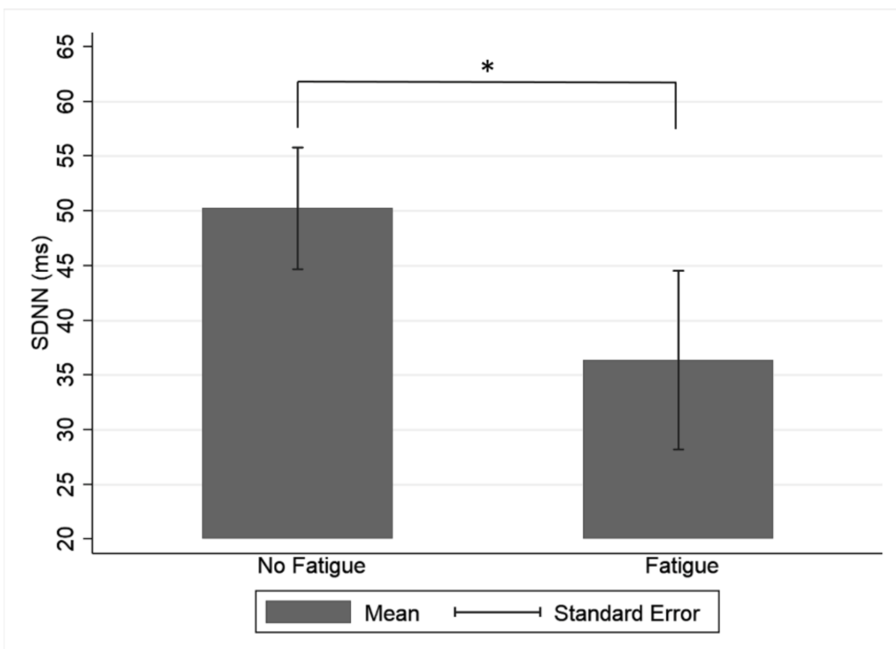

Figure 10. Difference in SDNN between no fatigue and fatigue groups (* p<0.05)

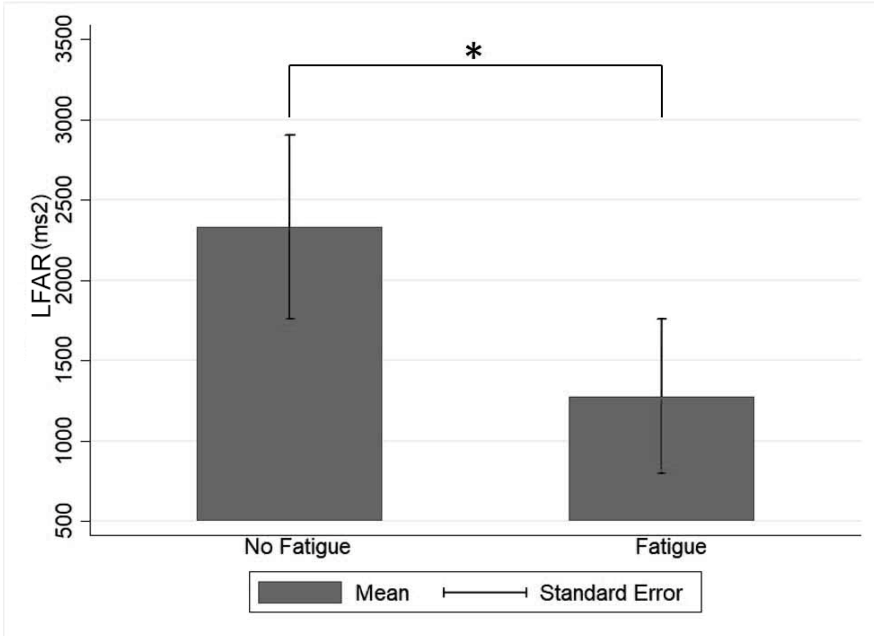

Figure 11. Difference in LFAR between no fatigue and fatigue groups (* $p<0.05$)

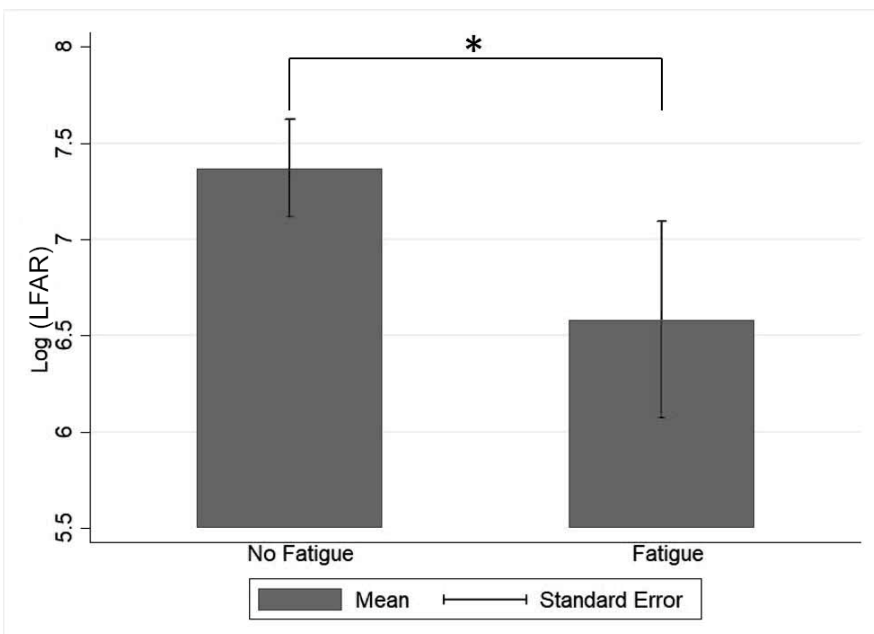

Figure 12. Difference in natural log-transformed LFAR between no fatigue and fatigue groups (* $p<0.05$)

Figures 7 through 12 present the mean difference of the selected predictor variables between not fatigued and fatigued groups. The standard error of the predictor variables for each group and statistical significance of the two-sided t-tests are also presented. Based on the analysis, the fatigue group had a higher perceived workload than the not fatigued group. Furthermore, the physical health score reported by the not fatigued group was higher than that of the fatigued group. The results of the SMWT were higher for the fatigued group than the not fatigued group, suggesting that the fatigued group had higher physical capacity than SMWT. However, another interpretation of these results is that participants in the fatigued group may have conducted their work more intensely with more intense SMWTs in reaction to being observed in an experiment, which eventually led to their increased fatigue. The SDNN was lower in the fatigued group than in the not fatigued group, indicating that that fatigue caused a decrease in parasympathetic system activity. Both LFAR and log-transformed LFAR were lower in the fatigued group compared with the mean values of the not fatigued group.

Based on the results in Tables 20 and 21, the initial model only included survey measurements presenting larger AIC and BIC when compared to the final model which additionally included demographic information, resting heart measurement and HRV measurement. Ultimately, the reduced model (after eliminating predictors whose association with fatigue was not statistically significant) was created to identify a simple prediction model with the fewest variables and lowest BIC. Because the absolute difference between the BIC statistics between models in the second- and third-row booths in Tables 20 (modeled with SDNN HRV measurement) and 21 (modeled with log-transformed LFAR HRV measurement) are larger than seven, the degrees of preference are strong.

Table 20. Results of AIC and BIC of variable selection

Model variables	AIC	BIC	Δ AIC (2)-(1)	Δ BIC (2)-(1)	Δ AIC (3)-(1)	Δ BIC (3)-(1)	Δ AIC (3)-(2)	Δ BIC (3)-(2)
(1) TLX, SF12PCS	77.1	84.3						
(2) TLX, SF12PCS, GENDER, BMI, SMWT, RESTHR, SDNN	63.7	82.8	-13.4	-1.5				
(3) TLX, SF12PCS, SMWT, SDNN	64.6	76.5			-12.5	-7.8	+0.9	-6.3

Table 21. Results of AIC and BIC of variable selection

Model variables	AIC	BIC	Δ AIC (2)-(1)	Δ BIC (2)-(1)	Δ AIC (3)-(1)	Δ BIC (3)-(1)	Δ AIC (3)-(2)	Δ BIC (3)-(2)
(1) TLX, SF12PCS	77.1	84.3						
(2) TLX, SF12PCS, GENDER, BMI, SMWT, RESTHR, Log (LFAR)	64.6	83.6	-12.5	-0.7				
(3) TLX, SF12PCS, SMWT, Log (LFAR)	64.9	76.8			-12.2	-7.5	+0.3	-6.8

5. Discussion

This is one of the few studies that has developed a model of fatigue among construction workers utilizing data from a variety of wearable sensors as well as survey data. The focus of this work was to identify a well-performing parsimonious model that can be used in practical work situations by construction managers to identify and mitigate worker fatigue in order to reduce the risk of subsequent workplace injuries and incidents. The research identified a few models that may have practical applications. For example, the best performing models that (a) only relies upon survey data, (b) a model that also includes data from wearable sensors, and (c) models that might work best for short-term vs chronic fatigue were each identified.

There are benefits to the design of the current experimental study. The dataset was collected under controlled lab conditions. Specifically, the participants were not tested under the condition of team task performance because the former study that collected data investigated task load and resources of individual worker. Noise and heat stress that can influence fatigue levels were controlled, and the participants performed material lifting work. Therefore, the findings may be generalized to apply to indoor construction workers who are not involved in the team process and use excessively noisy tools. Since the prediction model was developed for entry-level indoor construction workers, a potential future study may be to validate the model in workers with the same level of experience at a prefabrication shop that has continuous lifting and lowering of materials in its operation. Since the subjects conducted tasks in the experiment when they were available to come to the research facility, the time when the experiment was conducted varied across subjects. The circadian rhythm of workers is a known factor for influencing their increased risk of accidents at the jobsite (Folkard & Tucker, 2003); as such, the time of day when the subject participated in the experiment was recorded.

In developing these models, various factors that were found to be associated with fatigue were observed. Some of these factors have also be identified and discussed in previous studies. For instance, in the current study, the differences in age and BMI did not have any explanatory power as a variable differentiating the fatigue and no fatigue groups. This is consistent with the findings by Olsson, Roth, and Melin (2010) that reported no difference of BMI between stress-related fatigue group and healthy controls group.

The current study found that mean perceived workload among the participants in fatigued group is significantly higher than it among the participants in the not fatigued group. Rubio, Díaz, Martín, & Puente (2004) remarked that a variety of assessment tools for subjective mental workload should be validated and compared not only in laboratory-based experiments but also in real-world contexts. To date, there are limited studies available to use NASA-TLX for the subjective measurement of workload among the construction workers. Dey and Mann (2010) compared the sensitivity and diagnostic capacity between NASA-TLX and SWAT under authentic working conditions with an agricultural sprayer. Variations in working conditions, including different occupations, still remain to be compared and evaluated to share information about which types of tools will be appropriate for a specific industry. In particular, the management of occupational fatigue in shiftwork is important for preventing accidents and injuries caused by worker errors, and redesigning the work schedule and redistributing the workload are methods for coping with shiftwork (Rosa & Colligan, 1997). Thus, the perceived workload obtained by surveying the workers is important information to manage the occupational fatigue as suggested by the current research.

The not fatigued group reported a higher average of SF12 physical health summary score (i.e., SF12PCS) than the fatigued group. The SF12 physical summary score (i.e., SF12PCS) was significantly lower in the fatigued group in this current study. This result is consistent with the result of a cross-sectional survey of US workers conducted by Ricci, Chee, Lorandeau, and Berger (2007) that workers who have experienced health-related lost productive time and fatigue prevalence reported lower SF12 physical (i.e., SF12PCS) and mental (i.e., SF12MCS) health summary scores. Schnorpfeil et al. (2002) found a correlation between SF12 mental health summary score (i.e., SF12MCS) and vital exhaustion in a study of employees in airplane manufacturing plant. Another research for health care providers such as physicians and nurses also presented that the SF12 mental health survey score (i.e., SF12MCS) was associated with work-related burnout and stress (Goodman & Schorling, 2012). However, in the current research, it was found that the SF12 physical health summary score (i.e., SF12PCS) was more useful than the mental summary score (i.e., SF12MCS) as a predictor of occupational fatigue in entry-construction worker conducting material handling tasks.

In the current study, it was expected that participants with higher physical capacity measured through the six-minute walk test were expected to be less likely to show fatigue status because of their physical endurance for a given task. However, the findings were the opposite: the participants with higher physical capacity (i.e., greater walking distance) reported higher fatigue after conducting the simulated construction task. In other words, greater physical capacity was associated with reports by participants that they were in fatigue status after task execution. This opposite result may be attributed to other factors that affect the results of the six-minute walk test, such as the effect of anthropometrics on walking capacity (Chetta et al., 2006) and the different participants' body sizes which may affect their stride in the six-minute walk test. If this is true, anthropometric data for each employee should be collected for fatigue management.

There are several other types of physical performance and endurance tests other than six-minute walk test, such as the lift and carry test, timed up and go test, and the stair climb test (Bennell et al., 2011). Most of the assessments methods were originally developed for senior patients suffering from a performance-related disability, such as knee osteoarthritis (Rejeski et al., 1995) and lower back pain (Strand, Moe-Nilssen, & Ljunggren, 2002). It is important to develop and administer easily applicable physical performance tests for healthy worker populations to manage fatigue in occupational settings that expose workers to heavy workloads.

With respect to wearable sensor data, and HRV signals specifically, although RMSSD is the one of frequently used time domain HRV measurements, with SNDD used in occupational health research (Togo & Takahashi, 2009), SNDD is preferred to RMSSD for predicting fatigue caused by material handling activity. In the case of SDNN, as in previous studies, the fatigue group was lower than in the no fatigue group, indicating that the fatigue caused a decrease in parasympathetic system activity. In the current study, the LF value was lower in the fatigue group than in the no-fatigue group, and the association of this LF with the fatigue group probability was statistically significant. The results of this study are consistent with the results of a previous study showing that the people, who experienced stress-induced fatigue showed a pattern of decreased LF (Olsson et al., 2010). Subjects who were fatigued reported a decrease in LF. LF is an index of sympathetic nerve activation, and the decrease in sympathetic nerve activity means the loss of in vivo energy supply. In a study of 52 middle-aged male workers who measured vital exhaustion (VE) with HRV, VE group showed low LF and HF mean amplitudes (Watanabe et al., 2002). In a study of electronics manufacturing company workers, Park, et al. (2001) found that the log-transformed LF of workers decreased as the weekly working hours increased, and that this association was statistically significant ($p = 0.046$). In computer work, HRV HF decreased with mental stress, but there was no difference in LF (Hjortskov et al. 2004).

Because LF is known to be an index reflecting both sympathetic activity and parasympathetic activity, this study concluded that LF could have a contradictory result against previous research that showed the higher LF with the increased fatigue level (Leti & Bricout, 2013) due to participants in the study and design of the experiment. Both LF and log-transformed LF were lower in the fatigue group, which is consistent with previous research (Baumert et al., 2006; Olsson et al., 2010). LF is an indicator of RR interval change reflecting both sympathetic and parasympathetic activities (Makivić, Nikić Djordjević, & Willis, 2013). However, the normalized unit of LF could be considered to assess the sympathetic efferent activity, despite the fact that the LF components are affected by both sympathetic and parasympathetic activities (Tarvainen et al., 2017). According to the results of several studies, the normalized unit of LF was used instead of the raw value of LF to report overtraining and exercise intensities of athletes. As shown in Table 5, the normalized unit of LF was higher in the fatigue group than in the no fatigue group using both the FFT and AR methods.

In contrast, the value of HF is only used as an index of parasympathetic activity, and the value of HF is also reported to decrease with increasing fatigue or stress. Studies using a normalized unit of HF variable, reported its negative relationship with an increase in exercise intensities and consistent results. Therefore, although the use of LF variables showed lower AIC and BIC values in logistic regression, the value of HF should be used as a predictor of fatigue.

A potential limitation of the current study is an inaccurate assessment of awkward postures and its role in promoting fatigue. A positive association between awkward trunk posture and muscle fatigue has been shown by previous studies such as Keyserling, Brouwer, & Silverstein (1993). However, the impact of the award trunk posture on physical fatigue is still rare. In the current study, participants could respond to the questionnaire on the CIS8R survey without differentiating the muscle fatigue from the physical fatigue. This means that the participant who actually experienced muscle fatigue due to the awkward position of his/her neck, shoulders and back during material handling tasks could report that he/she experienced the physical fatigue. Ergonomic factors related to manual handling tasks, such as awkward posture and frequency of lifting are also known to cause occupational fatigue of construction workers. A number of methodologies for measuring the award posture through a wearable sensor have been studied, although to date, no study has been conducted to determine which metric is a predictor of fatigue.

Other than the stepwise logistic regression method with AIC and BIC used in the current research, other automated model selection methods may be used to solve the problem of data overfitting and to help select prediction variables. In our use of logistic regression, it was found that choice of the threshold on fatigue was also important. If the fatigue status is classified with a high threshold level based on the CIS survey, fatigue among the participants will be very rare, making it difficult to have enough cases of fatigue to fit a model. Similar work has been conducted to fit logistic regression models to fatigue data. For instance, Maman, Yazdi, Cavuoto, & Megahed (2017) modeled a physical fatigue logistic regression model using a penalized regression technique. Penalized regression, which includes Ridge, Lasso, ElasticNet, etc., is frequently used in big data analysis as coefficients are often underestimated. The penalized regression is a type of shrinkage estimation method intended to solve the overfitting problem by imposing penalization on a coefficient.

The results of this study show that workers may benefit from wearing an ECG sensor during working hours and that HRV analysis and information based on RR interval data are important for the prediction of occupational fatigue. Therefore, it is necessary to consider not only what kind of data is collected, but also what is required for fatigue management and which factors affect the acceptance of wearable technologies. In particular, this study has important implications for managers who manage and analyze worker data, but also for workers who may benefit from wearing sensor technologies. In order to increase the adoption of wearable technologies, it is necessary to inspire social influence (e.g., “Do co-workers think wearable technologies are a good idea?”), as well as perceived usefulness and perceived privacy risks for workers and managers (Choi, Hwang, & Lee, 2017). In addition, as the wearable device experience and job position act as moderating factors, it seems that workers and managers need different approaches toward training for the application of fatigue management and wearable technologies.

6. Conclusions

Occupational fatigue has been known to be a direct or indirect cause of accidents and injuries in construction sites. However, the management of occupational fatigue in the construction field has been limited by a dependence on subjective methods such as surveys and the problem of individual worker management. Wearable technology and the popularization of data processing methods have led to various research studies on the application of information and communications technology as an objective method for the management of worker fatigue at construction equipment companies and other innovative organizations. However, as different assisting tools and data become increasingly available, it is necessary to study what instruments and data are most efficient for measuring fatigue. This study was conducted to provide a solution to this research need. For the practical use of the wearable technology in the field, the management method should be considered in such a way that the acquisition and analysis of the data are simple and any form of wearable sensor discomfort is minimized.

Research on fatigue measurements and predictions of construction workers using wearable sensors was carried out using various types of sensor technology and measurement variables. Researchers have demonstrated that prediction of fatigue using wearable sensors is promising and will be a workplace fatigue prevention management tool. However, in the case of practical fatigue management using wearable sensors, there is no clear guideline on what type of sensor to use and which are the most informative sensor data variables. The process of data collection and processing needs to be simple enough for the safety professionals to actually apply.

As a physiologic indicator to predict fatigue, this research found that heart rate variability is more useful than other heart rate measurement variables. Among the HRV variables, fatigue can be predicted with SDNN, which provides relatively simple-to-process data, including calculation of variables. From a practical point of view, if the calculation of the RR interval is possible in real time, wearable sensors will have great value for fatigue prediction in real-world field applications. Because the task force guidelines focus on HRV in clinical research, further research is needed for use in occupational settings of HRV. In addition to material handling, construction workers perform a variety of activities and are exposed to environmental factors that affect fatigue, such as noise, vibration, and heat. Therefore, how these various factors influence the fatigue prediction model and the fatigue risk management plan of the construction worker should be assessed through future research on the control of these factors.

This research found that among the various physiological status and sleep measurements obtained from ActiGraph and the wearable ECG monitor, the LFAR measurement (non-normalized unit) obtained by the frequency domain method of heart rate variability is an important factor for fatigue prediction. Also, the research showed that fatigue was not precisely predicted only by variables collected from the wearable sensors. Additional information was also needed, including perceived workload using NASA-TLX, the physical health status of workers based on SF12 short version survey, and information regarding the physical capacity for work. According to the results of this research, the resting heart rate and heart rate recovery measured before and after the material handling and installation task did not improve the selection criteria of the prediction model. Also, sleep measurement was not useful. Therefore, this study contributed to planning the effective occupational fatigue management by reducing the need to collect unnecessary data and promoting the use of more information valuable wearable sensors to workers and managers.

References

Abdelhamid, T. S., & Everett, J. G. (2002). Physiological demands during construction work. *Journal of Construction Engineering and Management, 128*(5), 427-437. https://doi.org/10.1061/(ASCE)0733-9364(2002)128:5(427)

Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., & Suri, J. S. (2006). Heart rate variability: a review. *Medical and Biological Engineering and Computing*, *44*(12), 1031-1051. https://doi.org/10.1007/s11517-006-0119-0

ActiGraph (2012), ActiLife 6 User's Manual, Retrieved January 7, 2018 from http://actigraphcorp.com/wp-content/uploads/2015/11/SFT12DOC13-ActiLife-6-Users-Manual-Rev-A-110315.pdf

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716-723. https://doi.org/10.1109/TAC.1974.1100705

Aryal, A., Ghahramani, A., & Becerik-Gerber, B. (2017). Monitoring fatigue in construction workers using physiological measurements. *Automation in Construction*, *82*, 154-165. https://doi.org/10.1016/j.autcon.2017.03.003

Baumert, M., Brechtel, L., Lock, J., Hermsdorf, M., Wolff, R., Baier, V., & Voss, A. (2006). Heart rate variability, blood pressure variability, and baroreflex sensitivity in overtrained athletes. *Clinical Journal of Sport Medicine*, *16*(5), 412-417. https://doi.org/10.1097/01.jsm.0000244610.34594.07

Bennell, K., Dobson, F., & Hinman, R. (2011). Measures of physical performance assessments: Self-Paced Walk Test (SPWT), Stair Climb Test (SCT), Six-Minute Walk Test (6MWT), Chair Stand Test (CST), Timed Up & Go (TUG), Sock Test, Lift and Carry Test (LCT), and Car Task. *Arthritis Care & Research, 63*(S11), S350-70. https://doi.org/10.1002/acr.20538

Beurskens, A. J., Bültmann, U., Kant, I., Vercoulen, J. H., Bleijenberg, G., & Swaen, G. M. (2000). Fatigue among working people: validity of a questionnaire measure. *Occupational and Environmental Medicine, 57*(5), 353-357. http://dx.doi.org/10.1136/oem.57.5.353

Boardman, A., Schlindwein, F. S., & Rocha, A. P. (2002). A study on the optimum order of autoregressive models for heart rate variability. *Physiological Measurement, 23*(2), 325. https://doi.org/10.1088/0967-3334/23/2/308

Camm, A. J., Malik, M., Bigger, J. T., Breithardt, G., Cerutti, S., Cohen, R. J., ... & Lombardi, F. (1996). Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation, 93*(5), 1043-1065. https://doi.org/10.1161/01.CIR.93.5.1043

Caruso, C. C., & Hitchcock, E. M. (2010). Strategies for nurses to prevent sleep-related injuries and errors. *Rehabilitation nursing, 35*(5), 192-197. http://dx.doi.org/10.1002/j.2048-7940.2010.tb00047.x

Caterpillar Inc. (n.d.), CAT® Smartband: Wrist-worn fatigue assessment. Retrieved January 4, 2018 from https://www.cat.com/en_US/support/operations/frms/smartband.html

Centers for Disease Control and Prevention (2016). Hierarchy of Controls. Retrieved January 4, 2018 from https://www.cdc.gov/niosh/topics/hierarchy/default.html

Chan, M. (2011). Fatigue: the most critical accident risk in oil and gas construction. Construction *Management and Economics, 29*(4), 341-353. https://doi.org/10.1080/01446193.2010.545993

Chemla, D., Young, J., Badilini, F., Maison-Blanche, P., Affres, H., Lecarpentier, Y., & Chanson, P. (2005). Comparison of fast Fourier transform and autoregressive spectral analysis for the study of heart rate variability in diabetic patients. *International Journal of Cardiology*, *104*(3), 307-313. https://doi.org/10.1016/j.ijcard.2004.12.018

Chetta, A., Zanini, A., Pisi, G., Aiello, M., Tzani, P., Neri, M., & Olivieri, D. (2006). Reference values for the 6-min walk test in healthy subjects 20–50 years old. *Respiratory Medicine*, *100*(9), 1573-1578. https://doi.org/10.1016/j.rmed.2006.01.001

Cheuvront, S. N., Carter, R., & Sawka, M. N. (2003). Fluid balance and endurance exercise performance. *Current Sports Medicine Reports*, *2*(4), 202-208. Retrieved from http://journals.lww.com/acsm-csmr/Abstract/2003/08000/Fluid_Balance_and_Endurance_Exercise_Performance.6.aspx

Choi, B., Hwang, S., & Lee, S. (2017). What drives construction workers' acceptance of wearable technologies in the workplace?: Indoor localization and wearable health devices for occupational safety and health. *Automation in Construction*, *84,* 31-41. https://doi.org/10.1016/j.autcon.2017.08.005

Cole, R. J., Kripke, D. F., Gruen, W., Mullaney, D. J., & Gillin, J. C. (1992). Automatic sleep/wake identification from wrist activity. *Sleep*, *15*(5), 461-469. https://doi.org/10.1093/sleep/15.5.461

Dey, A., & Mann, D. D. (2010). Sensitivity and diagnosticity of NASA-TLX and simplified SWAT to assess the mental workload associated with operating an agricultural sprayer. *Ergonomics, 53*(7), 848-857. https://doi.org/10.1080/00140139.2010.489960

Dishman, R. K., Nakamura, Y., Garcia, M. E., Thompson, R. W., Dunn, A. L., & Blair, S. N. (2000). Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *International Journal of Psychophysiology*, *37*(2), 121-133. https://doi.org/10.1016/S0167-8760(00)00085-4

Folkard, S., & Tucker, P. (2003). Shift work, safety and productivity. *Occupational Medicine, 53*(2), 95-101. https://doi.org/10.1093/occmed/kqg047

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and science in sports and exercise*, *30*(5), 777-781.

Garza, J. L., Cavallari, J. M., Eijckelhof, B. H., Huysmans, M. A., Thamsuwan, O., Johnson, P. W., ... & Dennerlein, J. T. (2015). Office workers with high effort–reward imbalance and overcommitment have greater decreases in heart rate variability over a 2-h working period. *International Archives of Occupational and Environmental Health*, *88*(5), 565-575. https://doi.org/10.1007/s00420-014-0983-0

Goodman, M. J., & Schorling, J. B. (2012). A mindfulness course decreases burnout and improves well-being among healthcare providers. *The International Journal of Psychiatry in Medicine*, *43*(2), 119-128.

Grandjean, E., & Kroemer, K. H. (1997). *Fitting the task to the human: a textbook of occupational ergonomics*. CRC press, Boca Raton, FL

Hallowell, M. R. (2010). Worker fatigue. *Professional Safety, 55*(12), 18-26. Retrieved from http://elcosh.org/record/document/2087/d001102.pdf

Hammer, L. B., & Sauter, S. (2013). Total worker health and work–life stress. *Journal of Occupational and Environmental Medicine, 55*, S25-S29. https://doi.org/10.1097/JOM.0000000000000043

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, *52,* 139-183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hartmann, B., & Fleischer, A. G. (2005). Physical load exposure at construction sites. *Scandinavian Journal of Work, Environment & Health, 31*(Suppl. 2), 88-95. Retrieved from http://www.jstor.org/stable/pdf/40967468.pdf

Haslam, R. A., Hide, S. A., Gibb, A. G., Gyi, D. E., Pavitt, T., Atkinson, S., & Duff, A. R. (2005). Contributing factors in construction accidents. *Applied Ergonomics, 36*(4), 401-415. https://doi.org/10.1016/j.apergo.2004.12.002

Hilbe, J. M. (2017). *Logistic regression models.* CRC Press, Boca Raton, FL

Hjortskov, N., Rissén, D., Blangsted, A. K., Fallentin, N., Lundberg, U., & Søgaard, K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *European Journal of Applied Physiology*, *92*(1-2), 84-89. https://doi.org/10.1007/s00421-004-1055-z

Hsiao, H., & Simeonov, P. (2001). Preventing falls from roofs: a critical review. *Ergonomics, 44*(5), 537-561. https://doi.org/10.1080/00140130110034480

Hsu, D. J., Sun, Y. M., Chuang, K. H., Juang, Y. J., & Chang, F. L. (2008). Effect of elevation change on work fatigue and physiological symptoms for high-rise building construction workers. *Safety Science, 46*(5), 833-843. https://doi.org/10.1016/j.ssci.2007.01.011

Hursh, S. R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., ... & Eddy, D. R. (2004). Fatigue models for applied research in warfighting. *Aviation, Space, and Environmental Medicine*, *75*(Supplement 1), A44-A53. Retrieved from https://www.fatiguescience.com/wp-content/uploads/2016/09/SAFTE-Validation-in-US-Army-Soldiers.pdf

Hursh, S. R., Raslear, T. G., Kaye, A. S., & Fanzone Jr, J. F. (2006). Validation and calibration of a fatigue assessment tool for railroad work schedules, summary report (No. DOT/FRA/ORD-06/21). Retrieved from https://www.fra.dot.gov/eLib/details/L02535

Jaffe, R. S., & Fung, D. L. (1994). Constructing a heart-rate variability analysis system. *Journal of Clinical Monitoring and Computing*, *10*(1), 45-58. https://doi.org/10.1007/BF01651466

Janssen, T. W., Van Oers, C. A., Van der Woude, L. H., & Hollander, A. P. (1994). Physical strain in daily life of wheelchair users with spinal cord injuries. *Medicine and Science in Sports and Exercise*, *26*(6), 661-670. Retrieved from https://insights.ovid.com/medicine-science-sports-exercise/mespex/1994/06/000/physical-strain-daily-life-wheelchair-users-spinal/2/00005768

Kang, M. G., Koh, S. B., Cha, B. S., Park, J. K., Woo, J. M., & Chang, S. J. (2004). Association between job stress on heart rate variability and metabolic syndrome in shipyard male workers. *Yonsei Medical Journal, 45,* 838-846. https://doi.org/10.3349/ymj.2004.45.5.838

Keyserling, W. M., Brouwer, M., & Silverstein, B. A. (1993). The effectiveness of a joint labor-management program in controlling awkward postures of the trunk, neck, and shoulders: Results of a field study. International Journal of Industrial Ergonomics, 11(1), 51-65. https://doi.org/10.1016/0169-8141(93)90054-H

Klaassen, C.D. & Watkins, J.B. (2010). *Casarett & Doull's Essentials of Toxicology* (2ed.). Retrieved December 08, 2017 from http://accesspharmacy.mhmedical.com/content.aspx?bookid=449&sectionid=39910786

Lee, W. & Migliaccio, G. C. (2016). Physiological cost of concrete construction activities. *Construction Innovation, 16*(3), 281-306. https://doi.org/10.1108/CI-10-2015-0051

Lee, W., Lin, K. Y., Seto, E., & Migliaccio, G. C. (2017). Wearable sensors for monitoring on-duty and off-duty worker physiological status and activities in construction. *Automation in Construction, 83,* 341-353. https://doi.org/10.1016/j.autcon.2017.06.012

Lee, W., Seto, E., Lin, K. Y., & Migliaccio, G. C. (2017). An evaluation of wearable sensors and their placements for analyzing construction worker's trunk posture in laboratory conditions. *Applied Ergonomics, 65,* 424-436. https://doi.org/10.1016/j.apergo.2017.03.016

Lerman, S. E., Eskin, E., Flower, D. J., George, E. C., Gerson, B., Hartenbaum, N., ... & Moore-Ede, M. (2012). Fatigue risk management in the workplace. *Journal of Occupational and Environmental Medicine, 54*(2), 231-258. https://dx.doi.org/10.4103%2F0972-6748.160915

Leti, T., & Bricout, V. A. (2013). Interest of analyses of heart rate variability in the prevention of fatigue states in senior runners. *Autonomic Neuroscience, 173*(1), 14-21. https://doi.org/10.1016/j.autneu.2012.10.007

Li, G., & Chung, W. Y. (2013). Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier. *Sensors*, *13*(12), 16494-16511. http://dx.doi.org/10.3390/s131216494

Li, W., & Nyholt, D. R. (2001). Marker selection by Akaike information criterion and Bayesian information criterion. *Genetic Epidemiology, 21*(Suppl. 1), S272-7. http://dx.doi.org/10.1002/gepi.2001.21.s1.s272

Li, Z., Wang, C., Mak, A. F., & Chow, D. H. (2005). Effects of acupuncture on heart rate variability in normal subjects under fatigue and non-fatigue state. *European Journal of Applied Physiology*, *94*(5-6), 633-640. https://doi.org/10.1007/s00421-005-1362-z

Long, J. S., & J. Freese. 2006. *Regression Models for Categorical Dependent Variables Using Stata.* 2nd ed. Stata Press, College Station, TX

MacDonald, W. (2003). The impact of job demands and workload on stress and fatigue. *Australian Psychologist, 38*(2), 102-117. http://dx.doi.org/10.1080/00050060310001707107

Makivić, B., Nikić Djordjević, M., & Willis, M. S. (2013). Heart Rate Variability (HRV) as a Tool for Diagnostic and Monitoring Performance in Sport and Physical Activities. *Journal of Exercise Physiology Online*, *16*(3), 103-131. Retrieved from https://www.asep.org/asep/asep/JEPonlineJUNE2013_Willis.pdf

Maman, Z. S., Yazdi, M. A. A., Cavuoto, L. A., & Megahed, F. M. (2017). A data-driven approach to modeling physical fatigue in the workplace using wearable sensors. *Applied Ergonomics, 65,* 515-529. https://doi.org/10.1016/j.apergo.2017.02.001

Martin, A., Chalder, T., Rief, W., & Braehler, E. (2007). The relationship between chronic fatigue and somatization syndrome: a general population survey. *Journal of*

Psychosomatic Research, 63(2), 147-156. https://doi.org/10.1016/j.jpsychores.2007.05.007

McDonald, J. H. (2009). *Handbook of biological statistics* (Vol. 2, pp. 173-181). Sparky House Publishing, Baltimore, MD

McFadden, D. (1973). *Conditional logit analysis of qualitative choice behavior. In Frontiers in Econometrics* (Edited by P. Zarembka), 105-42. Academic Press, Cambridge, MA

Medicore (n.d.). Heart rate variability analysis system: Clinical information (Version 3.0). Retrieved January 7, 2018 from http://medi-core.com/download/HRV_clinical_manual_ver3.0.pdf

Murray, S. L., & Thimgan, M. S. (2016). *Human Fatigue Risk Management: Improving Safety in the Chemical Processing Industry*. Academic Press, Cambridge, MA

Murtaugh, P. A. (2009). Performance of several variable-selection methods applied to real ecological data. *Ecology Letters, 12*(10), 1061-1068. https://doi.org/10.1111/j.1461-0248.2009.01361.x

Nag, P. K., & Patel, V. G. (1998). Work accidents among shiftworkers in industry. International *Journal of Industrial Ergonomics, 21*(3-4), 275-281. https://doi.org/10.1016/S0169-8141(97)00050-4

Nardone, A., Tarantola, J., Giordano, A., & Schieppati, M. (1997). Fatigue effects on body balance. *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control, 105*(4), 309-320. https://doi.org/10.1016/S0924-980X(97)00040-4

Olsson, E. M., Roth, W. T., & Melin, L. (2010). Psychophysiological characteristics of women suffering from stress-related fatigue. *Stress and Health, 26*(2), 113-126. http://dx.doi.org/10.1002/smi.1271

Park, J., Kim, Y., Cho, Y., Woo, K. H., Chung, H. K., Iwasaki, K., ... & Hisanaga, N. (2001). Regular overtime and cardiovascular functions. *Industrial Health*, *39*(3), 244-249. https://doi.org/10.2486/indhealth.39.244

Park, K., Rosengren, K. S., Horn, G. P., Smith, D. L., & Hsiao-Wecksler, E. T. (2011). Assessing gait changes in firefighters due to fatigue and protective clothing. *Safety Science, 49*(5), 719-726. https://doi.org/10.1016/j.ssci.2011.01.012

Patel, M., Lal, S. K. L., Kavanagh, D., & Rossiter, P. (2011). Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Systems with Applications, 38*(6), 7235-7242. https://doi.org/10.1016/j.eswa.2010.12.028

Pichon, A., Roulaud, M., Antoine-Jonville, S., de Bisschop, C., & Denjean, A. (2006). Spectral analysis of heart rate variability: interchangeability between autoregressive analysis and fast Fourier transform. *Journal of Electrocardiology*, *39*(1), 31-37. https://doi.org/10.1016/j.jelectrocard.2005.08.001

Raftery, A.E. (1986). Choosing models for cross-classifications. *American Sociological Review, 51*(1), 145-146. Retrieved from https://www.csss.washington.edu/~raftery/Research/PDF/asr1986.pdf

Rejeski, W. J., Ettinger, W. H., Schumaker, S., James, P., Burns, R., & Elam, J. T. (1995). Assessing performance-related disability in patients with knee osteoarthritis. *Osteoarthritis and Cartilage, 3*(3), 157-167. https://doi.org/10.1016/S1063-4584(05)80050-0

Ricci, J. A., Chee, E., Lorandeau, A. L., & Berger, J. (2007). Fatigue in the US workforce: prevalence and implications for lost productive work time. *Journal of Occupational and*

Environmental Medicine, 49(1), 1-10. https://doi.org/10.1097/01.jom.0000249782.60321.2a

Rodahl, K. (1989). *The Physiology of Work*, Taylor & Francis, Abingdon United Kingdom

Roma, P. G., Hursh, S. R., Mead, A. M., & Nesthus, T. E. (2012). Flight attendant work/rest patterns, alertness, and performance assessment: Field validation of biomathematical fatigue modeling. FEDERAL AVIATION ADMINISTRATION OKLAHOMA CITY OK CIVIL AEROSPACE MEDICAL INST. Retrieved from https://www.fatiguescience.com/wp-content/uploads/2016/09/SAFTE-Validation-Aviation.pdf

Rosa, R. R., & Colligan, M. (1997). Plain language about shiftwork. Retrieved from https://stacks.cdc.gov/view/cdc/5177

Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology, 53*(1), 61-86. http://dx.doi.org/10.1111/j.1464-0597.2004.00161.x

Russell, C. A., Caldwell, J. A., Arand, D., Myers, L., Wubbels, P., & Downs, H. (2000). Validation of the Fatigue Science Readiband Actigraph and Associated Sleep/Wake Classification Algorithms. Archinoetics, LLC. Retrieved from https://www.fatiguescience.com/wp-content/uploads/2016/09/Readiband-Validation-Accuracy.pdf

Sauter, S. (1999). Stress... at work. The National Institute for Occupational Safety and Health (NIOSH). Retrieved December 10, 2017, from https://www.cdc.gov/niosh/docs/99-101/

Schmitt, L., Regnard, J., Desmarets, M., Mauny, F., Mourot, L., Fouillot, J. P., ... & Millet, G. (2013). Fatigue shifts and scatters heart rate variability in elite endurance athletes. *PloS One, 8*(8), e71588. https://doi.org/10.1371/journal.pone.0071588

Schnorpfeil, P., Noll, A., Wirtz, P., Schulze, R., Ehlert, U., Frey, K., & Fischer, J. E. (2002). Assessment of exhaustion and related risk factors in employees in the manufacturing industry–a cross-sectional study. *International Archives of Occupational and Environmental health, 75*(8), 535-540. https://doi.org/10.1007/s00420-002-0369-6

Seynnes, O., Fiatarone Singh, M. A., Hue, O., Pras, P., Legros, P., & Bernard, P. L. (2004). Physiological and functional responses to low-moderate versus high-intensity progressive resistance training in frail elders. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 59*(5), M503-M509.

Strand, L. I., Moe-Nilssen, R., & Ljunggren, A. E. (2002). Back Performance Scale for the assessment of mobility-related activities in people with back pain. *Physical Therapy, 82*(12), 1213-1223. https://doi.org/10.1093/ptj/82.12.1213

Swaen, G. M. H., Van Amelsvoort, L. G. P. M., Bültmann, U., & Kant, I. J. (2003). Fatigue as a risk factor for being injured in an occupational accident: results from the Maastricht Cohort Study. Occupational and Environmental Medicine, 60(suppl 1), i88-i92. http://dx.doi.org/10.1136/oem.60.suppl_1.i88

Swartz, A. M., Strath, S. J., Bassett, D. R., O'brien, W. L., King, G. A., & Ainsworth, B. E. (2000). Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Medicine & Science in Sports & Exercise, 32*(9), S450-S456.

Tanaka, H., Monahan, K. D., & Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology, 37*(1), 153-156. Retrieved from http://www.onlinejacc.org/content/accj/37/1/153.full.pdf

Tarvainen, M. P., Lipponen, J., Niskanen, J. P.& Ranta-aho, P.O. (2017), Kubios HRV ver. 3.0.2 User's guide. Retrieved from http://www.kubios.com/downloads/Kubios_HRV_Users_Guide.pdf

Tarvainen, M. P., Niskanen, J. P., Lipponen, J. A., Ranta-Aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV–heart rate variability analysis software. *Computer Methods and Programs in Biomedicine, 113*(1), 210-220. https://doi.org/10.1016/j.cmpb.2013.07.024

Techera, U., Hallowell, M., Marks, E., & Stambaugh, N. (2016) Measuring Occupational Fatigue: A Comprehensive Review and Comparison of Subjective and Objective Methods. In Construction Research Congress 2016 (pp. 2905-2915). https://doi.org/10.1061/9780784479827.289

Techera, U., Hallowell, M., Stambaugh, N., & Littlejohn, R. (2016). Causes and consequences of occupational fatigue: meta-analysis and systems model. *Journal of Occupational and Environmental Medicine, 58*(10), 961-973. https://doi.org/10.1097/JOM.0000000000000837

The Center for Construction Research and Training (2013), The Construction Chart Book, 5th ed., The Center for Construction Research and Training. Retrieved from https://www.cpwr.com/sites/default/files/publications/5th%20Edition%20Chart%20Book%20Final.pdf

Togo, F., & Takahashi, M. (2009). Heart Rate Variability in Occupational Health―A Systematic Review. *Industrial Health, 47*(6), 589-602. https://doi.org/10.2486/indhealth.47.589

Tran, Y., Wijesuriya, N., Tarvainen, M., Karjalainen, P., & Craig, A. (2009). The relationship between spectral changes in heart rate variability and fatigue. *Journal of Psychophysiology, 23*(3), 143-151. https://doi.org/10.1027/0269-8803.23.3.143

Tudor-Locke, C., Barreira, T. V., Schuna Jr, J. M., Mire, E. F., & Katzmarzyk, P. T. (2013). Fully automated waist-worn accelerometer algorithm for detecting children's sleep-period time separate from 24-h physical activity or sedentary behaviors. A*pplied Physiology, Nutrition, and Metabolism, 39*(1), 53-57. https://doi.org/10.1139/apnm-2013-0173

U.S. Chemical Safety and Hazard Investigation Board (2007, March). Investigation Report: refinery explosion and fire (REPORT NO. 2005-04-I-TX). Retrieved from http://www.csb.gov/assets/1/19/CSBFinalReportBP.pdf

van Buuren, S., Groothuis-Oudshoorn, K., Robitzsch, A., Vink, G., Doove, L., & Jolani, S. (2014). Package 'mice'. Retrieved from https://cran.r-project.org/web/packages/mice/mice.pdf

van Hoogmoed, D., Fransen, J., Bleijenberg, G., & van Riel, P. L. (2008). How to assess fatigue in rheumatoid arthritis: Validity and reliability of the checklist individual strength. *Arthritis and Rheumatism, 58*, S868-S868.

van Hoogmoed, D., Fransen, J., Bleijenberg, G., & Van Riel, P. (2010). Physical and psychosocial correlates of severe fatigue in rheumatoid arthritis. *Rheumatology, 49*(7), 1294-1302. https://doi.org/10.1093/rheumatology/keq043

Vercoulen JHMM, Alberts M, & Bleijenberg G. (1999). The checklist individual strength (CIS). *Gedragstherapie, 32*, 131–6.

Vrijkotte, T. G., Van Doornen, L. J., & De Geus, E. J. (2000). Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability. *Hypertension*, *35*(4), 880-886. https://doi.org/10.1161/01.HYP.35.4.880

Ware Jr, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care*, *34*(3), 220-233.

Watanabe, T., Sugiyama, Y., Sumi, Y., Watanabe, M., Takeuchi, K., Kobayashi, F., & Kono, K. (2002). Effects of vital exhaustion on cardiac autonomic nervous functions assessed by heart rate variability at rest in middle-aged male workers. *International Journal of Behavioral Medicine*, *9*(1), 68-75. https://doi.org/10.1207/S15327558IJBM0901_05

Williamson, A., Lombardi, D. A., Folkard, S., Stutts, J., Courtney, T. K., & Connor, J. L. (2011). The link between fatigue and safety. *Accident Analysis & Prevention*, *43*(2), 498-515. https://doi.org/10.1016/j.aap.2009.11.011

Yamazaki, S., Fukuhara, S., Suzukamo, Y., Morita, S., Okamura, T., Tanaka, T., & Ueshima, H. (2007). Lifestyle and work predictors of fatigue in Japanese manufacturing workers. *Occupational Medicine*, *57*(4), 262-269. https://doi.org/10.1093/occmed/kqm006

Yi, W., & Chan, A. P. (2014). Which environmental indicator is better able to predict the effects of heat stress on construction workers?. *Journal of Management in Engineering*, *31*(4), 04014063. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000284

Zhang, M., Sparer, E. H., Murphy, L. A., Dennerlein, J. T., Fang, D., Katz, J. N., & Caban-Martinez, A. J. (2015). Development and validation of a fatigue assessment scale for US construction workers. *American Journal of Industrial Medicine*, *58*(2), 220-228. https://doi.org/10.1002/ajim.22411

Zhu, Y., Jankay, R. R., Pieratt, L. C., & Mehta, R. K. (2017, September). Wearable Sensors and Their Metrics for Measuring Comprehensive Occupational Fatigue: A Scoping Review. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 61, No. 1, pp. 1041-1045). Sage CA: Los Angeles, CA: SAGE Publications. Retrieved from http://journals.sagepub.com/doi/pdf/10.1177/1541931213601744