Applied Nursing Research 62 (2021) 151504

Contents lists available at ScienceDirect

Applied Nursing Research

journal homepage: www.elsevier.com/locate/apnr

A classification algorithm to predict chronic pain using both regression and machine learning – A stepwise approach

Pao-Feng Tsai [a,*], Chih-Hsuan Wang [b], Yang Zhou [c], Jiaxiang Ren [c], Alisha Jones [d], Sarah O. Watts [a], Chiahung Chou [e,f], Wei-Shinn Ku [c]

[a] School of Nursing, Auburn University, Auburn, AL 36849, United States of America
[b] Department of Educational Foundations, Leadership, and Technology, College of Education, Auburn University, Auburn, AL 36849, United States of America
[c] Department of Computer Science and Software Engineering, Samuel Ginn College of Engineering, Auburn University, Auburn, AL 36849, United States of America
[d] Department of Speech, Language, and Hearing Sciences, College of Liberal Arts, Auburn University, Auburn, AL 36849, United States of America
[e] Department of Health Outcomes Research and Policy, Harrison School of Pharmacy, Auburn University, Auburn, AL 36849, United States of America
[f] Department of Medical Research, China Medical University Hospital, Taichung City 40447, Taiwan

ARTICLE INFO

Keywords:
Pain
Physical activity
Sleep
Depression
Machine learning

ABSTRACT

This secondary data analysis study aimed to (1) investigate the use of two sense-based parameters (movement and sleep hours) as predictors of chronic pain when controlling for patient demographics and depression, and (2) identify a classification model with accuracy in predicting chronic pain. Data collected by Oregon Health & Science University between March 2018 and December 2019 under the Collaborative Aging Research Using Technology Initiative were analyzed in two stages. Data were collected by sensor technologies and questionnaires from older adults living independently or with a partner in the community. In Stage 1, regression models were employed to determine unique sensor-based behavioral predictors of pain. These sensor-based parameters were used to create a classification model to predict the weekly recalled pain intensity and interference level using a deep neural network model, a machine learning approach, in Stage 2. Daily step count was a unique predictor for both pain intensity (75% Accuracy, F1 = 0.58) and pain interference (82% Accuracy, F1 = 0.59). The developed classification model performed well in this dataset with acceptable accuracy scores. This study demonstrated that machine learning technique can be used to identify the relationship between patients' pain and the risk factors.

1. Introduction

Approximately 11–40% of the United States adult population suffers from chronic pain conditions (Dahlhamer et al., 2018). Chronic pain disproportionally affects older adults. The prevalence of chronic pain that is bothersome was reported by at least 50% of the elderly population (Patel et al., 2013). Clinicians depend largely on patients' verbal reports when making treatment decisions regarding chronic pain in clinical practice. When patients fail to report properly, chronic pain may not be adequately treated. This is very likely to happen in those who have impaired cognition, believe pain is a normal process of aging, are afraid of potential addiction to analgesics, or have suffered from chronic pain for a long time (Connolly et al., 2011; Hadjistavropoulos et al., 2007; Herr, 2011), which may lead to further disability (Prkachin et al., 2007).

Non-verbal observational tools have been developed, which allow an observer to estimate a patient's pain using pain cues, such as behaviors, voice, words and/or physiological signs. There is little evidence that current non-verbal observational tools are sensitive enough to detect pain relief after analgesics treatment (Cohen-Mansfield & Lipson, 2008; Herr et al., 2006; Herr et al., 2010; Lane et al., 2003). Using osteoarthritis pain as an example, one of the best available non-verbal tools, the Non-Communicative Patient's Pain Assessment Instrument (NOPPAIN), can only make 3% prediction of pain in older adults with memory problems and osteoarthritis (Horgas et al., 2007). The low explained variance indicates that other cues to pain are needed to improve predictability. In addition, these non-verbal tools require extensive staff training to conduct observations accurately, which makes these tools not feasible for clinical use.

Research attempting to identify pain using sensor-based parameters,

* Corresponding author at: School of Nursing, Auburn University, 710 South Donahue Drive, Office 3221, Auburn, AL 36849-5505, United States of America.
E-mail address: tsaipaofeng@auburn.edu (P. Tsai).

https://doi.org/10.1016/j.apnr.2021.151504
Received 9 June 2021; Received in revised form 28 August 2021; Accepted 22 September 2021
Available online 28 September 2021

such as movement, sleep, heart health, and others, has received significant attention in the last decade. The advancement of modern technology, such as wearable devices, makes it possible to detect these sensor-based parameters in real time and objectively in patients with pain. Studies proposed to classify pain status using sensor-based parameters collected through wearable devices mainly identify acute pain, including lab-induced pain (Ben-Israel et al., 2013; Chu et al., 2017; Cowen et al., 2015; Jiang et al., 2017; Johnson et al., 2019; Martini et al., 2015; Yang et al., 2018; Yang et al., 2019). However, when working with patients with chronic pain, a unique set of parameters may differ from patients suffering from acute pain. This study focused on two objective behavioral parameters, movement and sleep hours, because they are associated with chronic pain and are easily detected by most wearable devices.

Fear of movement is a natural human response to avoid exaggerated pain (Lethem et al., 1983). Restricted movement may cause patients to withdraw from family and social life and to be absent from work that leads to other emotional problems, which further exacerbates pain, limits physical activity, and results in disability (Hapidou et al., 2012). Sleep is another potential parameter associated with chronic pain. More than 40% of individuals with chronic pain suffer from a sleep disorder (Mathias et al., 2018). Pain sufferers are 13 times more likely to have insomnia than those without chronic pain (Mathias et al., 2018). Also, after controlling for mood and catastrophizing, shorter sleep duration remains significantly associated with higher intensity of chronic pain (Roberts & Drummond, 2016). Finally, mood, such as depression, is a potential comorbidity of chronic pain and needs to be considered when studying chronic pain (Bair et al., 2003).

The machine learning technique has gained attention in recent years due to its ability to analyze big data to learn the relationship between patients' data and treatment outcomes. In turn, this will enable precise resource allocation leading to lower treatment costs and improved outcomes (Sevakula et al., 2020). Evidence shows that machine learning techniques, coupled with brain imaging, electromyography, electroencephalography, and/or other advanced sensor-based technologies, can identify patients at risk of having pain (Boissoneault et al., 2017; Chu et al., 2017; Jiang et al., 2017; Tavakolian & Hadid, 2018). Unfortunately, these expensive sensor technologies are not practical and/or available to use in community clinics to assess for pain. Therefore, this program of research aims to identify a low-cost sensor-based predicting model to be used in community clinics for identifying pain, especially in the elderly population. This pilot study was the first step of this program of research. It used a traditional statistical method and machine learning approach to achieve the following aims: (1) investigate the predictability of two sense-based parameters (movement and sleep) on chronic pain when controlling for demographics and depression, and (2) to identify a classification model with accuracy in predicting chronic pain using a deep neural network modeling, a machine learning approach.

2. Methods

2.1. Study design and sample

This study was a secondary data analysis project. The dataset was obtained through the Collaborative Aging Research using Technology (CART) Initiative. The CART has established a digital technology research platform to assess multiple wellness measurements and other activities longitudinally in older adult households. It was deployed to 232 US households with four diverse cohorts (African American, Hispanic, low-income, and predominantly rural-residing veterans) of older adults in Portland, OR; Chicago, IL; and Miami, FL. Original IRB approval was obtained from the four corresponding Institutional Review Boards at Oregon Health & Science University, the VA Portland Health Care System, Rush University Medical Center, and University of Miami & Weill Cornell Medicine. Study criteria were (1) older adults aged 62 years or older, (2) living independently or with a partner, (3) in a larger than a one-room apartment, (4) no dementia, (5) age and education adjusted Montreal Cognitive Assessment >18, (6) with current internet access or willingness to acquire access, and (7) with basic computer use experience (send/receive email). Exclusion criteria included (1) conditions that would limit physical participation at the entry to study, such as wheelchair-bound, (2) any uncontrolled medical condition, e.g., late-stage cancer, that was expected to preclude completion of the study, (3) more than two people living in the participant's residence (overnight visitors are acceptable) or (4) depression (Geriatric Depression Scale >5).

This current study only used data collected by Oregon Health & Science University because the available dataset at this site collected between March 2018 and December 2019 was the most complete one as compared to other sites. At this site, participants were of low income, predominantly living in Section 202 housing. This study was approved by the IRB of the authors' institution.

2.2. Procedure

This study included two stages. In Stage 1, a series of multiple linear regression analyses with a backward elimination procedure was used to select sensor-based predictors of verbally reported pain intensity and interference. Also, a multivariate regression analysis was used to determine if selected parameters were unique predictors of pain but not of depressive mood. In Stage 2, a final set of sensor-based parameters was used to create a classification model using a deep neural network modeling, a machine learning approach.

2.2.1. Stage 1. Selecting predictor variables using statistical analysis

Due to the enormous amount of available longitudinal sensor data, Stage 1 of this study only used four months (January, April, July, and October) of 2019 data to select the final set of parameters for Stage 2. The rationale to select those four months was that they distributed evenly across different seasons in 2019, and we assumed results from those four months would be sufficient to provide a general picture of the participants' wellbeing and activities.

Potential sensor-based predictor variables included numbers of steps and hours of sleep. A daily number of steps and daily hours of sleep obtained from the sensor data file were measured by Withings fitness tracker (Withings Inc., France). Daily data were averaged to create two major variables, daily step counts and daily sleep hours for the week, before merging with other data files for analysis.

Covariates included age, sex, and moods. Participants were asked to fill in their actual age by whole numbers. Sex was dummy coded before data analysis to indicate 1 = “male” and 0 = “female.” Two dichotomous items (felt blue and felt lonely in the last week), with 1 being yes and 2 being no, were obtained weekly. They were dummy coded to indicate 1 = “felt blue/lonely” and 0 = “no” and then combined to create an index of weekly mood score ranging from 0 to 2. A higher score indicated worse mood.

The dependent variables in the analysis were two pain variables; verbal report of weekly pain intensity and pain interference. The pain intensity scale ranged from 0 to 10, with 0 indicating no pain and 10 indicating the worst pain level in the last week. Pain interference with normal activities or work ranged between 1 and 5, with 1 indicating no interference and 5 indicating high interference.

Baseline data (age and sex), weekly report (mood, pain intensity, and pain interference), and sensor data (average daily step counts and average daily sleep hours of the week) from selected weeks (n = 16) were combined using the participant identification number. Data then were imported into the Statistical Package for Social Sciences (version 25) for data analyses. Descriptive statistics were used to describe the characteristics of the participants. The total number of participants was 77. However, the total number of participants used in the analysis each week varied (n = 43–59) because not all participants reported their weekly health updates or had sensor data available.

The first aim of this project was to select a unique set of sensor-based predictor variables that could be used as a proxy for identifying a moderate level of pain and pain interference. Therefore, a series of multiple regression analyses using the backward elimination procedure was conducted to explore the prediction model between two sensor-based predictor variables (numbers of steps and hours of sleep) and the two chronic pain outcomes (pain intensity and pain interference) while controlling for age, sex, and mood. Chronic pain and lower mood level are usually co-occurring conditions (Bair et al., 2003). To identify unique predictors for pain, the predictors identified by the multiple regression models were analyzed in a multivariate regression analysis, with number of steps, as a single predictor, to predict the linear combination of pain levels and mood.

2.2.2. Stage 2. Machine learning modeling for building a classification algorithm

2.2.2.1. Dataset description. The objective of Stage 2 was to predict the weekly recalled pain intensity and interference level based on the seven days of step counts in previous weeks, instead of based on an average daily steps of the week, as demonstrated in Stage 1. Stage 2 used the data collected between 3/30/2018 and 12/23/2019. The final dataset included weekly report of pain intensity and pain interference, and daily step count. We used identification number and report date as matching keys to merge all data into one unified dataset. After merging the data, each record contained a sequence of 7 daily step counts of an individual and his/her corresponding weekly pain intensity and weekly pain interference. There was a total of 1930 sequences in the entire dataset. We partitioned the original pain intensity (Range 0–10) and pain interference (Range 1–5) scores into two categories: "no or mild" vs "moderate or higher" level of pain intensity and pain interference. The splitting threshold of the pain intensity and pain interference scores were 3 and 2, respectively. About 432 and 458 sequences had a moderate or higher level of pain intensity and interference, respectively.

2.2.2.2. Training. Stage 1 identified the number of steps as a single parameter to predict pain intensity and pain interference. Stage 2 aimed to build a machine learning-based classification algorithm using the number of steps as a single parameter for identifying the individuals with moderate or higher pain intensity or pain interference. Also, the machine learning technique allowed for predicting the weekly pain intensity and pain interference levels based on seven daily step counts in previous weeks, instead of using an average step count as in Stage 1.

To evaluate the performance of the classification algorithm, we divided the entire dataset into three parts: training, validating, and testing sets. In this experiment, 80%, 10%, and 10% of data were used as the training set, validating set, and testing set, respectively. To make the training process smooth and stable, we standardized the data using the following formula before feeding the data into the classification algorithm.

$$Z = (x - \mu)/\sigma$$

where x is raw data, μ is the mean, σ is the standard deviation, and z is transformed data.

2.2.2.3. Machine learning method. Conventional (shallow) machine learning-based classification algorithms, like support vector machine (SVM) and Decision Tree, often ignore the correlations between samples' input data. On the other hand, the last generation of the deep learning-based classification algorithm, Recurrent Neural Network (RNN), fails to work well on long-sequence data. It cannot memorize all previous samples and forget some of the samples, resulting in poor performance of sequential prediction. Therefore, Long Short-Term Memory (LSTM) was utilized to solve the classification problem because it effectively predicted long sequences. It memorized previous samples (or aggregate the information of previous samples) in the sequence, and this allowed looking back to the previous context when predicting current samples (Hochreiter & Schmidhuber, 1997).

2.2.2.4. Implementation. Fig. 1 is an illustrative example of our LSTM-based prediction model for classifying the pain intensity level. We first used LSTM as an encoder to take daily steps as input. The model propagated and aggregated the data from the first day to the last day. After several recurrent forward propagations, the information of the entire sequence was encoded in the last hidden state unit H_{t+1}. And then, a fully connected neural layer (e.g., a rectangle in the figure) was employed to analyze the encoded information and output an actual number between 0 and 1, where 1 indicated a high level of pain intensity. We made use of a similar LSTM-based prediction model for classifying the pain interference level. Finally, the splitting thresholds of 0.3 for the pain intensity and 0.4 for the pain interference were applied on that scalar to predict whether a patient had a moderate or higher level of pain intensity and pain interference in corresponding weeks, respectively.

The classification task was essentially a binary classification problem with the binary cross-entropy loss as an objective function, i.e., the loss between ground-truth labels and predicted labels. The LSTM encoder consisted of two neural layers: the first layer had 64 neurons, and the second layer had 128 neurons. The size of the output layer was 1, and the Sigmoid function was used as the activation function of the last layer. The stochastic gradient descent method was used to optimize the parameters of neural networks. The model was implemented in Python and PyTorch, both popular programming languages and platforms to implement deep learning models.

2.2.2.5. Evaluation. As introduced earlier, our classification algorithm used the number of steps for predicting moderate or higher pain intensity or pain interference. The output of our model was a real number between 0 and 1. Here, we used the splitting thresholds of 0.3 for pain intensity and 0.4 for pain interference. Two popular metrics of Accuracy and F1 were used to evaluate the classification performance (Olson & Delen, 2008; Sasaki, 2007). Both Accuracy and F1 scores were bounded in the range of [0, 1]. A larger value of Accuracy or F1 indicated a better classification result. We used the following equations to calculate accuracy and F1 score.

$$\text{Accuracy} = \frac{\text{Number of Test Data That Are Correctly Predicted}}{\text{Number of All Test Data}}$$

$$\text{F1} = 2 \text{ x} \frac{\text{Positive Predictive Value x Sensitivity}}{\text{Positive Predictive Value} + \text{Sensitivity}}$$

where

Sensitivity = TP/(TP + FN), and

Positive predictive value = TP/(TP + FP)

where TP, FN, and FP can be defined as the following table

		Predicted class label	
		Class = 1	Class = 0
Actual class label	Class = 1	True positive (TP)	False negative (FN)
	Class = 0	False positive (FP)	True negative (TN)

In addition, sensitivity, specificity, positive predictive value, and negative predictive value were reported as indexes to evaluate the classification algorithm. Further, the receiver operating characteristic (ROC) curve was examined to evaluate how well the algorithm can make the classifications. A curve far away from the diagonal line is preferred. Then, the area under the curve range (AUC range) was examined, with a desirable AUC range of at least 0.70 (Meyers et al., 2006).

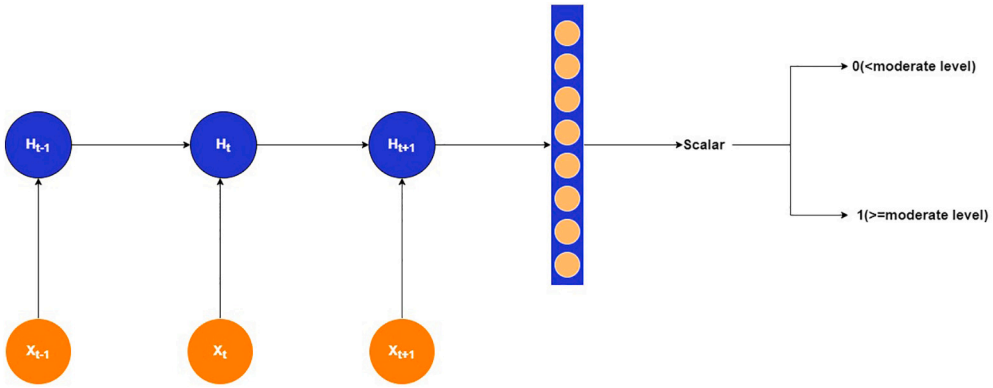

Fig. 1. LSTM-based Prediction Model for Level of Pain Intensity.

3. Results

3.1. Demographics

Seventy-seven participants with ages ranging from 62 to 89 years old (M = 71.83, SD = 6.08) provided baseline information (Table 1). Seventy-four percent ($n = 57$) of participants were females. The majority of the participants were white ($n = 74$, 96.10%) and only one of the participants reported as Hispanic origin (1.3%). The years of education ranged from 5 to 20 years (M = 15.16, SD = 2.60). A majority of them were divorced ($n = 42$, 54.54%), never married ($n = 17$, 22.08%), and widowed ($n = 8$, 10.39%). Average mood score, step counts per day and hours of sleep per day were 0.07 (SD = 0.31), 2813 (SD = 2363) and 8.48 (SD = 5.67), respectively. Pain intensity was an average of 2.04 (SD = 2.21) out of 10 and interference was an average of 1.86 (SD = 1) out of 5.

3.2. Multiple linear regression with backward procedures and five predictors

For the level of pain intensity (Table 2, Left Panel), 14 out of 16 regression models were statistically significant, except Week 29 and 42. The total variances which could be accounted for ranged from 8.2% to 30.3% for those significant models. Average daily steps of the week appeared to be the most dominant among all predictors since it statistically significantly predicted the pain intensity in 11 out of 16 weeks. The second important predictor to predict pain intensity was mood in 7 out of 16 significances. Age was a significant predictor of pain level in week 17, while sex and average daily sleep hours of the week did not predict pain intensity.

For pain interference (Table 2, Right Panel), 14 out of 16 regression models were statistically significant, except Weeks 29 and 42. The total variances which could be accounted for ranged from 9.9% to 34.4% for the significant models. The average daily steps of the week was also a dominant predictor. It statistically significantly predicted 10 out of 16 weeks of pain interference. Mood was another important predictor as well, with 7 out of 16 regression models demonstrating significance. The average daily sleep hours of the week was able to predict the level of pain interference in Week 2, while age and sex did not predict pain interference.

Table 1
Characteristics of the study sample ($n = 77$).

Variable	*n*	%	*M* (*SD*)
Age	77		71.83 (6.08)
Sex			
Male	20	25.97	
Female	57	74.03	
Race			
White	74	96.10	
Other	3	3.90	
Hispanic			
Yes	1	1.30	
No	76	98.70	
Years of education	77		15.16 (2.60)
Marriage status			
Married	4	5.19	
Widowed	8	10.39	
Divorced	42	54.54	
Separated	3	3.90	
Never married	17	22.08	
Unknown	3	3.90	
Mood			0.07 (0.31)
Step Count/day			2813 (2363)
Sleep (Hour/day)			8.48 (5.67)
Pain Intensity			2.04 (2.21)
Pain Interference			1.86 (1.00)

3.3. Multivariate regression analysis with simple regression follow-up

Results of multivariate regression analysis using step count as a single predictor indicated that 11 out of 16 weeks, the regression model for predicting both pain intensity and mood simultaneously were statistically significant, except Week 29 and Weeks 40–43 (Table 3, Left Panel). The Wilk's λ ranged from 0.74 to 0.89 for those significant models. Follow-up simple regression analyses indicated that average daily steps could successfully predict the level of pain intensity in the 11 weeks identified in the multivariate regression analysis. However, step count was not a significant predictor for mood.

For predicting the level of pain interference and mood simultaneously, results showed that 7 out of 16 weeks were statistically significant with the Wilk's λ ranging from 0.76 to 0.87 for those significant models. Simple regression follow-up analyses for those significant weeks indicated that step counts significantly predicted the level of pain interference in those seven weeks, and was also a significant predictor for Week 1, 17, and 28. However, it was not a significant predictor for mood in any of those weeks (Table 3, Right Panel).

3.4. Machine learning to identify classification model

Since step 1 only identified step count as the dominated sensor-based parameter for predicting pain variables, machine learning procedures only kept step count as the predictor. With a single predictor, step count, the classification model for pain intensity achieved 75% accuracy with F1 equal to 0.58 on test data. For pain interference, the classification model achieved 82% accuracy with 0.59 of F1 on test data. The AUC for these models were 0.7841 and 0.7921, indicating an acceptable performance to distinguish participants in different classes. Detailed evaluation indexes are presented on Table 4.

Table 2
Results of multiple regression with backward procedure analysis.

Week	*n*	Dependent variable: pain intensity						Dependent variable: pain interference					
		R^2 (%)	Step *p*	Sleep *p*	Mood *p*	Age *p*	Sex *p*	R^2 (%)	Step *p*	Sleep *p*	Mood *p*	Age *p*	Sex *p*
W1: 1/1–1/6	53	15.1*	0.005*	–	–	–	–	18.8*	0.036*	–	0.099	–	0.088
W2: 1/7–1/13	54	24.8*	0.003*	–	0.012*	–	–	34.4*	0.016*	0.029*	0.001*	–	–
W3: 1/14–1/20	55	17.6*	0.002*	–	–	–	–	18.8*	0.008*	0.079	–	–	–
W4: 1/21–1/27	55	20.2*	0.004*	0.072	–	–	–	19.2*	0.008*	0.053	–	–	–
W14: 4/1–4/7	56	11.3*	0.013*	–	–	–	–	13.1*	0.007*	–	–	–	–
W15: 4/8–4/14	59	15.1*	0.004*	–	–	–	–	21.6*	0.007*	–	0.018*	–	–
W16: 4/15–4/21	56	30.3*	0.032*	–	0.003*	0.056	–	29.5*	0.002*	–	0.003*	–	–
W17: 4/22–4/28	59	22.4*	0.031*	–	0.042*	0.044*	–	19.3*	0.020*	–	0.009*	–	–
W27: 7/1–7/7	43	25.1*	0.001*	–	–	–	–	24.3*	0.001*	–	–	–	–
W28: 7/8–7/14	51	12.3*	0.015*	–	–	–	–	10.5*	0.026*	–	–	–	–
W29: 7/15–7/21	48	6.4	0.091	–	–	–	–	6.9	0.077	–	–	–	–
W30: 7/22–7/28	48	26.4*	0.014*	–	0.008*	–	–	14.2*	0.088	–	0.064	–	–
W40: 9/30–10/6	54	8.2*	–	–	0.043*	–	–	12.4*	–	–	0.012*	–	–
W41: 10/7–10/13	55	10.3*	–	–	0.020*	–	–	9.9*	–	–	0.023*	–	–
W42: 10/14–10/20	56	5.9	–	–	0.080	–	–	3.3	–	–	0.193	–	–
W43: 10/21–10/27	53	17.6*	0.075	–	0.007*	–	–	13.4*	0.092	–	0.025*	–	–
Num. of significance out of 16 weeks		14	11	0	7	1	0	14	10	1	7	0	0

* $p < .05$.

Table 3
Results of multivariate regression analysis using step count as the single predictor with follow-up regression analysis.

Week	*n*	Dependent variables							Dependent variables						
		Wilk's λ	*F*	*df*	Pain Intensity		Mood		Wilk's λ	*F*	*df*	Pain Interference		Mood	
					R^2 (%)	*P*	R^2 (%)	*p*				R^2 (%)	*p*	R^2 (%)	*p*
W1: 1/1–1/6	53	0.85	4.41*	2,49	15.1	0.004*	0.8	0.529	0.92	2.14	2,49	8.0	0.042*	0.8	0.529
W2: 1/7–1/13	54	0.81	5.81*	2,50	14.1	0.006*	0.6	0.580	0.86	4.25*	2,50	9.3	0.026*	0.6	0.580
W3: 1/14–1/20	55	0.80	6.43*	2,52	17.6	0.001*	0.8	0.511	0.84	5.01*	2.52	13.3	0.006*	0.8	0.511
W4: 1/21–1/27	55	0.84	5.00*	2,52	14.6	0.004*	0.3	0.682	0.87	4.03*	2,52	12.5	0.008*	0.3	0.682
W14: 4/1–4/7	56	0.86	4.36*	2,53	11.3	0.011*	1.6	0.348	0.84	5.08*	2,53	13.1	0.006*	1.6	0.348
W15: 4/8–4/14	59	0.85	5.11*	2,56	15.1	0.002*	0.0	0.940	0.87	4.14*	2,56	12.5	0.006*	0.0	0.940
W16: 4/15–4/21	56	0.89	3.38*	2,53	11.1	0.012*	0.2	0.777	0.83	5.29*	2,53	15.8	0.002*	0.2	0.777
W17: 4/22–4/28	59	0.88	3.80*	2,56	11.1	0.010*	0.0	0.900	0.90	2.99	2.56	8.3	0.027*	0.0	0.900
W27: 7/1–7/7	43	0.74	7.07*	2,40	25.1	0.001*	1.9	0.376	0.76	6.47*	2,40	24.3	0.001*	1.9	0.376
W28: 7/8–7/14	51	0.87	3.60*	2,48	12.3	0.011*	0.9	0.502	0.89	3.09	2,48	10.5	0.020*	0.9	0.502
W29: 7/15–7/21	48	0.92	1.94	2,45	6.4	0.084	0.7	0.567	0.93	2.14	2,45	6.9	0.072	0.7	0.567
W30: 7/22–7/28	48	0.86	3.81*	2.45	12.4	0.014*	0.0	0.895	0.93	1.73	2,45	6.7	0.076	0.0	0.895
W40: 9/30–10/6	54	0.96	1.15	2,51	4.3	0.134	0.7	0.543	0.96	1.13	2,51	4.2	0.135	0.7	0.543
W41: 10/7–10/13	55	0.95	1.50	2,52	5.3	0.091	1.0	0.461	0.96	1.03	2,52	3.6	0.165	1.0	0.461
W42: 10/14–10/20	56	0.97	0.78	2,53	2.7	0.225	0.7	0.542	0.98	0.55	2,53	1.7	0.337	0.7	0.542
W43: 10/21–10/27	53	0.90	2.73	2,50	3.6	0.177	2.1	0.300	0.91	2.44	2,50	3.6	0.173	2.1	0.300

* $p < .05$.

Table 4
Evaluation results of the classification models.

	Accuracy	F1	AUC	Sensitivity	Specificity	PPV	NPV
Pain intensity	75%	0.58	0.7841	0.7083	0.7639	0.5000	0.8871
Pain interference	82%	0.59	0.7921	0.4800	0.9437	0.7500	0.8375

Note: AUC: Area Under the ROC Curve; PPV: Positive Predictive Value; NPV: Negative Predictive Value.

4. Discussion

This study found that the number of daily steps was a unique sensor-based predictor for both chronic pain outcomes, and the classification model performed well in this dataset with acceptable accuracy scores. According to the fear-avoidance model, fear of pain prevents an individual from performing physical activities, as restricting activities is believed to avoid further exaggerating pain conditions (Lethem et al., 1983). In the long run, the fear of movement decreases physical fitness and physical activity, increases social, occupational, and family functional disability, and causes depression (Thompson et al., 2010; Vlaeyen et al., 1995). Based on this theory, fear of pain is rather a consequence than an antecedent of pain severity. For the same reason, restricted movement is also a result of a high level of severity of pain. However, this study did not intend to tease out this causal relationship. Instead, as measured by number of steps, the restricted movement was used as a proxy or an indicator of the severity of pain, which we successfully demonstrated in this study.

This study failed to show hours of sleep as a proxy of pain. A study using the Sleep Heart Health Study dataset showed that the duration of total sleep time was associated with the level of pain (Weingarten et al., 2016). Another study found that when administering a once-a-day extended-release of morphine sulfate AVINZA to reduce pain, patients showed increased total sleep time (Rosenthal et al., 2007). Both studies used polysomnography for identifying total sleep time. On the other hand, the hours of sleep chosen for the current study were the total hours

of sleep epochs measured by Withings fitness tracker, which may explain the difference in study findings. Alternatively, a night of poor sleep intensified pain only in the following morning in patients with osteoarthritis (Whibley et al., 2019). The pain measures in this study were weekly reported data. To examine the association between pain measures and sleep hours, this study averaged seven-day sleep hours to obtain an averaged daily sleep hours of the week in Stage 1. This strategy might prevent the team from identifying an association between pain measure and sleep hours and thus sleep hours was not included in Stage 2.

As mentioned earlier, depression should be a consequence of chronic pain based on the fear-avoidance model (Lethem et al., 1983). It was also evidenced in a recent study showing that pain intensity and other variables significantly affected the level of depression (Dong et al., 2020). This study did identify mood, an indicator of depression, as the second significant covariate of both pain variables in Stage 1. Since the purpose of Stage 2 was to identify a predicting algorithm using sensor-based parameters, mood was not included in Stage 2 modeling.

Instead, mood was examined in Stage 1 to identify the unique ability of the steps in predicting pain intensity and interference because the number of steps could also be a proxy of depression. This postulate was based on study findings showing that physical exercise had a protective effect on developing depression in the elderly population (Thapa et al., 2020). The same study also found that sedentary behavior was strongly associated with depression symptoms (Del Pozo Cruz et al., 2020). Thus, a low number of steps could potentially indicate depression. Surprisingly, the number of steps was a unique predictor for chronic pain, but not for mood, in the current study. The potential reason could be that this study used a mood index created from two questions instead of a full depression scale used in other studies, although these two questions were included in both SF-12 and PROMISE scales. Further research will be needed to verify this finding.

This study showed acceptable accuracy, F1 and AUC values for the final predicting models. However, the positive predictive value for the model predicting pain intensity needs to be improved. In addition, the model predicting pain interference had a poor sensitivity, indicating additional or alternative sensor-based parameters should be identified.

The method used in this exploratory study included using the traditional statistical procedures to identify important predictors and then using machine learning to establish prediction models for pain intensity and pain inferences. The purpose of using traditional statistical model is to focus on inference, while machine learning aims to develop generalizable inferences from the data to make a more precise prediction (Bzdok et al., 2018). Hence, our approach combining both traditional statistical method and machine learning provides an efficient, focused, and practical method to develop the predictive model.

Limitations of this study include limited sensor-based parameters, small sample size, and a low number of subjects with a moderate or higher level of pain intensity and interference. Future research should expand on the variables included in this study, include larger sample size, and include more subjects with a moderate or higher level of pain. Also, different types of pain were lumped together when asking weekly questions about pain in this study, which might not accurately reflect patents' pain and reduced the predicting ability of the current classification model. Studies showed a higher level of reported osteoarthritis pain in the morning than in the evening. The most severe pain occurs between 2 and 6 pm in patients with pain, for example (Allen et al., 2009; Keefe et al., 2004). This suggests that pain may need to be measured daily and measured more than once.

If the proposed algorithm can be created successfully to generate a risk score, it can be built into an existing wearable device, such as Fitbit, or written as an app that can be downloaded to smartphones or other devices. A clinician can then use the device or the app as a "screening" tool for pain. In addition, the risk score may allow clinicians to identify which patients are most at risk and provide timely treatment (Sevakula et al., 2020). These functions will help inform treatment decisions to improve patients' health outcomes and quality of life. This study lays the foundation for a future study to assess pain in at-risk populations, such as older adults with severe cognitive impairment or impaired communication. Also, the machine learning model may be able to differentiate between various causes of changes in parameter values, such as cardiopulmonary cause vs musculoskeletal cause. Future studies will compare algorithms established for other causes, such as cardiopulmonary, against the results of this study examining pain in older adults.

Potential clinical implications of this study include the need to educate and counsel patients with chronic pain to continue moving. Clinicians should work with patients and caregivers to develop realistic plans to ensure patients consistently moving in their environments while maintaining safety. Lack of movement has been shown to restrict movement further, leading to muscle stiffness, more pain, and increased fall risks. These variables can contribute to patient isolation, which can further affect a patient's social and emotional well-being.

CRediT authorship contribution statement

Pao-Feng Tsai: Conceptualization, Methodology, Validation, Investigation, Writing-Original Draft, Writing-Review & Editing, Supervision. **Chih-Hsuan Wang**: Conceptualization, Methodology, Validation, Formal Analysis, Investigation, Data Curation, Writing-Original Draft, Writing-Review & Editing. **Yang Zhou**: Methodology, Writing-Review & Editing. **Jiaxiang Ren**: Methodology, Validation, Formal Analysis, Investigation, Data Curation, Writing-Original Draft, Writing-Review & Editing, Visualization. **Alisha Jones**: Writing-Original Draft. **Sarah O. Watts**: Writing-Original Draft. **Chiahung Chou**: Conceptualization, Writing-Review & Editing. **Wei-Shinn Ku**: Writing-Review & Editing.

Declaration of competing interest

The authors have no conflicts of interest to declare.

Acknowledgments

The Auburn University School of Nursing provided seed funding for this project. Data set was provided by the Collaborative Aging-in-place Research Using Technology (CART) initiative through funding from National Institutes of Health (U2CAG0543701, P30AG024978, UL1TR002369, U2CAG054397, P30AG008017, RF1AG22018, P30AG10161, and K01AG059839) and Veteran's Administration grant (IIR 17-144).

Appendix A

Classification table for pain intensity

		Predicted	
		Moderate to severe pain	No to mild pain
Actual	Moderate to severe pain	17	7
	No to mild pain	17	55

Classification table for pain interference

		Predicted	
		Moderate to severe interference	No to mild interference
Actual	Moderate to severe interference	12	13
	No to mild interference	4	67

References

Allen, K. D., Coffman, C. J., Golightly, Y. M., Stechuchak, K. M., & Keefe, F. J. (2009). Daily pain variations among patients with hand, hip, and knee osteoarthritis. *Osteoarthritis and Cartilage, 17*, 1275–1282.

Bair, M. J., Robinson, R. L., Katon, W., & Kroenke, K. (2003). Depression and pain comorbidity: A literature review. *Archives of Internal Medicine, 163*, 2433–2445.

Ben-Israel, N., Kliger, M., Zuckerman, G., Katz, Y., & Edry, R. (2013). Monitoring the nociception level: A multi-parameter approach. *Journal of Clinical Monitoring and Computing, 27*, 659–668.

Boissoneault, J., Sevel, L., Letzen, J., Robinson, M., & Staud, R. (2017). Biomarkers for musculoskeletal pain conditions: Use of brain imaging and machine learning. *Current Rheumatology Reports, 19*, 5.

Bzdok, D., Altman, N., & Krzywinski, M. (2018). Statistics versus machine learning. *Nature Methods, 15*, 233–234.

Chu, Y., Zhao, X., Han, J., & Su, Y. (2017). Physiological signal-based method for measurement of pain intensity. *Frontiers in Neuroscience, 11*, 279.

Cohen-Mansfield, J., & Lipson, S. (2008). The utility of pain assessment for analgesic use in persons with dementia. *Pain, 134*, 16–23.

Connolly, A., Gaehl, E., Martin, H., Morris, J., & Purandare, N. (2011). Underdiagnosis of dementia in primary care: Variations in the observed prevalence and comparisons to the expected prevalence. *Aging & Mental Health, 15*, 978–984.

Cowen, R., Stasiowska, M. K., Laycock, H., & Bantel, C. (2015). Assessing pain objectively: The use of physiological markers. *Anaesthesia, 70*, 828–847.

Dahlhamer, J., Lucas, J., Zelaya, C., Nahin, R., Mackey, S., DeBar, L., … Helmick, C. (2018). Prevalence of chronic pain and high-impact chronic pain among adults - United States, 2016. *Morbidity and Mortality Weekly Report, 67*, 1001–1006.

Del Pozo Cruz, B., Alfonso-Rosa, R. M., McGregor, D., Chastin, S. F., Palarea-Albaladejo, J., & Del Pozo Cruz, J. (2020). Sedentary behaviour is associated with depression symptoms: Compositional data analysis from a representative sample of 3233 US adults and older adults assessed with accelerometers. *Journal of Affective Disorders, 265*, 59–62.

Dong, H. J., Gerdle, B., Bernfort, L., Levin, L. A., & Dragioti, E. (2020). Pain catastrophizing in older adults with chronic pain: The mediator effect of mood using a path analysis approach. *Journal of Clinical Medicine, 9*, 2073.

Hadjistavropoulos, T., Herr, K., Turk, D. C., Fine, P. G., Dworkin, R. H., Helme, R., … Williams, J. (2007). An interdisciplinary expert consensus statement on assessment of pain in older persons. *The Clinical Journal of Pain, 23*, S1–43.

Hapidou, E. G., O'Brien, M. A., Pierrynowski, M. R., de Las Heras, E., Patel, M., & Patla, T. (2012). Fear and avoidance of movement in people with chronic pain: Psychometric properties of the 11-item tampa scale for kinesiophobia (TSK-11). *Physiotherapy Canada, 64*, 235–241.

Herr, K. (2011). Pain assessment strategies in older patients. *The Journal of Pain, 12*, S3–S13.

Herr, K., Bjoro, K., & Decker, S. (2006). Tools for assessment of pain in nonverbal older adults with dementia: A state-of-the-science review. *Journal of Pain and Symptom Management, 31*, 170–192.

Herr, K., Bursch, H., Ersek, M., Miller, L. L., & Swafford, K. (2010). Use of pain-behavioral assessment tools in the nursing home: Expert consensus recommendations for practice. *Journal of Gerontological Nursing, 36*(18–29), 30–31.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*, 1735–1780.

Horgas, A. L., Nichols, A. L., Schapson, C. A., & Vietes, K. (2007). Assessing pain in persons with dementia: Relationships among the non-communicative patient's pain assessment instrument, self-report, and behavioral observations. *Pain Management Nursing, 8*, 77–85.

Jiang, M., Mieronkoski, R., Rahmani, A. M., Hagelberg, N., S, S., & L, P. (2017). Ultra-short-term analysis of heart rate variability for real-time acute pain monitoring with wearable electronics. In *IEEE International Conference on Bioinformatics and Biomedicine* (pp. 1025–1032).

Johnson, A., Yang, F., Gollarahalli, S., Banerjee, T., Abrams, D., Jonassaint, J., … Shah, N. (2019). Use of Mobile health apps and wearable technology to assess changes and predict pain during treatment of acute pain in sickle cell disease: Feasibility study. *JMIR Mhealth Uhealth, 7*, Article e13671.

Keefe, F. J., Affleck, G., France, C. R., Emery, C. F., Waters, S., Caldwell, D. S., … Wilson, K. (2004). Gender differences in pain, coping, and mood in individuals having osteoarthritic knee pain: A within-day analysis. *Pain, 110*, 571–577.

Lane, P., Kuntupis, M., MaCdonald, S., McCarthy, P., Panke, J. A., Warden, V., & Volicer, L. (2003). A pain assessment tool for people with advanced Alzheimer's and other progressive dementias. *Home Healthcare Nurse, 21*, 32–37.

Lethem, J., Slade, P. D., Troup, J. D., & Bentley, G. (1983). Outline of a fear-avoidance model of exaggerated pain perception–I. *Behaviour Research and Therapy, 21*, 401–408.

Martini, C. H., Boon, M., Broens, S. J., Hekkelman, E. F., Oudhoff, L. A., Buddeke, A. W., & Dahan, A. (2015). Ability of the nociception level, a multiparameter composite of autonomic signals, to detect noxious stimuli during propofol-remifentanil anesthesia. *Anesthesiology, 123*, 524–534.

Mathias, J. L., Cant, M. L., & Burke, A. L. J. (2018). Sleep disturbances and sleep disorders in adults living with chronic pain: A meta-analysis. *Sleep Medicine, 52*, 198–210.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2006). *Applied multivariate research: Design and interpretation thousand oaks*. CA: Sage Publications.

Olson, D. L., & Delen, D. (2008). *Advanced data mining techniques*. Berlin Heidelberg: New York: Springer-Verlag.

Patel, K. V., Guralnik, J. M., Dansie, E. J., & Turk, D. C. (2013). Prevalence and impact of pain among older adults in the United States: Findings from the 2011 National Health and aging trends study. *Pain, 154*, 2649–2657.

Prkachin, K. M., Schultz, I. Z., & Hughes, E. (2007). Pain behavior and the development of pain-related disability: The importance of guarding. *The Clinical Journal of Pain, 23*, 270–277.

Roberts, M. B., & Drummond, P. D. (2016). Sleep problems are associated with chronic pain over and above mutual associations with depression and catastrophizing. *The Clinical Journal of Pain, 32*, 792–799.

Rosenthal, M., Moore, P., Groves, E., Iwan, T., Schlosser, L. G., Dziewanowska, Z., & Negro-Vilar, A. (2007). Sleep improves when patients with chronic OA pain are managed with morning dosing of once a day extended-release morphine sulfate (AVINZA): Findings from a pilot study. *Journal of Opioid Management, 3*, 145–154.

Sasaki, Y. (2007). *The truth of the F-measure* (https://people.cs.pitt.edu/~litman/courses/cs1671s20/F-measure-YS-26Oct07.pdf).

Sevakula, R. K., Au-Yeung, W. M., Singh, J. P., Heist, E. K., Isselbacher, E. M., & Armoundas, A. A. (2020). State-of-the-art machine learning techniques aiming to improve patient outcomes pertaining to the cardiovascular system. *Journal of the American Heart Association, 9*, Article e013924.

Tavakolian, M., & Hadid, A. (2018). Deep binary representation of facial expressions: A novel framework for automatic pain intensity recognition. In *The international conference on image processing* (pp. 1952–1956).

Thapa, D. K., Visentin, D. C., Kornhaber, R., & Cleary, M. (2020). Prevalence and factors associated with depression, anxiety, and stress symptoms among older adults: A cross-sectional population-based study. *Nursing & Health Sciences, 22*(4), 1139–1152.

Thompson, D. P., Oldham, J. A., Urmston, M., & Woby, S. R. (2010). Cognitive determinants of pain and disability in patients with chronic whiplash-associated disorder: A cross-sectional observational study. *Physiotherapy, 96*, 151–159.

Vlaeyen, J. W., Kole-Snijders, A. M., Boeren, R. G., & van Eek, H. (1995). Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain, 62*, 363–372.

Weingarten, J. A., Dubrovsky, B., Basner, R. C., Redline, S., George, L., & Lederer, D. J. (2016). Polysomnographic measurement of sleep duration and bodily pain perception in the sleep heart health study. *Sleep, 39*, 1583–1589.

Whibley, D., Braley, T. J., Kratz, A. L., & Murphy, S. L. (2019). Transient effects of sleep on next-day pain and fatigue in older adults with symptomatic osteoarthritis. *The Journal of Pain, 20*, 1373–1382.

Yang, F., Banerjee, T., Panaggio, M. J., Abrams, D. M., & Shah, N. R. (2019). Continuous pain assessment using ensemble feature selection from wearable sensor data. In *Proceedings IEEE international conference on bioinformatics and biomedicine* (pp. 569–576).

Yang, G., Jiang, M., Ouyang, W., Ji, G., Xie, H., Rahmani, A. M., … Tenhunen, H. (2018). IoT-based remote pain monitoring system: From device to cloud platform. *IEEE Journal of Biomedical and Health Informatics, 22*, 1711–1719.