

Original Paper

A Gamified Pain Management Intervention for Adults With Chronic Pain in Mainland China: Single-Arm Pre-Post Pilot Study With Machine Learning Predictive Modeling

Mun Yee Mimi Tse¹, Prof Dr; Jiafan He^{1,2}, MSc; Tyrone Tai On Kwok¹, PhD

¹School of Nursing and Health Sciences, Hong Kong Metropolitan University, Kowloon, China (Hong Kong)

²The Jockey Club Institute of Healthcare (IOH), Hong Kong Metropolitan University, Kowloon, China (Hong Kong)

Corresponding Author:

Jiafan He, MSc
School of Nursing and Health Sciences
Hong Kong Metropolitan University
1 Sheung Shing Street, Homantin
Kowloon 999007
China (Hong Kong)
Phone: 852 39708764
Email: s1350029@live.hkmu.edu.hk

Abstract

Background: The widespread prevalence of chronic pain (CP) significantly impacts daily functioning worldwide. In mainland China, maintaining engagement in biopsychosocial interventions remains challenging. Gamification, designed based on self-determination theory, can enhance motivation, while machine learning (ML) algorithms can assist clinicians in dynamically optimizing pain management.

Objective: This study aimed to (1) evaluate the preliminary effectiveness of a gamified pain management (GPM) program on CP and psychological outcomes and (2) identify key factors of significant pain improvements through the application of ML to guide intervention adjustments.

Methods: A single-arm, pre-post study was conducted with 16 participants with CP in mainland China, recruited via social media using convenience sampling. Participants engaged in a 10-week web-based GPM intervention consisting of education, physical activities, and gamified elements, including points, avatars, and feedback. Primary outcomes were pain intensity and interference measured by the Brief Pain Inventory. Secondary outcomes included anxiety, depression, and quality of life. Analysis included paired *t* tests, and ML models were trained to predict clinically meaningful pain reductions. Shapley additive explanations, least absolute shrinkage and selection operator regression, association rule mining, and Kaplan-Meier survival analysis were used to identify key predictors and optimal sessions and intervention durations across subgroups.

Results: A total of 16 participants were engaged, with a mean age of 27.63 (SD 9.584) years. Results from paired *t* tests reported significant improvements in pain intensity (decreased by 27.3%, 95% CI 1.061 to 3.064; $P=0.001$), pain interference (decreased by 27.3%, 95% CI 8.159-17.216; $P<0.001$), and psychological distress, including anxiety ($t_{15}=3.538$, 95% CI 0.969 to 3.906; $P=0.003$) and depression ($t_{15}=4.559$, 95% CI 2.230 to 6.145; $P<0.001$). The gradient boosting model demonstrated the highest predictive accuracy (area under the curve=0.89 and accuracy=0.82). Least absolute shrinkage and selection operator regression identified session 3 ($\beta=-0.45$, 95% CI -0.68 to -0.22 ; $P<0.001$) and session 5 ($\beta=-0.32$, 95% CI -0.59 to -0.05 ; $P=0.02$) as most predictive of clinical success, while association rule mining revealed effective session combinations for different patient subgroups. Time-to-event analyses indicated that individuals with low back pain and higher baseline severity required longer intervention durations for improvement (5 wk; $P=0.03$).

Conclusions: This pilot study presents an innovative method that combines ML with dynamic engagement data from a GPM program during interventions, rather than relying on static baseline data in prior studies. The results show preliminary efficacy and identify specific optimal session combinations and personalized treatment durations for different pain subgroups. These exploratory findings contribute to the field by providing a data-driven method for adaptive, personalized digital health interventions that move beyond one-size-fits-all strategies, potentially enabling clinicians to modify content and dosage to improve engagement and outcomes if validated in larger sample trials.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400094247; <https://www.chictr.org.cn/showprojEN.html?proj=245138>

JMIR Form Res 2026;10:e78823; doi: [10.2196/78823](https://doi.org/10.2196/78823)

Keywords: gamification; artificial intelligence; machine learning; pain management; chronic pain

Introduction

Problem

The Global Burden of Disease Study reported a high prevalence of chronic pain (CP) globally [1]. CP, defined as pain that persists for more than 3 months, significantly disrupts various aspects of daily life, including work, physical activity, and social engagement. It is frequently comorbid with anxiety and depression, sleep disturbances, and decreased productivity [2]. In mainland China, the prevalence of CP ranges from 50% to 70% and there is a preference for self-management [3,4], highlighting the urgent need for effective pain management. Despite the growing adoption of biopsychosocial approaches in clinical guidelines [5], individuals with CP face strong barriers to adhere to regular physical activity, including fear of pain and lack of personalized plans [6,7].

Literature Review

To tackle these challenges, gamification has emerged as an innovative strategy to enhance user engagement, motivation, and adherence to health interventions [8]. Drawing on theories such as self-determination theory (SDT), gamification incorporates game design elements into nongame contexts to motivate behavior change [9,10]. While various gamified interventions have been developed to reduce pain symptoms, the treatment effects have been inconsistent. Some studies have shown immediate efficacy for both adults and older adults [11,12], with evidence suggesting that individuals across age groups find these interventions enjoyable and accessible [13]. However, other randomized controlled trials reported nonsignificant effects on pain intensity [14,15]. This heterogeneity may stem from the complex and multidimensional nature of CP, as well as insufficient theoretical grounding and inability to tailor to specific populations [16], a challenge that causes treatments to struggle to maintain adherence over time. Therefore, this study developed a gamified pain management (GPM) program for adults in mainland China, grounded in SDT, to foster intrinsic motivation through multiple game elements such as avatar, choices, feedback, and challenges.

Building on the motivational benefits of gamification, integrating artificial intelligence (AI) can further enhance personalization and long-term adherence through data-driven adaptation by addressing barriers, including accessibility and sustained engagement. Numerous studies have demonstrated that AI can assist in treatment decision-making, dynamic prediction, and behavior guidance. For instance, rule-based recommendation systems can generate personalized exercise plans based on users' pain characteristics, improving intervention adaptability [17]. AI chatbots

provide psychological support and behavioral reminders through real-time interaction, enhancing user engagement and treatment adherence [18]. Machine learning (ML) models including random forests and extreme gradient boosting have been used to predict postoperative pain outcomes and determine opioid use, thereby assisting in drug management [19,20]. These advantages mark a shift from traditional, rule-based systems to dynamic, data-driven prediction models.

Unlike existing studies that primarily rely on static baseline data, our study uses dynamic real-time behavioral data recorded during the intervention engagement into ML models. Moreover, most studies focus on the binary questions of whether intervention reduces pain, but questions such as “how quickly improvement occurs” and “through which mechanisms improvement occurs” remain unexplored, thus lacking a deeper understanding of the time course and underlying processes of intervention effectiveness.

Aims and Objectives

In response to these gaps, the present pilot study aimed to (1) assess the effects of the GPM intervention on CP in adults in mainland China and (2) identify key factors that predict significant improvements in pain intensity using ML models.

Methods

Study Design

This single-arm, nonrandomized, pre-post pilot study aimed to evaluate the preliminary effects of the digital GPM intervention and to develop exploratory ML models predicting clinical pain reduction outcomes. Participants were recruited using convenience sampling through social media platforms and engaged with the intervention via a web-based platform. The study adhered to the TRIPOD-AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis-AI) reporting guidelines ([Checklist 1](#)).

Ethical Considerations

The work described was carried out in accordance with the guidelines of the Hong Kong Metropolitan University Research Ethics Committee (reference HE-SF2024/34) and was registered on the Chinese Clinical Trial Registry (registration number ChiCTR2400094247). All participants were provided with a detailed digital informed consent form including the study purpose and procedures, potential risks and benefits (no monetary compensation), confidentiality and voluntariness, and contact information for the principal investigator. To ensure privacy and confidentiality, all data were anonymized by assigning a unique reference number,

which was also used to log in to the game and surveys. All data are securely stored in an encrypted folder, and only the research team has access to the information. No images containing identifiable participant identities were obtained. No monetary compensation was provided for participation in this study.

Participants

No formal prior power calculation was performed due to the exploratory, feasibility-testing nature of this pilot study, and a sample size larger than 10 was deemed appropriate for the formative objectives [21]. Interested individuals were recruited through an online flyer via social media platforms including Redbook and WeChat. The flyer featured a QR code to access an online questionnaire for initial eligibility screening. To be eligible to participate in the study, the participants must meet all inclusion criteria: (1) aged 18 years or older; (2) able to read and understand the Mandarin Chinese language; (3) owners of smartphones with internet access; (4) have a previously diagnosed chronic noncancer pain condition lasting more than 3 months; (5) report a score of greater than or equal to 2 on the Visual Analog Scale [22]. Exclusion criteria were as follows: (1) patients with associated pathologies that make it impossible to perform physical exercise (myopathies, neurological diseases, cardiac disease, pregnancy, pulmonary diseases, infection, and fracture) [23]; (2) individuals undergoing surgery or invasive treatments in the last 3 months [24]; (3) concurrent participation (or participation within the preceding 3 mo) in a supervised exercise program or multidisciplinary treatment.

Intervention

The GPM was open for 10 weeks, allowing participants to complete the web-based intervention at their own pace. It

integrated gamification, education, and progressive exercise within a gamified framework informed by SDT. Prior to the first session, participants received web-based instructions outlining the intervention structure. The program was structured as an escape room game, where participants, via hero avatars, progressed through themed sessions by unlocking educational content, exercise tasks, quizzes, and relaxation activities.

The educational modules covered the following information via animated videos and notes: (1) the definition and mechanisms of pain, (2) distinctions between acute and CP, (3) the biopsychosocial impact of CP, and (4) pharmacological and (5) nonpharmacological management strategies [25]. Each progressive exercise module (5-7 min each) included strengthening and stretching [5] routines tailored to participants' primary pain locations (eg, neck, shoulder, back, knee, or head) (Table 1). Gamification elements included a points system (awarded for quiz accuracy and exercise completion), a hero avatar, choices and challenges, and positive or corrective feedback mechanisms. To progress, a minimum of 60 points per session was required. Hidden sessions embedded within the game provided opportunities for additional exercises and engagement.

Engagement data, including treatment duration (TD) (first and last login timestamps), quiz accuracy, and session-specific satisfaction scores, were automatically recorded by the GPM platform. Biweekly WeChat reminders were manually sent to promote adherence.

Table 1. Descriptive statistics of engagement and feedback metrics (N=16).

Variable	Mean (SD)	Range
Engagement variables		
Treatment duration (wk)	4.38 (2.19)	1-10
Number of quizzes correctly answered ^a	6.94 (1.53)	3-8
GUESS ^b	96.25 (12.69)	68-115
Cumulative session feedback scores ^c		
Understandability	33.94 (5.26)	25-40
Comprehensibility	33.94 (5.26)	25-40
Applicability	32.44 (7.56)	15-40

^aHigher quiz accuracy and satisfaction scores indicate greater understanding and engagement.

^bGUESS: Game User Experience Satisfaction Scale.

^cScores for understandability, comprehensibility, and applicability represent the cumulative sum of a participant's ratings across all 8 intervention sessions. Each item was rated from 0 to 5 per session, resulting in a possible cumulative range of 0 to 40 for each item.

Data Collection

Participants were recruited via social media posters containing a hyperlink to an online eligibility screening questionnaire [26], where detailed study information and consent form were provided. After providing informed consent, screening questions, including pain condition, medical history, and demographic criteria were embedded within the survey. The

survey system automatically submitted an incomplete survey if any exclusionary options were selected. Individuals who met the eligibility criteria were invited to voluntarily provide their WeChat ID to receive automatically generated access credentials for the intervention platform.

Baseline assessments (T0) were administered via the survey platform [26] and distributed to eligible participants through WeChat. Upon completion, participants were

provided with access credentials to the GPM intervention site [27]. Postintervention assessments (T1) were collected through the same platform immediately after the intervention was completed approximately 10 weeks after baseline.

Measurements

A list of outcome measures, including instruments and score ranges, is provided in Table 2.

Table 2. Summary of demographic, clinical, and intrainervention feature domains used as input variables for machine learning modeling.

Domain	Description
Demographic	<ul style="list-style-type: none"> Sex (male or female), age
Clinical	
Pain	<ul style="list-style-type: none"> Pain duration was assessed by demonstrating specific total months of pain experiences (divided into categories: 3 mo, 4 mo, 5 mo, 6-9 mo, 9-12 mo, and >12 mo) Pain site was provided with different body sites including head, neck, shoulder, low back, and knee
Functionality	<ul style="list-style-type: none"> Assessed by Brief Pain Inventory considering the total score (score range 0-70) and subscale scores (score range 0-10)
Mental health	<ul style="list-style-type: none"> Assessed by the 7-item General Anxiety Disorder Scale (score range 0-21) and the 9-item Patient Health Questionnaire (score range 0-27)
Fear-avoidance and catastrophizing	<ul style="list-style-type: none"> Evaluated by Fear Avoidance Beliefs Questionnaire (score range 0-96) and the total score and subscale scores (score range 0-42) Pain Catastrophizing Scale considering (score range 0-52) and the total score and subscale scores (score range 0-24)
Medication consumption	<ul style="list-style-type: none"> Considers opioid consumption (yes or no)
Quality of life	<ul style="list-style-type: none"> Evaluated by the 3-level version of EQ-5D^a considering the total score (score range 0-45) and subscale scores (score range 0-15)
Exercise adherence	<ul style="list-style-type: none"> Godin Leisure Time Exercise Questionnaire was used (score range 0-99)
Motivation	<ul style="list-style-type: none"> Game User Experience Satisfaction Scale considering the total score (score range 0-24) and subscale scores (score range 0-24)
Intrainervention variables	
Quiz accuracy	<ul style="list-style-type: none"> Number of correct answers
Adverse event	<ul style="list-style-type: none"> Considers adverse events (yes or no)
Intervention duration	<ul style="list-style-type: none"> The time span from the first to the last day of participation
Session satisfaction score	<ul style="list-style-type: none"> Summary score of understandability, comprehensibility, and applicability in each session (score range 0-15).

^aEQ-5D: EuroQol 5-dimension.

Primary Outcome

Pain intensity and pain interference were assessed using the Brief Pain Inventory, rated on a scale of 0 to 10. Pain severity was calculated based on the past 24 hours and the “current pain.” The pain interference includes the interference degree of pain on general activity, mood, walking, work, social relationships, sleep, and enjoyment of life. The Chinese scale has good reliability with a Cronbach α of 0.89 [28].

Secondary Outcomes

Secondary outcomes were as follows:

- Fear Avoidance Beliefs Questionnaire and Pain Catastrophizing Scale: Used to measure pain perception including attitude toward physical activity, work, rumination, magnification, and helplessness related to pain. The Cronbach coefficients of the Chinese version of the Fear Avoidance Beliefs Questionnaire are high, ranging from 0.75 to 0.85 [29], and Pain Catastrophizing Scale is 0.91 [30].
- Godin Leisure Time Exercise Questionnaire: Used to measure exercise adherence, including times of strenuous exercise, moderate exercise, and mild exercise per week via the website. The percentage

of exercises completed out of the total number of prescribed exercises is used to measure total exercise adherence. The Cronbach α is 0.64 and the κ index is 0.94 [31].

- Patient Health Questionnaire-9 items and Generalized Anxiety Disorder-7 items: Used to measure depression and anxiety. The Chinese version of the scale has excellent internal reliability; the Cronbach α of the Generalized Anxiety Disorder-7 items is 0.91 [32] and that of the Patient Health Questionnaire-9 items is 0.86 [33].
- The EuroQol 5-D 3-level: Used to assess an individual’s perceptions of quality of life including mobility, self-care, and usual activities. The Chinese version has been treated as the standard in China [34].
- Game User Experience Satisfaction Scale: Used to evaluate overall game satisfaction, with good internal reliability (Cronbach α =0.93) [35].
- Intrainervention variables (collected during or immediately after intervention [T1]):
 - Engagement metrics were assessed through session satisfaction scores (SAT) and quiz accuracy, recorded via the GPM platform.

- Satisfaction scores were derived from the sum of 3 items rated on a 5-point Likert scale following the completion of each session. These items evaluated the following: (1) Understandability: “Are the materials in this session easy for you to understand?” (2) Comprehensibility: “How much do you understand the materials in this session?” (3) Applicability: “To what extent do you think you would apply this session to future self-management?”
- Quiz accuracy was calculated as the proportion of correctly answered questions in each session’s embedded quiz.
- TD was automatically calculated as the total number of weeks between a participant’s first and last login on the GPM platform.

Statistical Analysis

To test aim 1, descriptive statistics for engagement metrics, including TD, quiz scores, and SAT were calculated. Paired *t* tests were conducted to compare differences between the two stages: baseline (T0) and postintervention (T1) across all outcome measures. The normality of difference scores was examined using the Shapiro-Wilk test. In case of violation of the normality assumption, the nonparametric Wilcoxon signed-rank test was applied. In addition, paired *t* tests were conducted within pain subgroups to examine the within-group intervention effects. The pain intensity level was categorized into two groups: mild-to-moderate (scores ranging from 3 to 6) and severe pain group (scores 7 or higher in Brief Pain Inventory) [36]. Given the small sample size and exploratory, hypothesis-generating nature of this study, no

multiple comparisons were applied. This decision was made to minimize the risk of type II errors in this initial investigation. A significant level of $P < .05$ was applied. The analyses were performed using SPSS Statistics Version 23.0 (IBM Inc) and Python 3.11.

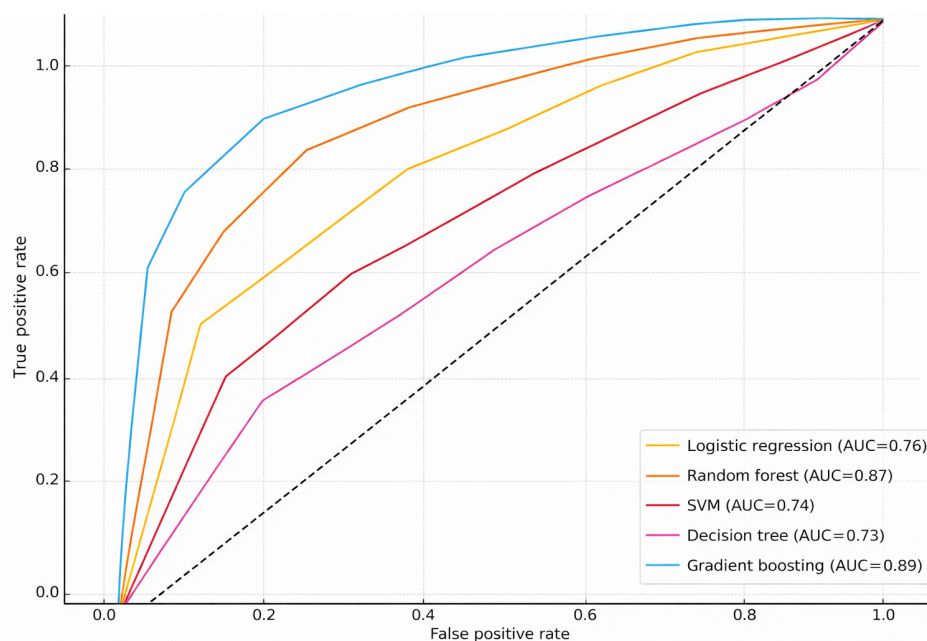
ML Analysis

ML modeling was exploratory and served as proof-of-concept analysis. The dataset comprised 53 features, including baseline demographic, clinical, and intrainervention engagement metrics (Table 2). The binary outcome variable, pain intensity, indicates a clinically significant improvement, defined as a $\geq 30\%$ reduction in pain intensity measured by the BPI [36]. Continuous predictors were *z*-score standardized, and categorical variables were one-hot encoded.

Predictive Modeling for Pain Reduction

Five supervised classification models were trained to predict clinical pain reduction. Due to the limited sample size, model training and hyperparameter tuning used a 5-fold stratified cross-validation procedure on the training data (80% training and 20% testing per iteration) to preserve the distribution of clinical pain reduction across folds. Accuracy, recall, specificity, precision, F_1 -score, and the area under the curve (AUC) were used to evaluate performance (Figure 1). To reduce overfitting, model training and evaluation relied on a 5-fold stratified cross-validation procedure, and all models were trained on standardized and encoded features only. Shapley additive explanations (SHAP) was used to identify the most influential predictors of intervention effectiveness from the best performing models (gradient boosting and random forest).

Figure 1. Receiver operating characteristic curve across models for predicting clinically significant pain reduction. AUC: area under the curve; SVM: support vector machine.



Session-Level Mechanisms of Clinical Improvement Analysis

A LASSO (least absolute shrinkage and selection operator) regression was performed to further examine each SAT's contribution to clinical pain reduction. This focused strategy sought to elucidate the role that session-specific engagement played in clinical improvement and to provide guidance for tailoring the order of interventions. A leave-one-out cross-validation approach was used to optimize training data use and guarantee reliable variable selection in the small-sample setting. In order to reduce overfitting, especially in high-dimensional, small-sample situations, this regularized regression technique enables variable selection and coefficient shrinkage. Sessions with the greatest predictive influence were identified by shrinking noninformative variables toward 0.

Association rule mining was then used to find potential interaction patterns for particular subgroups (SAT ≥ 12 coded as "high"), while LASSO regression identified the contribution of individual sessions. Rules linked to clinical improvement that satisfied high-confidence (≥ 0.75) and high-lift patterns (≥ 1.15) thresholds were extracted using the Apriori algorithm. Our threshold of greater than or equal to 1.15 was selected to find patterns where the session combination and clinical improvement co-occurred at least 15% more frequently than would be predicted by chance. Crucially, a lift value greater than or equal to 1 indicates a positive association. The goal of this exploratory analysis was to produce ideas for tailoring the sequencing of digital interventions.

Time-to-Improvement Analysis by Subgroup

Leave-one-out cross-validation was used with the gradient boosting model to examine the ideal intervention duration needed for clinical improvement across pain subgroups.

SHAP-informed patterns in model decision boundaries were used to estimate intervention day thresholds. The robustness and CIs of these thresholds were estimated using bootstrap sampling, which supports reliability under small sample constraints.

The duration of the intervention at which the postintervention assessment satisfied the criteria for clinically significant improvement ($\geq 30\%$ reduction in pain intensity score) was operationalized as the time to improvement. As a result, Kaplan-Meier survival curves were used for time-to-event analysis, with the time to clinical improvement serving as the survival time variable. To investigate variations in median improvement durations by pain site and severity, subgroup comparisons were carried out. Statistical significance was evaluated using log-rank tests. All results should be viewed as preliminary and hypothesis-generating rather than confirmatory, given the small sample size.

Results

Participant Demographics

Data collection procedures were completed online between March 2025 and August 2025. A total of 23 individuals accessed the online screening survey. Of these, 5 were excluded for not meeting eligibility criteria (eg, pain duration < 3 months, pain intensity < 2). Two participants withdrew prior to the intervention due to time conflicts, so no missing data was observed. And 16 participants were enrolled in the study, with a mean age of 27.63 (SD 9.58) years. The majority of participants were employed (15/16, 93.8%), and all had a higher education background. Pain duration varied, with 25% (4/16) of participants experiencing CP for over 1 year, while others reported pain durations between 4 and 12 months. At baseline, participants reported moderate pain intensity levels (mean 5.81, SD 1.38) and pain interference (mean 38.31, SD 12.41) (Table 3).

Table 3. Baseline demographics and clinical characteristics (N=16).

Variables	Values
Gender, n (%)	
Male	11 (68.8)
Female	5 (31.2)
Age, mean (SD)	27.63 (9.584)
Marriage, n (%)	
Single	12 (75)
Married	4 (25)
Occupation, n (%)	
Employed	15 (93.8)
Unemployed	1 (6.3)
Education background, n (%)	
Higher education	16 (100)
Duration of symptoms, n (%)	
3 mo	1 (6.3)
4 mo	2 (12.5)

Variables	Values
5 mo	3 (18.8)
6-9 mo	5 (31.3)
9-12 mo	1 (6.3)
>1 yr	4 (25)
Use of pain medication, n (%)	
Yes	3 (18.8)
No	13 (81.3)
Baseline scores, mean (SD)	
BPI ^a intensity	5.81 (1.377)
BPI interference	38.31 (12.408)
FABQ ^b	58.69 (13.189)
PCS ^c	35.81 (8.796)
PHQ-9 ^d	11.88 (4.674)
GAD-7 ^e	9.75 (4.946)
GLTEQ ^f	97.81 (84.680)
EQ-5D-3L ^g	3.06 (2.932)

^aBPI: Brief Pain Inventory.

^bFABQ: Fear-Avoidance Beliefs Questionnaire.

^cPCS: Pain Catastrophizing Scale.

^dPHQ-9: Patient Health Questionnaire-9 items.

^eGAD-7: Generalized Anxiety Disorder-7 items.

^fGLTEQ: Godin Leisure Time Exercise Questionnaire.

^gEQ-5D-3L: EuroQol 5-Dimension 3-Level.

Intervention Effectiveness and Outcome Variables

Table 4 presents the means and SDs of the key outcome measures at baseline and post intervention. Significant reductions were observed in pain intensity ($t_{15}=4.392$, 95% CI 1.061-3.064; $P<.001$) and pain interference ($t_{15}=5.971$, 95% CI 8.159-17.216; $P<.001$). Secondary outcomes, including pain catastrophizing and psychological distress, also demonstrated significant improvements post intervention. Pain-related catastrophizing ($t_{15}=3.190$, 95% CI 2.821-14.179; $P=.006$), depression ($t_{15}=4.559$, 95% CI 2.230-6.145; $P<.001$) and anxiety ($t_{15}=3.538$, 95% CI 0.969-3.906; $P=.003$) demonstrated significant improvement. Additionally, quality of life improved

significantly ($t_{15}=2.611$, 95% CI 0.172-1.703; $P=.02$). Fear-avoidance beliefs and exercise adherence showed no significance ($P>.05$, see Table 4 and Table S1 in Multimedia Appendix 1 for details).

Participants' engagement with the GPM varied. With a range of 1 to 10 weeks, the average TD was 4.38 (SD 2.19) weeks, demonstrating high flexibility and customized completion rates. The average quiz accuracy was roughly 75%, indicating a generally high level of understanding of the course material. Table 1 displays the cumulative scores for the session feedback items, which were consistently high (mean 33.44, SD 6.03, range 0-40) throughout all sessions, indicating favorable opinions of the sessions' comprehensibility and applicability (Tables 1 and 5).

Table 4. Preintervention and postintervention comparison of key outcomes (N=16).

Outcomes	Baseline, mean (SD)	Post intervention, mean (SD)	Statistics		
			<i>t</i> test (<i>df</i>)	95% CI	<i>P</i> value ^a
BPI ^b intensity	5.81 (1.377)	3.75 (1.342)	4.392 (15)	1.061-3.064	.001
BPI interference	38.31 (12.408)	25.63 (12.387)	5.971 (15)	8.159-17.216	<.001
PCS ^c	35.81 (8.796)	27.31 (13.200)	3.190 (15)	2.821-14.179	.006
FABQ ^d	58.69 (13.189)	53.81 (16.586)	1.440 (15)	-2.342-12.092	.17
GLTEQ ^e	97.81 (84.680)	101.00 (93.047)	-0.237 (15)	-31.890-25.515	.82
PHQ-9 ^f	11.88 (4.674)	7.69 (3.860)	4.559 (15)	2.230-6.145	<.001
GAD-7 ^g	9.75 (4.946)	7.31 (4.316)	3.538 (15)	0.969-3.906	.003
EQ-5D-3L ^h	3.06 (2.932)	2.13 (2.579)	2.611 (15)	0.172-1.703	.02

^aPaired *t* tests were used to compare preintervention and postintervention scores.

^bBPI: Brief Pain Inventory.

^cPCS: Pain Catastrophizing Scale.

^dFABQ: Fear-Avoidance Beliefs Questionnaire.

^eGLTEQ: Godin Leisure Time Exercise Questionnaire.

^fPHQ-9: Patient Health Questionnaire-9 items.

^gGAD-7: Generalized Anxiety Disorder-7 items.

^hEQ-5D-3L: EuroQol 5-Dimension 3-Level.

Table 5. Session-specific satisfaction scores of each session (N=16).

Session	Description		Understandability	Comprehensibility	Applicability
	Education	PA ^a	Mean (SD)	Mean (SD)	Mean (SD)
Session 1	Definition of pain	Warm and balancing	4.06 (0.772)	4.01 (0.771)	4.00 (0.966)
Session 2	Biopsychosocial aspects of CP ^b	Stretching ^c	4.19 (0.750)	4.19 (0.750)	4.06 (0.998)
Session 3	Mechanism of CP	Core strength training	4.38 (0.719)	4.37 (0.719)	4.19 (0.981)
Session 4	Mechanism of CP	Stretching ^c	4.31 (0.793)	4.32 (0.763)	4.13 (1.204)
Session 5	Pharmacological intervention	Core strength training	4.25 (0.856)	4.25 (0.856)	3.88 (1.360)
Session 6	Nonpharmacological management	Core strength training	4.38 (0.719)	4.31 (0.732)	4.06 (1.237)
Session 7	PA	Stretching ^c	4.19 (0.981)	4.19 (0.985)	4.13 (0.957)
Session 8	Goal setting strategies	Core strength training	4.19 (1.047)	4.22 (1.041)	4.00 (1.211)

^aPA: physical activity.

^bCP: chronic pain.

^cStretching for head, neck or shoulder, low back, and knee.

Predictive Modeling and Key Determinants of Treatment Response

Gradient boosting and random forest models exhibited superior performance compared to other models. Specifically, gradient boosting achieved an AUC of 0.89, F_1 -score of 0.82, accuracy of 0.83, precision of 0.82, and recall of 0.81, demonstrating predictive accuracy for pain intensity changes. The random forest model also performed well (AUC=0.87, F_1 -score=0.80, accuracy=0.81, precision=0.80, and recall=0.79) (Table 6).

Figures 2 and 3 illustrate SHAP value analysis, identifying the most influential predictors of treatment response.

Pain interference with general activity (BPI_GenA1) emerged as the strongest predictor, suggesting that participants with greater pain interference experienced more significant reductions in pain intensity after the GPM intervention. Pain catastrophizing (rumination) and age emerged as the second most significant predictors, with variations in these measures substantially affecting intervention effectiveness. Fear-avoidance belief demonstrated a significantly negative effect on prediction. Features including depression, TD, engrossment in the game (GUESS_Eng), and pain site are also identified as highly influential factors. These patterns suggest that participants with greater pain impact and consistent engagement derive the most benefit from GPM.

Table 6. Performance comparison of machine learning models for predicting clinically significant pain reduction (N=16).

Model	AUC ^a	Accuracy	Recall	Specificity	Precision	F_1 -score ^b	% positive ^c
Gradient boosting	0.89	0.83	0.81	0.85	0.82	0.82	50%
Random forest	0.87	0.81	0.79	0.83	0.8	0.8	
Logistic regression	0.76	0.72	0.7	0.74	0.71	0.71	
SVM ^d	0.74	0.7	0.68	0.72	0.69	0.69	
Decision tree	0.73	0.69	0.67	0.71	0.68	0.68	

^aAUC: area under the curve.

^b F_1 -score: the harmonic mean of precision and recall.

^c% positive: clinical positive outcome (reduction of 30% compared to the baseline in pain intensity).

^dSVM: support vector machine.

Figure 2. Average absolute Shapley additive explanations values from gradient boosting and random forest models for predicting pain reduction. BPI: Brief Pain Inventory; FABQ: Fear-Avoidance Beliefs Questionnaire; GUESS: Game User Experience Satisfaction Scale; PCS: Pain Catastrophizing Scale; PHQ: Patient Health Questionnaire.

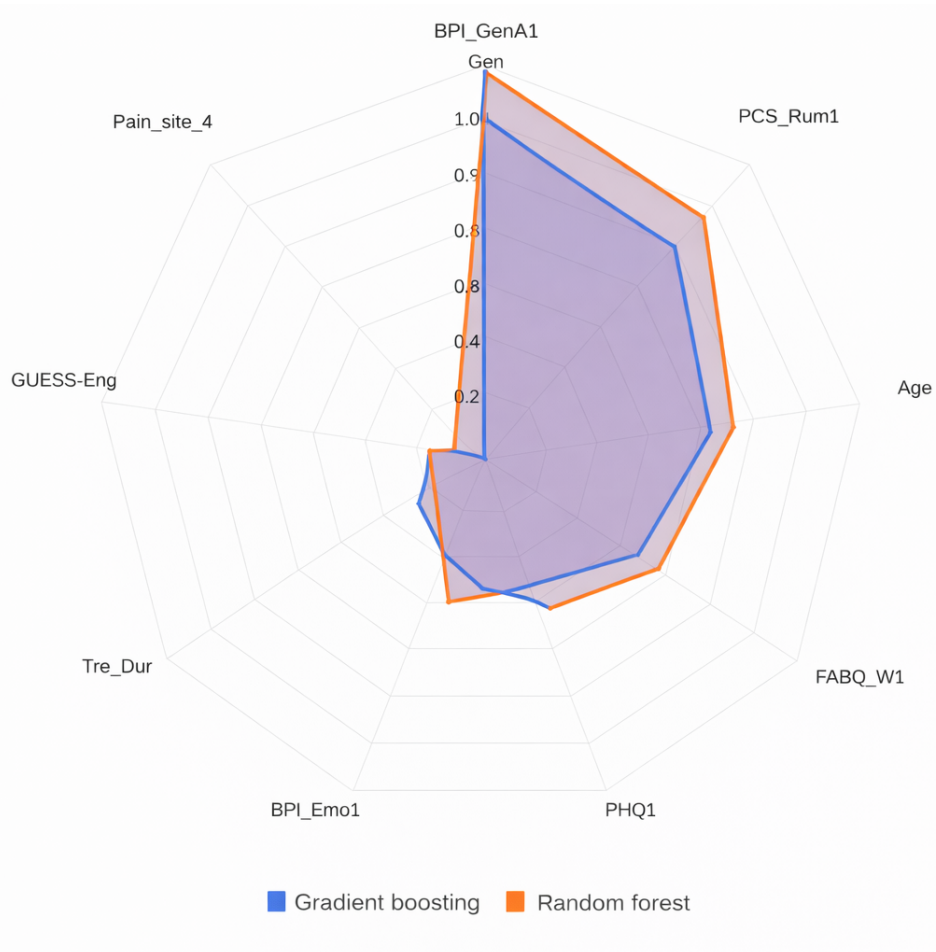
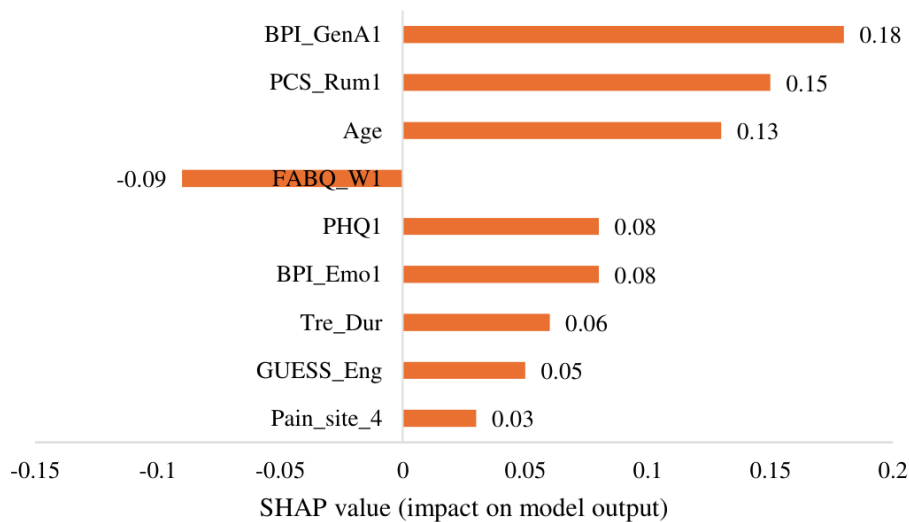


Figure 3. SHAP value for gradient boosting model predicting pain reduction. BPI: Brief Pain Inventory; FABQ: Fear-Avoidance Beliefs Questionnaire; GUESS: Game User Experience Satisfaction Scale; PCS: Pain Catastrophizing Scale; SHAP: Shapley additive explanations.



Session Level Mechanisms of Clinical Improvement

We conducted further analysis to examine session-level engagement patterns. Table 7 shows LASSO regression

results linking SAT to pain improvement. Session 3 and 5 were significant predictors ($P < .05$), indicating that higher satisfaction in these sessions was associated with greater reductions in pain intensity. However, other sessions were

assigned near-zero coefficients, suggesting minimal influence on treatment outcomes (Table 7).

Apriori association rule analysis identified effective combinations of session engagement associated with clinical improvement. The combination of high satisfaction in session 3 and session 5 exhibited the strongest association with clinical improvement (support: 0.38, confidence: 0.83, and lift: 1.65), particularly among participants with severe pain.

The combination of sessions 4 and 6 was associated with improvement in the mild-to-moderate pain (MMP) subgroup (confidence: 0.78). Additionally, the combination of sessions 3, 5, and 7 had the highest confidence (0.85) and lift (1.70), suggesting a particularly potent effect for participants with severe low back pain (LBP). Other session combinations analysis is provided in Table 8 and Figure 4.

Table 7. Least absolute shrinkage and selection operator regression coefficients identifying key session-specific satisfaction scores associated with clinical pain improvement (N=16).

Session	Coefficient	SE	z score	P value ^a	95% CI
SAT 1 ^b	-0.05	0.19	-0.26	.79	-0.42 to 0.32
SAT 2	-0.18	0.16	-1.13	.26	-0.49 to 0.13
SAT 3	-0.45	0.12	-3.75	<.001	-0.68 to -0.22
SAT 4	-0.12	0.17	-0.71	.48	-0.45 to 0.21
SAT 5	-0.32	0.14	-2.29	.02	-0.59 to -0.05
SAT 6	-0.08	0.18	-0.44	.66	-0.43 to 0.27
SAT 7	-0.28	0.15	-1.87	.06	-0.57 to 0.01
SAT 8	-0.03	0.20	-0.15	.88	-0.42 to 0.36

^aStatistical significance was determined using the least absolute shrinkage and selection operator regression. Negative coefficients indicate higher satisfaction associated with greater pain reduction.

^bSAT: session satisfaction scores.

Table 8. Association rule patterns of session satisfaction combinations associated with clinically significant pain improvement by pain subgroup (N=16).

Session combination ^a	Support	Confidence	Lift	Subgroup
High SAT ^b 3 and 5	0.38	0.83	1.65	Severe pain
High SAT 4 and 6	0.29	0.78	1.55	MMP ^c
High SAT 1 and 8	0.25	0.75	1.50	Other pain
High SAT 2 and 4	0.27	0.77	1.53	N/SP ^d
High SAT 2 and 7	0.31	0.80	1.60	LBP ^e
High SAT 3, 5, and 7	0.19	0.85	1.70	Severe and LBP
Comprehensibility and applicability	0.22	0.81	1.68	Severe pain
High SAT 6 and applicability	0.17	0.78	1.52	MMP
High SAT 3 and comprehensibility	0.18	0.82	1.63	LBP
High SAT 5 and understandability	0.20	0.79	1.58	N/SP

^aData was analyzed by using Apriori association rules to identify frequent item sets of session satisfaction associated with clinical improvement. Higher lift (>1.5) indicates strong positive association between session combination and clinical improvement.

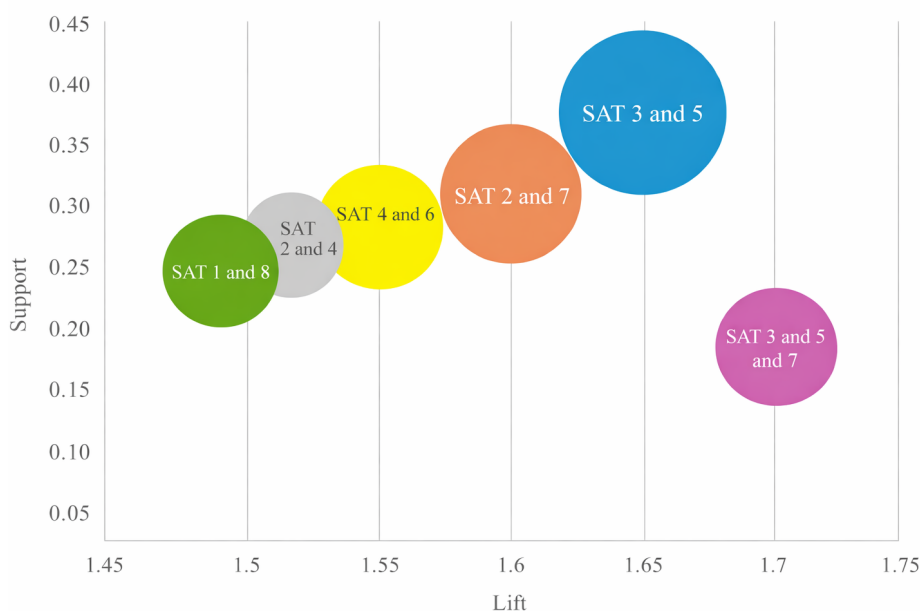
^bSAT: session satisfaction scores.

^cMMP: mild-to-moderate pain.

^dN/SP: neck or shoulder pain.

^eLBP: low back pain.

Figure 4. The bubble diagram of the grouping matrix for the 6 association rules. SAT: session satisfaction scores.



Time-to-Improvement Analysis Across Subgroups

Table 9 presents within-group effects across different pain subgroups. Significant reductions in pain intensity were observed from baseline to post intervention in the MMP ($t_{10}=3.130$, 95% CI 0.367-2.179; $P=.01$), severe pain group

($t_4=4.750$, 95% CI 1.579, 6.021; $P=.009$), and neck or shoulder pain (N/SP) group ($t_6=5.925$, 95% CI 1.075-2.925; $P=.002$), indicating that the intervention had a statistically significant effect across multiple subgroups (Figures 5 and 6). These differences indicate that treatment response may depend on pain site and baseline intensity.

Table 9. Within-group pre-to-post changes in pain intensity across pain subgroups (N=16).

Group	Participants, n (%)	Baseline	Post intervention	Statistics	95% CI	P value ^a
		Mean (SD)	Mean (SD)	t test (df)		
MMP ^b	11 (68.75)	5.09 (0.94)	3.82 (1.17)	3.130 (10)	0.367 to 2.179	.01
Severe pain	5 (31.25)	7.40 (0.55)	3.60 (1.82)	4.750 (4)	1.579 to 6.021	.009
Other pain	4 (25)	5.75 (2.22)	2.75 (0.96)	2.121 (3)	-1.501 to 7.501	.12
N/SP ^c	7 (43.75)	5.86 (1.07)	3.86 (1.07)	5.925 (6)	1.075 to 2.925	.002
LBP ^d	5 (31.25)	5.80 (1.30)	4.40 (1.67)	1.510 (4)	-1.175 to 3.975	.21

^aPaired t tests were used to compare preintervention and postintervention scores.

^bMMP: mild-to-moderate pain.

^cN/SP: neck or shoulder pain.

^dLBP: low back pain.

Figure 5. Comparison of Brief Pain Inventory intensity scores for mild-to-moderate pain and severe pain across time.

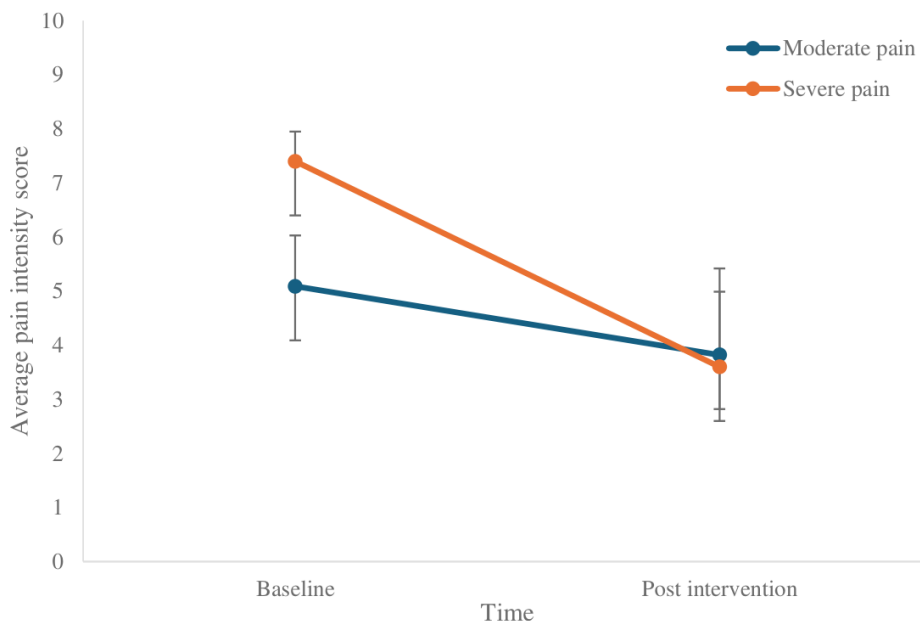
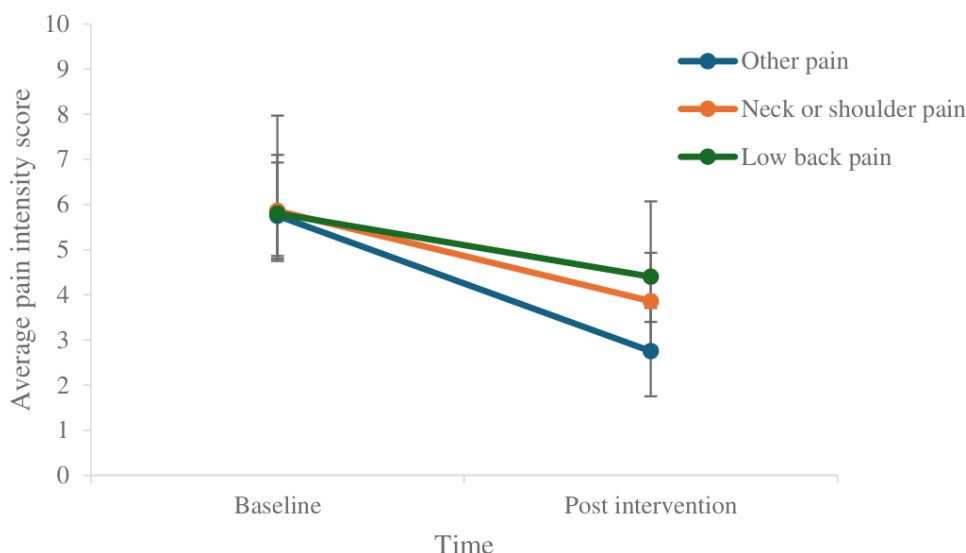


Figure 6. Comparison of Brief Pain Inventory intensity scores for different pain site groups across time.



SHAP values indicated that participants in the severe pain group required longer treatment (≥ 5.0 wk) for significant benefit, with session 3 identified as the top session-level predictor, while TD dominated overall predictions. An increase to 80% probability of clinically significant improvement benefit from a minimum of 5 weeks in the severe group and 3.6 weeks in the MMP group. LBP benefited after 4.6 weeks and N/SP after 4 weeks, suggesting that dose response varied by condition (Table 10).

Kaplan-Meier survival curves were generated for participants stratified by baseline pain intensity and pain site,

corroborating findings with SHAP-based threshold analysis. The median time to improvement was longer for participants with severe pain (5 wk) compared to those with MMP (3.6 wk), and the difference was significant (log-rank $P=.03$) (Table 11, Figure 7). The median time to improvement was 4.6 weeks for LBP, which was longer than N/SP (3.4 wk) and other pain group (4 wk) (Table 10, Figure 8). Bootstrap validation provided support for the stability of these estimates (Table 12). However, these findings require confirmation in larger samples.

Table 10. SHAP^a-informed thresholds for optimal treatment duration by subgroup (N=16).

Group	Optimal TD ^b (wk)	Probability increase (%)	Key predictors (SHAP values)
Severe pain	≥5.0	+80	SAT ^c 3 (0.45), TD (0.30)
MMP ^d	≥3.6	+70	SAT 4 (0.28), SAT 6 (0.22)
LBP ^e	≥4.6	+75	SAT 7 (0.35), TD (0.25)
N/SP ^f	≥4.0	+65	SAT 2 (0.40), SAT 4 (0.20)
Other pain	≥3.4	+60	SAT 1 (0.30), SAT 8 (0.25)

^aSHAP: Shapley additive explanations.

^bTD: treatment duration.

^cSAT: session satisfaction scores.

^dMMP: mild-to-moderate pain.

^eLBP: low back pain.

^fN/SP: neck or shoulder pain.

Table 11. Kaplan-Meier survival estimates of time to clinically significant improvement by subgroup (N=16).

Group	TD ^a (wk), median	3-wk PCI ^b (%)	6-wk PCI (%)	Log-rank <i>P</i>
MMP ^c	3.6	50	90	0.03
Severe pain	5	20	70	0.03
LBP ^d	4.6	25	80	0.04
N/SP ^e	4	40	85	0.05
Other pain	3.4	55	75	0.15

^aTD: treatment duration.

^bPCI: proportion with clinical improvement.

^cMMP: mild-to-moderate pain.

^dLBP: low back pain.

^eN/SP: neck or shoulder pain.

Figure 7. Kaplan-Meier survival curves for time to clinical improvement by pain intensity group.

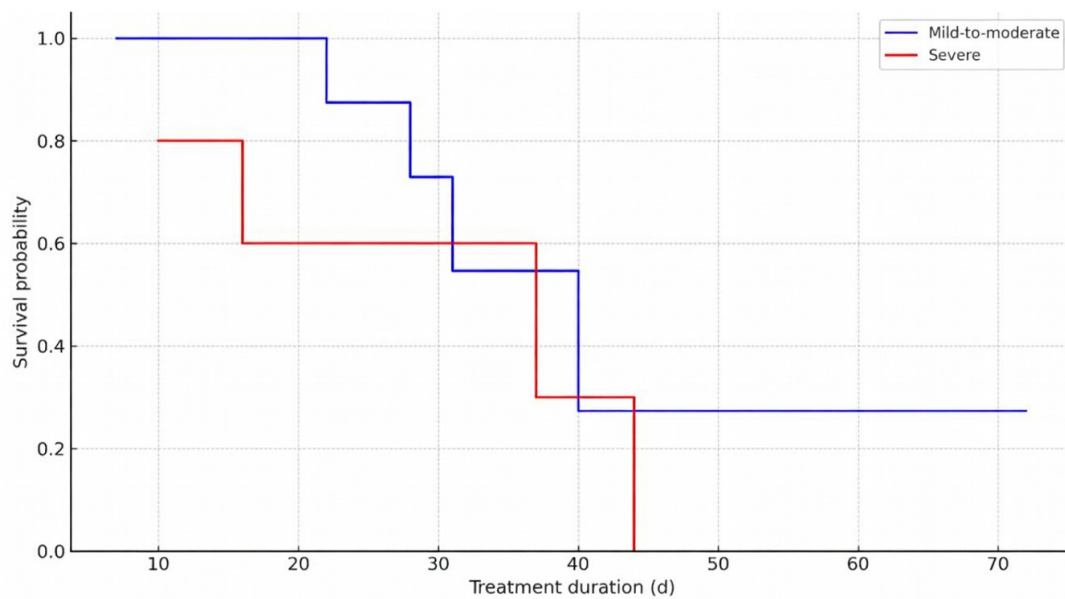
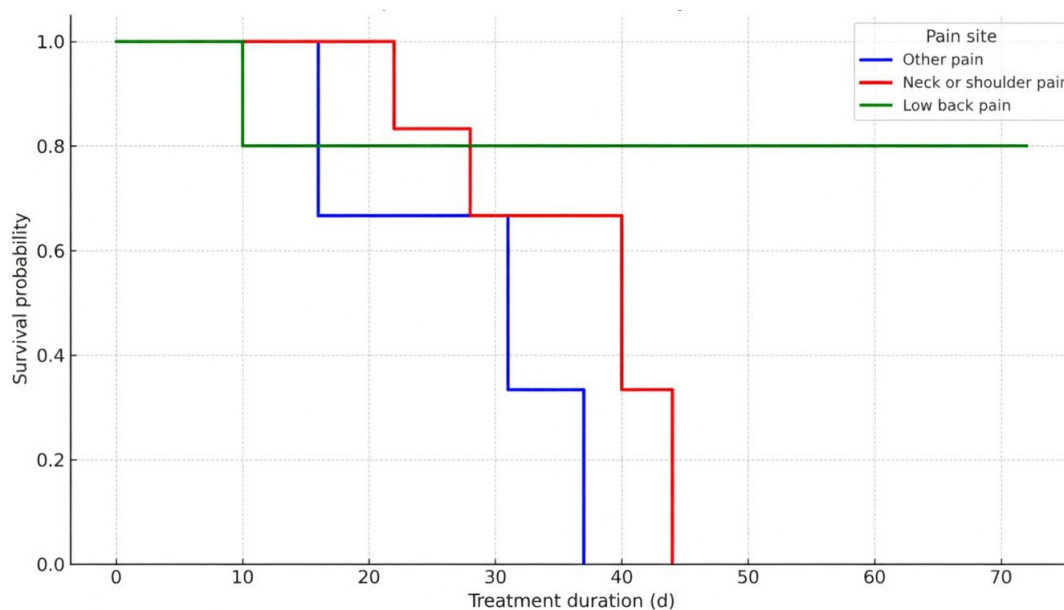


Figure 8. Kaplan-Meier survival curves for time to clinical improvement by pain site group.**Table 12.** Bootstrap validation of median time-to-improvement estimates by pain subgroup (N=16).

Group	KM ^a survival analysis (wk), median	Bootstrap ^b (wk), median	95% CI
Severe pain	5.0	4.9	4.4-5.4
MMP ^c	3.6	3.7	3.1-4.1
LBP ^d	4.6	4.4	4.0-5.0
N/SP ^e	4.0	3.9	3.4-4.4
Other pain	3.4	3.6	3.0-4.3

^aKM: Kaplan-Meier.

^bBootstrap validation, using 1000 iterations, was conducted to assess the robustness and reliability of the median time to clinical improvement.

^cMMP: mild-to-moderate pain.

^dLBP: low back pain.

^eN/SP: neck or shoulder pain.

Discussion

Principal Findings

This pilot study examined the preliminary effectiveness of GPM and explored predictive factors of clinically significant improvement using ML models, demonstrating significant effectiveness. The ML model analysis revealed that baseline pain interference, pain catastrophizing, fear-avoidance beliefs, and game engagement were identified as key predictors. The findings also highlight the optimal sessions and intervention duration by pain intensity and site subgroups, suggesting the potential for personalized pathways to achieve optimal outcomes.

Interpretation

The GPM intervention demonstrated significant improvement in pain and psychological outcomes, supporting that the integration of game elements can effectively address both physical and psychosocial dimensions of CP [9,10]. The superior performance of the gradient boosting classifier in predicting clinically significant pain reduction, consistent with previous studies, suggests that treatment outcomes are

multifactorial [37-39]. The strongest predictor was baseline pain interference rather than pain intensity, supported by the Health Belief Model that individuals perceiving higher functional impairment may benefit more from self-management [40]. Furthermore, our findings demonstrated the role of cognitive and affective factors. Rumination of pain catastrophizing was a key predictor, magnifying the negative cognitive loop of pain that is related to disability [41,42]. Similarly, the impact of fear-avoidance beliefs, especially work-related fear-avoidance, suggests that maintaining daily function may be more critical than purely physical rehabilitation [43]. These findings highlight that pain management should primarily address the barriers which are strongly related to daily lives rather than pain intensity score.

Our findings also identified session 3 and session 5 as key contributors to clinical pain recovery. Both of these sessions introduced core strength training, which was more physically demanding [44,45]. Psychoeducational content focusing on pain reconceptualization, including the alarm system and safe analgesic use, may foster self-efficacy by simultaneously providing patients with tangible tools to reduce pain-related fear and misunderstanding [46-49]. The combination of sessions 3, 5, and 7 further suggests that sequencing-targeted

exercises with specific educational content may be more effective than a general exercise program. In addition, this combination was most potent for severe LBP, further suggesting that a specific sequence and combination of content are critical. The SHAP-informed modeling and Kaplan-Meier analysis further indicated that individuals with severe pain or LBP required a longer intervention period to achieve a high probability of clinical benefit compared to those with MMP or N/SP. These findings suggest an adaptive intervention strategy that patients with severe LBP could be guided toward a curriculum emphasizing core strength and pain neuroscience education (such as sessions 3 and 5) and encouraged to adhere to at least 5 weeks, while patients with neck or shoulder pain might achieve similar outcomes with a shorter intensive program.

This study established and extended previous studies on digital GPM. First, despite the existing gamified interventions [11,14,50] (Table S2 in [Multimedia Appendix 1](#)), they often reported heterogeneous effects and seldom tracked intrainervention engagement behaviors and time-to-improvement thresholds. Our study addresses this by integrating self-tracking function and satisfaction, quiz accuracy recorded functions, allowing us to analyze the high impact intervention components. Second, while recent literature on the application of ML in CP management reveals promise, most studies relied on baseline patient data rather than intrainervention dynamics, lacking behavioral, and engagement metrics [51]. Other studies focused on predictive accuracy including diagnosis or referral need [11,19,51-54], without detailing modeling or treatment engagement, session-by-session behaviors or time-to-improvement. Our study incorporated both baseline and continuously recorded variables into the ML algorithms to offer an adaptive framework for future personalized GPM intervention.

Limitations

This study has several limitations that should be acknowledged. First, the small sample size limits the statistical power and generalizability of the findings [55]. The risk of overfitting in the ML models is considerable given the high feature-to-sample ratio, although cross-validation was used to mitigate this. These results should therefore be considered preliminary and hypothesis-generating for future larger-scale studies. As the employed participants may exhibit higher motivation to reduce work-related inequities and return to work [56], larger and more diverse samples are needed to confirm the results and understand the intervention's broader applicability. Additionally, the absence of an external validation dataset limits the assessment of the ML model's generalizability. Future studies should validate model performance using independent datasets. Moreover, given the high feature-to-sample ratio, there remains a risk of overfitting despite cross-validation. Dimensionality reduction techniques or regularization methods should be considered in future models. Moreover, reliance on self-reported measures,

such as pain intensity, interference, and engagement-related behaviors, may introduce subjective bias. Incorporating objective monitoring tools or physiological measures could complement self-report data and enhance data accuracy. In addition, recruitment via social media may introduce selection bias, as individuals with higher digital literacy differ from the broader CP population. Future studies should aim for more inclusive participant recruitment strategies to enhance the generalizability of the findings.

Implications

While the findings should be interpreted with caution due to the small sample size, this study highlights the potential of gamified interventions and its accompanying ML models in managing CP. The most immediate implications stem from the ML models, which demonstrated potentially predictive accuracy. With a larger sample size, clinicians could use these models as a decision-support tool to personalize pain management strategies, ensuring more effective and tailored treatment plans. To implement the proposed gradient boosting model in clinical practice, the model should be integrated into a digital platform where clinicians input patient data, incorporating data quality check by using valid scale and halt predictions for missing or invalid inputs to ensure reliability. The clinician role would be to accurately administer the required patient-reported outcome measures and interpret the probability of clinical improvement. Furthermore, the preliminary findings underscore the necessity of integrating psychological support within gamified programs. Addressing factors such as pain catastrophizing and fear-avoidance beliefs appears as crucial as targeting physical symptoms for improving outcomes. Finally, the analysis of intrainervention metrics and session-specific effects across multiple pain subgroups offers a foundation for future research, enabling the development of tailored interventions with optimized treatment lengths and content sequencing. Future research should also evaluate the feasibility of GPM through controlled trials with multiple populations.

Conclusion

In conclusion, this pilot study highlights the potential of leveraging ML models to analyze dynamic data from GPM to inform adaptive and personalized CP management. The study transcends the traditional reliance on static baseline data and overall efficacy assessments of previous studies, instead integrating continuously recorded behavioral metrics to identify key predictors, optimal session combinations, and subgroup-specific TDs. These exploratory findings present a process-oriented, data-driven paradigm that could potentially assist health care providers in customizing intervention content, sequencing, and dosage to enhance adherence in clinical practice. Future larger-scale controlled trials are necessary to validate these exploratory models and convert them into clinical decision-support systems that could improve patient involvement and outcomes.

Acknowledgments

The authors would like to thank all of the doctors and clinicians who assisted in revising the intervention design for their support and consultations. They would also like to thank Ziji Chen, who was involved in data collection. The authors used the generative AI tool ChatGPT (GPT-4) to refine the language and improve the readability of the manuscript. The content was reviewed and approved by all authors.

Funding

The authors declared no financial support was received for this work.

Data Availability

The deidentified participant data and analytical code generated during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: MMYT, TTOK

Data Curation: JH

Formal analysis: JH

Investigation: JH

Methodology: JH

Project administration: MMYT

Resources: JH

Software: JH

Supervision: MMYT, TTOK

Validation: JH

Visualization: JH

Writing - review & editing: MMYT, TTOK

Writing - original draft: JH

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed preintervention and postintervention comparisons of pain and psychological outcomes.

[\[DOCX File \(Microsoft Word File\), 21 KB-Multimedia Appendix 1\]](#)

Checklist 1

TRIPOD-AI checklist.

[\[PDF File \(Adobe File\), 294 KB-Checklist 1\]](#)

References

1. GBD 2017 Mortality Collaborators. Global, regional, and national age-sex-specific mortality and life expectancy, 1950–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*. Nov 2018;392(10159):1684–1735. [doi: [10.1016/S0140-6736\(18\)31891-9](https://doi.org/10.1016/S0140-6736(18)31891-9)]
2. Turk DC, Monarch ES. Biopsychosocial perspective on chronic pain. In: Turk DC, Gatchel RJ, editors. *Psychological Approaches to Pain Management: A Practitioner's Handbook*. 2nd ed. The Guilford Press; 2022:3–29. URL: <https://www.guilford.com/excerpts/turk3.pdf?t=1> [Accessed 2026-03-29]
3. He J, Tse MMY, Kwok TTO, Wu TCM, Tang S. Exploring the pain situation, pain impact, and educational preferences of pain among adults in Mainland China, a cross-sectional study. *Healthcare (Basel)*. Jan 31, 2025;13(3):289. [doi: [10.3390/healthcare13030289](https://doi.org/10.3390/healthcare13030289)] [Medline: [39942478](https://pubmed.ncbi.nlm.nih.gov/39942478/)]
4. Ma Y, Wu X, Shen S, et al. Relationship between locomotive syndrome and musculoskeletal pain and generalized joint laxity in young Chinese adults. *Healthcare (Basel)*. Feb 10, 2023;11(4):532. [doi: [10.3390/healthcare11040532](https://doi.org/10.3390/healthcare11040532)] [Medline: [36833063](https://pubmed.ncbi.nlm.nih.gov/36833063/)]
5. WHO guideline for non-surgical management of chronic primary low back pain in adults in primary and community care settings. World Health Organization; 2023. URL: <https://iris.who.int/server/api/core/bitstreams/29f3a0de-6c68-49b5-b8ed-e04f53022edd/content> [Accessed 2026-03-13]
6. Aiyegbusi OL, Cruz Rivera S, Oliver K, et al. The opportunity for greater patient and public involvement and engagement in drug development and regulation. *Nat Rev Drug Discov*. May 2023;22(5):337–338. [doi: [10.1038/d41573-023-00031-x](https://doi.org/10.1038/d41573-023-00031-x)] [Medline: [36854837](https://pubmed.ncbi.nlm.nih.gov/36854837/)]

7. Powell SM, Larsen CA, Phillips SM, Pellegrini CA. Exploring beliefs and preferences for reducing sedentary behavior among adults with symptomatic knee osteoarthritis or knee replacement. *ACR Open Rheumatol*. Jan 2021;3(1):55-62. [doi: [10.1002/acr2.11216](https://doi.org/10.1002/acr2.11216)] [Medline: [33400397](https://pubmed.ncbi.nlm.nih.gov/33400397/)]
8. Sailer M, Homner L. The gamification of learning: a meta-analysis. *Educ Psychol Rev*. Mar 2020;32(1):77-112. [doi: [10.1007/s10648-019-09498-w](https://doi.org/10.1007/s10648-019-09498-w)]
9. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". Presented at: MindTrek '11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; Sep 28-30, 2011; Tampere, Finland. URL: <https://dl.acm.org/doi/10.1145/2181037.2181040> [Accessed 2026-03-13] [doi: [10.1145/2181037.2181040](https://doi.org/10.1145/2181037.2181040)]
10. Hamari J, Koivisto J. Social motivations to use gamification: an empirical study of gamifying exercise. Presented at: ECIS 2013 - Proceedings of the 21st European Conference on Information Systems; Jun 6-8, 2013; Utrecht, Netherlands. URL: <https://research.aalto.fi/en/publications/social-motivations-to-use-gamification-an-empirical-study-of-gami/> [Accessed 2026-03-13]
11. Vermeir JF, White MJ, Johnson D, Crombez G, Van Ryckeghem DML. Gamified web-delivered attentional bias modification training for adults with chronic pain: randomized, double-blind, placebo-controlled trial. *JMIR Serious Games*. Jan 16, 2025;13:e50635. [doi: [10.2196/50635](https://doi.org/10.2196/50635)] [Medline: [39819575](https://pubmed.ncbi.nlm.nih.gov/39819575/)]
12. Saragih ID, Suarilah I, Saragih IS, Lin YK, Lin CJ. Efficacy of serious games for chronic pain management in older adults: a systematic review and meta-analysis. *J Clin Nurs*. Mar 2024;33(3):1185-1194. [doi: [10.1111/jocn.17012](https://doi.org/10.1111/jocn.17012)] [Medline: [38291564](https://pubmed.ncbi.nlm.nih.gov/38291564/)]
13. Gajardo Sánchez AD, Murillo-Zamorano LR, López-Sánchez J, Bueno-Muñoz C. Gamification in health care management: systematic review of the literature and research agenda. *Sage Open*. Oct 2023;13(4). [doi: [10.1177/21582440231218834](https://doi.org/10.1177/21582440231218834)]
14. Harvie DS, Kelly J, Kluver J, Deen M, Spitzer E, Coppeters MW. A randomized controlled pilot study examining immediate effects of embodying a virtual reality superhero in people with chronic low back pain. *Disabil Rehabil Assist Technol*. Apr 2024;19(3):851-858. [doi: [10.1080/17483107.2022.2129846](https://doi.org/10.1080/17483107.2022.2129846)] [Medline: [36256688](https://pubmed.ncbi.nlm.nih.gov/36256688/)]
15. Mueller J, Niederer D, Tenberg S, Oberheim L, Moesner A, Mueller S. Acute effects of game-based biofeedback training on trunk motion in chronic low back pain: a randomized cross-over pilot trial. *BMC Sports Sci Med Rehabil*. Nov 13, 2022;14(1):192. [doi: [10.1186/s13102-022-00586-z](https://doi.org/10.1186/s13102-022-00586-z)] [Medline: [36372870](https://pubmed.ncbi.nlm.nih.gov/36372870/)]
16. Rütth M, Schmelzer M, Burtiak K, Kaspar K. Commercial exergames for rehabilitation of physical health and quality of life: a systematic review of randomized controlled trials with adults in unsupervised home environments. *Front Psychol*. 2023;14:1155569. [doi: [10.3389/fpsyg.2023.1155569](https://doi.org/10.3389/fpsyg.2023.1155569)] [Medline: [37333591](https://pubmed.ncbi.nlm.nih.gov/37333591/)]
17. Marcuzzi A, Klevanger NE, Aasdahl L, et al. An artificial intelligence-based app for self-management of low back and neck pain in specialist care: process evaluation from a randomized clinical trial. *JMIR Hum Factors*. Jul 9, 2024;11:e55716. [doi: [10.2196/55716](https://doi.org/10.2196/55716)] [Medline: [38980710](https://pubmed.ncbi.nlm.nih.gov/38980710/)]
18. Anan T, Kajiki S, Oka H, et al. Effects of an artificial intelligence-assisted health program on workers with neck/shoulder pain/stiffness and low back pain: randomized controlled trial. *JMIR Mhealth Uhealth*. Sep 24, 2021;9(9):e27535. [doi: [10.2196/27535](https://doi.org/10.2196/27535)] [Medline: [34559054](https://pubmed.ncbi.nlm.nih.gov/34559054/)]
19. Amorim P, Paulo JR, Silva PA, Peixoto P, Castelo-Branco M, Martins H. Machine learning applied to low back pain rehabilitation – a systematic review. *Int J Digit Health*. 2021;1(1):10. [doi: [10.29337/ijdh.34](https://doi.org/10.29337/ijdh.34)]
20. Marcuzzi A, Nordstoga AL, Bach K, et al. Effect of an artificial Intelligence-based self-management app on musculoskeletal health in patients with neck and/or low back pain referred to specialist care: a randomized clinical trial. *JAMA Netw Open*. Jun 1, 2023;6(6):e2320400. [doi: [10.1001/jamanetworkopen.2023.20400](https://doi.org/10.1001/jamanetworkopen.2023.20400)] [Medline: [37368401](https://pubmed.ncbi.nlm.nih.gov/37368401/)]
21. Hertzog MA. Considerations in determining sample size for pilot studies. *Res Nurs Health*. Apr 2008;31(2):180-191. [doi: [10.1002/nur.20247](https://doi.org/10.1002/nur.20247)] [Medline: [18183564](https://pubmed.ncbi.nlm.nih.gov/18183564/)]
22. Ersek M, Turner JA, Cain KC, Kemp CA. Results of a randomized controlled trial to examine the efficacy of a chronic pain self-management group for older adults [ISRCTN11899548]. *Pain*. Aug 15, 2008;138(1):29-40. [doi: [10.1016/j.pain.2007.11.003](https://doi.org/10.1016/j.pain.2007.11.003)] [Medline: [18086516](https://pubmed.ncbi.nlm.nih.gov/18086516/)]
23. Ibrahim AA, Akindede MO, Ganiyu SO. Effectiveness of patient education plus motor control exercise versus patient education alone versus motor control exercise alone for rural community-dwelling adults with chronic low back pain: a randomised clinical trial. *BMC Musculoskelet Disord*. Feb 23, 2023;24(1):142. [doi: [10.1186/s12891-022-06108-9](https://doi.org/10.1186/s12891-022-06108-9)] [Medline: [36823567](https://pubmed.ncbi.nlm.nih.gov/36823567/)]
24. Chimenti RL, Post AA, Rio EK, et al. The effects of pain science education plus exercise on pain and function in chronic achilles tendinopathy: a blinded, placebo-controlled, explanatory, randomized trial. *Pain*. Jan 1, 2023;164(1):e47-e65. [doi: [10.1097/j.pain.0000000000002720](https://doi.org/10.1097/j.pain.0000000000002720)] [Medline: [36095045](https://pubmed.ncbi.nlm.nih.gov/36095045/)]
25. McMahon SB, Koltzenburg M, Tracey I, Turk D. Wall & Melzack's Textbook of Pain E-Book: Expert Consult - Online and Print. Elsevier Health Sciences; 2013. ISBN: 9780702053740

26. Wenjuanxing. URL: <https://www.wjx.cn/> [Accessed 2026-03-17]
27. ivx. URL: <https://www.ivx.cn/> [Accessed 2026-03-23]
28. Wang S, Yao S, Shang L, Xu C, Ma J. Validation of the Chinese version of the Brief Pain Inventory in patients with knee osteoarthritis. *J Orthop Surg Res*. 2023;18(720). [doi: [10.1186/s13018-023-04218-1](https://doi.org/10.1186/s13018-023-04218-1)]
29. Pei LB, Xia JJ, Yan JL. Cross-cultural adaptation, reliability and validity of the Chinese version of the Fear Avoidance Beliefs Questionnaire. *J Int Med Res*. 2010;38(6):1985-1996. [doi: [10.1177/147323001003800612](https://doi.org/10.1177/147323001003800612)] [Medline: [21227002](https://pubmed.ncbi.nlm.nih.gov/21227002/)]
30. Xiong Y MD, Xu X, Wei X, et al. Validation of a simplified chinese version of the Pain Catastrophizing Scale and an exploration of the factors predicting catastrophizing in pain clinic patients. *Pain Physician*. Nov 4, 2015;18(6):E1059-E1072. [Medline: [26606019](https://pubmed.ncbi.nlm.nih.gov/26606019/)]
31. Godin G, Shephard RJ. Godin Leisure-Time Exercise Questionnaire. *Med Sci Sports Exerc*. Jun 1997;29:36-38. [doi: [10.1097/00005768-199706001-00009](https://doi.org/10.1097/00005768-199706001-00009)]
32. Zeng QZ, He YL, Liu H, et al. Reliability and validity of Chinese version of the Generalized Anxiety Disorder 7-item (GAD-7) scale in screening anxiety disorders in outpatients from traditional Chinese internal department. *Chin Ment Health J*. 2013;27(3):163-168. [doi: [10.3969/j.issn.1000-6729.2013.03.001](https://doi.org/10.3969/j.issn.1000-6729.2013.03.001)]
33. Wang W, Bian Q, Zhao Y, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry*. 2014;36(5):539-544. [doi: [10.1016/j.genhosppsych.2014.05.021](https://doi.org/10.1016/j.genhosppsych.2014.05.021)] [Medline: [25023953](https://pubmed.ncbi.nlm.nih.gov/25023953/)]
34. Luo N, Li M, Liu GG, Lloyd A, de Charro F, Herdman M. Developing the Chinese version of the new 5-level EQ-5D descriptive system: the response scaling approach. *Qual Life Res*. May 2013;22(4):885-890. [doi: [10.1007/s11136-012-0200-0](https://doi.org/10.1007/s11136-012-0200-0)] [Medline: [22622627](https://pubmed.ncbi.nlm.nih.gov/22622627/)]
35. Phan MH, Keebler JR, Chaparro BS. The development and validation of the Game User Experience Satisfaction Scale (GUESS). *Hum Factors*. Dec 2016;58(8):1217-1247. [doi: [10.1177/0018720816669646](https://doi.org/10.1177/0018720816669646)] [Medline: [27647156](https://pubmed.ncbi.nlm.nih.gov/27647156/)]
36. Jensen MP, Tomé-Pires C, de la Vega R, Galán S, Solé E, Miró J. What determines whether a pain is rated as mild, moderate, or severe? The importance of pain beliefs and pain interference. *Clin J Pain*. May 2017;33(5):414-421. [doi: [10.1097/AJP.0000000000000429](https://doi.org/10.1097/AJP.0000000000000429)] [Medline: [27584819](https://pubmed.ncbi.nlm.nih.gov/27584819/)]
37. Lv S, Sun N, Hao C, Li J, Li Y. Development and validation of machine learning models for predicting post-cesarean pain and individualized pain management strategies: a multicenter study. *BMC Anesthesiol*. Apr 10, 2025;25(1):170. [doi: [10.1186/s12871-025-03034-w](https://doi.org/10.1186/s12871-025-03034-w)] [Medline: [40211131](https://pubmed.ncbi.nlm.nih.gov/40211131/)]
38. Oh AR, Park J, Shin SJ, et al. Prediction model for postoperative atrial fibrillation in non-cardiac surgery using machine learning. *Front Med (Lausanne)*. 2022;9:983330. [doi: [10.3389/fmed.2022.983330](https://doi.org/10.3389/fmed.2022.983330)] [Medline: [36703881](https://pubmed.ncbi.nlm.nih.gov/36703881/)]
39. Zhao P, Xue J, Xu X, Wang L, Chen D. Logistic model and gradient boosting machine model for physical therapy of lumbar disc herniation. *Comput Math Methods Med*. 2022;2022:4799248. [doi: [10.1155/2022/4799248](https://doi.org/10.1155/2022/4799248)] [Medline: [35602348](https://pubmed.ncbi.nlm.nih.gov/35602348/)]
40. Heelas L, Soni A, Barker K. Do baseline patient reported outcome measures predict changes in self-reported function, following a chronic pain rehabilitation programme? *Br J Pain*. Dec 2023;17(6):532-545. [doi: [10.1177/20494637231190190](https://doi.org/10.1177/20494637231190190)] [Medline: [37974636](https://pubmed.ncbi.nlm.nih.gov/37974636/)]
41. Bartley EJ, Rhudy JL. The influence of pain catastrophizing on experimentally induced emotion and emotional modulation of nociception. *J Pain*. May 2008;9(5):388-396. [doi: [10.1016/j.jpain.2007.11.015](https://doi.org/10.1016/j.jpain.2007.11.015)] [Medline: [18226969](https://pubmed.ncbi.nlm.nih.gov/18226969/)]
42. Meints SM, Stout M, Abplanalp S, Hirsh AT. Pain-related rumination, but not magnification or helplessness, mediates race and sex differences in experimental pain. *J Pain*. Mar 2017;18(3):332-339. [doi: [10.1016/j.jpain.2016.11.005](https://doi.org/10.1016/j.jpain.2016.11.005)] [Medline: [27908838](https://pubmed.ncbi.nlm.nih.gov/27908838/)]
43. Jakobsen MD, Vinstrup J, Andersen LL. Work-related fear-avoidance beliefs and risk of low-back pain: prospective cohort study among healthcare workers. *J Occup Rehabil*. Sep 2025;35(3):547-555. [doi: [10.1007/s10926-024-10221-y](https://doi.org/10.1007/s10926-024-10221-y)] [Medline: [39103730](https://pubmed.ncbi.nlm.nih.gov/39103730/)]
44. Gatti R, Faccendini S, Tettamanti A, Barbero M, Balestri A, Calori G. Efficacy of trunk balance exercises for individuals with chronic low back pain: a randomized clinical trial. *J Orthop Sports Phys Ther*. Aug 2011;41(8):542-552. [doi: [10.2519/jospt.2011.3413](https://doi.org/10.2519/jospt.2011.3413)]
45. Chang WD, Lin HY, Lai PT. Core strength training for patients with chronic low back pain. *J Phys Ther Sci*. Mar 2015;27(3):619-622. [doi: [10.1589/jpts.27.619](https://doi.org/10.1589/jpts.27.619)] [Medline: [25931693](https://pubmed.ncbi.nlm.nih.gov/25931693/)]
46. Nijs J, Paul van Wilgen C, Van Oosterwijck J, van Ittersum M, Meeus M. How to explain central sensitization to patients with “unexplained” chronic musculoskeletal pain: practice guidelines. *Man Ther*. Oct 2011;16(5):413-418. [doi: [10.1016/j.math.2011.04.005](https://doi.org/10.1016/j.math.2011.04.005)] [Medline: [21632273](https://pubmed.ncbi.nlm.nih.gov/21632273/)]
47. Lepri B, Romani D, Storari L, Barbari V. Effectiveness of pain neuroscience education in patients with chronic musculoskeletal pain and central sensitization: a systematic review. *Int J Environ Res Public Health*. Feb 24, 2023;20(5):4098. [doi: [10.3390/ijerph20054098](https://doi.org/10.3390/ijerph20054098)] [Medline: [36901108](https://pubmed.ncbi.nlm.nih.gov/36901108/)]

48. Cardenas DD, Jensen MP. Treatments for chronic pain in persons with spinal cord injury: a survey study. *J Spinal Cord Med.* 2006;29(2):109-117. [doi: [10.1080/10790268.2006.11753864](https://doi.org/10.1080/10790268.2006.11753864)] [Medline: [16739554](https://pubmed.ncbi.nlm.nih.gov/16739554/)]
49. Davies LEM, Koerkamp EAJG, Koster ES, et al. Patients' perspectives about the role of primary healthcare providers in long-term opioid therapy: a qualitative study in Dutch primary care. *Br J Gen Pract.* Jul 2024;74(744):e475-e481. [doi: [10.3399/BJGP.2023.0547](https://doi.org/10.3399/BJGP.2023.0547)] [Medline: [38499298](https://pubmed.ncbi.nlm.nih.gov/38499298/)]
50. Sun Y, Xian Y, Lin H, Sun X. Enhancing the management of non-specific neck pain through gamification: design and efficacy of a health application. *Bioengineering (Basel).* Jun 23, 2024;11(7):640. [doi: [10.3390/bioengineering11070640](https://doi.org/10.3390/bioengineering11070640)] [Medline: [39061722](https://pubmed.ncbi.nlm.nih.gov/39061722/)]
51. Fundoiano-Hershcovitz Y, Pollak K, Goldstein P. Personalizing digital pain management with adapted machine learning approach. *Pain Rep.* 2023;8(2):e1065. [doi: [10.1097/PR9.0000000000001065](https://doi.org/10.1097/PR9.0000000000001065)] [Medline: [37731749](https://pubmed.ncbi.nlm.nih.gov/37731749/)]
52. Lian Y, Shi Y, Shang H, Zhan H. Predicting treatment outcomes in patients with low back pain using gene signature-based machine learning models. *Pain Ther.* Feb 2025;14(1):359-373. [doi: [10.1007/s40122-024-00700-8](https://doi.org/10.1007/s40122-024-00700-8)] [Medline: [39722081](https://pubmed.ncbi.nlm.nih.gov/39722081/)]
53. C Areias A, G Moulder R, Molinos M, et al. Predicting pain response to a remote musculoskeletal care program for low back pain management: development of a prediction tool. *JMIR Med Inform.* Nov 19, 2024;12:e64806. [doi: [10.2196/64806](https://doi.org/10.2196/64806)] [Medline: [39561359](https://pubmed.ncbi.nlm.nih.gov/39561359/)]
54. Skoric J, Janmohamed T, Lumsden-Ruegg H, Clarke H, Katz J, Rahman QA. Predicting functional improvement in chronic pain using machine learning and digital health data from the Manage My Pain App. *IEEE J Biomed Health Inform.* Oct 23, 2025:1-12. [doi: [10.1109/JBHI.2025.3624093](https://doi.org/10.1109/JBHI.2025.3624093)] [Medline: [41129438](https://pubmed.ncbi.nlm.nih.gov/41129438/)]
55. Almutiri R, Alhabeeb S, Alhumud S, Ullah Khan R. A survey of machine learning for big data processing. *J Big Data.* 2022;4(2):97-111. [doi: [10.32604/jbd.2022.028363](https://doi.org/10.32604/jbd.2022.028363)]
56. Ostelo RWJG, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976).* Jan 1, 2008;33(1):90-94. [doi: [10.1097/BRS.0b013e31815e3a10](https://doi.org/10.1097/BRS.0b013e31815e3a10)] [Medline: [18165753](https://pubmed.ncbi.nlm.nih.gov/18165753/)]

Abbreviations

AI: artificial intelligence

AUC: area under the curve

CP: chronic pain

GPM: gamified pain management

GUESS: Game User Experience Satisfaction Scale

LASSO: least absolute shrinkage and selection operator

LBP: low back pain

ML: machine learning

MMP: mild-to-moderate pain

N/SP: neck or shoulder pain

PHQ-9: Patient Health Questionnaire-9 items

SAT: session satisfaction scores

SDT: self-determination theory

SHAP: Shapley additive explanations

TD: treatment duration

TRIPOD-AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis-AI

Edited by Stefano Brini; peer-reviewed by Alexandros Gazis, Geisa Guimaraes de Alencar, Jiaqi Li; submitted 10.Jun.2025; final revised version received 26.Dec.2025; accepted 26.Dec.2025; published 07.Apr.2026

Please cite as:

Tse MYM, He J, Kwok TTO

A Gamified Pain Management Intervention for Adults With Chronic Pain in Mainland China: Single-Arm Pre-Post Pilot Study With Machine Learning Predictive Modeling

JMIR Form Res 2026;10:e78823

URL: <https://formative.jmir.org/2026/1/e78823>

doi: [10.2196/78823](https://doi.org/10.2196/78823)

License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.