Medical and Health Care Applications, Vol. 211, 2026, 291–298
https://doi.org/10.54941/ahfe1007492

Machine Learning-Based Identification of Non-Responders to A 12-Month Digital Self-Management in Knee Osteoarthritis

Yingyi Li, Xiaoyi Wang, and Calvin Or

Department of Data and Systems Engineering, The University of Hong Kong, Hong Kong, China

ABSTRACT

Knee osteoarthritis (KOA) places a significant burden on individuals and healthcare systems worldwide. Although self-management programs (SMPs) offer accessible support for KOA management, individual responses vary. Therefore, early identification of those unlikely to show substantial improvement with self-management is important to adapt treatment plans promptly, introduce more effective interventions when needed, and ultimately improve patient outcomes by avoiding prolonged use of strategies that do not produce responses. This study aims to develop a machine learning-based approach to identify potential non-responders to a digital SMP using characteristics at program entry. Data were obtained from a previously conducted 12-month app-based SMP for KOA. Responders were defined as individuals whose arthritis self-efficacy (ASE) and health-related quality of life (HRQoL) outcomes both improved. Non-responders included those who showed improvement in only one outcome, no improvement in either, or deterioration in one or both outcomes. After excluding participants who did not complete the SMP or who underwent knee-related surgery and/or hospitalization during the study period, 57 participants were included in the analysis. Body mass index, presence of non-musculoskeletal comorbidities, ASE score, HRQoL index, and hip range of motion were the input features for model development. Gradient boosting decision tree achieved the best performance, with an AUC of 0.822 and balanced sensitivity and specificity in identifying non-responders. These findings present the feasibility of using machine learning to early identify individuals with limited expected benefit from digital self-management. Such identification may facilitate more efficient, tailored, and proactive strategies for managing KOA. Future research should prioritize external validation in larger and more diverse cohorts.

Keywords: Knee osteoarthritis, Digital self-management, Machine learning, Early identification

INTRODUCTION

Knee osteoarthritis (KOA) is a leading chronic musculoskeletal disease worldwide, commonly resulting in pain, functional limitations, and reduced quality of life (Cui et al., 2020; Duong et al., 2023). Individuals with mild KOA often experience frustrating knee pain during daily activities and work, while those with more severe KOA may develop progressive mobility impairment and require long-term care. Consequently, KOA imposes a substantial and

Received February 3, 2026; Revised April 4, 2026; Accepted April 21, 2026; Available online July 20, 2026

growing burden on affected individuals, caregivers, and healthcare systems (Bannuru et al., 2019; Cui et al., 2020).

With the rapid advances in digital technologies, digital self-management programs (SMPs) have emerged as accessible approaches to support long-term osteoarthritis management (Eysenbach, 2001; Slater et al., 2016; Gogovor et al., 2017; Chen, Kalun Or and Chen, 2020). These digital SMPs usually deliverexercise guidance, educational materials, and health-related content through mobile apps, the internet, virtual reality equipment, and other digital technologies, enabling individuals to engage in rehabilitation with limited clinical supervision (Osborne, Spinks and Wicks, 2004; Chen and Or, 2021; Or et al., 2025). Existing studies suggest that SMPs can be effective in improving physical function and quality of life (Dahlberg et al., 2020; Safari, Jackson and Sheffield, 2020; Shah et al., 2022). Digital SMPs have been recommended as one promising application of KOA management.

However, evidence also indicates substantial heterogeneity in individual responses to SMPs (Smith et al., 2013; Kroon et al., 2014). While many individuals benefit from digital SMPs, others experience limited or no meaningful improvement. Despite this challenge, most previous studies have evaluated intervention effectiveness retrospectively (Kroon et al., 2014; Babatunde et al., 2023), with less attention to identifying individuals unlikely to respond prior to intervention. If such individuals cannot be identified early, they may continue to participate in SMPs with limited improvement, potentially delaying access to alternative treatments and leading to prolonged suboptimal management.

Therefore, early identification of patients who are unlikely to be suitable for SMPs is critical for informing clinical decision-making in KOA management. Accordingly, this study aims to develop a machine learning-based model to identify non-responders to a digital SMP using characteristics at program entry, enabling earlier identification of individuals less suitable for SMPs and supporting more efficient KOA management.

METHODS

Study Population and Responder Classification

Data used in this study were derived from our previously conducted 12-month app-based SMP for individuals with KOA, which recruited participants from four public hospitals (n = 80). Eligible participants were aged ≥ 50 years, had radiographically confirmed (Kellgren-Lawrence grade 2-3), were able to complete the physical function assessments, communicate in Chinese, and provide written informed consent. Individuals with inflammatory arthritis, recent knee surgery or trauma, concurrent structured knee interventions, or cognitive impairment were excluded. Prior to the 12-month digital SMP, demographic characteristics and comorbidity profiles were collected. Participants were required to complete self-reported assessments of knee pain, arthritis self-efficacy (ASE), and health-related quality of life (HRQoL), as well as eight standardized tests of lower-limb physical function (see Table 1), which have been widely recommended for evaluating functional limitations in individuals with KOA (Dobson et al., 2013). These measures encompassed symptoms, self-management capacity, perceived overall well-being, and physical function relevant to self-management in KOA, thereby supporting responder classification and model development. The same assessments were conducted at the 12-month follow-up.

To evaluate individual response to self-management, responder classification was based on changes in ASE and HRQoL, as these measures directly reflect individuals' capacity to manage their condition and their perceived overall well-being. Knee pain and physical function measures were not used as classification criteria because they are influenced by symptom fluctuations and short-term contextual factors, such as temporary painkiller use or testing-day conditions, and may not reliably indicate whether an individual has developed effective and sustainable self-management skills.

Accordingly, given the study focus on early identification of potential non-responders, a conservative classification criterion was adopted. Participants were classified as responders if both ASE and HRQoL improved over the study period (ΔASE > 0 AND ΔHRQoL > 0). All other change patterns were classified as non-responders, including improvement in only one outcome, no improvement in either outcome, or deterioration in one or both outcomes. After excluding participants who did not complete the SMP (n = 18) or underwent knee-related surgery and/or hospitalization during the study period (n = 5), 57 participants were included in the final analysis and subsequent model development.

Input Feature Selection

A filter-based feature selection approach was applied as an exploratory screening step to reduce dimensionality and improve interpretability. Participant characteristics at SMP entry (baseline characteristics), including demographic and clinical characteristics, as well as assessments of lower-limb physical function, knee pain, ASE, and HRQoL, were compared between responders and non-responders. Normality of continuous variables was assessed prior to analysis. Independent-samples t-tests were used for continuous variables that met normality assumptions, otherwise Mann-Whitney U-tests were applied. Chi-square tests were used for categorical variables. Variables with statistically significant between-group differences ($p < 0.05$) were selected as input features for subsequent model development. The baseline characteristics of the two groups and the statistical results are summarized in Table 1.

Table 1: Baseline characteristics of responders and non-responders, with statistically significant between-group differences highlighted in bold.

Characteristics	Non-responder n (%) or Mean ± SD	Responder n (%) or Mean ± SD	p[1]
Female/Male	23 (65.7%)/ 12 (34.3%)	14 (63.6%)/ 8(36.4%)	N/A[2]
Age (years)	64.91 ± 6.32	64.86 ± 7.37	0.979
BMI (kg/m^2)	24.46 ± 4.25	26.71 ± 4.05	0.049*
Single-leg stance (second)	43.71 ± 32.57	38.92 ± 30.80	0.576
15-second stepping (repetition)	12.14 ± 3.19	10.50 ± 3.86	0.103
30-second chair stand (repetition)	11.34 ± 4.02	10.09 ± 4.48	0.292
Fastest pace walking (m/s)	1.28 ± 0.27	1.16 ± 0.31	0.132
Knee ROM (degree)	127.77 ± 26.21	124.96 ± 19.61	0.431
Hip ROM (degree)	132.79 ± 27.72	125.09 ± 16.21	0.032*
Maximum quadriceps strength (kg)	14.47 ± 6.29	13.75 ± 6.79	0.692
Maximum hamstrings strength (kg)	8.42 ± 4.07	9.22 ± 5.57	0.935
Knee pain in past 3 months (0-10)	4.57 ± 2.10	5.00 ± 1.80	0.456
Knee pain in walking (0-10)	3.63 ± 2.28	3.73 ± 1.80	0.857
Knee pain in stairs (0-10)	5.00 ± 2.54	5.77 ± 1.90	0.197
ASE score (1-10)	6.83 ± 1.87	5.50 ± 1.36	0.003**
HRQoL index (0-1)	0.89 ± 0.11	0.78 ± 0.14	0.002**
Presence of non-musculoskeletal comorbidities	16 (72.7%)	14 (40.0%)	0.033*
History of lower-limb problems	12 (54.5%)	23 (65.7%)	0.573

[1] * *$p < 0.05$, ** $p < 0.01$.*

[2] *No statistical test was performed.*

Model Development and Validation

For model development, non-responders were the target category and labelled as the positive class (label = 1), while responders were labelled as the negative class (label = 0). Multiple machine learning models were applied to identify non-responders, including logistic regression (LR) and support vector machines (SVM) as baseline classifiers, as well as tree-based models such as decision trees (DT), random forests (RF), and gradient boosting decision trees (GBDT). These models were selected to balance interpretability and robustness while enabling non-linear modelling in small-sample conditions.

A Bayesian optimization approach (Optuna) was applied to optimize hyper-parameters for each candidate classifier independently, using a repeated 5-fold cross-validation framework to enable fair performance comparison across models. Receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) were computed for each classifier to determine the best-performing model. Model performance was evaluated using the AUC, sensitivity, specificity, F1-score, and balanced accuracy, with particular emphasis on sensitivity for detecting non-responders.

RESULTS

Participant Characteristics and Input Features

Participant characteristics are summarized in Table 1. Baseline comparisons between responders and non-responders identified five characteristics that differed significantly between the two groups, including body mass index (BMI), hip range of motion (ROM), ASE score, HRQoL index, and the presence of non-musculoskeletal comorbidities. These characteristics were subsequently used as input features for model development.

Best-Performing Model and Performance

Table 2 summarizes the performance of all candidate classifiers evaluated using repeated 5-fold cross-validation. Among these models, the GBDT achieved the highest AUC and the most favourable balance between sensitivity and specificity, and was therefore identified as the best-performing model.

Specifically, the GBDT model achieved an AUC of 0.822, a sensitivity of 0.781, a specificity of 0.687, an F1-score of 0.784, and a balanced accuracy of 0.734 for detecting non-responders (Table 2).

Table 2: Performance metrics of all candidate classifiers for non-responder identification.

Model	AUC	Sensitivity	Specificity	F1-score	Balanced Accuracy
GBDT	**0.822**	**0.781**	**0.687**	**0.784**	**0.734**
RF	0.761	0.638	0.627	0.677	0.632
LR	0.750	0.629	0.723	0.683	0.676
DT	0.729	0.695	0.643	0.700	0.690
SVM	0.729	0.581	0.690	0.643	0.635

DISCUSSION

This study demonstrates the feasibility of a machine learning-based approach for identifying individuals with KOA who may be unlikely to benefit from 12-month digital self-management using baseline characteristics, including physical function, comorbidity, health-related quality of life, and self-management confidence.

Among the evaluated machine learning models, the GBDT model showed the strongest performance in identifying individuals at high risk of non-response. The relatively high sensitivity indicates that the GBDT was able to correctly identify a substantial proportion of individuals who were unlikely to benefit from self-management at an early stage. Moreover, the model maintained a balanced trade-off between sensitivity and specificity. Although sensitivity was prioritized, F1-score and balanced accuracy exceeded 0.7, indicating that non-responders were not detected at the expense of misclassifying potential responders.

Implications

Beyond the stable model performance, the demonstrated feasibility of the proposed machine learning-based approach has clinical implications. The input features used for identification in this study are readily available in routine care settings. These variables, including demographic information, comorbidities, simple physical function assessments, and self-reported measures of arthritis self-efficacy and health-related quality of life, can be collected without specialized equipment or additional clinical burden, which indicates that the proposed early identification approach is cost-effective and scalable in clinical practice.

Furthermore, making an initial judgment about the suitability of self-management based on easily obtainable information is valuable for supporting clinical decision-making. When potential non-responders are not identified early and continue to engage in self-management programs with limited gains, this may reduce the care efficiency and postpone consideration of alternative or clinician-guided options. Therefore, early identification of individuals at risk of non-response supports integrating early stratification and decision-making into digital self-management for KOA. This approach may allow potential responders to continue self-management with greater confidence, while enabling clinicians to proactively consider personalized or clinically supervised treatments for those at higher risk of limited benefit. Building on this, incorporating early identification strategies into digital care has the potential to improve care efficiency, reduce unnecessary delays, and promote a more patient-centered approach to the KOA management.

Limitations and Future Work

Several limitations should be noted. First, the sample size was relatively modest, which may influence the generalizability of the results and the stability of the model. Second, this study employed an exploratory approach using baseline characteristics for early identification, and further validation with external datasets would be beneficial to assess the robustness and applicability of these findings. Third, this study was conducted within a single digital SMP. Further validation in larger and more diverse populations, as well as across different types of digital intervention platforms, should be conducted. Finally, response was defined based on changes in arthritis self-efficacy and quality of life. While this may not capture all aspects of individual trajectories, it offers a practical and conceptually aligned basis for early identification. Future studies could include longitudinal measures of objective functional indicators, such as lower-limb physical function, and utilize multi-time-point data to enhance early prediction models.

CONCLUSION

This study provides a machine learning-based approach to detecting the potential non-responders to KOA self-management. By identifying non-responders early, the proposed approach may support more timely consideration of alternative or clinician-guided treatments. While the findings demonstrate feasibility, further external validation in larger and more diverse populations is required. This work will contribute to the development of more efficient and proactive strategies for KOA management.

ACKNOWLEDGMENT

This study was supported by a grant from the Health and Medical Research Fund, Food and Health Bureau, The HKSAR (PI: Calvin Or; project number: 17181291).

REFERENCES

Babatunde, O.O. et al. (2023) 'Interventions for supporting self-management in people with musculoskeletal pain, including focus on varying levels of health literacy: A systematic overview of reviews', Health Open Research, 5, p. 2. Available at: https://doi.org/10.12688/healthopenres.13151.1.

Bannuru, R.R. et al. (2019) 'OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis', Osteoarthritis and Cartilage, 27(11), pp. 1578–1589. Available at: https://doi.org/10.1016/j.joca.2019.06.011.

Chen, T., Kalun Or, C. and Chen, J. (2020) 'A systematic review and meta-analysis of randomized controlled trials to evaluate technology-supported exercise programs for knee health', Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 64(1), pp. 639–640. Available at: https://doi.org/10.1177/1071181320641145.

Chen, T. and Or, C.K. (2021) 'Development and pilot test of a machine learning-based knee exercise system with video demonstration, real-time feedback, and exercise performance score', Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 65(1), pp. 1519–1523. Available at: https://doi.org/10.1177/1071181321651109.

Cui, A. et al. (2020) 'Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies', EClinicalMedicine, 29–30, p. 100587. Available at: https://doi.org/10.1016/j.eclinm.2020.100587.

Dahlberg, L.E. et al. (2020) 'Improving osteoarthritis care by digital means - Effects of a digital self-management program after 24- or 48-weeks of treatment', PLOS ONE, 15(3), p. e0229783. Available at: https://doi.org/10.1371/journal.pone.0229783.

Dobson, F. et al. (2013) 'OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis', Osteoarthritis and Cartilage, 21(8), pp. 1042–1052. Available at: https://doi.org/10.1016/j.joca.2013.05.002.

Duong, V. et al. (2023) 'Evaluation and Treatment of Knee Pain: A Review', JAMA, 330(16), pp. 1568–1580. Available at: https://doi.org/10.1001/jama.2023.19675.

Eysenbach, G. (2001) 'What is e-health?', Journal of Medical Internet Research, 3(2), p. e833. Available at: https://doi.org/10.2196/jmir.3.2.e20.

Gogovor, A. et al. (2017) 'Informing the development of an Internet-based chronic pain self-management program', International Journal of Medical Informatics, 97, pp. 109–119. Available at: https://doi.org/10.1016/j.ijmedinf.2016.10.005.

Kroon, F.P. et al. (2014) 'Self-management education programmes for osteoarthritis', Cochrane Database of Systematic Reviews. Edited by Cochrane Musculoskeletal Group, 2014(1). Available at: https://doi.org/10.1002/14651858.CD008963.pub2.

Or, C.K. et al. (2025) 'Technological Surrogate Physiotherapy to Improve Knee Health Through Exercise: Human-Computer Interaction to Build Trust and Acceptance Notwithstanding Pain', International Journal of Human–Computer Interaction, 41(17), pp. 10834–10845. Available at: https://doi.org/10.1080/10447318.2024.2438289.

Osborne, R.H., Spinks, J.M. and Wicks, I.P. (2004) 'Patient education and self-management programs in arthritis', The Medical Journal of Australia, 180(S5), pp. S23-26. Available at: https://doi.org/10.5694/j.1326-5377.2004.tb05909.x.

Safari, R., Jackson, J. and Sheffield, D. (2020) 'Digital Self-Management Interventions for People With Osteoarthritis: Systematic Review With Meta-Analysis', Journal of Medical Internet Research, 22(7), p. e15365. Available at: https://doi.org/10.2196/15365.

Shah, N. et al. (2022) 'Applications of Digital Health Technologies in Knee Osteoarthritis: Narrative Review', JMIR Rehabilitation and Assistive Technologies, 9(2), p. e33489. Available at: https://doi.org/10.2196/33489.

Slater, H. et al. (2016) 'Use of eHealth technologies to enable the implementation of musculoskeletal Models of Care: Evidence and practice', Best Practice & Research Clinical Rheumatology, 30(3), pp. 483–502. Available at: https://doi.org/10.1016/j.berh.2016.08.006.

Smith, T.O. et al. (2013) 'Self-management Programmes for People with Osteoarthritis: A Systematic Review and Meta-analysis', Current Rheumatology Reviews, 9(3), pp. 165–175. Available at: https://doi.org/10.2174/157339710903140130121859.