IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 14 March 2024, accepted 26 March 2024, date of publication 1 April 2024, date of current version 9 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3383786

RESEARCH ARTICLE

Multivariate Time-Series Cluster Analysis for Multiple Functional Domains to Identify Recovery Patterns of Patients With Fragility Hip Fracture After Surgery

CHANYOUNG PARK[1], HONGBUM KIM[1], JUNGWON SUH[2], JINHEE KO[2], JUN HWAN CHOI[3], SANG YOON LEE[4], JAEWON BEOM[5], JAE-YOUNG LIM[5], BO RYUN KIM[2], AND HYO KYUNG LEE[1], (Member, IEEE)

[1]School of Industrial and Management Engineering, Korea University, Seoul 02841, Republic of Korea
[2]Department of Physical Medicine and Rehabilitation, Korea University College of Medicine, Seoul 02841, Republic of Korea
[3]Department of Rehabilitation Medicine, Regional Rheumatoid and Degenerative Arthritis Center, Jeju National University Hospital, Jeju National University College of Medicine, Jeju 63241, Republic of Korea
[4]Department of Rehabilitation Medicine, Seoul National University Boramae Medical Center, Seoul 07061, Republic of Korea
[5]Department of Rehabilitation Medicine, Seoul National University Bundang Hospital, Seoul National University, College of Medicine, Seongnam-si, Gyeonggi-do 13620, Republic of Korea

Corresponding authors: Hyo Kyung Lee (hyokyunglee@korea.ac.kr) and Bo Ryun Kim (brkim08@korea.ac.kr)

This work was supported in part by the Brain Korea21 (BK21) FOUR funded by the Ministry of Education of Korea and National Research Foundation of Korea, and in part by Korean Academy of Geriatric Rehabilitation Medicine Grant and the National Research Foundation of Korea (NRF) Grant funded by Korean Government through [Ministry of Science and ICT (MSIT)] under Grant RS-2022-00165597.

ABSTRACT Patients who have undergone hip fracture surgery have the primary goal of recovering their premorbid level of function across diverse functional domains, including walking ability, balance, cognitive function, emotional well-being, frailty, and activities of daily living. As the speed and level of recovery can vary substantially across functional domains and individuals, the varying recovery patterns of different outcome measures should be considered when designing rehabilitation plans for patients. However, the lack of knowledge of recovery trajectories and their variations in hip fracture patients impedes such efforts. In this study, we develop a multivariate time-series clustering algorithm to analyze the recovery patterns and identify patient groups with similar recovery patterns across multiple functional outcomes. Five distinct recovery patterns were observed that exhibit varying maximum recovery levels and speeds. These findings demonstrated the significance of utilizing multiple outcome measures concurrently to assess the patient's recovery level. Recovery patterns are identified to exhibit variations across different domains, revealing contrasting trends between walking ability and cognitive outcomes. Furthermore, we present predictions on the trajectory of recovery during the post-acute phase solely based on the acute-phase information. This approach facilitates the early identification of patient groups with an unfavorable prognosis for recovery.

INDEX TERMS Machine learning, patient rehabilitation, pattern clustering, time series analysis.

I. INTRODUCTION

Hip fractures are a major public health problem worldwide with the annual global incidence expected to increase to 2.6 million by 2025 [1]. Owing to the aging population, the numbers are expected to rise steeply, and fragile hip fractures are known to be a major cause of mortality and morbidity in the older adult population [2], [3]. A substantial proportion of patients with fragility hip fractures undergo surgical intervention. Subsequently, postsurgical interventions are primarily centered around rehabilitating patients to swiftly attain stability and pain-free lower extremities [4]. Surgical treatments, together with post-surgical rehabilitation, consume a major portion of healthcare

The associate editor coordinating the review of this manuscript and approving it for publication was Juan A. Lara.

time and resources, leading to a significant economic burden worldwide.

To optimize the management of fragility hip fractures, a considerable amount of the literature has focused on rehabilitative endeavors aimed at reinstating older adult patients to their pre-injury level of function. A review of studies on functional recovery after hip fractures reported that maximum functional recovery occurs primarily within the first six months after the fracture. Notably, only 40–70% of patients regain their ability to perform activities of daily living (ADL), and only 40–60% of patients regain their pre-fracture walking ability [5]. Furthermore, in most older adult patients, hip fractures severely affect not only physical function, but also various other areas, including health status, quality of life, and mental function [6]. In another study, patients exhibited a gradual reduction in dependence across the majority of functional domains during the first year following the fracture. However, the recovery time for depression, upper limb function, and cognitive function extended up to approximately four months, while lower extremity function took up to 12 months to fully recover. Therefore, it can be inferred that the duration of recovery can vary substantially depending on the specific functional area [7].

Several studies have been conducted to investigate the factors that influence post-surgical functional outcomes in patients with fragility hip fractures, and the following key factors have been reported: age, pre-fracture functional status, pre-existing comorbidity, fracture site, type of surgery, delay in operation, functional level at discharge, malnutrition, frailty, and cognitive status [8], [9], [10], [11], [12], [13], [14]. Hence, the speed and level of recovery following a hip fracture can vary substantially among individuals depending on their preoperative health status, surgical approach, age, and pre-fracture functional status. Such individual-level factors and varying recovery patterns across multiple functional domains should be considered when designing optimal rehabilitation treatment plans for patients. At present, the lack of knowledge about recovery trajectories and their diversities in patients with hip fractures hinders these endeavors.

Research on recovery after hip fracture surgery has encountered limitations in adequately capturing these variations. A majority of previous studies have concentrated on individual functional outcomes and compared recovery levels at fixed time points among different patient groups. In general, recovery patterns were compared by dividing groups by demographic information such as sex [15], [16], [17] and age [18], [19], while other studies have categorized patient groups based on frailty [20], cognition [21] or ADL [22], which are considered primary outcomes of the rehabilitation process. There also exists studies that have placed focus on analyzing the recovery trajectories among groups that received rehabilitation compared to control groups [23]. Consequently, the comprehensive understanding of the overall recovery trajectory remains deficit, as the analysis of recovery levels at each time point is conducted independently from each other.

Research has also explored the analysis of recovery patterns post hip fracture surgery by tracking the temporal progression of a single metric. These studies typically chose ADL as the main outcome, creating different groups to assess variations among them [24], [25]. Some studies delved into the patterns of mental aspects such as cognition, emotions, and depression [26], [27], [28], [29]. Nevertheless, these studies are limited to examining individual functional outcomes and do not capture the integrated recovery patterns across multiple outcome measures.

A few studies have examined the recovery trajectory using a combination of measures. Dakhil et al. [30] identified four distinct groups for ADL trajectories using instrumental ADL (iADL) and personal ADL (pADL) with the growth mixture model. Coló-Emeric et al. [31] compared patient characteristics by dividing 10 outcome measures at two, six, and 12 months post-operation into low, intermediate, and high resilience groups using latent class profile analysis (LCPA). However, while one study grouped subjects individually without considering multiple indicators simultaneously, the other study overlooked the recovery level during the acute rehabilitation period, which is vital for comprehending maximum recovery potential. Moreover, both studies solely relied on outcome measures within restricted functional domains such as ADLs and walking ability.

In addition, existing research predominantly employs a statistical approach in which models are constructed under restrictive statistical assumptions (mandating that the entire distribution should be represented as a sum of multiple distributions), a condition that might not universally hold true. In general, statistical methodologies, such as LCPA when applied to longitudinal datasets, face a disadvantage in terms of scalability. As the analysis aims to uncover more intricate patterns by including additional time points and variables, the number of parameters that need to be estimated grows exponentially. This increase in parameters not only makes the model more complex but also demands a significantly larger sample size for accurate estimation.

To overcome the limitations posed by conventional statistical methodologies, we aimed to achieve the following: (1) develop a multivariate time-series clustering algorithm that comprehensively accounts for the recovery trajectories of multiple functional outcome measures, (2) identify patient groups characterized by similar recovery patterns and compute the functional outcome recovery trajectory curves for each group, and (3) predict future recovery trajectories solely utilizing the outcomes obtained during the acute-phase.

II. METHOD

The overall research framework is comprised of three steps which are conducted sequentially as depicted in Fig. 1. First, missing values from the time points where data were not collected are adequately filled in to obtain a complete dataset (A. Data Imputation). Second, utilizing a multivariate time-series clustering method on the complete dataset, common patterns of recovery across multiple functional outcomes

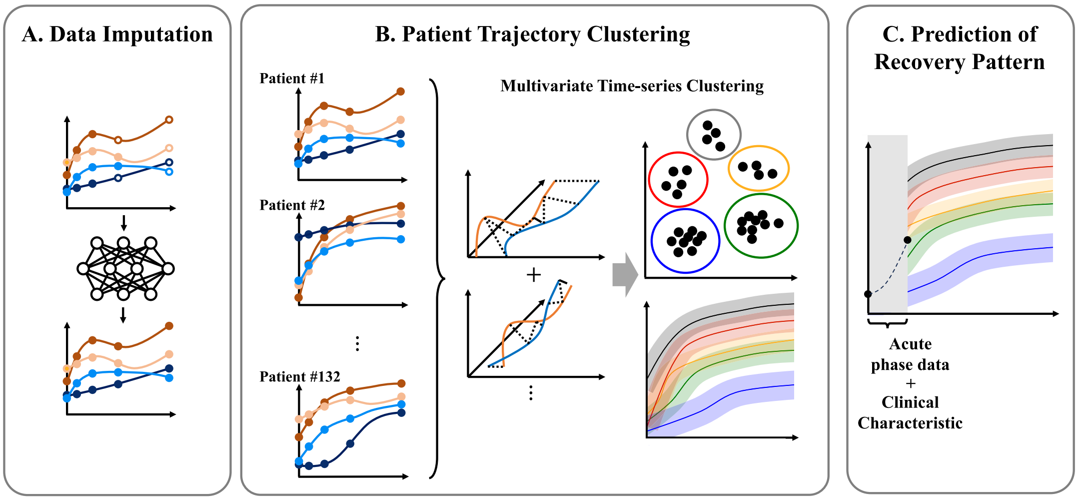


FIGURE 1. **Framework of the proposed methodology. (A) Data Imputation: Missing values are adequately filled in using deep learning based imputation method. (B) Patient Trajectory Clustering: Recovery trajectories of multiple functional outcomes are concurrently utilized to group patients into similar recovery pattern groups. (C) Prediction of Recovery Pattern: Using the acute rehabilitation phase information, predictions are made regarding the pattern of the recovery progression that will be observed in the post-acute phase.**

are derived (B. Patient Trajectory Clustering). Finally, exclusively utilizing the acute rehabilitation phase information, predictions are made on the pattern of recovery progression in the post-acute phase (C. Prediction of Recovery Pattern). This study was approved by the Institutional Review Board of the Korea University Hospital (IRB No. 2022AN0259).

A. DATA DESCRIPTION

Data was collected from a previous fragility fracture integrated rehabilitation management (FIRM) study [32]. FIRM is a 10-day comprehensive rehabilitation program designed for patients undergoing hip fracture surgery. Every patient participated in a total of 10 physical therapy sessions twice a day, with each session lasting for 60 minutes, and also engaged in four occupational therapy sessions [33]. This study included 211 patients who underwent surgery for fragility hip fractures and were subsequently admitted to the Department of Rehabilitation Medicine between February 2017 and February 2019 at three different locations, including Seoul National University Bundang Hospital, Chung-Ang University Hospital, and Jeju National University Hospital. The patient selection criteria were as follows: (1) age 65 years or older, (2) experienced an acute unilateral hip joint fracture (femoral neck, intertrochanteric, or subtrochanteric), and (3) successful hip joint surgery (reduction and fixation, bipolar hemiarthroplasty, or total hip arthroplasty). Patients who underwent surgery for reasons unrelated to hip joint fractures, such as osteoarthritis, hip joint infection, and avascular necrosis; or patients with other fractures unrelated to acute unilateral hip joint fractures, such as femoral shaft fractures, patellar fractures, isolated fractures of the greater trochanter, pathological fractures caused by tumors, or complex multiple fractures; patients who had a history of neurodegenerative diseases or unstable cardiopulmonary conditions; patients who underwent revision surgery; patients with severe cognitive impairment; and patients who refused to participate in the clinical trials, were excluded.

The following demographic data were collected on admission to the rehabilitation department: age, sex, fracture location and laterality, surgical type, and underlying diseases. Functional outcomes were measured at five different predetermined stages of patient care: postoperative days (POD) at one week (1 W), three weeks (3 W), three months (3 M), six months (6 M), and 12 months (12 M). Functional outcome measurements consisted of both clinical outcome assessments (COA) and patient reported outcome measures(PROMs), including walking ability (Koval walking ability scores, Functional independence Measure-locomotion, Functional Ambulatory Category, Modified Rivermead mobility index, Berg balance scale), activities of daily living(Korean version of the modified Barthel index, Korean version of the Instrumental ADL), cognition(Korean version of the MiniMental State Examination), emotion (Korean version of the Geriatric Depression Scale), quality of life (Euro Quality of Life Questionnaire 5-dimension Index Score), and frailty (Korean version of the FRAIL scale). Detailed information on the outcome measures is presented in Table 1.

In general, when patients were admitted to or visited the hospital, all functional outcomes listed in Table 1 were measured. If a patient did not visit the hospital during the assessment period, only the outcomes that could be confirmed via telephone (KOVAL, FAC, EQ5D, K-IADL, and K-FRAIL) were collected. In addition, missing values arose during the data collection period due to reasons such as

TABLE 1. Description of functional outcome measurements where ↗ indicates a higher score representing better outcome and ↘ indicates lower score representing better outcome. Abbreviations: ADL: Activities of Daily Living, QoL: Quality of Life, KOVAL: Koval Walking Ability Scores; FIM-locomotion: Functional Independence Measure-Locomotion; FAC: Functional Ambulatory Category; MRMI: Modified Rivermead mobility index; BBS: Berg balance scale; K-MMSE: Korean version of the MiniMental State Examination; K-GDS: Korean version of the Geriatric Depression Scale; EQ-5D: Euro Quality of Life Questionnaire 5-dimension Index Score; K-MBI: Korean version of the modified Barthel index; K-IADL: Korean version of the Instrumental ADL; K-FRAIL: Korean version of the FRAIL scale.

Domain		Description	Range
Clinical Outcome Assessment			
Walking Ability	KOVAL [34]	Evaluate the dependency on walking indoors and outdoors	1-7(↘)
	FIM-locomotion [35]	Evaluate a participant's independent ability to walk or use a wheelchair over a distance of 45m	1-7(↗)
	FAC [36]	Evaluate walking ability	0-5(↗)
	MRMI [37]	Evaluate changes in a patient's level of mobility in routine clinical practice, comprising a range of activities from turning over in bed to running	0-40(↗)
	BBS [38]	A patient's balancing ability while performing common functional tasks in everyday life	0-56(↗)
Activities of daily living	K-MBI [39]	Tool for assessing basic ADL in stroke patients	0-100(↗)
	K-IADL [40]	Tool for evaluating instrumental activities of daily living	0-3.7(↘)
Cognition	K-MMSE [41]	Evaluate cognitive function including memory, orientation to place and time, naming, reading, copying, writing, and the ability to follow a three-stage command	0-30(↗)
Patient Reported Outcome Measures			
Emotion	K-GDS [42]	A self-report scale designed to assess depression	0-30(↘)
Quality of life	EQ-5D [43]	Evaluate self-reported quality of life and measure general health status	-0.066-0.95(↗)
Frailty	K-FRAIL [44]	Screening tool for frailty status using a simple 5-item questionnaire	0-5(↘)

cohort dropout, death, or the conclusion of the study period. Since the purpose of this study is to analyze the recovery trajectory of functional outcomes, patients who did not visit the hospital more than twice to measure the outcomes or who did not confirm the outcomes via telephone are excluded. Thus, out of the 211 patient records, data from 132 patients are ultimately selected for analysis.

B. MISSING DATA IMPUTATION

To analyze the recovery trajectories of various functional outcomes for the 132 patients, it was imperative to address the missing values effectively as an initial step. The dataset contained a total of 20.6% missing values where outcome measures exhibited varying missing rates depending on whether they could be collected remotely or not. Outcome measures that could be collected via telephone (KOVAL, FAC, EQ5D, K-IADL, K-FRAIL) showed lower missing rate of about 7% while those that required patient visit to the hospital showed higher missing rate of around 30%. Hence, to make use of the outcome measures exhibiting low missingness as much as possible, imputation method that exploits the relationship between multiple time-series dataset is selected to fill in the missing values. Specifically, we used the multidirectional Recurrent Neural Networks (M-RNN) [45] algorithm, which is known to be highly effective in handling missing values in multivariate time-series setting. M-RNN is based on a deep learning architecture that exploits a combination of Bidirectional Recurrent Neural Networks (Bi-RNN) and Fully Connected Neural Networks (FCNN).

The multivariate time-series dataset used in this study has two unique characteristics. First, as is common with time-series data, each functional outcome measurement displays temporal correlation across consecutive time steps. In other words, the value of a functional outcome at a particular time point is influenced by the value of the same functional outcome at another time point. Second, there exists a correlation between different functional outcome measurements taken at the same time point. This is an interesting property that commonly exists in longitudinal health datasets, where changes in various clinical outcome measurements are tracked over time. Different clinical outcomes measured at the same time point may be related because all clinical outcomes are dependent on the patient's underlying health status at that time. The M-RNN addresses these characteristics by considering the correlations within and across outcome measures. Fig. 2 shows that M-RNN first employs Bi-RNN to interpolate the missing value of each outcome measure across time, and then employs FCNN to impute missing values across multiple outcome measures for the same time point. Hence, while traditional data processing methods typically consider either interpolation or imputation, the M-RNN considers both interpolation and imputation methods to address missing values. Additionally, because the intervals between patient visits were consistently measured for all patients, these periods were incorporated into the Bi-RNN part of the M-RNN by assigning weights during the training process, thus enabling effective missing-value interpolation.

C. RECOVERY PATTERN CLUSTERING ACROSS MULTIPLE OUTCOME MEASURES

Recovery patterns of fragility hip fracture encompass multiple domains, including physical health, cognition,

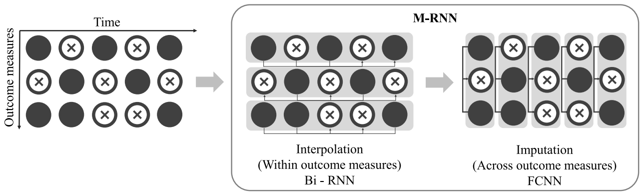


FIGURE 2. **Structure of M-RNN algorithm. A bidirectional recurrent neural network interpolates missing values by considering temporal correlations within outcome measures, followed by a fully connected neural network that performs imputation considering correlations across outcome measures.**

psychosocial health, and multidimensional outcomes, such as health-related quality of life. For this, we develop a multivariate time-series clustering algorithm to identify patient groups with similar recovery patterns. Specifically, we employed a distance-based clustering approach that quantifies the distance between each pair of time-series data points. This methodology clusters patients with comparable time-series data points, signifying that the recovery trajectories of distinct outcome measures exhibit similar patterns.

1) K-MEDOIDS CLUSTERING

K-medoids is a distance-based clustering method that is similar to the commonly used k-means algorithm. Both methods categorize the dataset into groups and attempt to minimize the distance between points in a cluster and a point that is designated as the "center" of that cluster. While k-means sets the cluster "center" as the mean value of the points within the cluster, k-medoids select the "center" from the actual data points within the cluster. Consequently, k-medoids exhibit a higher level of robustness because they are less affected by outliers, since it utilizes representative points within the cluster. This approach ensures that noise data or outliers exert minimal influence while assessing distances or similarities between clusters. Furthermore, k-medoids offer the advantage of accommodating alternative distance measures or similarity assessment methods beyond the traditional Euclidean distance, as the distances between the centers and points within the cluster are pre-computed. Among the various available distance measures, one can leverage the dynamic time warping (DTW) algorithm, which is specifically designed for pattern matching within time-series data.

2) MIXED MULTIVARIATE-DYNAMIC TIME WARPING (MM-DTW)

Dynamic Time Warping is an algorithm that compares the similarity between two time-series datasets with different speeds. In contrast to the Euclidean distance, DTW allows many-to-one comparisons between two time-series datasets to create the best possible alignment, exploiting temporal distortions between them. This facilitates the computation of similarity between two recovery trajectories that exhibit overall similarity in the recovery pattern but may possess a slight temporal shift. As a result, the approach remains insensitive to temporal shifts. Adopting DTW for recovery pattern matching presents the capability of identifying similar patterns of recovery while accounting for small differences in speed owing to individual characteristics.

Algorithm 1 MM-DTW

Input: $X = \left[x_{l,i}\right]_{L\times m}$ and $Y = \left[y_{l,j}\right]_{L\times n}$ are the multivariate time-series data of sample X and Y consisting of L variables and having sequence length of m and n, respectively.
Output: MM-DTW(X, Y)
Calculate pairwise-correlation for all L variables and construct correlation matrix $\left[\rho_{i,j}\right]_{L\times L}$ where $i, j = 1, \cdots, L$.
Convert correlation matrix to distance matrix $M_d = \left[m_{i,j}\right]_{L\times L}$ where $m_{i,j} = 1 - \rho_{i,j}^2 \ \forall\, i, j = 1, \cdots, L$.
Create K variable subsets S_k using k-medoids on matrix M_d.
for k = 1 to K **do**
  $X_k = \left[x_{p,i}\right]_{p\in S_k}, Y_k = \left[y_{p,i}\right]_{p\in S_k} \ \forall\, k = 1, \cdots, K$
  $D_k = DTWdistance(X_k, Y_k)$
end for
return $\sum_{k=1}^{K} D_k$

Although DTW is commonly used for univariate time-series, it can be extended to multivariate time-series. There are two types of multivariate DTW, depending on how we integrate the DTWs of multiple variables, including independent DTW, which treats each variable independently, and dependent DTW, which compares all variables under the assumption that the variables are related to each other. In other words, independent DTW calculates the DTW for each variable separately and then sums the variable-wise DTW to compute the overall distance between two consecutive time-series datasets. On the other hand, dependent DTW constructs a distance matrix between the elements of the two time-series using values from all dimensions simultaneously and performs DTW on the matrix. Both methods have advantages and disadvantages, and the appropriate method should be selected based on the data characteristics, which depends on whether the variables are related to each other or not. According to the experimental results, dependent DTW is a suitable choice when the variables exhibit similar temporal patterns, whereas independent DTW performs better when the variables are independent of each other [46].

In a longitudinal health dataset consisting of multiple time-series data, assessing different clinical outcomes of the patient, the temporal flow of the time-series may appear similar. Therefore, in this study, we develop a multivariate DTW algorithm that calculates distances using dependent DTW for measurements with similar temporal flows, and independent DTW for measurements with independent temporal flows. This requires the identification of measurements with similar temporal flows and measurements with independent temporal flows. Hence, we propose the following algorithm, mixed Multivariate- DTW (MM-DTW), in which variable-wise correlations are first computed to identify variables with similar patterns over time. The correlation matrix $[\rho_{i,j}]$ is then transformed into a distance matrix $M_d = [m_{i,j}]$ using $m_{i,j} = 1 - \rho_{i,j}^2$ [47] where $\rho_{i,j}$ is the Pearson's correlation coefficient of variables i and j. To identify variables exhibiting similar temporal pattern, we use these computed distance $m_{i,j}$ values to perform clustering analysis. Specifically, by using the k-medoids algorithm, K distinct variable subsets S_k can be identified where variables that exhibit similar temporal patterns belong to the same subset. Within each subset, DTW is performed to measure the distances between the variables (D_k). Finally, the distances calculated from all the subsets are aggregated to derive the final similarity measures.

Upon computation of the similarity measures for all patient pairs, k-medoids is performed to cluster patient groups into k distinct recovery patterns. Note that before performing k-medoids for subset clustering and patient group clustering, the optimal k should be selected which is determined by performing both the elbow and silhouette methods.

D. PREDICTING POST-ACUTE RECOVERY PATTERN

Rehabilitation after hip fracture begins in the acute care setting, but most cases occur in post-acute care settings, which can vary among nursing care facilities, home health, and outpatient settings [48]. Disparities in post-acute rehabilitation care are commonly observed, necessitating an evidence-based approach to match the degree of assistance required for each patient. Identifying susceptible patient groups whose recovery is expected to be limited can help target rehabilitation services more efficiently. Doing so requires an accurate prediction of how recovery will progress in the post-acute phase by exclusively utilizing the information obtained during the acute care phase. Thus, based on the patient's clinical characteristics and outcomes during the acute care phase, we aimed to develop a machine learning-based prediction model. This model would predict the recovery pattern cluster to which a patient is likely to transition in the post-acute phase. Here, the recovery pattern cluster refers to the patient groups that we identify in Section II-C. This type of predictive capability could enhance clinical decision-making by providing clinicians with the ability to counsel patients and caregivers on the anticipated recovery pattern, thereby enabling them to better plan their needs.

III. RESULT

A. DATA IMPUTATION RESULTS

To accurately compare the performance of the missing-value algorithms, 23 patients without missing values were selected from the dataset. Subsequently, we randomly masked certain time-point values to predict values for which the ground truth is known. For comparison with the M-RNN, multiple imputation by chained equations (MICE) [49] which is a commonly used multivariate imputation approach, is selected as a representative missing value handling method. Furthermore, the prevalence of missing values at different time points varied in the dataset, exhibiting percentages of 0.5%, 0.8%, 22.3%, 28.1%, and 51.3% at one week, three weeks, three months, six months, and 12 months, respectively. Missing value masks are generated based on these percentages, and five-fold cross-validation is performed. Different algorithms are compared using the RMSE (Root Mean Square Error) metric.

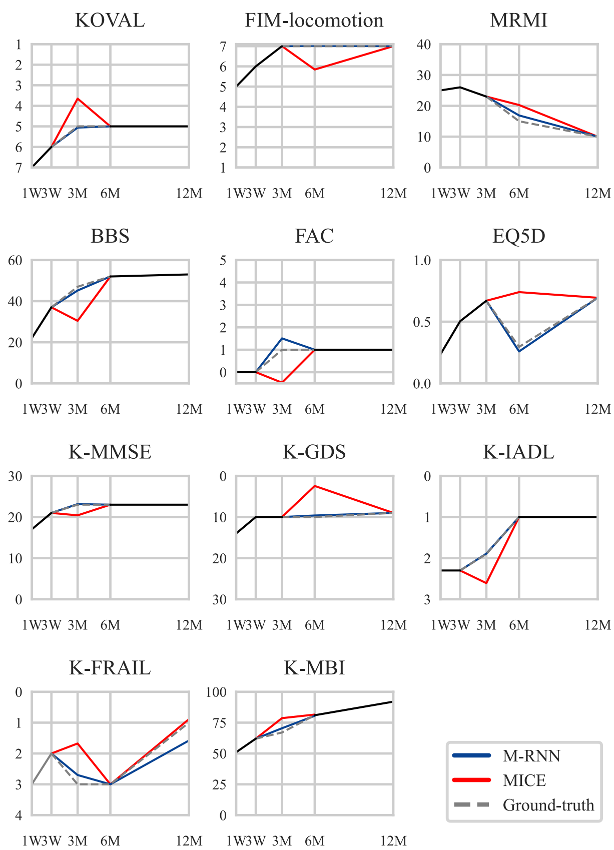


FIGURE 3. Illustrative examples of imputed results comparing M-RNN vs. MICE methods.

Prior to the performance comparison, we aim to find the optimal hyperparameter settings by considering the inclusion of number of deltas (time interval between visits), missing method (randomly mask according to real missing ratio), and clinical characteristics (age, sex, height, and weight). According to the experimental results, the optimal hyperparameter setting is [Delta, Missing Method, Clinical Characteristic = (NO, Select, NO)] where the details of which are shown in Table A3. The corresponding RMSE value for this combination is 0.588 ± 0.005 for M-RNN, which outperforms the RMSE of 0.932 obtained with MICE. Fig. 3 shows that MICE exhibits limitations in effectively adding missing values owing to its linear combination characteristics, whereas M-RNN demonstrates an enhanced ability to learn and fill missing value patterns. It can be observed that M-RNN performs relatively well even for cases

TABLE 2. Demographic and clinical characteristics for each patient group. Values represent the mean (±standard deviation) or number of patients. [a]chi-square test; [b]one-way ANOVA; [c]Fisher exact test. Abbreviations: FIRM: Fragility Fracture Integrated Rehabilitation Management; KOVAL: Koval walking ability scores; FAC: Functional Ambulatory Category.

		Total	HIGH	MID 1	MID 2	MID 3	LOW	P-value
N		132	30	36	12	24	30	
Male		29	8	10	2	4	5	0.703[a]
Female		103	22	26	10	20	25	
Age		81.05(7.00)	77.10(6.28)	80.69(6.50)	83.58(6.84)	82.50(7.78)	83.27(6.28)	0.003[b]
Height		155.69(8.04)	159.20(6.37)	156.00(7.16)	155.83(8.69)	154.37(9.26)	152.85(8.36)	0.035[b]
Weight		55.34(10.13)	61.92(7.38)	53.85(9.96)	53.35(10.36)	53.82(10.21)	52.55(10.33)	0.002[b]
Body mass index		22.80(3.64)	24.53(3.42)	22.11(3.68)	21.91(3.01)	22.46(3.23)	22.50(3.98)	0.054[b]
Fracture type	Femur neck	58	17	18	4	10	9	
	Intertrochanteric	68	11	15	8	13	21	0.181[c]
	Subtrochanteric	6	2	3	0	1	0	
Operation type	Bipolar hemiarthroplasty	60	13	14	2	12	19	
	Total hip relacement arthroplasty	9	3	5	1	0	0	0.058[c]
	Reduction and internal fixation	63	14	17	9	12	11	
Premorbid mobility status	Independent ambulation	84	26	29	6	14	9	
	Cane ambulation	32	4	6	2	7	13	
	Walker ambulation	14	0	1	3	2	8	< 0.001[c]
	Crunch ambulation	1	0	0	0	1	0	
	Wheelchair ambulation	1	0	0	1	0	0	
Group	FIRM	68	19	22	5	11	11	0.169[a]
	Control	64	11	14	7	13	19	
Premorbid KOVAL		1.80(1.29)	1.13(0.35)	1.42(0.94)	2.75(2.09)	1.63(0.82)	2.72(1.51)	< 0.001[b]
Premorbid FAC		4.58(0.77)	4.93(0.25)	4.67(0.76)	4.33(0.89)	4.63(0.58)	4.14(1.03)	0.002[b]

in which two successive values are missing, whereas MICE tends to either overestimate or underestimate the values. With the validated M-RNN, we filled in the actual missing values to obtain the complete dataset.

B. RECOVERY PATTERN CLUSTERING ANALYSIS

First, two variable subsets are derived by clustering variables with similar temporal flows. The first subset consists of the KOVAL, FIM-locomotion, FAC, EQ-5D, K-IADL, and K-FRAIL, whereas the second subset comprise the MRMI, BBS, K-MMSE, K-GDS, and K-MBI. This implies that the outcome measures within the same subtype exhibited similar temporal trends, whereas those in different subtypes display distinct temporal patterns. An interesting finding is that within the same subset, various types of outcomes that measure different functional domains are present. Although these measures evaluated different areas, it can be inferred that the recovery patterns over time are similar.

Subsequently, using MM-DTW to calculate the distances and perform k-medoids, the recovery patterns of the patients are subtyped into five distinct clusters. These clusters represent patient groups showing similar recovery patterns based on the concurrent consideration of various functional outcome measures. The five clusters are labeled as follows: a high recovery level cluster (HIGH), a low recovery level cluster (LOW), and three intermediate clusters (MID 1, MID 2, and MID 3) representing different levels of recovery progress in terms of overall functional outputs.

A comparison of the demographic and clinical characteristics of the five patient groups are shown in Table 2. The average age tends to be lower in the groups with a higher overall level of recovery, where the observed differences between the groups are statistically significant ($p = 0.003$). In addition, the group with the highest overall recovery level (HIGH) shows the highest values for height, weight, and BMI. The observed differences between the groups in terms of height ($p = 0.035$) and weight ($p = 0.002$) are statistically significant. In the context of the sex ratio, the difference between the groups is not statistically significant ($p = 0.703$). Significant differences among the groups are observed in premorbid mobility status where the group with higher recovery levels demonstrated a higher proportion of independent ambulation before hip fracture, while the group with lower recovery levels exhibit a higher proportion of walking aid usage, such as canes or walkers. However, no significant differences are observed in terms of fracture and operation types. Finally, while the difference lacks statistical significance, a prevalent pattern emerges indicating a greater proportion of the FIRM treatment group within the cohorts characterized by higher overall recovery levels.

For each functional outcome measure, the average recovery trajectory for each patient group is shown in Fig. 4 along with the overall mean trajectory, which represents the average recovery trajectory across all patients. The horizontal axis represents the elapsed time after surgery, while the vertical axis is appropriately ordered to indicate recovery when values are plotted above in the figure. The average trajectory of all patients is indicated using a dashed line, and the average trajectories of the five patient groups are indicated by black, red, orange, green, and blue solid lines in the order of HIGH, MID 1, MID 2, MID 3, and LOW, respectively. The shaded areas represent a standard deviation of ± 1 from the mean.

First, the overall mean recovery trajectories for all functional outcomes showed rapid recovery at three months post-operation, followed by a relatively gradual recovery or maintenance from three to 12 months. This observation aligns with previous literature, in which most hip fracture patients were reported to recover in the first three months post surgery [22], [25]. Upon examination of the recovery trajectories of each cluster, it becomes apparent that no cluster follows an identical path as the mean recovery trajectory (indicated by the dashed line) across all functional outcomes. This suggests that considerable information is lost when a simple estimation of the overall average movement is performed, as recovery patterns among patients vary substantially. When comparing the recovery patterns among the clusters, we observed that for most outcomes, the top two clusters (HIGH in black and MID 1 in red) and the bottom cluster (LOW in blue) were consistently ranked in the same order, while the recovery trajectories for MID 2 (orange) and MID 3 (green) did not follow a specific pattern. In the context of KOVAL and FIM-locomotion, which measure walking dependency, or BBS, which assesses balance and the risk of falling, the orange cluster (MID 2) exhibit better recovery patterns in comparison to the green cluster (MID 3), but in the context of the frailty index K-FRAIL and daily living performance K-IADL index, the green cluster performed better than the orange cluster. Of particular interest is the K-MMSE, which exhibit a significantly distinct pattern in comparison to other outcome measures and is in the order of HIGH (black), MID3 (green), MID1 (red), LOW (blue), and MID2 (orange).

C. POST-ACUTE RECOVERY PATTERN PREDICTION

Five recovery patterns are derived in Section III-B which exhibit distinctive pathways in terms of both recovery speed and maximum level of recovery. Using acute-phase recovery information, we developed a machine learning-based prediction model capable of identifying the recovery pattern to which a patient is likely to transition in the post-acute phase. The inputs utilized for the model were demographics, clinical characteristics and acute-phase functional outcomes. In this, the acute phase is further divided into the following time points: 1) before receiving any form of intervention (using only one week of information); and 2) before discharge from the hospital (using both one and three weeks of information). All prediction models are evaluated through five-fold cross validation.

Extreme Gradient Boosting (XGBoost) is selected as the machine learning classifier, and the results of predicting five recovery patterns using exclusively one-week information are accuracy: 0.633±0.133, AUROC: 0.770±0.132, and f1-score: 0.614±0.134. The results of using one and three weeks of information are accuracy: 0.722±0.115, AUROC: 0.826±0.121, and f1-score: 0.706±0.116. Considering that multi-class classification is acknowledged as a challenging endeavor, the task of directly classifying a patient into five distinct categories may prove excessively difficult, particularly considering the relatively small size of the dataset. Hence, we opt for a two-step hierarchical approach, first grouping MID 1, MID 2, and MID 3 clusters into a single MID cluster. This grouping allows us to classify the data into HIGH, MID, and LOW clusters. Second, another predictive model was developed to classify the three MID clusters. This two-step approach led to a significant improvement in the prediction results, where the results of using one and three weeks of information findings in the first-step model are, accuracy: 0.852±0.074, AUROC: 0.889±0.076, and f1-score: 0.852±0.069, and the results of the second-step model were, accuracy: 0.786±0.071, AUROC: 0.838±0.121, and f1-score: 0.757±0.089. The results indicate that employing solely the acute-phase information enables a reasonable prediction of the patient's recovery progress in the post-acute phase. Thus, by predicting the expected recovery patterns, patient populations that are vulnerable to recovery can be identified early in the acute-phase and be targeted more effectively.

IV. DISCUSSION

A. RECOVERY PATTERN COMPARISON

Among the five groups identified in Section III-B, several interesting patterns are observed, which could be largely attributed to the differences in recovery speed, recovery levels in different functional domains of the outcome measurement, and recovery to the premorbid functional status.

1) RECOVERY SPEED DIFFERENCE

Among the five patient groups, HIGH and MID 1 are the two groups that exhibit the best maximum recovery levels. However, the recovery rate varies between the two groups, resulting in different recovery levels at different time points. While the HIGH group consistently shows the best performance across all functional outcomes at all time points, MID 1 group exhibits poor functional outcomes during the acute recovery period (at POD one week and three weeks), but demonstrates rapid recovery afterward up to three months. Specifically, the average outcomes of the MID 1 group are similar to those of the LOW group at one week and similar to those of the MID 2 and 3 groups at three weeks, which is then followed by a rapid recovery.

Since the difference in the acute-phase recovery period may have contributed to the patient's status before receiving surgery or rehabilitation, we further analyze walking ability measured before hip fracture (premorbid walking ability) and cognitive function measured before receiving rehabilitation (one week K- MMSE). Although there is no statistically significant difference in premorbid walking ability between the HIGH and MID 1 groups (KOVAL $p = 0.122$, FAC $p = 0.070$), cognitive function (K-MMSE) measured at 1 week varies substantially between the two groups ($p < 0.001$). Patients with hip fractures and lower baseline cognitive function (such as dementia) tend to exhibit higher

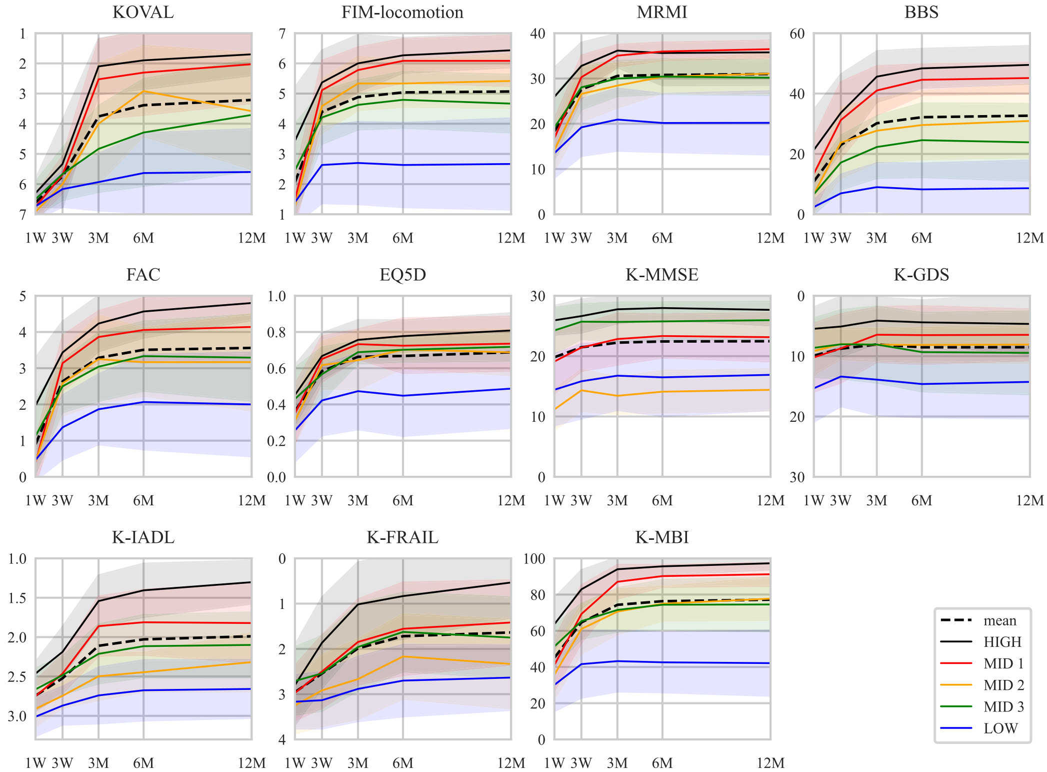


FIGURE 4. Recovery trajectories of each outcome measure for all patients (mean: dashed line) and for each of the 5 patient groups (HIGH: black, MID 1: red, MID 2: orange, MID 3: green, LOW: blue). Trajectory lines represent the mean and shaded areas represent ±1 standard deviation from the mean.

rates of delirium [50]. Hence, the lower baseline cognitive function of patients at MID 1 may have led to an initial underestimation of function due to challenges in functional assessment, potentially influenced by factors like delirium. Thus, cognitive function could play a key role in the observed discrepancies in recovery speeds among the groups.

2) RECOVERY LEVEL DIFFERENCE BASED ON FUNCTIONAL DOMAIN

MID 2 and MID 3 are the groups with moderate recovery that exhibit inconsistency in the ordering between the two groups and depend on the functional domain of the outcome measure. MID 2 show better recovery compared to the premorbid state but has the highest average age and lowest BMI, and thus is at a higher risk of frailty. In addition, the cognitive outcome (K-MMSE) score is the lowest among all groups. As a result, the K-IADL, which is a comprehensive outcome of daily living activities, is also low. In contrast, the MID 3 group has high premorbid walking ability but slower overall recovery. Compared with MID 2, MID 3 patients are less frail ($p = 0.030$) and exhibit higher cognitive function ($p < 0.001$), which is advantageous for K-IADL recovery. However, the level of walking ability and balance outcomes are similar to or worse than those of the MID 2 group. Thus, it can be inferred that recovery from hip fracture varies across different domains, with large discrepancies observed between the recovery of cognitive outcomes and walking abilities.

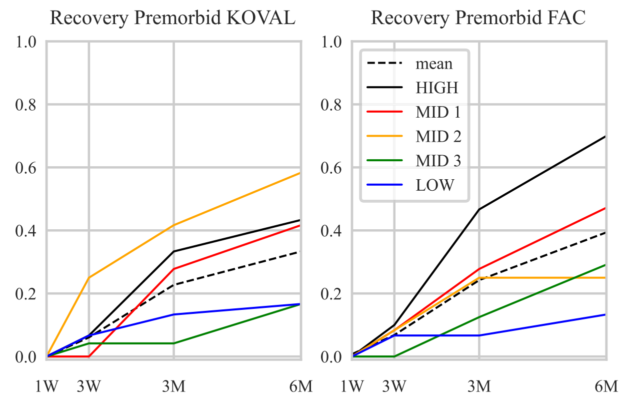


FIGURE 5. Ratio of patients recovering to their premorbid walking ability (KOVAL, FAC) within six months.

3) RECOVERY OF PREMORBID WALKING ABILITY

One of the primary goal of rehabilitation for hip fracture patients is the rapid recovery of their premorbid walking ability. Among the various functional outcomes, KOVAL and FAC, which are commonly considered key measures for assessing walking ability, are used to assess recovery to a premorbid status.

The proportion of patients recovering their premorbid walking ability within six months for all patients and each patient group is shown in Fig. 5. Among all patients, 22.7 and

33.3% recovered their premorbid walking ability at three and six months, respectively based on the KOVAL scale. The corresponding recovery rates based on the FAC scale are 24.3 and 39.4%. It is observed that KOVAL evaluates a lower recovery rate than FAC. This is expected, as KOVAL evaluates the dependency of walking outdoors and indoors separately, whereas FAC does not differentiate between outdoor and indoor activities. Patients with hip fractures tend to have psychosocial barriers and physical limitations in outdoor activities, which may have led to a low assessment of outdoor dependency and, hence, poor recovery in terms of KOVAL.

By comparing recovery rates across groups, it can be seen that as the recovery level of a group increases (from LOW to HIGH), the percentage of recovery for FAC also increases, and this trend is consistently observed at both three and six months. In contrast, the KOVAL recovery do not match the order of the recovery levels of the groups, with the MID 2 group exhibiting a significantly higher recovery rate. This can be attributed to the fact that the MID 2 group started with the worst premorbid KOVAL score, which means that recovery of premorbid walking ability is relatively easier and quicker. It is observed that 25% of the patients with MID 2 had already recovered their premorbid walking ability at three weeks. As patients with MID 2 shows the lowest performance in terms of cognitive function (K-MMSE), this may have led to an initial underestimation of premorbid walking ability owing to poor functional assessment. Thus, it can be inferred that cognitive function plays a key role in determining the possibility of recovery to a premorbid status.

B. VALIDITY OF MULTIPLE FUNCTIONAL OUTCOME MEASURES

In this section, we aim to further validate why multiple outcome measures should be concurrently utilized to assess patients' recovery levels using feature importance and cluster analysis examples.

1) IMPORTANT FEATURES OF PREDICTION MODELS

In order to interpret the prediction results of machine learning models, it is possible to derive feature importance which is a measure of the relative importance of input variables that represents which variables have the greatest impact on the prediction results. Feature importance provides insights regarding the influence among input variables and how the model performs predictions, thus enabling interpretation of the machine learning model. By deriving the important features for each of the two-step prediction models, it is possible to understand important functional outcome measures used at each step.

The first-step model, which classifies HIGH, MID, and LOW groups use K-MBI and FAC at POD three weeks as the most important features, followed by BBS, K-MMSE, K-MBI, K-IADL, and FAC measured at one week (Fig. 6 top panel). By categorizing the features by measurement period (one and three weeks) and functional domain (clinical characteristics, walking ability, activities of daily living, cognition, emotion, quality of life, and frailty), the walking ability outcomes at one and three weeks contribute the most, followed by the activity of daily living outcomes at three weeks and one week.

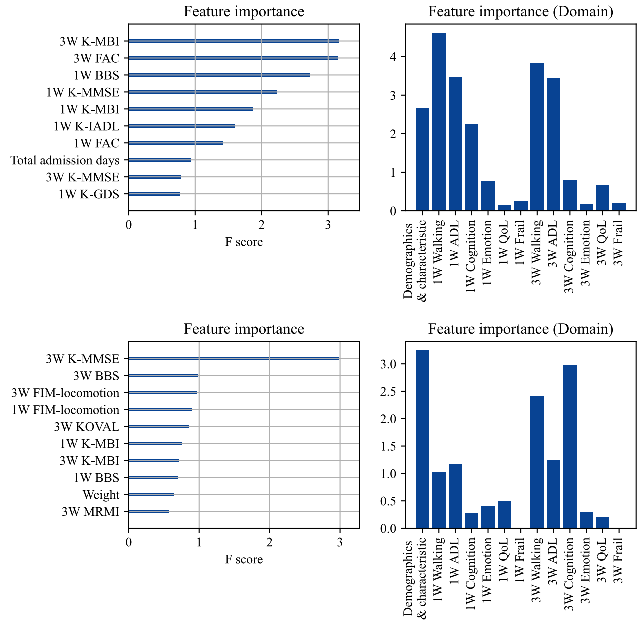


FIGURE 6. Feature importance of cluster prediction models for first step model (top) and second step model (bottom).

The second model, which classifies the three MID groups, relies mainly on three weeks of K-MMSE value, followed by BBS, FIM-locomotion measured at three weeks, and FIM-locomotion measured at one week (Fig. 6 bottom panel). Walking ability and cognition at three weeks and patients' clinical characteristics are important in classifying the three MID groups, followed by walking ability and ADL at one week, and mental status measured at three weeks. Measurement of mental status at one week do not contribute significantly to the model. Additionally, frailty outcomes at one and three weeks do not contribute to the classification of the three groups.

Comparing the results of the two models, the important features used in the two steps differ in terms of when and what they measure. Walking ability and ADLs are important for predicting the overall level of recovery (HIGH/MID/LOW), while mentality and clinical characteristics are important for predicting the recovery of intermediate patients (approximately 55% of all patients). These results demonstrate the importance of assessing a patient's level of recovery and designing rehabilitation programs based on a combination of indicators from multiple domains. A single measurement is not sufficient to accurately predict a patient's recovery level.

2) USING SINGLE VS. MULTIPLE OUTCOME MEASURES

Additionally, we discuss the problems that could arise when assessing recovery in hip fracture patients using a

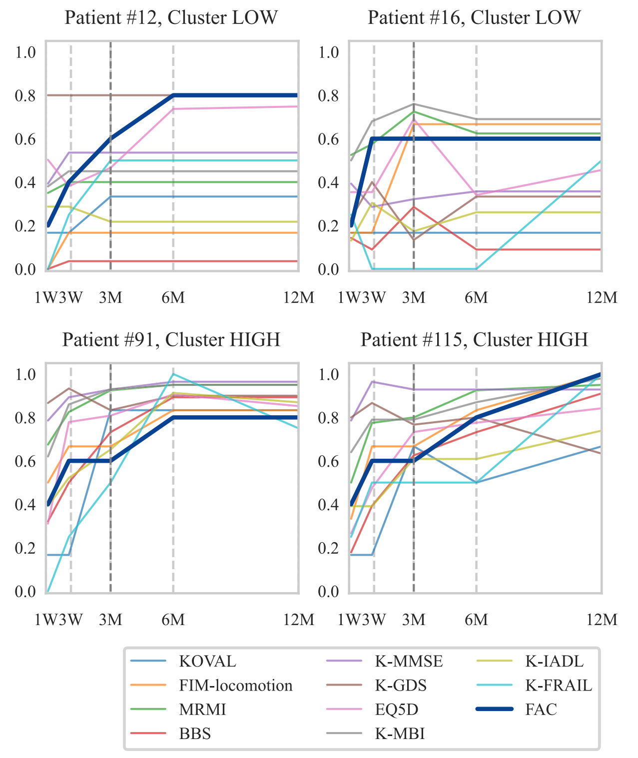


FIGURE 7. Recovery trajectory of four patients for all outcomes. FAC score of all patients evaluates same level at three months, but patients 12 and 16 are classified as low group and patents 91 and 115 classified as high group.

single outcome measure. For illustrative purposes, we use two key performance measures to assess walking ability, including KOVAL and FAC. Similar to the post-acute recovery prediction model, a model that predicts the level of recovery of KOVAL at three months using only basic clinical characteristics and the KOVAL score in the acute phase is developed. A corresponding model for FAC is developed that utilize basic clinical characteristics and acute-phase FAC scores to predict the recovery level of FAC within three months. The model performance is as follows: the KOVAL prediction model exhibit accuracy: 0.753±0.091, AUROC: 0.858±0.111, and f1-score: 0.743±0.083. FAC prediction model exhibit accuracy: 0.748±0.073, AUROC: 0.880±0.122, and f1-score: 0.738±0.069. Despite using information from the acute phase to predict the future recovery levels of the same outcome measures, the prediction accuracy is lower than that of the previous recovery pattern prediction model. This result demonstrates that patients' recovery cannot be explained by their walking ability alone. Hence, it can be inferred that a patient's walking ability recovery is not only dependent on walking ability but also on other functional outcomes.

Recovery trajectories of the four patients are shown in Fig. 7 which highlights the limitation of using single outcome measures to assess recovery. All four patients recover to the same FAC value of three at the post-surgery evaluation at three months, indicating that patients can walk unaided on flat ground. However, when all outcome measures are used to perform clustering, patients 12 and 16 are classified into the LOW group, whereas patients 91 and 115 are classified into the HIGH group. Most outcome measures tend to recover poorly in the LOW group, while the HIGH group show satisfactory recovery in most outcomes. This suggests that we need to consider multiple functional aspects when analyzing a patient's functional recovery.

V. CONCLUSION

In this study, we identify five distinct patterns of recovery after hip fracture surgery using 11 different outcome measures encompassing multiple functional domains. The derived recovery patterns vary in maximum recovery levels and recovery speed, and show different recovery levels depending on the functional outcome of interest. The results clearly demonstrate the importance of utilizing multiple outcome measures concurrently to assess a patient's recovery level, as a single measure alone is not sufficient to capture holistic information on the recovery progress. We further provide predictions on how recovery will progress in the post-acute phase using only acute-phase information, which enables early identification of patient groups with a poor prognosis of recovery.

We provide an end-to-end pipeline for clustering analysis, starting from missing data imputation using deep learning architectures to the early prediction of the identified clusters using machine learning methods. The proposed methodology can be extended to other multivariate time-series settings beyond hip fractures or the medical domain because the proposed methods (M-RNN and MM-DTW) are specialized algorithms that are generally well-suited for any multivariate time-series datasets. Furthermore, unlike statistical models commonly utilized in earlier studies, this approach can accommodate a broader range of measures and patterns as it relies less on distributional assumptions and parameters. This flexibility enables the identification of diverse patterns using multiple variables even in small-scale longitudinal studies with frequent missing data.

However, this study has several limitations. First, the dataset size was limited because patients had to be recruited to keep track of and collect multiple outcome measures continuously for a year. The second limitation arises since the data was collected from patients enrolled in a high-intensity rehabilitation program called the FIRM; hence, all recovery pattern groups tend to improve or maintain recovery over time, and we do not observe a decline in recovery. However, given that the majority of hip fracture patients in clinical practice undergo rehabilitation, the findings of this study may help identify patients with relatively unfavorable recovery despite rehabilitation, which may provide room for further intervention. Finally, the prediction model can only predict the average recovery patterns, rather than the exact recovery level for each patient. However, predicting the exact recovery level for each outcome may be less significant because providing a holistic view of recovery by considering the recovery of multiple outcomes simultaneously would be more informative. That said, being able to predict the general trend of a patient's recovery (e.g., walking ability scores are expected to recover well while activities of daily living might not recover well) could be more informative to both the patient and clinicians rather than predicting to what score each functional outcome score will recover to.

By studying the recovery process from multiple functional domains, the proposed method can serve as a clinical decision support system for designing personalized rehabilitation interventions. First, clinicians can advise patients and caregivers about the expected patterns of recovery, allowing them to make additional preparations, if necessary. Second, by providing advice on expected recovery patterns, patient populations that are vulnerable to recovery can be identified and rehabilitation services and other therapeutic interventions can be targeted more effectively. Finally, the overall clinical understanding of hip fracture recovery is expected to improve by identifying the important factors that contribute to the differences in recovery between groups.

REFERENCES

[1] B. Abrahamsen, T. van Staa, R. Ariely, M. Olson, and C. Cooper, "Excess mortality following hip fracture: A systematic epidemiological review," *Osteoporosis Int.*, vol. 20, no. 10, pp. 1633–1650, Oct. 2009.

[2] N. Veronese and S. Maggi, "Epidemiology and social costs of hip fracture," *Injury*, vol. 49, no. 8, pp. 1458–1460, Aug. 2018.

[3] I. Quevedo, J. C. Ormeño, B. Weissglas, and C. Opazo, "Epidemiology and direct medical cost of osteoporotic hip fracture in Chile," *J. Osteoporosis*, vol. 2020, pp. 1–7, Mar. 2020.

[4] F. Leung, T. W. Lau, K. Kwan, S. P. Chow, and A. W. C. Kung, "Does timing of surgery matter in fragility hip fractures?" *Osteoporosis Int.*, vol. 21, no. S4, pp. 529–534, Dec. 2010.

[5] S. M. Dyer, M. Crotty, N. Fairhall, J. Magaziner, L. A. Beaupre, I. D. Cameron, and C. Sherrington, "A critical review of the long-term disability outcomes following hip fracture," *BMC Geriatrics*, vol. 16, pp. 1–18, Dec. 2016.

[6] K. Alexiou, A. Roushias, S. Varitimidis, and K. Malizos, "Quality of life and psychological consequences in elderly patients after a hip fracture: A review," *Clin. Interventions Aging*, vol. 13, pp. 143–150, Jan. 2018.

[7] J. Magaziner, W. Hawkes, J. R. Hebel, S. I. Zimmerman, K. M. Fox, M. Dolan, G. Felsenthal, and J. Kenzora, "Recovery from hip fracture in eight areas of function," *Journals Gerontol. Ser. A, Biol. Sci. Med. Sci.*, vol. 55, no. 9, pp. 498–507, Sep. 2000.

[8] T.-L. Tam, K.-K. Tsang, and K.-B. Lee, "Development of a prognostic model to predict post-operative mobility of patients with fragility hip fractures: A retrospective cohort study," *Int. J. Orthopaedic Trauma Nursing*, vol. 38, Aug. 2020, Art. no. 100770.

[9] D. Lee, J. Y. Jo, J. S. Jung, and S. J. Kim, "Prognostic factors predicting early recovery of pre-fracture functional mobility in elderly patients with hip fracture," *Ann. Rehabil. Med.*, vol. 38, no. 6, pp. 827–835, 2014.

[10] P. Wang, Y. Li, H. Zhuang, H. Yu, S. Cai, H. Xu, Z. Chen, J. Lin, and X. Yao, "Anti-osteoporosis medications associated with decreased mortality after hip fracture," *Orthopaedic Surg.*, vol. 11, no. 5, pp. 777–783, Oct. 2019.

[11] M. Di Monaco, C. Castiglioni, E. De Toma, L. Gardin, S. Giordano, and R. Tappero, "Handgrip strength is an independent predictor of functional outcome in hip-fracture women: A prospective study with 6-month follow-up," *Medicine*, vol. 94, no. 6, p. e542, 2015.

[12] L. Feng, S. C. Scherer, B. Y. Tan, G. Chan, N. P. Fong, and T. P. Ng, "Comorbid cognitive impairment and depression is a significant predictor of poor outcomes in hip fracture rehabilitation," *Int. Psychogeriatrics*, vol. 22, no. 2, pp. 246–253, Mar. 2010.

[13] J. González-Zabaleta, S. Pita-Fernandez, T. Seoane-Pillado, B. López-Calviño, and J. L. Gonzalez-Zabaleta, "Comorbidity as a predictor of mortality and mobility after hip fracture," *Geriatrics Gerontol. Int.*, vol. 16, no. 5, pp. 561–569, May 2016.

[14] G. Bellelli, M. Noale, F. Guerini, R. Turco, S. Maggi, G. Crepaldi, and M. Trabucchi, "A prognostic model predicting recovery of walking independence of elderly patients after hip-fracture surgery. An experiment in a rehabilitation unit in northern Italy," *Osteoporosis Int.*, vol. 23, no. 8, pp. 2189–2200, Aug. 2012.

[15] D. L. Orwig, D. S. Abraham, M. C. Hochberg, A. Gruber-Baldini, J. M. Guralnik, A. R. Cappola, J. Golden, G. E. Hicks, R. R. Miller, B. Resnick, M. Shardell, R. S. Sterling, R. Bajracharya, and J. Magaziner, "Sex differences in recovery across multiple domains among older adults with hip fracture," *Journals Gerontol., Ser. A*, vol. 77, no. 7, pp. 1463–1471, Jul. 2022.

[16] A. M. Rathbun, M. Shardell, D. Orwig, J. R. Hebel, G. E. Hicks, T. Beck, M. C. Hochberg, and J. Magaziner, "Differences in the trajectory of bone mineral density change measured at the total hip and femoral neck between men and women following hip fracture," *Arch. Osteoporosis*, vol. 11, no. 1, pp. 1–9, Dec. 2016.

[17] Y. Huang, D. Orwig, H. Hayssen, W. Lu, A. L. Gruber-Baldini, N. Chiles Shaffer, J. Magaziner, and J. M. Guralnik, "Longitudinal characteristics of physical frailty and its components in men and women post hip fracture," *J. Amer. Geriatrics Soc.*, vol. 72, no. 1, pp. 170–180, Jan. 2024.

[18] N. K. Latham, D. U. Jette, R. L. Warren, and C. Wirtalla, "Pattern of functional change during rehabilitation of patients with hip fracture," *Arch. Phys. Med. Rehabil.*, vol. 87, no. 1, pp. 111–116, Jan. 2006.

[19] F. J. Ortiz-Alonso, M. Vidan-Astiz, M. Alonso-Armesto, M. Toledano-Iglesias, L. Alvarez-Nebreda, F. Branas-Baztan, and J. A. Serra-Rexach, "The pattern of recovery of ambulation after hip fracture differs with age in elderly patients," *Journals Gerontol. Ser. A, Biol. Sci. Med. Sci.*, vol. 67, no. 6, pp. 690–697, Jun. 2012.

[20] L. de Munter, C. L. P. van de Ree, O. P. van der Jagt, T. Gosens, H. A. A. M. Maas, and M. A. C. de Jongh, "Trajectories and prognostic factors for recovery after hip fracture: A longitudinal cohort study," *Int. Orthopaedics*, vol. 46, no. 12, pp. 2913–2926, Dec. 2022.

[21] Y. Young, K. Xiong, and R. M. Pruzek, "Longitudinal functional recovery after postacute rehabilitation in older hip fracture patients: The role of cognitive impairment and implications for long-term care," *J. Amer. Med. Directors Assoc.*, vol. 12, no. 6, pp. 431–438, Jul. 2011.

[22] C. A. Jones, G. S. Jhangri, D. H. Feeny, and L. A. Beaupre, "Cognitive status at hospital admission: Postoperative trajectory of functional recovery for hip fracture," *Journals Gerontol. Ser. A, Biol. Sci. Med. Sci.*, vol. 72, no. 1, pp. 61–67, Jan. 2017.

[23] Y.-J. Wang, P.-H. Lin, L.-N. Peng, M.-J. Kao, S. Y. Lin, K.-H. Huang, Y.-F. Chen, and M.-R. Liang, "Effectiveness of post-acute rehabilitation on functional outcome after hip fracture," *Taiwan J. Phys. Med. Rehabil.*, vol. 47, no. 1, pp. 11–20, 2019.

[24] J. J. Aarden, M. van der Esch, R. H. H. Engelbert, M. van der Schaaf, S. E. de Rooij, and B. M. Buurman, "Hip fractures in older patients: Trajectories of disability after surgery," *J. Nutrition, Health Aging*, vol. 21, no. 7, pp. 837–842, Jul. 2017.

[25] M.-Y. Tseng, Y.-I.-L. Shyu, and J. Liang, "Functional recovery of older hip-fracture patients after interdisciplinary intervention follows three distinct trajectories," *Gerontologist*, vol. 52, no. 6, pp. 833–842, Dec. 2012.

[26] S. J. E. Beishuizen, B. C. van Munster, A. de Jonghe, A. Abu-Hanna, B. M. Buurman, and S. E. de Rooij, "Distinct cognitive trajectories in the first year after hip fracture," *J. Amer. Geriatrics Soc.*, vol. 65, no. 5, pp. 1034–1042, May 2017.

[27] H. P. Wang, J. Liang, L. M. Kuo, C. Y. Chen, and Y.-I.-L. Shyu, "Trajectories of nutritional status and cognitive impairment among older Taiwanese with hip fracture," *J. nutrition, health aging*, vol. 21, no. 1, pp. 38–45, Jan. 2017.

[28] H.-Y. Liu, C.-T. Yang, M.-Y. Tseng, C.-Y. Chen, C.-C. Wu, H.-S. Cheng, Y.-E. Lin, and Y.-I.-L. Shyu, "Trajectories in postoperative recovery of elderly hip-fracture patients at risk for depression: A follow-up study," *Rehabil. Psychol.*, vol. 63, no. 3, pp. 438–446, Aug. 2018.

[29] P. Cristancho, E. J. Lenze, M. S. Avidan, and K. S. Rawson, "Trajectories of depressive symptoms after hip fracture," *Psychol. Med.*, vol. 46, no. 7, pp. 1413–1425, May 2016.

[30] S. Dakhil, I. Saltvedt, J. S. Benth, P. Thingstad, L. O. Watne, T. B. Wyller, J. L. Helbostad, F. Frihagen, L. G. Johnsen, and K. Taraldsen, "Longitudinal trajectories of functional recovery after hip fracture," *PLoS ONE*, vol. 18, no. 3, Mar. 2023, Art. no. e0283551.

[31] C. Colón-Emeric, H. E. Whitson, C. F. Pieper, R. Sloane, D. Orwig, K. M. Huffman, J. P. Bettger, D. Parker, D. M. Crabtree, A. Gruber-Baldini, and J. Magaziner, "Resiliency groups following hip fracture in older adults," *J. Amer. Geriatrics Soc.*, vol. 67, no. 12, pp. 2519–2527, Dec. 2019.
[32] S. Y. Lee, J. Beom, B. R. Kim, S.-K. Lim, and J.-Y. Lim, "Comparative effectiveness of fragility fracture integrated rehabilitation management for elderly individuals after hip fracture surgery: A study protocol for a multicenter randomized controlled trial," *Medicine*, vol. 97, no. 20, 2018, Art. no. e10763.
[33] J. H. Choi, B. R. Kim, K. W. Nam, S. Y. Lee, J. Beom, S. Y. Lee, M. J. Suh, and J.-Y. Lim, "Effectiveness of a home-based fragility fracture integrated rehabilitation management (FIRM) program in patients surgically treated for hip fractures," *J. Clin. Med.*, vol. 10, no. 1, p. 18, Dec. 2020.
[34] K. J. Koval, M. L. Skovron, G. B. Aharonoff, S. E. Meadows, and J. D. Zuckerman, "Ambulatory ability after hip fracture: A prospective study in geriatric patients," *Clin. Orthopaedics Rel. Res.*, no. 310, pp. 150–159, 1995.
[35] D. Kidd, G. Stewart, J. Baldry, J. Johnson, D. Rossiter, A. Petruckevitch, and A. J. Thompson, "The functional independence measure: A comparative validity and reliability study," *Disability Rehabil.*, vol. 17, no. 1, pp. 10–14, Jan. 1995.
[36] F. M. Collen, D. T. Wade, and C. M. Bradshaw, "Mobility after stroke: Reliability of measures of impairment and disability," *Int. Disability Stud.*, vol. 12, no. 1, pp. 6–9, Jan. 1990.
[37] S. Lennon and L. Johnson, "The modified rivermead mobility index: Validity and reliability," *Disability Rehabil.*, vol. 22, no. 18, pp. 833–839, Jan. 2000.
[38] K. Berg, S. Wood-Dauphinee, and J. I. Williams, "The balance scale: Reliability assessment with elderly residents and patients with an acute stroke," *Scandin. J. Rehabil. Med.*, vol. 27, no. 1, pp. 27–36, 1995.
[39] H. Y. Jung, B. K. Park, H. S. Shin, Y. K. Kang, S. B. Pyun, N. J. Paik, S. H. Kim, T. H. Kim, and T. R. Han, "Development of the Korean version of modified Barthel index (K-MBI): Multi-center study for subjects with stroke," *J. Korean Acad. Rehabil. Med.*, vol. 31, no. 3, pp. 283–297, 2007.
[40] C. W. Won, K. Y. Yang, Y. G. Rho, S. Y. Kim, E. J. Lee, J. L. Yoon, K. H. Cho, H. C. Shin, B. R. Cho, and J. R. Oh, "The development of Korean activities of daily living (K-ADL) and Korean instrumental activities of daily living (K-IADL) scale," *J. Korean Geriatrics Soc.*, vol. 6, no. 2, pp. 107–120, 2002.
[41] Y. Kang, N. A. Duk, and S. Hahn, "A validity study on the Korean mini-mental state examination (K-MMSE) in dementia patients," *J. Korean Neurological Assoc.*, vol. 15, no. 2, pp. 300–308, 1997.
[42] I.-K. Jung, D.-I. Kwak, S.-H. Joe, and H.-S. Lee, "A study of standardization of Korean form of geriatric depression scale (KGDS)," *J. Korean Geriatric Psychiatry*, vol. 1, no. 1, pp. 61–72, 1997.
[43] The EuroQol Group, "EuroQol—A new facility for the measurement of health-related quality of life," *Health Policy*, vol. 16, no. 3, pp. 199–208, 1990.
[44] H.-W. Jung, H.-J. Yoo, S.-Y. Park, S.-W. Kim, J.-Y. Choi, S.-J. Yoon, C.-H. Kim, and K.-I. Kim, "The Korean version of the FRAIL scale: Clinical feasibility and validity of assessing the frailty status of Korean elderly," *Korean J. Internal Med.*, vol. 31, no. 3, pp. 594–600, May 2016.
[45] J. Yoon, W. R. Zame, and M. van der Schaar, "Estimating missing data in temporal data streams using multi-directional recurrent neural networks," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 5, pp. 1477–1490, May 2019.
[46] M. Shokoohi-Yekta, B. Hu, H. Jin, J. Wang, and E. Keogh, "Generalizing DTW to the multi-dimensional case requires an adaptive approach," *Data Mining Knowl. Discovery*, vol. 31, no. 1, pp. 1–31, Jan. 2017.
[47] V. Solo, "Pearson distance is not a distance," 2019, *arXiv:1908.06029*.
[48] J. K. Freburger, G. M. Holmes, and L. E. Ku, "Postacute rehabilitation care for hip fracture: Who gets the most care?" *J. Amer. Geriatrics Soc.*, vol. 60, no. 10, pp. 1929–1935, Oct. 2012.
[49] I. R. White, P. Royston, and A. M. Wood, "Multiple imputation using chained equations: Issues and guidance for practice," *Statist. Med.*, vol. 30, no. 4, pp. 377–399, Feb. 2011.
[50] J. Wu, Y. Yin, M. Jin, and B. Li, "The risk factors for postoperative delirium in adult patients after hip fracture surgery: A systematic review and meta-analysis," *Int. J. Geriatric Psychiatry*, vol. 36, no. 1, pp. 3–14, Jan. 2021.

CHANYOUNG PARK received the B.S. degree in electronic and electrical engineering from Hongik University, Seoul, South Korea, in 2022. He is currently pursuing the M.S. degree with the School of Industrial and Management Engineering, Korea University, Seoul.

HONGBUM KIM received the B.S. degree from the Department of Industrial Management Engineering, Hankuk University of Foreign Studies, Seoul, South Korea, in 2021. He is currently pursuing the M.S. degree with the School of Industrial and Management Engineering, Korea University, Seoul.

His current research interest includes clinical decision support systems using artificial intelligence.

JUNGWON SUH received the B.S. degree in biochemical engineering from Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2016, and the M.D. degree from Korea University College of Medicine, Seoul, South Korea, in 2020.

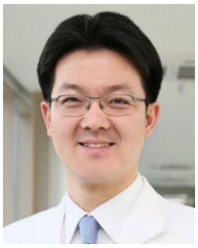
In 2021, he was an Intern with the Department of Physical Medicine and Rehabilitation, Korea University Anam Hospital. He has been in residency training, since 2021.

JINHEE KO was born in South Korea, in 1982. She received the B.S. degree from Jeju Halla University, Jeju, South Korea, in 2017, and the M.S. degree from Jeju National University, Jeju, in 2020.

From 2016 to 2020, she was a Clinical Research Coordinator and an Occupational Therapist with Jeju National University Hospital, Jeju. Since 2021, she has been a CRC and Occupational Therapist with the Department of Physical Medicine and Rehabilitation, Korea University Anam Hospital. Her research interests include cardiac and geriatric rehabilitation.

JUN HWAN CHOI was born in South Korea, in 1981. He received the B.S. and M.S. degrees from Jeju National University College of Medicine, Jeju, South Korea, in 2007 and 2014, respectively.

From 2011 to 2014, he was a Resident Doctor with the Department of Rehabilitation Medicine, Jeju National University Hospital, Jeju. Since 2022, he has been an Assistant Professor with the Department of Rehabilitation Medicine, Jeju National University College of Medicine. He has authored more than ten SCI(E) articles. His main research interests include geriatric and orthopedic rehabilitation through clinical outcome measures and gait analysis. He is also interested in cardiac rehabilitation through digital healthcare.



SANG YOON LEE was born in South Korea, in 1980. He received the B.S., M.S., and Ph.D. degrees from the College of Medicine, Seoul National University, Seoul, South Korea, in 2005, 2010, and 2016, respectively.

From 2005 to 2010, he was a Resident Doctor with the Department of Rehabilitation Medicine, Seoul National University Hospital, Seoul. Since 2017, he has been a Clinical Professor with the Department of Rehabilitation Medicine, Seoul National University Boramae Medical Center. He is the author of more than 60 SCI(E) articles. His majority is the musculoskeletal rehabilitation, including interventional physiatry and sports medicine. His clinical and research interests include sarcopenia, degenerative spinal disorder, and regenerative medicine with mesenchymal stem cells.

JAEWON BEOM received the M.D. degree from Seoul National University College of Medicine, in 2004, and the Ph.D. degree from the Department of Biomedical Engineering, Seoul National University, in 2016.

He is currently a Clinical Associate Professor with the Department of Rehabilitation Medicine, Seoul National University Bundang Hospital. His main research interests include musculoskeletal and geriatric rehabilitation, critical care rehabilitation, and digital rehabilitation.

JAE-YOUNG LIM was born in South Korea, in 1969. He received the B.S., M.S., and Ph.D. degrees from the College of Medicine, Seoul National University, Seoul, South Korea, in 1994, 1998, and 2006, respectively.

From 1995 to 1999, he was a Resident Doctor with the Department of Rehabilitation Medicine, Seoul National University Hospital, Seoul. Since 2003, he has been a Professor with the Department of Rehabilitation Medicine, Seoul National University Bundang Hospital, Seoul National University, College of Medicine. As a specialist in musculoskeletal and geriatric rehabilitation, his major research interest is to investigate the key mechanisms of functional restoration and adaptation after injury, chronic diseases, or aging in humans. He has contributed seminal research findings on muscle quality to assess frailty and sarcopenia as well as conducted modality-specific intervention studies for the old population. He has published more than 200 peer-reviewed articles in reputed international journals and written dozens of book chapters. His research interests include the development of integrated rehabilitation methods for fall prevention and fragility fracture care in older adults. He is the Editor-in-Chief of *Annals of Geriatric Medicine and Research* and an Associate Editor of *PM&R*.

BO RYUN KIM was born in South Korea, in 1979. She received the B.S. and M.S. degrees from the University of Ulsan College of Medicine, Seoul, South Korea, in 2004 and 2010, respectively, and the Ph.D. degree from Jeju National University College of Medicine, Jeju, South Korea, in 2014.

From 2006 to 2010, she was a Resident Doctor with the Department of Rehabilitation Medicine, Asan Medical Center, Seoul. Since 2020, she has been an Associate Professor with the Department of Physical Medicine and Rehabilitation, Korea University Anam Hospital. She is the author of more than 60 SCI(E) articles. Her majority is geriatric, cardiac, and pediatric rehabilitation. Her clinical and research interests include cardiac and geriatric rehabilitation fields connected with wearable technology, including motion tracking and digital healthcare.

HYO KYUNG LEE (Member, IEEE) received the B.S. degree from the Department of Information and Industrial Engineering, Yonsei University, Seoul, South Korea, in 2012, the M.S. degree in industrial and systems engineering from Georgia Institute of Technology, Atlanta, GA, USA, in 2013, and the Ph.D. degree from the Department of Industrial and Systems Engineering, University of Wisconsin–Madison, Madison, WI, USA, in 2019.

She was an Assistant Professor with the University of Pittsburgh, from 2019 to 2021. She is currently an Assistant Professor with the School of Industrial and Management Engineering, Korea University, Seoul. Her current research interests include healthcare decision-making and prediction models for precision medicine.