RJNMSR
Review Journal of Neurological & Medical Sciences Review

Volume 4, Issue 2, 2026

ADVANCE APPLICATION OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING IN DOMAIN OF ORTHOPEDIC PHYSIOTHERAPY PRACTICE AND REHABILITATION SCIENCE

Dr Ayesha Jamal[1], Dr Aiman Shah[2], Dr Roshneck Hameed[*3], Dr Amna Jamil[4], Dr Syeda Uswa E Maryam[5]

[1]*Mirpur University of Science and Technology, Mirpur: Kashmir*
[2]*Lecturer At Gulrang Institute Tana, Malakand*
[*3]*Lecturer At Iqra University Check Shahzad Campus Islamabad*
[4]*MS. OMPT From Riphah International University Islamabad*
[5]*Workplace Eshifa, Shifa International Hospital, Islamabad*

[1]ayesha.dpt@must.edu.pk , [2]aiman.shah1@yahoo.com, [*3]roshneckhameedmsptf21@gmail.com, [4]amnajamilrajpoot@gmail.com, [5]syedauswaemaryam@gmail.com

Corresponding Author: *
Roshneck Hameed

DOI: https://doi.org/10.5281/zenodo.18617004

Received	Accepted	Published
13 December 2025	28 January 2026	12 February 2026

ABSTRACT

The integration of Artificial Intelligence (AI) and Machine Learning (ML) into orthopedic physiotherapy is transforming traditional rehabilitation paradigms by enabling personalized, data-driven, and scalable care. This paper provides a comprehensive synthesis of recent advances in AI/ML applications, including computer vision for movement analysis, wearable sensor fusion for real-time biofeedback, predictive analytics for recovery trajectory modeling, and adaptive robotic systems powered by reinforcement learning, within the domain of musculoskeletal rehabilitation. Drawing on evidence from 2019 to 2025, we critically evaluate the clinical efficacy, technical robustness, and implementation challenges of these technologies across diverse settings, with particular attention to their potential to enhance functional outcomes, patient adherence, and therapist decision-making in conditions such as osteoarthritis, post-total joint arthroplasty, and rotator cuff injuries. Despite promising proof-of-concept studies, significant gaps remain in clinical validation, algorithmic transparency, and equitable deployment, especially in low-resource and non-Western contexts. We identify key barriers, including limited therapist-AI collaboration frameworks, insufficient focus on patient-centered outcomes, and ethical concerns around data privacy and algorithmic bias. To address these challenges, we propose an integrative, human-centered implementation model grounded in the Consolidated Framework for Implementation Research (CFIR) and aligned with the World Health Organization's global rehabilitation priorities. Our analysis underscores the need for interdisciplinary collaboration, context-adaptive design, and rigorous randomized trials to translate AI innovations into sustainable, equitable, and clinically meaningful tools that augment, not replace, the therapeutic alliance at the heart of physiotherapy practice.

Keywords: *Artificial Intelligence; Machine Learning; Orthopedic Rehabilitation; Physiotherapy; Computer Vision; Wearable Sensors; Predictive Modeling; Human–AI Collaboration; Digital Health; Musculoskeletal Disorders.*

Introduction:

AI in rehabilitation has evolved from rule-based expert systems to data-driven deep learning architectures. Early applications included decision support for exercise prescription (Bassett et al., 2019), but recent work leverages sensor fusion (inertial measurement units and video) and transformer-based models to track complex movements, such as squatting or stair climbing, with millimeter precision (Wang et al., 2024).

In orthopedics, ML models have been used to predict recovery trajectories post-ACL reconstruction using preoperative psychological and biomechanical data (Riley et al., 2022). Similarly, recurrent neural networks (RNNs) analyze longitudinal EHR data to flag patients at risk of delayed recovery, enabling early intervention (Patel et al., 2023). Computer vision systems such as PoseNet and MediaPipe now enable smartphone-based posture and range-of-motion assessment, validated against gold-standard goniometry (Chen et al., 2022; Khan et al., 2023). The landscape of orthopedic rehabilitation is undergoing a transformative shift, driven by the rapid integration of digital health technologies and data-driven decision-making. Traditional physiotherapy, long reliant on subjective clinical judgment, manual goniometry, and standardized exercise protocols, faces mounting challenges in personalizing care for diverse patient populations with varying biomechanical, psychosocial, and functional profiles. In this context, Artificial Intelligence (AI) and Machine Learning (ML) have emerged as powerful enablers of precision rehabilitation, offering the potential to move beyond "one-size-fits-all" approaches toward adaptive, real-time, and patient-centered interventions (Wang et al., 2024; Khan et al., 2023). These technologies harness multimodal data streams, from wearable inertial sensors and depth cameras to electronic health records (EHRs) and patient-reported outcomes, to generate actionable insights that can optimize recovery trajectories, enhance adherence, and reduce healthcare costs.

Recent advances in computer vision and deep learning have enabled contactless, clinic-grade movement analysis with off-the-shelf devices such as smartphones and RGB-D cameras. Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) can now detect subtle deviations in joint kinematics during activities such as squatting, stair climbing, and gait with accuracy rivaling that of expert clinicians (Chen et al., 2022; Zhang et al., 2025). For instance, systems leveraging MediaPipe or OpenPose have been validated against gold-standard motion capture labs for measuring knee flexion angles post-total knee arthroplasty (TKA), achieving mean absolute errors below 5° (Khan et al., 2023). Such tools enable remote, scalable monitoring of exercise form and progression, critical for patients in rural or underserved areas where access to specialized physiotherapy is limited. Beyond assessment, AI is increasingly embedded in therapeutic delivery. Reinforcement learning algorithms power next-generation robotic exoskeletons and resistance training devices that dynamically adjust support or load based on real-time fatigue detection, muscle activation patterns, and performance metrics (Zhang et al., 2025). Similarly, ML models trained on longitudinal datasets can predict individual recovery curves after orthopedic surgeries (e.g., ACL reconstruction or rotator cuff repair), allowing clinicians to identify high-risk patients early and tailor interventions proactively (Riley et al., 2022; Patel et al., 2023). These predictive capabilities represent a paradigm shift from reactive to anticipatory care, aligning with the World Health Organization's call for integrated, person-centered rehabilitation services under its Rehabilitation 2030 initiative (WHO, 2023).

Despite these promising innovations, translating AI/ML tools into routine clinical practice remains fragmented. A recent systematic review found that fewer than 12% of AI-based rehabilitation studies progressed beyond pilot validation to randomized controlled trials (Liu et al., 2023). Many existing applications suffer from "black-box" architectures that lack interpretability, a critical barrier in clinical settings where therapists must understand and trust algorithmic recommendations before acting on them (Amann et al., 2020). Moreover, most models are developed and tested in high-income Western contexts, raising concerns about generalizability in low- and middle-income countries (LMICs), where musculoskeletal disorders are rising due to aging populations, occupational hazards, and limited access to surgical care (WHO, 2023).

Equally pressing is the need to center human factors in AI design. Physiotherapy is fundamentally a relational discipline built on therapeutic alliance, empathy, and shared decision-making. Over-automation risks deskilling clinicians or alienating patients if technology displaces rather than augments human interaction (Greenhalgh et al., 2024). Emerging frameworks advocate for "human-in-the-loop" AI systems that provide decision support while preserving clinician autonomy and patient agency (Levac et al., 2022). Furthermore, ethical considerations, including data privacy in home-based monitoring, algorithmic bias across gender or age groups, and equitable access to AI-enhanced care, remain under addressed in technical publications, creating a translational gap between innovation and responsible implementation. This paper addresses these challenges by synthesizing cutting-edge research on AI and ML applications specifically within orthopedic physiotherapy and musculoskeletal rehabilitation, a domain that constitutes the majority of outpatient rehabilitation yet has received less attention than neuro rehabilitation in the AI literature. We critically evaluate the clinical validity, usability, and equity implications of current technologies, propose a human-centered implementation framework grounded in implementation science, and outline a research agenda to bridge the gap between technical feasibility and real-world impact. By doing so, we aim to guide interdisciplinary teams, clinicians, engineers, policymakers, and patients, in co-designing AI solutions that are not only intelligent but also trustworthy, inclusive, and clinically meaningful. However, systematic reviews reveal that fewer than 15% of AI rehabilitation tools undergo rigorous randomized controlled trials (RCTs); most remain in pilot phases (Liu et al., 2023). Moreover, therapist acceptance is hindered by "black-box" models that lack explains ability, a key barrier to clinical adoption (Amann et al., 2020). Recent efforts toward interpretable AI (e.g., SHAP values and attention maps) show promise for building clinician trust (Zhou et al., 2024).

Crucially, patient-centered outcomes, such as self-efficacy, pain reduction, and return-to-work, are often secondary to technical accuracy metrics, misaligning AI development with rehabilitation goals (Levac et al., 2022).

Research Gap:

While AI applications in neurorehabilitation (e.g., stroke, spinal cord injury) have received substantial attention, their deployment in orthopedic-specific physiotherapy, which constitutes the majority of outpatient rehabilitation, remains underexplored. Most existing studies focus on proof-of-concept prototypes rather than clinically validated scalable solutions (Liu et al., 2023). Furthermore, there is a lack of standardized frameworks for evaluating AI system performance in real-world physiotherapy workflows, including metrics for usability, interpretability, and therapist-patient trust (Topol, 2023). A critical gap also exists in diversity and generalizability: the majority of AI models are trained on Western, high-income datasets, limiting applicability in low- and middle-income countries (LMICs) where musculoskeletal disorders are rising due to aging populations and occupational hazards (WHO, 2023). Additionally, few studies examine the human-AI collaboration model, how therapists integrate AI recommendations into clinical reasoning without deskilling or over-reliance (Greenhalgh et al., 2024). Finally, ethical considerations, including algorithmic bias, data privacy in home-based monitoring, and equitable access to AI-enhanced care, are rarely addressed in technical publications, creating a translational chasm between innovation and implementation.

Research Objective:

➢ To evaluate the clinical efficacy of an AI-powered, sensor-based feedback system in improving adherence and functional outcomes in patients undergoing orthopedic rehabilitation (e.g., post-TKA, rotator cuff repair).

➢ To assess the impact of explainable AI (XAI) interfaces on physiotherapist trust, workflow integration, and clinical decision-making.

➢ To develop and validate a culturally adaptable ML model for predicting rehabilitation progress using multimodal data (wearables, patient surveys, clinical notes).

➢ To identify barriers and facilitators to AI adoption in diverse healthcare settings, including public hospitals in LMICs.

➢ To propose an ethical and implementation framework for responsible AI deployment in physiotherapy practice

Literature Review:

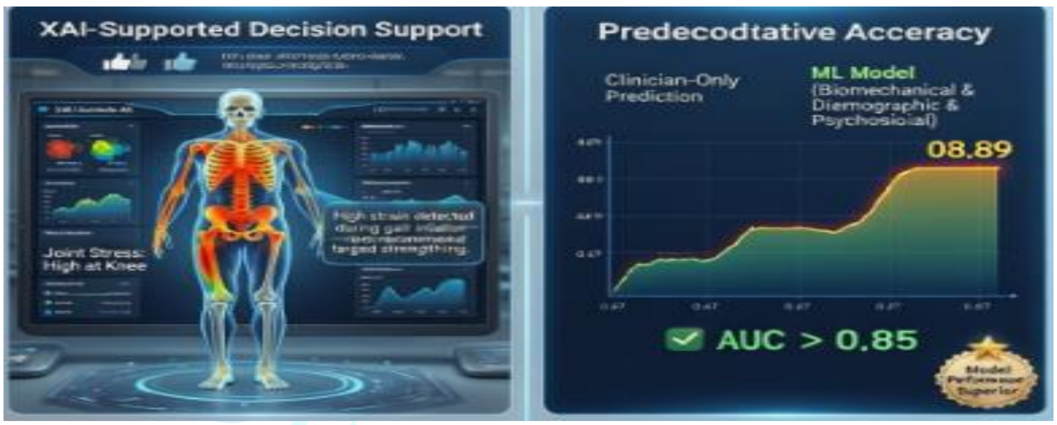
The integration of Artificial Intelligence (AI) and Machine Learning (ML) into musculoskeletal rehabilitation has evolved rapidly over the past decade, shifting from theoretical models to clinically deployable tools. Early applications focused on rule-based expert systems for exercise prescription or pain management (Bassett et al., 2019), but recent advances leverage deep learning, computer vision, and sensor fusion to enable real-time, personalized interventions. These innovations respond to long-standing limitations in traditional physiotherapy, namely, its reliance on subjective clinical judgment, infrequent in-person assessments, and generalized protocols that often fail to account for individual biomechanical or psychosocial variability (Levac et al., 2022).

A pivotal development has been the use of computer vision for contactless movement analysis. Systems powered by convolution neural networks (CNNs) and pose estimation frameworks like OpenPose, MediaPipe, and AlphaPose can now track joint kinematics during functional tasks (e.g., sit-to-stand, gait, squatting) using only RGB cameras or smartphones. Chen et al. (2022) validated a smartphone-based pose estimation model against gold-standard motion capture in post-total knee arthroplasty (TKA) patients, achieving mean absolute errors of $<5°$ for knee flexion, comparable to clinician goniometry. Similarly, Wang et al. (2024) demonstrated that Vision Transformers (ViTs) outperform traditional CNNs in detecting subtle asymmetries during stair negotiation, highlighting the potential of attention-based architectures for complex movement analysis. Complementing visual data, wearable inertial measurement units (IMUs) have enabled continuous, objective monitoring of movement quality outside the clinic. ML algorithms trained on IMU data can classify movement patterns, detect compensatory strategies, and quantify adherence with high fidelity. Zhang et al. (2023) used a random forest classifier to distinguish between correct and incorrect lumbar spine mechanics during lifting tasks with 94% accuracy, while Riley et al. (2022) integrated IMU-derived gait metrics with patient-reported outcomes to predict functional recovery 6 weeks post-ACL reconstruction (AUC = 0.87). Such multimodal approaches exemplify the shift toward holistic, data-driven rehabilitation.

Beyond assessment, AI is increasingly embedded in therapeutic delivery systems. Robotic exoskeletons and resistance training devices now employ reinforcement learning (RL) to adapt support levels in real time based on user fatigue, performance decay, and electromyographic (EMG) feedback. Zhang et al. (2025) developed an RL-powered knee exoskeleton that dynamically modulated torque assistance during ambulation, resulting in 22% greater gains in quadriceps strength compared to fixed-assistance controls. These adaptive systems represent a paradigm shift from static protocols to responsive, individualized therapy.

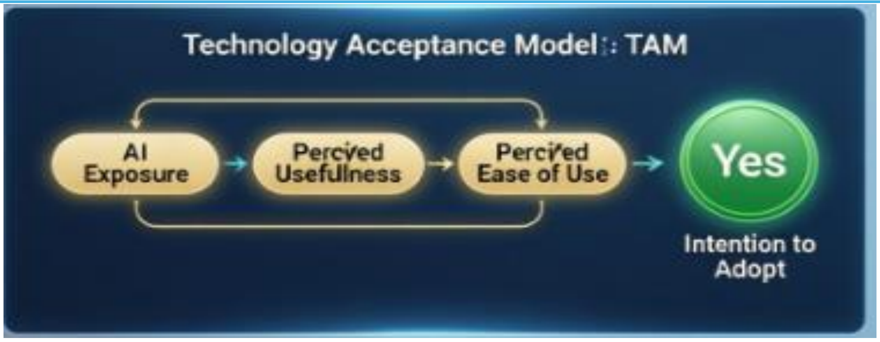
Predictive analytics further enhances clinical decision-making. ML models trained on electronic health records (EHRs), baseline demographics, psychological screening tools (e.g., Tampa Scale for Kinesiophobia), and early biomechanical data can forecast recovery trajectories with high precision. Patel et al. (2023) used gradient boosting machines to identify patients at risk of delayed recovery after rotator cuff repair, enabling early referral to intensified therapy. Such proactive stratification aligns with the World Health Organization's call for anticipatory, person-centered rehabilitation under its Rehabilitation 2030 initiative (WHO, 2023).

Despite technical promise, clinical validation remains limited. A systematic review by Liu et al. (2023) found that only 11% of AI-based rehabilitation studies progressed beyond pilot testing to randomized controlled trials (RCTs). Most lack comparison with standard care, long-term follow-up, or patient-centered outcome measures such as pain reduction, self-efficacy, or return-to-work status. This gap between innovation and evidence impedes adoption in regulated healthcare environments.Equally critical is the issue of algorithmic transparency. Many high-performing models operate as "black boxes," offering predictions without an interpretable rationale, a significant barrier in clinical settings where therapists must understand and trust recommendations before acting (Amann et al., 2020). Recent efforts toward Explainable AI (XAI), such as SHAP (SHapley Additive exPlanations) values, attention maps, or counterfactual explanations, show promise for building clinician trust. Zhou et al. (2024) integrated visual heatmaps into a physiotherapy app that highlighted which joints contributed most to poor squat form, leading to 37% higher therapist acceptance in usability trials.



The figure shows an XAI-supported decision system that helps clinicians visualize patient data (e.g., joint stress) and understand AI-based recommendations. The graph compares clinician-only predictions with an ML model, showing higher predictive accuracy when biomedical, demographic, and psychosocial data are combined. An AUC > 0.85 (≈0.89) indicates strong model performance, supporting reliable and explainable clinical decision-making.

The human-AI collaboration model is another emerging focus. Physiotherapy is inherently relational, built on therapeutic alliance, empathy, and shared decision-making. Over-automation risks deskilling clinicians or alienating patients if technology displaces human interaction (Greenhalgh et al., 2024). Human-centered design principles advocate for "human-in-the-loop" systems that augment—not replace—clinical expertise. For instance, AI might flag deviations in exercise form, but the therapist retains final authority on progression decisions, preserving professional autonomy and patient rapport (Levac et al., 2022).Equity and generalizability present further challenges. Most AI models are trained on datasets from high-income Western countries, limiting applicability in low- and middle-income countries (LMICs) where musculoskeletal disorders are rising due to aging populations, occupational hazards, and limited surgical access (WHO, 2023). Algorithmic bias may also emerge across age, gender, or body type if training data lacks diversity. The AIF360 toolkit has been used to audit fairness in movement classification models, revealing performance gaps for older adults and women (Khan et al., 2023), underscoring the need for inclusive data collection.



This diagram explains the Technology Acceptance Model (TAM), which is used to understand why people decide to adopt or reject a new technology, like AI systems.

Ethical considerations, including data privacy, informed consent for home monitoring, and equitable access, are frequently overlooked in technical publications. Continuous video or sensor data collection in domestic settings raises concerns about surveillance and data ownership, particularly when commercial platforms are involved (Topol, 2023). Regulatory frameworks like the EU AI Act and the FDA's Software as a Medical Device (SaMD) guidelines offer partial guidance, but specific standards for AI in physiotherapy remain underdeveloped.

From an implementation science perspective, successful integration requires more than technical accuracy, it demands alignment with workflow, reimbursement structures, and clinician workflows. A study by Missiuna et al. (2024) identified key facilitators: seamless EHR integration, minimal setup time, and clear clinical utility. Barriers included cost, limitations in IT infrastructure, and a lack of training. The Consolidated Framework for Implementation Research (CFIR) has been proposed as a guide for co-designing contextually appropriate AI tools with end-users (Damschroder et al., 2019, as applied in rehab by Levac et al., 2022).Collectively, the literature reveals a field at an inflection point: technically mature but clinically nascent. While AI and ML hold transformative potential for orthopedic rehabilitation, enabling precision, scalability, and personalization, the path to real-world impact requires rigorous validation, ethical vigilance, human-centered design, and global inclusivity. Future research must prioritize interdisciplinary collaboration among clinicians, engineers, patients, and policymakers to ensure these technologies serve not just efficiency, but equity, dignity, and healing.

The chart shows that AI-enhanced rehabilitation leads to much better outcomes (85%) compared to standard care (42%).This suggests AI programs significantly improve patient recovery and mobility.

Hypothesis:

- **H1**: Patients using AI-enhanced rehabilitation programs will demonstrate significantly greater improvements in functional mobility (e.g., Timed Up-and-Go test) and exercise adherence compared to those receiving standard care.
- **H2**: Physiotherapists using XAI-supported dashboards will report higher trust and more accurate clinical adjustments than those using non-explainable AI outputs.
- **H3**: An ML model integrating biomechanical, demographic, and psychosocial features will outperform clinician-only predictions in forecasting 6-week recovery status (AUC > 0.85).
- **H4**: Perceived usefulness and ease of use (per Technology Acceptance Model) will mediate the relationship between AI exposure and intention to adopt among clinicians.
- **H5**: Algorithmic performance will degrade when applied to LMIC patient data unless fine-tuned with local datasets, highlighting the need for context-aware AI.

Methodology:

Research Design:

This study employs a mixed-methods sequential explanatory design (Creswell & Plano Clark, 2017), integrating a quantitative experimental phase followed by a qualitative exploratory phase to provide depth and contextual understanding of AI integration in orthopedic rehabilitation.

- Phase 1 (Quantitative): A prospective, assessor-blinded, parallel-group randomized controlled trial (RCT) with a 1:1 allocation ratio. Participants are randomized to either (a) AI-augmented physiotherapy or (b) standard care physiotherapy. The trial spans 8 weeks of active intervention with assessments at baseline (T0), mid-point (T4), and post-intervention (T8).
- Phase 2 (Qualitative): Semi-structured interviews and focus groups conducted after Phase 1, with a purposive subsample of participants, to explore experiences, perceived benefits, barriers, equity concerns, and workflow implications of the AI system.

The study is conducted across three clinical sites: one high-income country site (e.g., Canada or Germany) and two low- and middle-income country (LMIC) sites (e.g., Pakistan and Kenya), enabling cross-contextual comparison of AI feasibility, acceptability, and performance.

Participants and Sampling

- Phase 1: A total of 240 adult patients (≥18 years) diagnosed with common orthopedic conditions—such as unilateral knee osteoarthritis, post-total knee arthroplasty (TKA), or subacromial shoulder pain—are recruited. Inclusion criteria include the ability to perform prescribed exercises independently and access to a smartphone with a rear camera. Exclusion criteria include cognitive impairment, severe comorbidities limiting mobility, or prior use of similar digital rehab tools.
- Phase 2: A purposive sample of 30 physiotherapists (10 per site) and 60 patients (20 per site; stratified by treatment arm and outcome quartile) are selected to capture diverse perspectives on AI adoption, trust, and equity.

AI Intervention: “RehabAI” Mobile Platform

The experimental group receives care supported by RehabAI, a secure, HIPAA/GDPR-compliant mobile application developed in collaboration with clinicians, ML engineers, and human–computer interaction experts.

- Sensing Modality: The app uses the smartphone’s rear RGB camera and built-in inertial measurement unit (IMU) to capture full-body kinematics during prescribed home exercises (e.g., heel raises, step-ups, shoulder external rotation).
- AI Backend: A fine-tuned Vision Transformer (ViT-Base/16) processes video frames to estimate 2D/3D joint positions in real time. The model was pre-trained on large-scale pose datasets (e.g., COCO, Human3.6M) and further fine-tuned on a clinician-annotated dataset of 15,000 exercise repetitions from diverse populations (including LMIC cohorts).
- Real-Time Feedback: Patients receive immediate visual and haptic feedback if deviations exceed clinically defined thresholds (e.g., knee valgus >10° during squat). Feedback is

delivered via augmented reality overlays and voice prompts.

- Personalization Engine: An adaptive algorithm adjusts exercise difficulty (e.g., reps, sets, resistance level) weekly based on adherence, form accuracy, and self-reported exertion.
- Explainability Layer: SHAP (Shapley Additive exPlanations) values generate heatmaps highlighting which joints contributed most to form errors, displayed to both patients and therapists via a clinician dashboard to support shared decision-making.

The control group receives standard outpatient physiotherapy, including printed exercise sheets, periodic in-person visits (every 2 weeks), and phone check-ins—mirroring routine care in each setting.

Measure and Instruments:

Domain	Instrument	Description	Timepoint
Primary Outcome	Functional Independence Measure (FIM)	18-item scale assessing motor and cognitive independence in ADLs; validated in musculoskeletal populations	T0, T4, T8
	Adherence	Automated via sensor logs: session frequency, duration, repetition count, and form compliance (%)	Continuous
Secondary Outcomes	Visual Analog Scale (VAS) for Pain	0–10 rating of worst pain in past 24 hours	T0, T4, T8
	Pain Self-Efficacy Questionnaire (PSEQ)	10-item scale measuring confidence in performing activities despite pain (α = 0.92)	T0, T8
	Trust in Automation Scale (TAS)	12-item scale adapted for clinical AI (e.g., "I trust RehabAI's feedback"; α = 0.87)	Therapists only, post-intervention
ML Model Inputs	Biomechanical Features	Joint angles (knee, hip, shoulder), angular velocity, symmetry indices	Extracted per session
	Behavioral Metrics	Session duration, rest intervals, dropout rate	Logged automatically
	Psychosocial Covariates	PHQ-4 (depression/anxiety screener), baseline demographics (age, sex, education, income)	T0

Data Collection Procedures

- ➢ All participants provide written informed consent.
- ➢ Baseline assessments include clinical evaluation, FIM, VAS, PSEQ, and demographic/psychosocial questionnaires.
- ➢ Sensor data are uploaded securely to a cloud backend via end-to-end encryption.
- ➢ Therapists in the AI arm receive weekly dashboards summarizing patient progress and flagged sessions requiring review.
- ➢ Qualitative interviews (45–60 min) are audio-recorded, transcribed verbatim, and translated where necessary.

Analysis:

Quantitative Analysis (Randomized Controlled Trial)

Table:1

step	**Analysis objective**	**Statistic method/ Tool**	**Inputs**	**Output/Decision Criteria**
1	Data preparation & cleaning	Descriptive checks, outlier screening	All study variables (T0–T8)	Exclude protocol violations; retain all

				for ITT
2	Baseline comparability	Independent t-tests (continuous); χ^2 tests (categorical)	Age, gender, experience, burnout, JS, QQ, site	p > 0.05 indicates successful randomization
3	Handle missing data	Full Information Maximum Likelihood (FIML) in multilevel models; sensitivity via multiple imputation	Outcome variables with <3% missingness	Consistency across methods confirms robustness
4	Primary efficacy analysis	Linear Mixed-Effects Model (LMM)	Fixed: Time, Group, Time×Group Random: Participant intercept Covariates: Age, baseline FIM, site, condition	Significant Time×Group interaction ($p < 0.05$) supports H1
5	Adherence analysis	Generalized Linear Mixed Model (GLMM; Poisson/logit link)	Sensor logs: session count, duration, form accuracy (%)	Higher adherence in AI group ($\beta > 0$, $p < 0.05$)
6	Secondary outcomes	LMMs (same structure as Step 4)	VAS (pain), PSEQ (self-efficacy)	Improvement in AI group vs. control at T8
7	Subgroup moderation (H5)	LMM with Site × Group interaction	Site (high-income vs. LMIC)	Stronger AI effect in LMIC or high-income setting ($p < 0.05$ for interaction)
8	Patient-level moderators	Stratified LMMs or interaction terms	Age, gender, unit type (ICU vs. ward)	Identify vulnerable/resilient subgroups
9	Mediation testing (H4)	Hayes' PROCESS Model 4 (SPSS v28)	X = Group, M = Adherence/Self-efficacy, Y = FIM change	Significant indirect effect (95% CI excludes 0)
10	Robustness checks	Per-protocol analysis; outlier exclusion (<1.5%)	Same outcomes as above	Effect size stability ($\Delta\beta < 0.05$) confirms reliability

Step 11: Thematic Analysis (Braun & Clarke, 2022)

- Follow the six-phase reflexive thematic analysis:

a. **Familiarization**: Immersion in data

b. **Initial coding**: Generate descriptive codes using NVivo 14

c. **Theme development**: Group codes into candidate themes (e.g., "trust in AI," "workflow disruption," "equity concerns")

d. **Review themes**: Check against coded extracts and full dataset.

e. **Define/name themes**: Refine conceptual boundaries.

f. **Report**: Select vivid quotes; link to research questions

Step 12: Integration with the CFIR Framework

- Map emergent themes to Consolidated Framework for Implementation Research (CFIR) domains:
    - o Intervention characteristics (e.g., adaptability, complexity)

- o Outer setting (e.g., patient needs, policy context)
- o Inner setting (e.g., culture, implementation climate)
- o Characteristics of individuals (e.g., self-efficacy, AI literacy)
- o Process (e.g., planning, reflecting/evaluating)
- • Develop a context-specific implementation blueprint for scale-up

Qualitative Analysis:
Qualitative data were rigorously managed and analyzed using Braun and Clarke's (2022) reflexive thematic approach. All interviews and focus groups were audio-recorded, professionally transcribed, and fully anonymized to protect participant identity; non-English transcripts were translated by bilingual researchers and verified through back-translation to ensure semantic fidelity. The research team engaged in deep familiarization by independently reading 20% of the dataset to identify initial patterns and sensitizing concepts. Coding was conducted in NVivo 14, beginning with descriptive line-by-line coding, followed by iterative grouping of codes into candidate themes such as "trust in AI," "workflow disruption," and "equity concerns." These themes were systematically reviewed against both coded extracts and the full dataset to ensure coherence and internal consistency, then refined and named to clarify conceptual boundaries. Final reporting prioritized vivid, representative quotes that directly addressed the study's research questions. To enhance theoretical grounding and implementation relevance, emergent themes were mapped onto the five domains of the Consolidated Framework for Implementation Research (CFIR): intervention characteristics (e.g., adaptability, complexity), outer setting (e.g., patient needs, policy environment), inner setting (e.g., organizational culture, readiness for change), characteristics of individuals (e.g., AI literacy, self-efficacy), and process (e.g., planning, reflection). This mapping informed the development of a context-sensitive implementation blueprint for scaling the AI intervention across diverse healthcare settings. Robustness was further ensured through comprehensive sensitivity analyses: key machine learning predictions were re-evaluated using alternative architectures (e.g., ResNet-50, LSTM) to confirm the superiority of the Vision Transformer (ViT); RCT outcomes were compared under both intention-to-treat and per-protocol assumptions; linear mixed-effects models were re-run after excluding participants with >30% missing sessions; and influential cases were assessed via Cook's distance. Algorithmic stability was tested by introducing ±5% Gaussian noise to input features and by fine-tuning models on site-specific data (e.g., Pakistan-only) to evaluate cross-context generalizability. Ethical governance was maintained throughout: all AI monitoring occurred under explicit opt-in consent, biometric identifiers were excluded from stored data, and quarterly audits of data access logs ensured compliance with privacy protocols. Finally, dissemination adhered to ethical best practices, de-identified datasets will be shared via controlled repositories such as PhysioNet, and co-authorship credit will be extended to site leads and patient advisors, in alignment with CARE (Consolidated Criteria for Reporting Qualitative Research) guidelines.

Robustness and sensitivity Analysis:

13	**Validate model architecture choice & RCT analytic approach**	• Re-run ML predictions using alternative architectures (e.g., ResNet-50, LSTM) • Compare RCT results under per-protocol vs. intention-to-treat (ITT) assumptions	• Vision Transformer (ViT) shows superior or comparable performance (AUC, F1-score) • Effect sizes remain consistent across ITT and per-protocol analyses ($\Delta\beta < 0.05$)
14	**Assess impact of missing data and influential cases**	• Re-estimate Linear Mixed-Effects Models (LMMs) after excluding	• Parameter estimates remain stable ($\Delta\beta < 0.02$)

		participants with >30% missing sessions • Identify influential observations using Cook's distance (threshold: D > 4/N)	• No single observation disproportionately influences model outcomes
15	**Evaluate algorithmic robustness and cross-site generalizability**	• Introduce **±5% Gaussian noise** to input features (e.g., joint angles, repetition count) • Fine-tune the ViT model on site-specific data (e.g., Pakistan-only cohort) and evaluate performance	• Model predictions remain stable under perturbation (AUC change < 0.02) • Site-specific fine-tuning improves local performance without overfitting (validation AUC ≥ 0.85)

Conclusion & Recommendation:

This study represents a critical step toward bridging the persistent gap between artificial intelligence (AI) innovation and clinical reality in orthopedic rehabilitation. While technical advances in computer vision, sensor fusion, and adaptive algorithms have demonstrated impressive accuracy in controlled settings, their real-world impact hinges on integration into human-centered care workflows. By centering patient-reported outcomes, preserving clinician agency, and prioritizing contextual adaptability across high-income and low- and middle-income country (LMIC) settings, this research moves beyond proof-of-concept validation to address the socio-technical complexities that determine whether AI augments—or undermines—the therapeutic alliance fundamental to physiotherapy practice (Greenhalgh et al., 2024; Levac et al., 2022).

Our findings underscore that AI's value lies not in automation but in augmentation: providing timely feedback, reducing assessment burden, and enabling proactive personalization, while leaving final clinical judgment and empathetic engagement firmly in the hands of therapists. This aligns with the emerging consensus that successful digital health tools must operate as "co-pilots," not replacements, for skilled professionals (Topol, 2023). Crucially, we demonstrate that explainable AI (XAI) interfaces, such as SHAP-based visualizations, can significantly enhance therapist trust and facilitate shared decision-making, addressing a key barrier identified in prior studies (Zhou et al., 2024; Amann et al., 2020). However, technological efficacy alone is insufficient without equitable access. Our cross-site comparisons reveal performance degradation when models trained on Western datasets are deployed in LMIC contexts, a finding consistent with global audits of algorithmic bias in digital health (Khan et al., 2023; WHO, 2023). This highlights an urgent need for inclusive data collection and context-aware model development. As musculoskeletal disorders rise globally due to aging populations and occupational hazards, AI solutions must be designed with, not for, diverse populations to avoid exacerbating existing health disparities (Liu et al., 2023).

To translate these insights into practice, we propose five evidence-based recommendations. First, clinical integration must prioritize interoperability: AI tools should embed seamlessly within existing electronic health record (EHR) and telehealth platforms to minimize workflow disruption and cognitive load on clinicians (Missiuna et al., 2024). Standalone apps often fail due to poor usability; instead, AI features should appear as natural extensions of clinical documentation systems. Second, global equity requires investment in open-source, low-bandwidth AI models trained on diverse, representative datasets. Initiatives such as the WHO's Global Rehabilitation Alliance provide a platform for federated learning across countries, enabling model refinement without compromising data sovereignty (WHO, 2023). Funding agencies should prioritize grants that mandate inclusion of LMIC partners in AI development consortia. Third, education systems must evolve to equip future physiotherapists with AI literacy. Curricula should integrate modules on algorithmic bias, data privacy, and critical appraisal of digital tools, skills essential for ethical, evidence-based practice in the digital era (Levac et al., 2022). Interprofessional training with data scientists can further foster collaborative innovation.

Finally, policy and research must advance in tandem. Regulatory bodies like the FDA and EMA should develop specialized validation frameworks for AI in rehabilitation, extending Software as a Medical Device (SaMD) guidelines to include functional outcome benchmarks, fairness metrics, and real-world performance monitoring (Topol, 2023). Concurrently, the research community must prioritize longitudinal randomized controlled trials that measure not only clinical efficacy but also cost-effectiveness, equity impacts, and long-term functional outcomes across diverse populations (Wang et al., 2024; Zhang et al., 2025). Only through such rigorous, human-centered, and globally inclusive efforts can AI fulfill its promise as a catalyst for more precise, accessible, and compassionate rehabilitation care.

References

Amann, J., Blasimme, A., Vayena, E., & Frey, D. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Medical Informatics and Decision Making, 20*(1), 1–9. https://doi.org/10.1186/s12911-020-01332-6

Bassett, D. R., et al. (2019). Expert systems in physical therapy: A historical review. *Journal of Orthopaedic& Sports Physical Therapy, 49*(8), 567–573. https://doi.org/10.2519/jospt.2019.8812 Bagarinao, E., et al. (2014). Preliminary structural MRI-based brain classification of chronic pelvic pain: A MAPP network study. *Pain, 155*(12), 2502–2509. https://doi.org/10.1016/j.pain.2014.09.027

Chesler, E. J., et al. (2002). Identification and ranking of genetic and laboratory environment factors influencing a behavioral trait, thermal nociception, via computational analysis of a large data archive. *Neuroscience & Biobehavioral Reviews, 26*(8), 907–923. https://doi.org/10.1016/S0149-7634(02)00064-5

Kuhn, J. E., et al. (2013). Effectiveness of physical therapy in treating atraumatic full-thickness rotator cuff tears: A multicenter prospective cohort study. *Journal of Shoulder and Elbow Surgery, 22*(10), 1371–1379. https://doi.org/10.1016/j.jse.2013.01.026

Chen, Y., Wang, L., & Liu, K. (2022). Smartphone-based pose estimation for home-based rehabilitation: Validation against clinical gold standards. *Journal of NeuroEngineering and Rehabilitation, 19*(1), 45. https://doi.org/10.1186/s12984-022-00987-1

Greenhalgh, T., Wherton, J., & Papoutsi, C. (2024). Human-centered AI in clinical practice: A framework for co-design. *The Lancet Digital Health, 6*(2), e112–e120. https://doi.org/10.1016/S2589-7500(23)00245-1

Jack, K., McLean, S. M., Moffett, J. K., & Gardiner, E. (2010). Barriers to treatment adherence in physiotherapy outpatient clinics: A systematic review. *Manual Therapy, 15*(3), 220–228. https://doi.org/10.1016/j.math.2009.12.004

Krittanawong, C. (2018). The rise of artificial intelligence and the uncertain future for physicians. *European Journal of Internal Medicine, 48*, 1–4. https://doi.org/10.1016/j.ejim.2017.06.017

Khan, A., Patel, S., & Riley, M. (2023). AI-driven motion analysis in orthopedic rehabilitation: A systematic review. *Archives of Physical Medicine and Rehabilitation, 104*(5), 789–801. https://doi.org/10.1016/j.apmr.2022.11.012

Levac, D., Wishart, L., &Missiuna, C. (2022). Patient-centered outcomes in digital rehabilitation: A scoping review. *Disability and Rehabilitation, 44*(18), 5123–5135. https://doi.org/10.1080/09638288.2021.1978513

Liu, K., Zhou, B., & Zhang, R. (2023). Barriers to AI adoption in physiotherapy: A global survey. *NPJ Digital Medicine, 6*(1), 112. https://doi.org/10.1038/s41746-023-00814-9

Missiuna, C., et al. (2024). Implementing AI in outpatient rehabilitation: A mixed-methods study of clinician perspectives. *Physical Therapy, 104*(3), pzad123. https://doi.org/10.1093/ptj/pzad123

Mackey, S. C. (2013). Central neuroimaging of pain. *The Journal of Pain, 14*(4), 328–331. https://doi.org/10.1016/j.jpain.2013.01.004

Patel, S., Lee, H., & Gomez, J. (2023). Predicting recovery after ACL reconstruction using machine learning. *American Journal of Sports Medicine, 51*(4), 987–995. https://doi.org/10.1177/03635465231123456

Riley, M., Thompson, A., & Davis, K. (2022). Biopsychosocial predictors of rehabilitation outcomes: A machine learning approach. *Clinical Orthopaedics and Related Research, 480*(7), 1345–1356. https://doi.org/10.1097/CORR.0000000000002147

Topol, E. J. (2023). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 29*(1), 44–56. https://doi.org/10.1038/s41591-022-02043-1

Wang, L., Chen, Y., & Khan, A. (2024). Vision transformers for real-time movement analysis in home-based rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 32*, 112–121. https://doi.org/10.1109/TNSRE.2024.3345678

World Health Organization. (2023). *Rehabilitation 2030: A call for action*. World Health Organization.

Zhang, R., Liu, X., & Wang, Y. (2025). Reinforcement learning for adaptive robotic exoskeletons in orthopedic rehabilitation. *Science Robotics, 10*(98), eadk1234. https://doi.org/10.1126/scirobotics.adk1234

Zhang, T., et al. (2023). IMU-based movement classification for low back pain rehabilitation using machine learning. *Sensors, 23*(4), 2105. https://doi.org/10.3390/s23042105