

Review

A Review of Research on the Applications of Large Models in Each Functional Module of the Entire Rehabilitation Process

Tingting Bai ¹, Kaiwen Jiang ², Yixuan Yu ¹, Shuyan Qie ³, Congxiao Wang ³, Boyuan Wang ^{4,5} and Wenli Zhang ^{2,*} 

¹ School of Art and Design, Beijing University of Technology, Beijing 100124, China; baitingting@emails.bjut.edu.cn (T.B.); yxyx@emails.bjut.edu.cn (Y.Y.)

² School of Information Science and Technology, Beijing University of Technology, Beijing 100124, China; kecin@emails.bjut.edu.cn

³ Department of Rehabilitation, Beijing Rehabilitation Hospital, Capital Medical University, Beijing 100144, China

⁴ Beijing Xiaotangshan Hospital, Beijing 102211, China

⁵ Center for Medical Artificial Intelligence Technology Innovation, Zhuhai Fudan Innovation Research Institute, Zhuhai 519031, China

* Correspondence: zhangwenli@bjut.edu.cn

Abstract

Population ageing and chronic disease are increasing demand for rehabilitation, while resources remain limited. This review does not report an implemented end-to-end system; instead, it proposes a modular workflow framework for applying large AI foundation models across rehabilitation. Organised into four stages—assessment, prescription, execution, and monitoring—we summarise recent evidence and highlight techniques most suitable at each stage. In assessment, multimodal models can enable more continuous and objective functional measurement from heterogeneous sensor and imaging data. In prescription, large language models can support evidence-informed, personalised plan formulation by synthesising guidelines and patient context. In execution, vision–language–sensor models can provide real-time feedback for telerehabilitation and adherence support. In monitoring, longitudinal and cross-setting data integration can facilitate risk prediction and early warning for safety and long-term management. We also discuss practical adaptation options (e.g., parameter-efficient fine-tuning) and propose a clinimetric-oriented evaluation framework to assess validity, reliability, and generalisability. By mapping AI capabilities to concrete workflow tasks, the framework provides a theoretical foundation and roadmap for reproducible research and future translation toward a universal rehabilitation model.

Keywords: rehabilitation; large language models; vision and multimodal models; digital rehabilitation; artificial intelligence



Academic Editors: Lidia Bajenaru and Virginia Sandulescu

Received: 24 December 2025

Revised: 31 January 2026

Accepted: 9 February 2026

Published: 12 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

In recent years, global population aging and the rising burden of chronic non-communicable diseases have shifted rehabilitation from an “optional” adjunct into an essential health service across the life course. The World Health Organization’s Rehabilitation 2030 initiative estimates that approximately 2.4 billion people live with health conditions that may benefit from rehabilitation, consistent with the International Classification of Functioning, Disability and Health (ICF) framing of functioning limitations and disability [1–5]. In China, accelerated ageing coupled with high chronic disease prevalence

has created substantial unmet needs, including around 180 million older adults, tens of millions living with chronic diseases, and more than ten million new cases annually involving functional limitations and rehabilitation needs [3,6].

To orient the reader, this review aims to map how foundation models (LLMs, vision models, and multimodal models) can support key rehabilitation tasks by organizing current evidence around four workflow modules—Assessment, Prescription, Execution, and Monitoring. Accordingly, this paper reviews the technical pathways and application landscapes of large models in empowering rehabilitation. We further discuss methodological and regulatory considerations to support safe, effective, and sustainable clinical translation. In this context, this review is intended for researchers and practitioners working on the development and clinical translation of foundation models for rehabilitation at the intersection of rehabilitation medicine, biomedical engineering, and artificial intelligence.

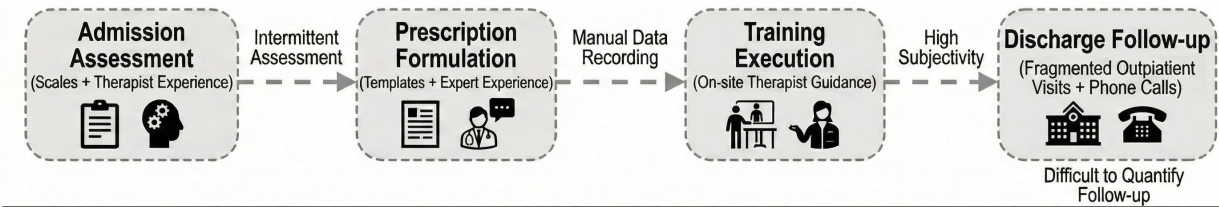
In this paper, “rehabilitation” refers to research-oriented clinical rehabilitation, primarily including neurological, orthopedic and geriatric rehabilitation, as well as rehabilitation for chronic disease-functional limitations and disability. Across these domains, common goals include functional recovery and participation-oriented life reconstruction, typically delivered through therapist-led workflows. In resource-constrained settings, major bottlenecks persist across functional assessment, intervention decision-making, and out-of-hospital follow-up. Assessments remain highly dependent on therapist experience and scale-based scoring, which can be time-consuming, subjective, and insufficiently sensitive to continuous changes [7–9]. Prescriptions are often template-driven, making it difficult to adapt plans to evolving patient states and goals. Follow-up is frequently hindered by poor adherence and fragmented data, limiting evidence accumulation for outcome evaluation and pathway optimization [10–15]. As summarised in Figure 1, conventional rehabilitation is often episodic and fragmented—with intermittent, scale-based assessment, template-driven prescription, manual data recording, and poorly quantifiable post-discharge follow-up—whereas a foundation-model-enabled paradigm aims to connect Assessment → Decision → Action → Re-evaluation through multimodal sensing, real-time feedback, and data loopback for iterative optimisation.

Foundation models offer new capabilities to reconnect this fragmented chain [15–17]. LLMs can parse electronic health records, rehabilitation logs and support traceable, guideline-linked reasoning through natural language interaction, enabling personalised plan drafting under clinician oversight [18–21]. Vision and multimodal models can extract kinematics (e.g., pose, range of motion, and movement patterns) from video, imaging and wearable sensors, enabling richer representation of functional status [16,22]. Combined with real-time feedback, these representations can support prognosis modelling and adaptive parameter adjustment within an “assessment–decision–action–reassessment” loop [11,14,15].

However, converting model capabilities into clinical utility remains challenging. Key issues include data governance and privacy, transferability across populations and settings, fairness and robustness, interpretability and human factors, as well as real-world evidence generation and economic evaluation [23–28]. In rehabilitation, acceptability to therapists, patients, and organisations is a first-order determinant of adoption. Given the long time horizon, repeated interactions, and safety-sensitive nature of rehabilitation, tools may fail to be implemented if perceived usefulness is low, workflow burden is high, trust is limited, or usability is poor—even when technical accuracy is adequate. Therefore, acceptability constructs (e.g., affective attitude, burden, intervention coherence, ethicality, perceived effectiveness) should be considered when interpreting current evidence and designing future systems [27,29]. More broadly, while medical AI regulation is increasingly adopting a full-lifecycle approach (development, evaluation, deployment, and post-market moni-

toring) [27,28], highly contextualized rehabilitation scenarios require an evaluation and reporting system that is guided by clinical goals, centered on tasks, and deeply embedded within therapist and patient workflows [10,30–33].

Traditional Rehabilitation Model



Large Model Empowered Closed-Loop Model

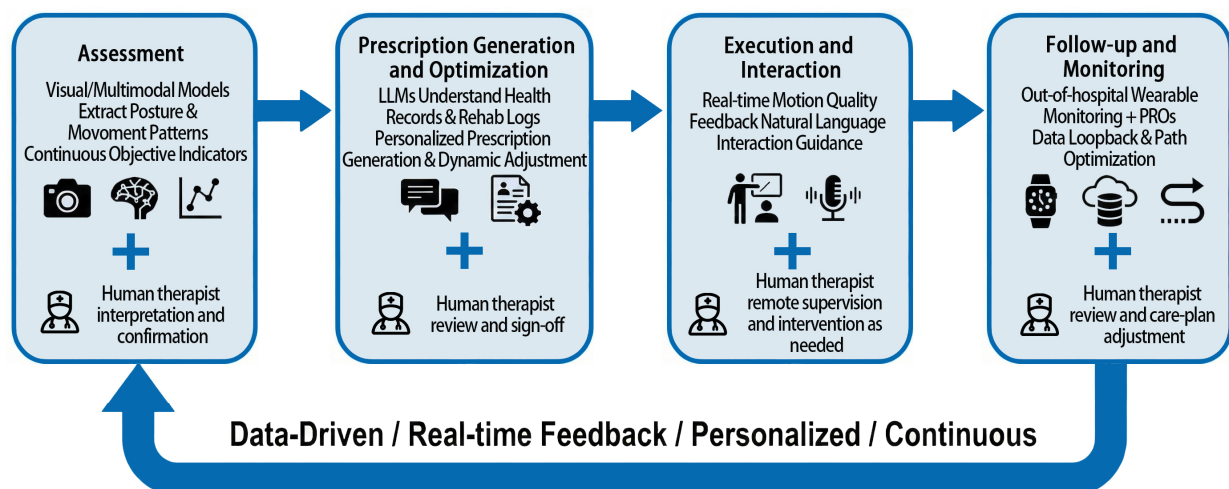


Figure 1. Comparison of conventional rehabilitation workflows and a foundation-model-enabled human-in-the-loop closed-loop paradigm, where foundation models support information processing and decision-making, and therapists supervise, validate, and retain final clinical responsibility.

Compared with medical AI reviews organised by model taxonomy or disease categories, our workflow-module framework aligns directly with rehabilitation practice (Assessment–Prescription–Execution–Monitoring), enabling task-level linkage among clinical objectives, input modalities, and stage-specific risks. The remainder of this paper is organised as follows: Section 2 summarizes core rehabilitation tasks and the data ecosystem; Section 3 reviews representative applications and methodological strategies for LLMs, vision models, and multimodal models; Section 4 proposes a rehabilitation-oriented evaluation framework; Section 5 discusses policy and ethical considerations and Section 6 concludes with future directions for translation and deployment.

2. Key Rehabilitation Tasks and Data Ecosystem

To address the practical bottlenecks in rehabilitation, this section summarises core workflow tasks and the data ecosystem that supports them. We organise the evidence into four modules: Assessment, Prescription Generation and Optimization, Execution and Interaction, and Follow-up and Monitoring, which together describe how information is collected, transformed into decisions, delivered as interventions, and evaluated over time. These modules are connected by multimodal data, including clinical text (EHRs, therapist notes, and patient-reported outcomes), images/videos, and time-series sensing signals (e.g., IMU/EMG and physiological measures). The subsections below review current practice

and challenges in each module, and highlight where foundation models are most likely to add value [10,23].

2.1. Assessment

High-fidelity motion analysis remains a cornerstone for objective functional assessment, yet its clinical accessibility is limited. Marker-based three-dimensional motion capture (MoCap), often treated as a reference standard for detailed kinematic analysis, provides accurate and reliable measures but typically requires dedicated laboratory infrastructure and trained personnel, restricting routine deployment in day-to-day rehabilitation [34]. Consequently, many routine settings still rely on therapist observation and scale-based scoring, while objective motion capture is often confined to high-resource centres or specific indications [35]. This creates a trade-off between feasibility and precision: manual observation is scalable but subjective and intermittent, whereas laboratory-grade measurement is precise but costly and difficult to scale. Manual video annotation, when used, further increases workload and introduces error-prone steps.

Recent advances in markerless Human Pose Estimation (HPE) and wearable sensing offer a feasible pathway toward more objective and continuous assessment outside specialised laboratories. Deep learning-based vision models can extract skeletal key points and kinematic indicators (e.g., joint angles, velocities, symmetry indices) from RGB/RGB-D videos, enabling continuous indicators that may better distinguish true recovery from compensatory patterns than discrete scale scores. In parallel, wearables such as IMUs and EMGs can capture limb dynamics and muscle activation, complementing vision-derived kinematics and improving robustness across viewpoints and occlusions [34,36–38].

Foundation models further extend this capability by enabling multimodal fusion and longitudinal modelling. Multimodal architectures can integrate vision and sensor streams to represent functional status holistically and support downstream estimation (e.g., motor function, ADL capability) and prognosis. Transformer-style sequence modelling is particularly relevant for learning temporal trends from long-range time-series and detecting clinically meaningful change patterns over time [38]. In addition, LLMs can extract function-relevant information from clinical narratives (EHRs, therapist notes, assessment reports) to generate structured summaries and highlight inconsistencies or missing elements, supporting therapist review rather than replacing judgment. Overall, the assessment-module value proposition is to improve objectivity, continuity, and interpretability of functional measurement, thereby supplying high-frequency quantitative inputs for downstream decision-making.

Despite this promise, using markerless and wearable data for rehabilitation foundation/functional models depends on two conditions: (i) clinimetric soundness of the measurement pipeline (validity, reliability, and sensitivity to change), and (ii) sustained real-world adoption—especially among older adults—because longitudinal data are only informative when adherence is maintained. In free-living settings, consumer devices may exhibit device-specific bias, motion artefacts, and non-trivial missingness; without explicit data-quality monitoring (calibration, artefact detection, and missing-data handling), downstream inference can be systematically biased. From a human-factors perspective, continued use is shaped by perceived value, ease of use, comfort, and trust; therefore, closed-loop systems should incorporate user-centred design, training/support, and low-burden protocols to maintain adherence and representativeness of longitudinal streams [39–46].

2.2. Prescription Generation and Optimization

Rehabilitation prescriptions are traditionally derived from clinical guidelines and therapist expertise, often implemented through template-like choices of exercise type, inten-

sity, and frequency. Due to the difficulty in fully incorporating patient-specific functional level, goals, comorbidities, and time-varying tolerance, prescriptions may be insufficiently personalised, adjusted too slowly, and underutilise multimodal feedback. Such mismatches between prescribed load and evolving capability (e.g., over-challenging programmes that increase fear or pain) can undermine self-efficacy and adherence, which are closely linked to clinically meaningful outcomes in rehabilitation populations [47,48].

LLMs offer a mechanism to connect “knowledge–decision–feedback” in a more systematic way. With retrieval augmentation, LLMs can draft guideline-consistent, structured prescriptions from diagnosis, assessment findings, and patient goals, and can provide transparent rationales and safety caveats for therapist review. Early studies suggest generative models can produce prescriptions that are reasonably structured and aligned with mainstream principles, while still showing limitations in nuanced strategy selection and true individual specificity [49–51].

To move from drafting to real personalisation, it is necessary to couple large models with real-time patient data in a closed loop to build a data-driven adaptive prescription system. Here, a “data-driven adaptive prescription system” denotes a framework in which: (i) multimodal data are collected continuously (pose/kinematics, wearables/IMU/EMG, physiological signals, PROs), (ii) patient state and short-term risk are estimated (fatigue, pain, movement quality, adherence), (iii) a decision policy updates between-session dose/intensity/frequency under explicit safety constraints and clinician override, and (iv) effects are re-evaluated in subsequent sessions (Perception–Decision–Feedback) [23,33,47,48,52–60]. The policy layer may be implemented via rules, supervised learning, or sequential decision learning (including reinforcement learning) to optimise long-horizon objectives (functional gains, adherence) while enforcing guardrails.

2.3. Execution and Interaction

During execution, clinical practice faces two prominent challenges: first, in-hospital training requires high-intensity one-on-one guidance from therapists, whose resources are limited; second, home-based training lacks professional supervision, making it difficult to guarantee movement quality and training adherence. Existing digital rehabilitation systems and robotic assistive devices mostly rely on preset programs, lacking real-time personalized feedback and humanized interaction, and thus struggle to replace the on-site guidance and psychological support provided by therapists.

LLMs can enhance execution quality by combining perception + interaction. Vision and sensor models enable real-time detection of movement deviations, compensatory patterns, and physiological signs of fatigue (e.g., pose estimation, joint trajectories, heart rate, EMG) [5]. Naqvi et al. [13] noted that Large Language Models can act as virtual therapists, providing concrete, actionable corrective suggestions and safety reminders via voice or text, such as giving customized feedback on insufficient joint angles or center-of-gravity shifts. Simultaneously, LLMs can assume the role of a coach or companion, offering encouragement and explaining the significance of training, thereby enhancing motivational support.

Nevertheless, patient acceptance of AI-mediated coaching is not universal. Adoption may be constrained by preferences for human contact, trust and privacy concerns, perceived stigma, and the cognitive/physical burden of interacting with digital agents—especially in older or more impaired populations—suggesting that hybrid care models with clear human fallback and participatory, user-centered design are often necessary [40,61].

Recent work also demonstrates feasibility for LLM-driven feedback generation when skeletal features are transformed into structured prompts and the LLM is constrained to produce targeted, safety-aware guidance. For example, Tang et al. [62] reported a pipeline

that extracts kinematic descriptors from joint sequences and uses prompt strategies (e.g., few-shot prompting/role prompting) to generate movement-quality feedback on public datasets. More broadly, multimodal foundation models can align vision and language, mapping identified specific movement issues to corresponding natural language prompts and visual markers, thereby simulating the comprehensive guidance process of “observing, speaking, and demonstrating” performed by a therapist. Regarding interaction design, factors such as interface usability, language comprehensibility, and emotional friendliness must also be considered to reduce the psychological burden on patients. Overall, the value of large models in the execution and interaction phase lies in extending professional guidance remotely and embedding emotional support, thereby improving training quality and persistence.

2.4. Follow-Up and Monitoring

Rehabilitation often spans weeks to months (or longer) for chronic diseases and severe impairment populations; post-discharge management is critical for consolidating gains and preventing decline. Yet follow-up is frequently undermined by low adherence, fragmented data chains, and poor interoperability between in-clinic and at-home systems. Daily status between visits is difficult to capture systematically, self-recorded data are often unstructured, and device/institution heterogeneity complicates longitudinal aggregation and reuse [24,63,64].

Foundation-model-enabled monitoring aims to build a continuous data ecosystem across settings. Wearables and connected smart devices can passively collect behavioral metrics (e.g., steps, activity intensity, sleep) alongside patient-reported outcomes (PROs), potentially supplemented by home IoT signals capturing environment and activity context. These multi-source time-series data, analyzed by architectures adept at long-sequence modeling (such as Transformers), can be used to delineate functional recovery trajectories and predict risks of major events such as falls and readmissions, providing a basis for earlier intervention. In parallel, machine learning models trained on large-scale inpatient rehabilitation data have shown promising performance in prognosis-related prediction tasks (e.g., such as discharge destination), offering tools for risk stratification and personalized follow-up strategies.

Evidence from recent clinical trials and syntheses also supports the clinical relevance of model-enabled remote rehabilitation at scale. For musculoskeletal disorders, a systematic review and meta-analysis integrating 37 clinical trials reported that internet-based telerehabilitation was generally more favorable than minimal or usual-care comparators across multiple outcomes, although effects varied by condition and comparator intensity [19]. Complementary health-economic evidence suggests telerehabilitation can be cost-effective in several settings, but estimates depend on local context and implementation pathways [65]. In neurorehabilitation, telerehabilitation has also shown measurable benefits: a meta-analysis of randomized controlled trials during the COVID-19 period suggested that telerehabilitation can improve balance-related outcomes in stroke patients compared with traditional models [66], and recent randomized trials have begun to combine internet-based programs with wearable device-assisted training to support post-discharge recovery [67]. Beyond motor outcomes, randomized evidence has further emerged for cognitive telerehabilitation after stroke, indicating that AI-supported home programs can be evaluated with rigorous trial designs alongside therapist-supervised rehabilitation [18]. In parallel, wearable sensor-based monitoring has matured as an objective substrate for longitudinal follow-up: recent reviews describe how inertial and multimodal wearables—often paired with machine learning—support post-stroke rehabilitation assessment and exercise evaluation [68], while also enabling fall-risk assessment and model development for earlier

risk identification and targeted intervention in both community-dwelling older adults and neurorehabilitation populations [69–71]. Finally, postoperative orthopedics has also started to adopt app- and sensor-supported remote rehabilitation in randomized controlled settings (e.g., after total knee arthroplasty), illustrating a broader clinical pathway for scaling follow-up and adherence management beyond single disease categories [72].

Concurrently, LLMs may serve as conversational virtual rehabilitation assistants for selected low-risk follow-up tasks (e.g., periodic check-ins, structured collection of PROs, and answering frequently asked questions), while maintaining clear escalation pathways to clinicians for ambiguous or high-risk issues [73]. Although early studies in rehabilitation education/Q&A have reported encouraging user-rated satisfaction and perceived empathy in text-based interactions, these findings are heterogeneous and context-dependent, and they do not necessarily imply clinical correctness or safety [19,74]. Recent healthcare evaluations of ChatGPT-like systems have also highlighted variable accuracy, risks of hallucination and overconfident errors, and the need for retrieval grounding, safety guardrails, and human oversight before integration into patient-facing workflows [65,75].

At the system level, the follow-up module is also the primary source of real-world evidence for iterative improvement. As summarised in Figure 2, longitudinal data accumulated during monitoring (e.g., PROs, context signals, and care trajectories) can be curated to identify effective training protocols and early predictive indicators, and then reintegrated to refine upstream components—improving assessment (e.g., functional scoring and risk stratification), enhancing prescription generation and optimization (individualized plans and progression rules), and informing execution and interaction (real-time feedback and safety support). This “follow-up → learning → optimization” pathway aligns with a Data–Knowledge–Action learning health system view of rehabilitation delivery [76]. In summary, foundation model-enabled monitoring can shift follow-up from episodic documentation to more continuous, risk-aware management embedded in daily life.

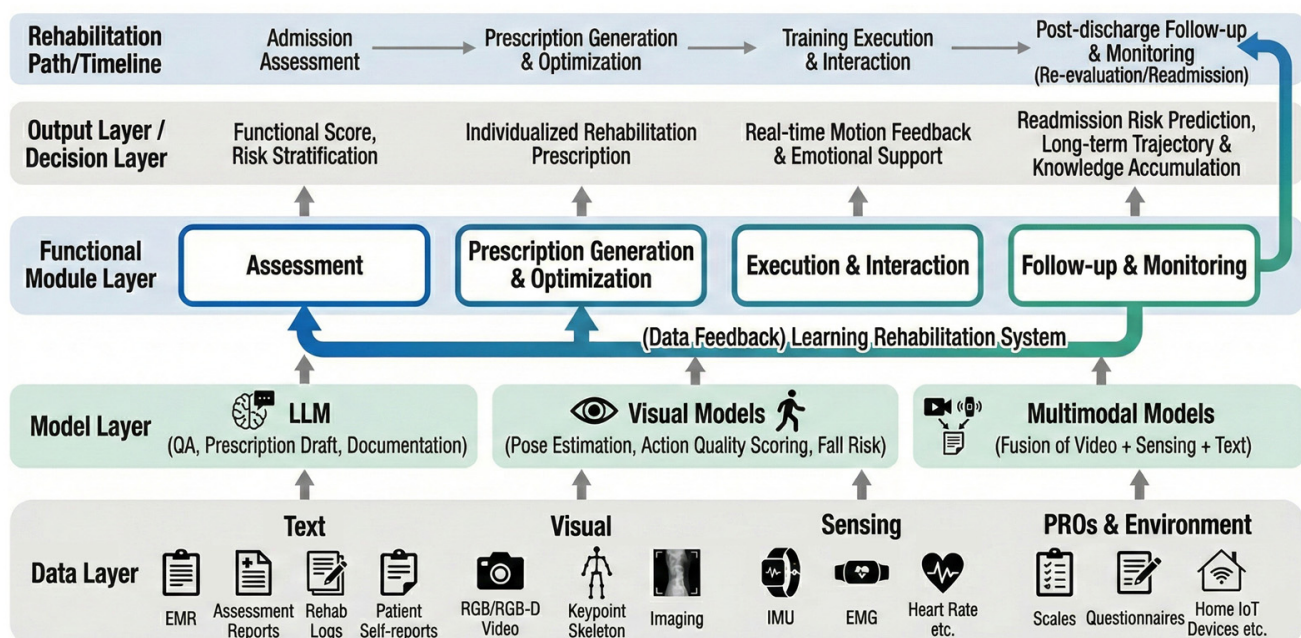


Figure 2. Closed-Loop Framework Diagram of the Four Core Functional Modules and Data Flow in Rehabilitation Foundation Models.

2.5. Therapist–LLM Collaboration for Refining Assessments and Conclusions

While large models enable high-frequency measurement and longitudinal monitoring, translating these outputs into defensible rehabilitation assessments requires an explicit

therapist-in-the-loop collaboration paradigm. In practice, the LLM should function as a drafting and sensemaking layer: it aggregates multimodal signals (e.g., video/skeleton features, wearable time-series, and PROs) together with clinical narratives into a structured assessment synopsis, and—where applicable—grounds key claims with retrievable evidence (e.g., via retrieval-augmented generation) and communicates uncertainty to support review rather than replace judgment. The therapist then performs clinimetric adjudication by verifying plausibility against standardized scales and expected recovery trajectories, reconciling discordant signals (e.g., sensor trends versus subjective reports), and contextualizing interpretations with comorbidities, goals, and environmental constraints before finalizing conclusions and triggering downstream actions (e.g., prescription adjustments or escalation for in-person evaluation). Recent work on clinical deployment of LLMs emphasizes that human verification should be embedded as a system design element—supported by standardized review checklists, auditable interaction logs, and risk-stratified escalation rules for safety-critical outputs—rather than treated as a post hoc safeguard [77–79]. Importantly, therapist feedback can be operationalized as a continuous refinement signal: edited summaries, corrected labels, and rejected recommendations inform prompt updates, knowledge-base curation, or parameter-efficient fine-tuning, and are paired with recurring local validation to monitor performance drift at each deployment site [80]. Evidence from rehabilitation-oriented chatbot evaluations further suggests that patient-perceived “humanization” may diverge from expert-rated clinical correctness, reinforcing the need for therapist oversight and iterative alignment before any model-derived conclusions are communicated to patients or used to trigger care escalation [73,81].

Consistent with the closed-loop view in Figure 2, this therapist–model collaboration is the mechanism that converts longitudinal signals into actionable, auditable decisions: monitoring data are summarised and interpreted, decisions are reviewed and documented, and resulting actions generate new evidence that feeds the next cycle. In this sense, therapist-in-the-loop design is not only a safety requirement but also an enabling condition for learning-oriented improvement over time [82,83].

3. Current Status and Methodologies of Large Model Applications in Rehabilitation

This section reviews the types of large models currently applied in the rehabilitation domain and representative research efforts. It further summarizes the methodological strategies used to implement these applications. We discuss the typical roles and advancements of Large Language Models (LLMs), Vision Models, and Multimodal Models in rehabilitation scenarios, followed by an overview of common adaptation and training techniques—including Retrieval-Augmented Generation (RAG), Parameter-Efficient Fine-Tuning (PEFT), and Self-Supervised Pre-training—highlighting the trade-offs involved in rehabilitation AI development.

3.1. Applications of Large Language Models in Rehabilitation

Leveraging their exceptional natural language understanding and generation capabilities, LLMs have been widely adopted for rehabilitation-related textual tasks [84]. First, regarding knowledge acquisition and Q&A, LLMs serve as rehabilitation knowledge assistants, providing medical staff with evidence-based guidelines and summaries of training strategies, while answering patients’ questions regarding home rehabilitation and health education. Wenjie He et al. [74] found that the general-purpose model ChatGPT v4.0 outperformed the specialized Chinese model ERNIE Bot in accuracy and empathy when responding to Chinese-language web-based patient questions. However, they emphasized that for fine-grained rehabilitation knowledge and localized regulatory norms, integration

with specialized knowledge bases remains necessary. Second, in clinical decision support and documentation, LLMs assist in summarizing complex medical records and assessment reports, generating draft treatment recommendations, and automatically drafting discharge summaries, stage assessments, and personalized educational materials, thereby reducing the documentation burden [85]. Third, concerning patient communication and psychological support, LLM-based dialogue systems can provide emotional companionship and motivational support during long-term rehabilitation, enhancing adherence. However, strict guardrails are required to prevent the models from overstepping boundaries and providing independent diagnostic or treatment decisions [86].

Recent rehabilitation-specific evaluations further illustrate the potential clinical value of LLMs—particularly for patient education and clinician-facing information support—while highlighting the need for expert oversight. In vestibular rehabilitation education, Arbel et al. compared ChatGPT and Google Gemini against clinician responses and reported that LLMs can provide useful educational content on selected tasks but require clinical review for safety-critical recommendations [87]. In sports rehabilitation, McBee et al. used a structured “PanelGPT” simulation to synthesize interdisciplinary perspectives, suggesting a pathway to standardize patient education and return-to-play guidance [88]. In physiotherapy-oriented evaluations, Bilika et al. discussed ChatGPT-assisted clinical reasoning workflows and highlighted governance requirements for safe use [89], while Sawamura et al. further showed that both responses and cited references for physical therapy clinical questions may be inaccurate or unverifiable [90]. Preliminary musculoskeletal-care investigations also explored ChatGPT’s role in decision support and training (e.g., musculoskeletal clinical decision-making and physical-therapy education settings) [85,91]. Finally, emerging clinical evidence in knee osteoarthritis suggests that ChatGPT-assisted education materials may improve patient education outcomes in small clinical cohorts, although readability may still fall short of health-literacy recommendations—underscoring the need for clinician editing and patient-tailored simplification before deployment [92,93].

Concurrently, the application of LLMs in rehabilitation faces challenges such as factual hallucinations, privacy compliance, and ethical boundaries. Here, hallucination denotes generated statements containing fabricated or clinically incorrect facts (e.g., nonexistent history, wrong contraindications, or spurious citations) not supported by the input or verifiable evidence [94]. In rehabilitation decision support, such errors may mislead goal setting, exercise dosing, or risk screening. Common mitigation strategies include evidence grounding via RAG, constrained/structured prompting, uncertainty reporting, and mandatory clinician verification for safety-critical recommendations [56]. To improve response reliability, Xiong et al. [95] proposed NurRAG, a nursing Q&A system based on Retrieval-Augmented Generation (RAG). By retrieving evidence-based guidelines, databases, or patient history during the inference phase and generating answers based on these retrieved results, NurRAG helps reduce the risk of fabrication and enhances traceability. Regarding privacy, Huang et al. [94] proposed that in highly sensitive scenarios like healthcare, regulatory requirements must be met through local deployment, dedicated medical large models, and strict data governance mechanisms. Overall, LLMs demonstrate significant potential in rehabilitation, with future developments expected to yield more specialized models for specific sub-domains (e.g., speech therapy, orthopedic rehabilitation) [96].

3.2. Applications of Vision Large Models in Rehabilitation

Vision-based models are central to rehabilitation because they enable the “datafication” of movement—transforming video into kinematics and quality metrics that can support both assessment and training. In practice, most rehabilitation vision systems rely on task-specific deep models (e.g., pose estimation and action-quality scoring), with

a growing trend toward adopting large-scale pretraining and transformer backbones to improve generalization.

Sardari et al. showed that deep learning-based pose estimation and motion analysis can extract 2D/3D skeleton sequences from video, enabling objective quantification of joint angles and trajectories that correlate with established clinical scales (e.g., Fugl–Meyer) [97]. Mourchid et al. reported that learning-based architectures—such as Spatio-Temporal Graph Convolutional Networks (ST-GCN) and attention-enhanced transformers—can outperform simple rule/threshold heuristics for scoring rehabilitation movement quality [98]. Here, “threshold-based methods” refer to fixed-cutoff rules (e.g., ROM or repetition count exceeding a preset threshold). While interpretable, they are sensitive to noise, viewpoint variation, and compensatory strategies. In contrast, ST-GCN/transformer models exploit spatiotemporal dependencies and inter-joint coordination patterns, producing continuous quality scores more aligned with therapist ratings [59]. These approaches enable a more refined quantification of movement quality metrics, driving the assessment paradigm from “task completion” to “quality quantification”.

In terms of risk control and interaction, vision models can identify compensatory movements and gait patterns associated with high fall risk during training, issuing timely warnings to improve safety [99]. In Human–Computer Interaction (HCI) and Virtual/Augmented Reality (VR/AR) scenarios, pose tracking serves as input for immersive gamified training and rehabilitation robot control, allowing patient movements to be mapped to virtual environments or exoskeletons in real-time [97]. Furthermore, Cardenas et al. [100] noted that beyond wearable and smart assistant frameworks, vision large models can assist in analyzing post-operative X-ray and MRI imaging. This provides a quantitative basis for assessing anatomical changes and fracture/lesion healing, thereby supporting rehabilitation planning with structural and healing information.

Across these use cases, the main limitations are domain shift (camera placement, lighting, occlusion, clothing), calibration/standardization, and clinical validity. Accordingly, vision systems benefit from explicit data-quality monitoring, multi-view/robust pose pipelines, and clinically grounded evaluation (validity, reliability, sensitivity to change) before they can be trusted for closed-loop decision support.

3.3. Applications of Multimodal Large Models in Rehabilitation

Rehabilitation is inherently a highly multimodal data process: functional status emerges from the interaction of movement (video/skeleton), physiology (wearables), context and symptoms (PROs), and clinical narratives. Multimodal large models aim to learn joint representations across these sources, enabling more holistic assessment, risk prediction, and closed-loop intervention. In assessment scenarios, multimodal models can fuse kinematic metrics, physiological signals, and subjective reports (e.g., pain diaries) to output comprehensive scores or risk predictions that closer reflect true functional status. In this review, “personalized risk” refers to an individualized, time-horizon-specific probability estimate of clinically relevant adverse outcomes (e.g., falls, readmission, functional decline, exacerbation, or non-adherence) conditioned on baseline characteristics and longitudinal multimodal signals. “Risk stratification” maps these probabilities into actionable strata (e.g., low/medium/high) using pre-specified thresholds aligned with clinical workflows and escalation pathways. In prescription optimization and closed-loop interventions, models can simultaneously utilize data from cameras and wearable devices to identify movement deformation or excessive fatigue, linking with language modules to generate personalized adjustment suggestions, thus realizing an integrated “Perception–Decision–Feedback” loop [100].

As illustrated in Figure 3, these systems can be conceptualized as a three-layer stack: at the bottom layer, heterogeneous raw inputs are ingested (e.g., video/imagery, wearable time-series signals such as IMU/EMG/heart rate, and text/speech including medical records and patient-reported questionnaires); at the middle layer, these inputs are transformed into feature and representation spaces via traditional handcrafted descriptors, deep representations (e.g., skeleton sequence embeddings and temporal Transformer features), and multimodal alignment representations that bind action and language (e.g., “action + voice guidance”); and at the top layer, these shared representations support the four core functional modules—assessment, prescription, execution & interaction, and follow-up & monitoring—which correspond to scoring/prognosis, plan generation/adaptation, error/compensation detection with feedback, and long-term trajectory modeling with risk warning.

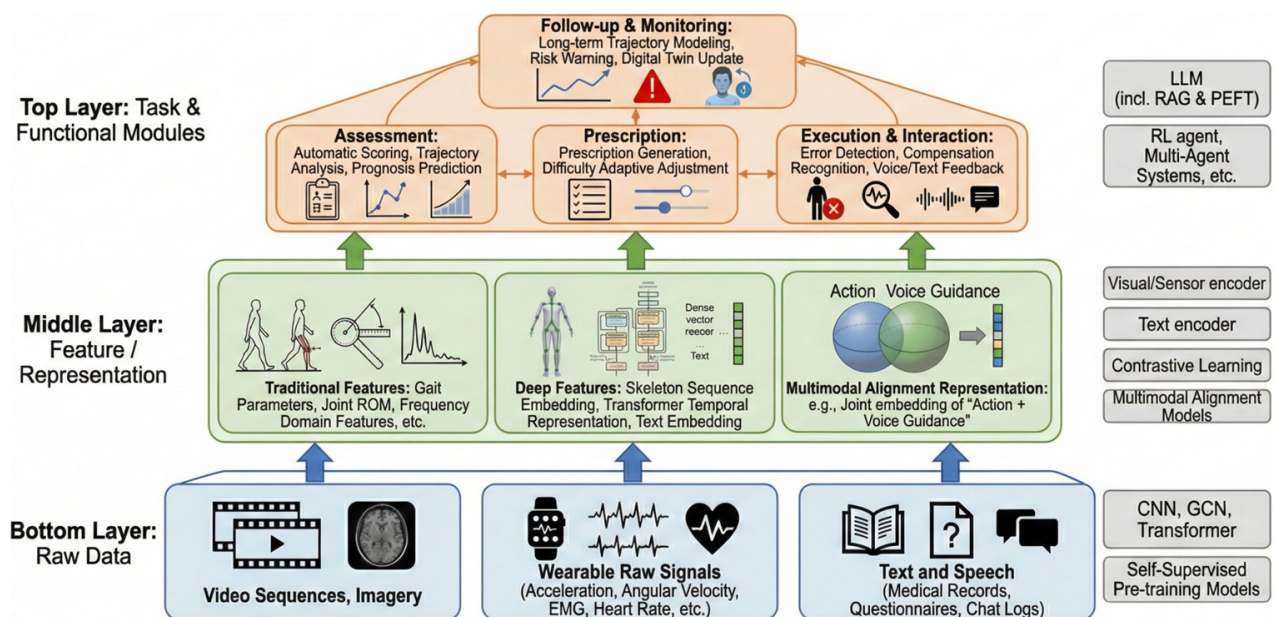


Figure 3. Hierarchical Architecture Diagram of Multimodal Rehabilitation Data.

Conversational multimodal assistants represent a significant recent direction. Zakka et al. highlighted the emerging paradigm where assistants can “see,” “read,” and “converse,” enabling symptom assessment and conservative advice based on patient-uploaded images and narratives [78]. In rehabilitation, “indirect multimodal” approaches are also used—e.g., visualizing time-series sensor data into images and feeding them into vision-language models to leverage existing pretrained capabilities. Xiong et al. further argued that multimodal fusion is a prerequisite for a “rehabilitation digital twin,” integrating structural imaging, functional measures, environment, and psychosocial context to forecast trajectories and optimize strategies [101].

Despite the promise, multimodal large models impose higher requirements on data standardization, interoperability, privacy-preserving sharing, and compute, and their real-world robustness is sensitive to missingness, device drift, and cross-site distribution shifts—issues that must be explicitly addressed in evaluation and governance.

3.4. Adaptation and Training Strategies for Large Models

Given the specialized nature and high reliability requirements of rehabilitation contexts, the direct application of general-purpose large models often fails to meet clinical demands, necessitating the adoption of appropriate adaptation and training strategies—RAG,

PEFT, and self-supervised pretraining—selected according to task type, data availability, and governance constraints.

RAG (retrieval at inference time) improves factual grounding and traceability by conditioning generation on retrieved evidence (guidelines, manuals, institutional protocols, or patient-specific records) [101]. This approach is applicable to use cases prioritizing interpretability and regulatory compliance.

PEFT (lightweight task adaptation) methods such as LoRA and adapters tune a small subset of parameters to align models with domain terminology, structured outputs, and local workflows, making them practical under limited data/compute [102].

Self-supervised pretraining (representation learning) learns generalizable motion or multimodal representations from large unlabeled rehabilitation data (videos, skeleton sequences, wearable signals), improving label efficiency for downstream assessment tasks [103].

These three strategies are not mutually exclusive; real-world systems often adopt a combined approach. For instance, a system might first conduct multimodal self-supervised pre-training based on multi-institutional data to acquire a foundation model for the rehabilitation domain. It then adapts to specific tasks via PEFT and integrates RAG during deployment to access authoritative knowledge bases, thereby enhancing interpretability and safety. In rehabilitation, the “best” combination is ultimately determined by clinical utility and system reliability—especially under therapist-in-the-loop review and lifecycle governance.

3.4.1. Retrieval-Augmented Generation (RAG): Practical Rehabilitation Cases

Rehabilitation applications are often knowledge-intensive and safety-critical, where responses must be traceable to authoritative sources (e.g., device manuals, clinical guidelines, and patient education materials). RAG mitigates hallucinations by retrieving external evidence at inference time and conditioning generation on that evidence.

A practical rehabilitation case is the RehabCoach conversational agent (“Matthias”) for stroke survivors, which uses RAG to answer questions about an upper-limb rehabilitation device (ReHandyBot). In the evaluation, the system generated multiple answers per question under different prompts and data sources (device manual vs. conversational operation dataset). Expert-blinded assessment indicated that RAG answers grounded in conversational operation data were rated more concise and better overall than those grounded solely in the manual, highlighting that retrieval corpus selection (not only the base LLM) is a key determinant of performance in device-assisted rehabilitation Q&A [104].

In musculoskeletal/physical therapy education and consultation scenarios, RAG-style pipelines are also commonly implemented by embedding domain textbooks or curated professional content into a vector database and retrieving relevant passages to support question answering. For example, a Llama2-based physical therapy QA system indexed multiple musculoskeletal diagnosis and treatment textbooks, and the fine-tuned/retrieval-enabled model produced more domain-specific answers than the base model when responding to common physical therapy questions [105]. This case illustrates the suitability of RAG for rehabilitation tasks where correctness and source grounding are prioritized over stylistic adaptation.

3.4.2. Parameter-Efficient Fine-Tuning (PEFT): Practical Rehabilitation Cases

PEFT methods (e.g., LoRA, adapters) update only a small number of trainable parameters to adapt general foundation models to specialized tasks, reducing compute and data requirements compared with full fine-tuning [33]. This is particularly relevant in rehabilitation institutions where compute budgets and labeled data are limited.

A practical case in the physical therapy domain showed that fine-tuning Llama2-13B on ~1.2 million tokens extracted from musculoskeletal diagnosis and therapeutic exercise textbooks consistently improved the specificity and clinical relevance of generated answers compared with the base model. While that study used domain textbooks, the same PEFT paradigm can be applied to rehabilitation-oriented corpora such as therapy notes, discharge summaries, and rehabilitation guideline QA pairs, enabling task customization (e.g., “rehabilitation Q&A” and “rehabilitation report generation”) without updating the entire model.

Compared with RAG, PEFT can improve generation behavior (terminology usage, reasoning patterns, template adherence) even when retrieval is unavailable; however, it requires careful governance of training data quality and periodic re-training when guidelines change, whereas RAG can update knowledge by refreshing the retrieval corpus [102].

3.4.3. Self-Supervised Learning for Rehabilitation Movement Understanding: Practical Cases

For movement-centric rehabilitation tasks (e.g., exercise quality assessment, motion classification, and progress tracking), labels are expensive and inter-rater variability is unavoidable. Self-supervised learning (SSL) addresses this by learning representations from large-scale unlabeled motion data, such as skeleton sequences, videos, or wearable sensor signals, and then transferring them to downstream assessment tasks [103].

A representative practical case is SSL-Rehab, which pretrains a foundation model on 3D skeleton sequences via masked-motion self-supervision and combines self-supervised pretraining with low-rank adaptation for efficient downstream transfer. The method reports improved generalization across public rehabilitation exercise datasets (e.g., KIMORE and UI-PRMD) and supports robust assessment under limited labels, which is consistent with the rehabilitation need for scalable home-based monitoring [106].

In addition, self-supervised or contrastive objectives can be integrated with spatio-temporal graph backbones for rehabilitation exercise quality assessment. Karlov et al. proposed supervised contrastive learning with hard/soft negatives on an ST-GCN backbone and validated improvements across multiple rehabilitation movement quality datasets. These cases collectively indicate that SSL is most effective when the core task is learning motion representations rather than retrieving factual knowledge.

3.4.4. Comparative Analysis Across Rehabilitation Task Types

In rehabilitation, the “effectiveness” of RAG, PEFT, and SSL depends strongly on task type. As summarized in Table 1, these strategies differ in typical tasks, data requirements, key strengths, and key limitations. Knowledge-intensive tasks (patient education, device operation Q&A, guideline-based decision support) benefit most from RAG due to traceability and rapid knowledge updating [104]. Tasks requiring domain-specific language style and structured outputs (rehabilitation documentation, domain-specific consultation) benefit from PEFT/LoRA-style adaptation. Movement-centric tasks (exercise quality scoring, progress estimation) benefit most from SSL representation learning on multimodal motion data.

In practice, hybrid systems are increasingly common: SSL pretraining provides motion encoders, PEFT aligns multimodal reasoning and output formats, and RAG injects up-to-date guideline evidence at deployment to enhance interpretability and safety.

Table 1. Comparison of Adaptation and Training Strategies for Large Models in Rehabilitation Domain.

Strategy	Typical Tasks	Data Requirement	Key Strengths	Key Limitations
RAG	Device operation Q&A; patient education; guideline-grounded counseling	Needs curated retrievable knowledge base (manuals, guidelines, education materials)	Traceable evidence; mitigates hallucination via grounding [101]	Retrieval quality bottleneck; limited improvement in style/behavior without tuning
PEFT (LoRA/Adapter)	Rehab Q&A templates; rehab documentation/report generation; domain-specific consultation	Needs domain dataset (QA pairs, notes, templates); smaller than full FT	Improves terminology/style and structured outputs; low compute compared with full FT [33,102]	Requires governance of training data; may lag behind newest guidelines unless combined with RAG
Self-supervised learning	Exercise quality assessment; motion classification/scoring; progress tracking	Requires large-scale unlabeled motion data; few labels for downstream	Strong motion representations; better label efficiency; suits home-based monitoring [21,106]	Less direct for factual QA; needs careful clinical validation and robustness checks

3.5. Core Limitations of Current LLM Applications in Rehabilitation

Across the rehabilitation workflow, existing LLM deployments remain early-stage and exhibit recurring limitations that hinder safe translation. First, evidence and validation gaps persist: many studies are small, context-specific, and focused on intermediate endpoints (e.g., satisfaction or text quality), while prospective clinical impact, transparent reporting, and recurring local validation remain limited for heterogeneous rehabilitation settings [107]. Second, grounding and hallucination risks are non-trivial even in seemingly low-risk education or counseling; LLMs may generate overconfident errors or fabricated references, which can mislead exercise dosing, contraindication screening, and goal setting. Third, privacy, governance, and integration constraints restrict large-scale use of sensitive longitudinal rehabilitation data and complicate continuous learning across sites and devices, necessitating stricter lifecycle controls and auditing [108,109]. Fourth, equity, accountability, and human factors remain under-addressed: healthcare LLMs can reproduce biased content and show subgroup performance disparities, and rehabilitation populations may be particularly vulnerable without participatory design, explicit accountability, and human-override pathways [108,110,111]. Collectively, these limitations motivate the rehabilitation-oriented evaluation framework and the governance/ethics considerations.

4. Evaluation Framework for Rehabilitation-Oriented Large Models

To be clinically deployable, rehabilitation-oriented large models must be both “clinically beneficial to patients” and “systemically reliable and scalable.” Therefore, it is essential to construct a comprehensive evaluation framework based on two dimensions: clinimetric properties and machine learning metrics. The former addresses whether the model output is genuinely meaningful for clinical practice, while the latter addresses whether the model, as an engineering system, is stable, safe, and generalizable; both are indispensable.

Because rehabilitation depends on longitudinal monitoring and continuous data streams, model performance alone is insufficient for translation. We recommend treating data-quality and implementation indicators as core endpoints, as they condition real-world validity and safety. Reports should include wear-time distributions, missingness rates, device/sensor failure rates, known error sources (measurement bias, motion artifacts,

occlusion/viewpoint shifts), and mitigation strategies (quality filtering, wear-threshold definitions, uncertainty-aware modeling, and periodic re-validation against reference standards when feasible). In parallel, usability and acceptability outcomes (e.g., System Usability Scale, comfort, perceived burden, intention/continuation of use) should be treated as primary endpoints because they determine whether longitudinal monitoring is sustainable—particularly in older adults and functionally impaired users.

4.1. Clinical Utility: Clinimetric Evaluation (Validity, Reliability, Responsiveness, Safety)

At the level of clinical utility, evaluation can be conducted using clinimetrics, focusing on validity, reliability, Minimal Clinically Important Difference (MCID), and safety, with endpoints prioritizing functional benefit and participation in daily life rather than statistical significance alone.

Validity assesses whether the model output truly reflects the target functional status, such as the consistency between pose estimation or gait analysis results and gold-standard scales or motion capture systems. Reliability emphasizes the consistency of results across repeated measurements, different devices, and varying environmental contexts. This is typically quantified by metrics such as the Intra-class Correlation Coefficient (ICC) and Coefficient of Variation (CV) to determine the model's suitability for longitudinal monitoring [112]. MCID links statistical changes to “whether the patient truly perceives improvement.” This requires the model not only to detect change but also to identify whether clinically meaningful thresholds are reached, avoiding the interpretation of random fluctuations as progress or, conversely, the oversight of significant improvements [113]. Safety evaluation focuses on whether model intervention increases patient risk—for instance, whether erroneous feedback could induce dangerous movements or delay treatment. Systematic monitoring of adverse events in clinical trials and real-world applications is required to ensure error rates do not exceed acceptable human levels and that error-correction mechanisms are in place [114]. Overall, clinical utility evaluation should be patient-centered, prioritizing Patient-Reported Outcomes (PROs) and functional benefits in daily life over mere statistical significance.

To operationalize reliability (ICC) and MCID in LLM-enabled rehabilitation scenarios, the evaluation protocol should explicitly define the “measurement target” (e.g., patient/session/video segment), the “rater” (e.g., clinician panel and/or the LLM instance), and the rating design (complete vs. incomplete) before selecting an ICC form. Updated ICC selection guidelines based on generalizability theory recommend treating raters as random effects in most clinical settings and provide practical options for unbalanced or incomplete rating designs, which are common in real-world rehabilitation services [115]. For LLM pipelines that output continuous functional estimates (e.g., predicted scale scores or movement-quality scores), test-retest reliability can be implemented by running the identical pipeline twice (or multiple times) on the same inputs under a stable-condition window; ICC should be reported together with the standard error of measurement (SEM) and minimal detectable change (MDC) so that longitudinal changes are interpreted against measurement error rather than noise [116,117]. MCID should then be implemented as a “responder” criterion using anchor-based approaches whenever feasible (e.g., patient global rating of change or clinician global impression), complemented by distribution-based sensitivity checks, following recent recommendations on constructing credible anchor-based minimal important difference values [118,119].

In practice, for an LLM-assisted telerehabilitation coaching study, investigators can report the proportion of patients exceeding published MCID thresholds for the selected outcomes (e.g., stroke-related scales summarized in recent reviews) and require that the observed change exceed MDC to be considered both statistically reliable and clinically meaningful [120].

4.2. Engineering and Deployment Reliability: Generalization, Robustness, Fairness, Efficiency, Interpretability

At the technical performance level, the engineering reliability of large models must be assessed regarding generalization, robustness, fairness, and efficiency / real-time capability. As summarized in Figure 4, these engineering/system dimensions are not isolated technical checks; rather, they function as cross-cutting conditions that determine whether clinimetric targets—validity, reliability, MCID, and safety—remain stable across institutions, devices, and real-world deployment constraints.

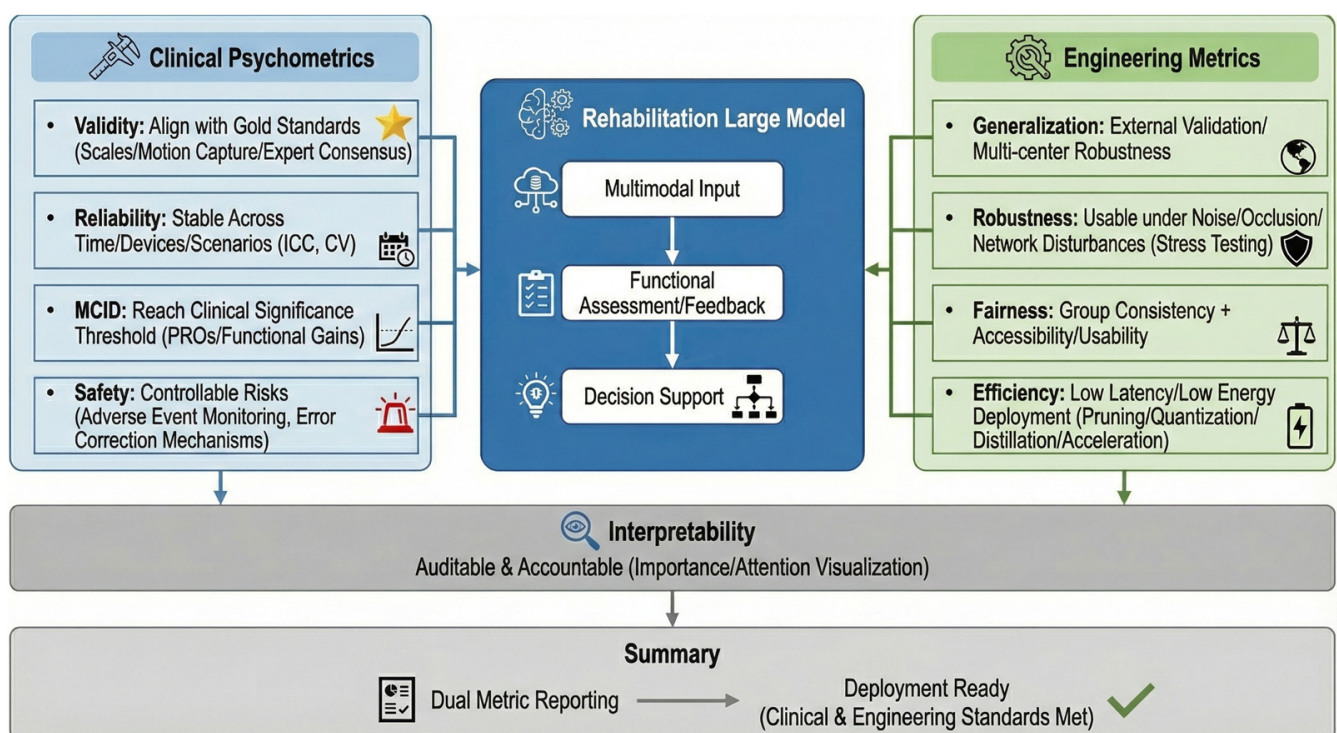


Figure 4. Hierarchical Architecture of Multimodal Rehabilitation Data.

Generalization: Rehabilitation data is highly heterogeneous, with significant differences in distribution across institutions, populations, and devices. Therefore, models must maintain stable performance in external validation sets and multicenter testing to avoid the scenario of being “excellent on the training set but failing in clinical practice.”

Robustness: This requires the model to provide acceptable results despite interference such as noise, occlusion, non-standard poses, or network fluctuations. Stress testing is necessary to verify the model’s sensitivity to input perturbations.

Fairness: This emphasizes consistent performance across populations of different ages, genders, ethnicities, and functional levels to prevent data bias from exacerbating health inequalities [110,121,122]. Additionally, the usability of interfaces and interactions for people with disabilities must be considered.

Efficiency and Real-time Capability: These directly impact deployability and accessibility. Since rehabilitation applications often require low-latency feedback and run on resource-constrained home or mobile terminals, techniques such as pruning, quantization, distillation, and dedicated hardware acceleration should be employed to control inference time and energy consumption while maintaining performance [32,33].

Interpretability: Spanning both clinical and technical dimensions, this requires the model to provide the rationale for its conclusions in a human-understandable manner to support therapist review and accountability (e.g., explaining why functional decline or increased fall risk is predicted via feature importance or attention visualization).

In conclusion, the evaluation of rehabilitation-oriented large models should adopt a “Dual-Metric Reporting” mode: on one hand, systematically presenting clinimetric results (validity, reliability, MCID, and safety) to demonstrate “proven benefit to patients and clinical decision-making”; on the other hand, reporting engineering metrics (generalization, robustness, fairness, and efficiency) to ensure “stable and reliable operation in real-world scenarios.” Only when standards are met in both the clinical utility and technical performance dimensions will rehabilitation large models possess the sufficient conditions to transition from experimental environments to routine clinical practice.

5. Policy and Ethical Considerations

5.1. Risk-Based Regulation and Lifecycle Control

As medical AI that directly impacts patient safety, Rehabilitation large models can influence functional assessment, prescription decisions, and behavior change over long horizons, often extending into home environments. They therefore must be integrated into existing medical device and digital health regulatory pathways using risk-based, intended-use-driven classification. Low-risk systems that provide general health information may follow lighter pathways, whereas systems performing functional assessment, risk stratification, or prescription/parameter recommendations that can influence clinical decisions require stricter evaluation, clinical validation, and post-market surveillance. This aligns with a Total Product Lifecycle (TPLC) approach spanning design controls, quality systems, premarket review, and continuous post-market monitoring [123,124].

For adaptive or frequently updated models, controlled change management becomes essential. Regulatory guidance increasingly operationalizes this via a Predefined Change Control Plan (PCCP)—predefining the types of modifications, verification/validation procedures, and regression testing needed to keep updates transparent, traceable, and verifiable, while linking changes to post-market performance monitoring [125].

5.2. Evidence Generation, Transparency, Accountability, and Fairness

For evidence generation, rehabilitation large models should undergo ethics review and prospective studies before routine deployment, using endpoints aligned with rehabilitation goals (functional scale improvement, ADL participation, adherence, resource utilization, and care burden). Safety and effectiveness should be demonstrated as non-inferior to standard care where appropriate, followed by continuous real-world evidence (RWE) monitoring to assess long-term performance, drift, and rare adverse events. This is particularly important in rehabilitation, where outcomes depend on sustained engagement and repeated interactions, and short-term offline metrics may not reflect real-world benefit.

Governance for clinical adoption requires transparency and accountability: developers should disclose data characteristics, intended-use boundaries, performance, and known failure modes; patients should receive understandable disclosure and consent mechanisms; therapists should be supported by explainable collaborative interfaces that present evidence,

comparable cases, confidence/uncertainty, and escalation recommendations—reinforcing human-in-the-loop/ultimate human responsibility. EU guidance on MDR/IVDR–AI Act interplay also emphasizes logging, documentation, and human oversight expectations for high-risk medical AI to enable quality control and post-market processes.

Fairness is central because rehabilitation populations vary widely in age, functional impairment, comorbidity burden, and digital literacy. Under-representation across key dimensions can yield subgroup performance disparities and inequitable recommendations. Health-AI ethics and risk-management guidance highlights the need for bias checks, subgroup reporting, and ongoing local validation, with accessibility and user burden treated as first-order considerations.

For risk prediction/stratification outputs (e.g., falls, readmissions, adherence drop-off), clinical utility hinges on calibration, external validation, and ongoing drift monitoring. Reporting and risk-of-bias assessment should follow emerging guidance such as TRIPOD + AI and PROBAST + AI, including structured disclosure of data sources, missingness handling, threshold selection, calibration performance, and subgroup differences. Risk should be communicated as probability with uncertainty to support triage and escalation, not to replace clinician judgment.

5.3. Cross-Jurisdiction Landscape and Rehabilitation-Specific Ethical Dilemmas

Across jurisdictions, the U.S. commonly regulates many clinically functional systems under the SaMD paradigm and emphasizes evidence generation, risk management, and post-market monitoring through the AI/ML SaMD Action Plan, GMLP, and a TPLC approach; PCCP guidance further operationalizes controlled update pathways. The EU AI Act (Regulation (EU) 2024/1689) introduces obligations for high-risk AI and can apply when AI is used in or as a safety component of regulated medical devices; MDCG 2025-6 clarifies documentation and post-market alignment between MDR/IVDR and the AI Act. China follows a risk-based medical device classification system (Class I/II/III) and, for home-based rehabilitation using video/wearables, must also comply with personal information protection requirements—particularly for health-related sensitive data, which requires stricter protections and consent mechanisms.

Rehabilitation’s long-horizon, home-extended, and vulnerable population nature makes several ethical dilemmas especially salient:

- “Virtual therapist” responses can be overconfident or misleading; safety-critical recommendations require grounding, uncertainty disclosure, and clinician escalation channels. WHO guidance on generative AI/LMMs in health highlights transparency, risk control, and meaningful human oversight.
- Motion assessment/prognosis models may fail more often in older adults and under-represented or high-severity subgroups, motivating continuous local validation, drift monitoring, and conservative fallback-to-human rules.
- Adherence and early warning benefit from continuous data capture but increase privacy and psychological burden; implement data minimization, privacy-by-design, granular consent/withdrawal, and auditable access controls.
- Without enforced versioning, audit logs, and standardized change documentation, frequent updates can blur responsibility; PCCP-linked post-market surveillance helps preserve traceability and safety.

5.4. Outlook: Governed Innovation and Equitable Translation

Looking forward, rehabilitation large models possess significant potential for technological and clinical integration. Firstly, leveraging multimodal perception and long-term memory capabilities, large models are poised to evolve into personalized rehabilitation

assistants spanning hospitalization and home care. By dynamically perceiving patient functional and emotional states, they can adaptively adjust training content and intensity while providing movement correction and emotional support, thereby enhancing adherence and user experience. Secondly, combined with digital twins and reinforcement learning, these models are expected to construct updatable “Virtual Rehabilitation Entities” for each patient. This allows for the simulation of different rehabilitation strategies and assistive device configurations in a virtual environment, with optimized results fed back into real-world interventions, reducing trial-and-error costs and improving decision-making precision. Thirdly, large models can be deeply coupled with emerging technologies such as Brain–Computer Interfaces (BCI), Augmented Reality (AR), intelligent robotics, and wearable devices. By forming multi-device collaborative human–machine systems through unified data and communication standards, their function will extend from “providing suggestions” to “collaborative execution,” supporting higher-intensity and more complex rehabilitation training [126].

Simultaneously, the large-scale application of rehabilitation large models must balance inclusivity with ethical constraints. On one hand, barriers to entry should be lowered through open-source models, shared datasets, and low-cost terminals, prioritizing accessibility and usability for the elderly and individuals with disabilities to avoid exacerbating health inequalities. Continuous monitoring with wearables may amplify inequities if older adults or low-literacy groups experience higher burden, lower adoption, or early discontinuation. Evidence suggests that long-term integration of wearables in older adults’ daily life depends on a support structure (training, troubleshooting, and motivational scaffolding) and on devices meeting expectations of reliability and accuracy, rather than on feature abundance alone. On the other hand, new issues such as patient over-reliance on AI, blurred doctor-patient boundaries, and the privacy and psychological burdens brought by continuous monitoring must be addressed. These challenges should be managed through dynamically updated governance frameworks and feedback mechanisms from professional societies and regulatory bodies [127].

As illustrated in Figure 5, the clinical translation of large models in rehabilitation can be conceptualized as a holistic framework bridging the paradigm of “Regulation-Evidence-Governance” with “Future Application Pathways.” The upper layer is constituted by three foundational pillars: risk stratification and full-lifecycle regulation, evidence generation and continuous evaluation, and transparency and accountability governance. The middle layer emphasizes the core principles of patient-centeredness, safety and equity, and integration into clinical workflows. Furthermore, it mandates human-in-the-loop verification for critical decisions and the maintenance of auditable human–computer interaction records. Finally, the lower layer delineates three distinct application pathways: personalized rehabilitation assistants, virtual rehabilitation entities, and multi-device collaborative execution. In conclusion, the future development of rehabilitation large models depends not only on algorithmic and hardware advances, but also on validated sensing fidelity and sustained user adoption/adherence of wearable-based data streams. Only under the premises of safety, fairness, and patient-centeredness can their technical advantages be translated into sustainable clinical and societal value.

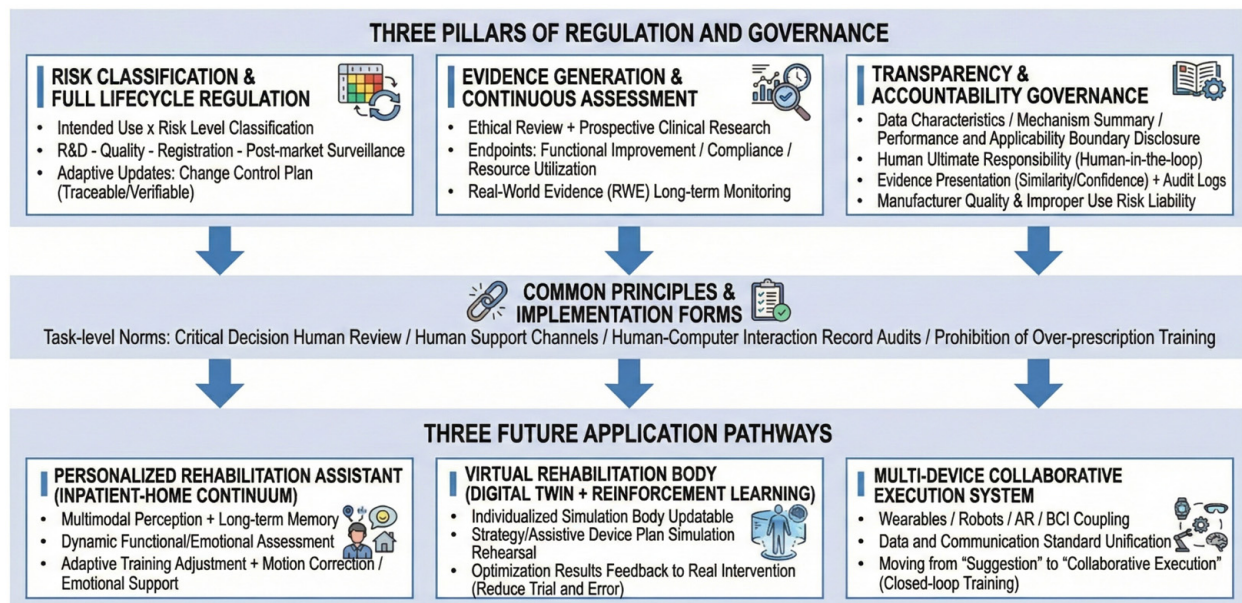


Figure 5. Regulatory pillars and future application pathways for foundation models in rehabilitation.

6. Conclusions

As an emerging frontier at the intersection of artificial intelligence and rehabilitation, large rehabilitation models are reshaping traditional paradigms across the entire workflow—from assessment and prescription formulation to training execution and follow-up monitoring. They provide novel technical pathways to address long-standing issues such as insufficient objectivity, lack of personalization, and weak continuity between hospital and home settings. Centering on this theme, this paper first outlined the critical challenges and needs within rehabilitation services, defining data flows and the specific links where large models empower the four key stages: assessment, intervention, interaction, and monitoring. Subsequently, we reviewed representative advancements in Large Language Models (LLMs), vision models, and multimodal models within rehabilitation scenarios, discussing the roles of technologies such as Retrieval-Augmented Generation (RAG), fine-tuning, and self-supervised learning. On this basis, a rehabilitation-oriented evaluation framework was proposed, emphasizing the simultaneous use of clinimetric properties to assess the genuine significance for patient outcomes and engineering and system metrics to guarantee model reliability and scalability. Furthermore, from regulatory and ethical perspectives, we discussed key issues including risk stratification, clinical validation, transparency, explainability, and accountability, advocating for the construction of compliance and ethical alignment mechanisms centered on clinical scenarios. Finally, the paper looked forward to future directions such as intelligent personalized rehabilitation assistants, digital twin-driven precision rehabilitation, deep interdisciplinary fusion, and inclusive applications, while noting the need for continued attention to emerging ethical and social challenges.

Although fields such as neurorehabilitation, orthopedic rehabilitation, geriatric rehabilitation, and chronic disease rehabilitation differ in their service populations and clinical focus, they share a high degree of similarity in the task workflow of “Assessment–Intervention–Monitoring.” Consequently, all these domains can achieve closed-loop optimization through the application of large models. However, significant heterogeneity exists regarding functional goals, data modalities, feedback granularity, and personalized needs, necessitating the employment of targeted model adaptation strategies to balance these differences.

Overall, large models are more likely to serve as “capability amplifiers” for rehabilitation professionals rather than replacements. By undertaking high-load data processing and monitoring tasks, they enable therapists to devote more energy to clinical decision-making and humanistic care, while ensuring patients receive more continuous and personalized support. Realizing this vision relies on interdisciplinary collaboration across medicine, engineering, and policy, as well as rigorous validation and prudent governance based on evidence-based medicine. Only by strictly adhering to the core goals of improving patient function and quality of life, and placing technical innovation under the framework of clinical evidence and ethical norms, can large rehabilitation models truly translate into sustainable and scalable clinical value.

Author Contributions: Conceptualization, T.B., W.Z. and S.Q.; methodology, T.B., C.W. and B.W.; Formal analysis, T.B.; investigation, K.J. and Y.Y.; resources, W.Z., S.Q. and C.W.; data curation, T.B., K.J. and Y.Y.; writing—original draft preparation, T.B. and K.J.; writing—review and editing, S.Q., C.W. and B.W.; visualization, T.B., K.J. and Y.Y.; supervision, S.Q., C.W. and B.W.; project administration, W.Z.; funding acquisition, W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially supported by the National Key R&D Program of China 2025YFE0117700.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Glossary

The Glossary Table defines key terms and abbreviations involved in this paper.

Term/Abbrev	Definition
AI/Models and Methods	
Transformer	Self-attention-based neural network architecture widely used as the backbone of modern foundation models in NLP, vision, and multi-modal learning.
Large Language Model (LLM)	Transformer-based foundation model trained on large-scale text corpora for natural language understanding and generation.
Vision Model	Deep neural network for image/video understanding (e.g., recognition, segmentation, or pose estimation).
Multimodal Large Model/MLLM	Foundation model that jointly learns from multiple modalities (e.g., text + images/video + audio/sensors) to enable cross-modal representation and reasoning.
Prompt/Prompt Engineering	Natural-language instructions and structured examples used to steer model behavior; includes patterns such as zero-shot/few-shot and role prompting.
In-Context Learning (ICL)	Model adapts behavior from demonstrations provided in the prompt without parameter updates.
Retrieval-Augmented Generation (RAG)	Pipeline that retrieves external evidence and conditions generation on retrieved context to improve grounding/traceability.
Parameter-Efficient Fine-Tuning (PEFT)	Fine-tuning strategy that updates only a small subset of parameters (e.g., LoRA/adapters/prefix tuning) to adapt an LLM efficiently.
Hallucination (factual)	Generated statements that are unsupported by provided evidence or contradict reliable references (high-risk in healthcare).
Rehabilitation Data and Sensing	
EHR/EMR	Electronic health/medical records: structured and unstructured clinical documentation used for care and decision support.

Pose Estimation/ HPE	Human Pose Estimation: estimating body keypoints/skeleton from images/videos for kinematic analysis.
Range of Motion (ROM)	Joint motion extent (e.g., degrees) used as a functional assessment indicator.
MoCap	Marker-based motion capture providing high-precision kinematics in lab/clinic settings.
RGB-D	Color + depth video used for improved kinematic estimation in rehabilitation settings.
ADL	Activities of Daily Living: functional abilities in everyday tasks (often an outcome endpoint).
PROs	Patient-Reported Outcomes: patient self-reported health/function measures used for monitor
Evaluation/Regulation	and Clinical Workflow
ICC/CV	Intraclass correlation coefficient/coefficient of variation for reliability quantification.
MCID	Minimal clinically important difference: the smallest change perceived as beneficial/meaningful by patients.
RWE	Real-world evidence: evidence generated from routine practice and observational data streams.
Human-in-the-loop/Guardrails	Design pattern where clinicians retain final responsibility; system includes constraints/refusals/escalation for high-risk outputs.

References

1. World Health Organization. International Classification of Functioning, Disability and Health (ICF). Available online: <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health> (accessed on 29 January 2026).
2. World Health Organization. Rehabilitation 2030: A Call for Action. Meeting Report. 7 February 2017. Available online: <https://www.who.int/publications/m/item/rehabilitation-2030-a-call-for-action> (accessed on 29 January 2026).
3. World Health Organization. Landmark Resolution on Strengthening Rehabilitation in Health Systems. News Release. 27 May 2023. Available online: <https://www.who.int/news/item/27-05-2023-landmark-resolution-on-strengthening-rehabilitation-in-health-systems> (accessed on 29 January 2026).
4. Cieza, A.; Causey, K.; Kamenov, K.; Hanson, S.W.; Chatterji, S.; Vos, T. Global Estimates of the Need for Rehabilitation Based on the Global Burden of Disease Study 2019: A Systematic Analysis for the Global Burden of Disease Study 2019. *Lancet* **2020**, *396*, 2006–2017. [CrossRef] [PubMed]
5. Gimigliano, F.; Negrini, S. The World Health Organization “Rehabilitation 2030: A Call for Action”. *Eur. J. Phys. Rehabil. Med.* **2017**, *53*, 155–168. [CrossRef] [PubMed]
6. Lv, Y.; Fan, L.; Zhou, J.; Ding, E.; Shen, J.; Tang, S.; He, Y.; Shi, X. Burden of Non-Communicable Diseases Due to Population Ageing in China: Challenges to Healthcare Delivery and Long Term Care Services. *BMJ* **2024**, *387*, e076529. [CrossRef] [PubMed]
7. United Nations, Department of Economic and Social Affairs (UN DESA). World Population Ageing 2023: Challenges and Opportunities of Population Ageing in the Least Developed Countries. Available online: <https://desapublications.un.org/publications/world-population-ageing-2023-challenges-and-opportunities-population-ageing-least> (accessed on 29 January 2026).
8. World Health Organization. Noncommunicable Diseases. Fact Sheet. Updated 25 September 2025. Available online: <https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases> (accessed on 29 January 2026).
9. National Bureau of Statistics of China. Main Data of the Seventh National Population Census. News Release. 11 May 2021. Available online: https://www.stats.gov.cn/english/PressRelease/202105/t20210510_1817185.html (accessed on 29 January 2026).
10. Alshami, A.; Nashwan, A.; AlDardour, A.; Qusini, A. Artificial Intelligence in Rehabilitation: A Narrative Review on Advancing Patient Care. *Rehabilitación* **2025**, *59*, 100911. [CrossRef]
11. Bao, J.; Zhou, L.; Liu, G.; Tang, J.; Lu, X.; Cheng, C.; Jin, Y.; Bai, J. Current State of Care for the Elderly in China in the Context of an Aging Population. *Biosci. Trends* **2022**, *16*, 107–118. [CrossRef]
12. Tian, T.; Zhu, L.; Fu, Q.; Tan, S.; Cao, Y.; Zhang, D.; Wang, M.; Zheng, T.; Gao, L.; Volontovich, D.; et al. Needs for Rehabilitation in China: Estimates Based on the Global Burden of Disease Study 1990–2019. *Chin. Med. J.* **2025**, *138*, 49–59. [CrossRef]

13. Geng, F.; Liu, Z.; Yan, R.; Zhi, M.; Grabowski, D.C.; Hu, L. Post-Acute Care in China: Development, Challenges, and Path Forward. *J. Am. Med. Dir. Assoc.* **2024**, *25*, 61–68. [\[CrossRef\]](#)
14. Tu, W.-J.; Hua, Y.; Yan, F.; Bian, H.; Yang, Y.; Lou, M.; Kang, D.; He, L.; Chu, L.; Zeng, J.; et al. Prevalence of Stroke in China, 2013–2019: A Population-Based Study. *Lancet Reg. Health West. Pac.* **2022**, *28*, 100550. [\[CrossRef\]](#)
15. Ma, Q.; Li, R.; Wang, L.; Yin, P.; Wang, Y.; Yan, C.; Ren, Y.; Qian, Z.; Vaughn, M.G.; McMillin, S.E.; et al. Temporal Trend and Attributable Risk Factors of Stroke Burden in China, 1990–2019: An Analysis for the Global Burden of Disease Study 2019. *Lancet Public Health* **2021**, *6*, e897–e906. [\[CrossRef\]](#)
16. Tu, W.-J.; Zhao, Z.; Yin, P.; Cao, L.; Zeng, J.; Chen, H.; Fan, D.; Fang, Q.; Gao, P.; Gu, Y.; et al. Estimated Burden of Stroke in China in 2020. *JAMA Netw. Open* **2023**, *6*, e231455. [\[CrossRef\]](#)
17. Deslauriers, S.; Déry, J.; Proulx, K.; Laliberté, M.; Desmeules, F.; Feldman, D.E.; Perreault, K. Effects of Waiting for Outpatient Physiotherapy Services in Persons with Musculoskeletal Disorders: A Systematic Review. *Disabil. Rehabil.* **2021**, *43*, 611–620. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Laver, K.E.; Adey-Wakeling, Z.; Crotty, M.; Lannin, N.A.; George, S.; Sherrington, C. Telerehabilitation Services for Stroke. *Cochrane Database Syst. Rev.* **2020**, *2020*, CD010255. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Molina-Garcia, P.; Mora-Traverso, M.; Prieto-Moreno, R.; Díaz-Vásquez, A.; Antony, B.; Ariza-Vega, P. Effectiveness and Cost-Effectiveness of Telerehabilitation for Musculoskeletal Disorders: A Systematic Review and Meta-Analysis. *Ann. Phys. Rehabil. Med.* **2024**, *67*, 101791. [\[CrossRef\]](#)
20. Barger, S.; Castellini, G.; Vitale, J.A.; Guida, S.; Banfi, G.; Gianola, S.; Pennestrì, F. Effectiveness of Telemedicine for Musculoskeletal Disorders: Umbrella Review. *J. Med. Internet Res.* **2024**, *26*, e50090. [\[CrossRef\]](#)
21. Amin, J.; Ahmad, B.; Amin, S.; Siddiqui, A.A.; Alam, M.K. Rehabilitation Professional and Patient Satisfaction with Telerehabilitation of Musculoskeletal Disorders: A Systematic Review. *BioMed Res. Int.* **2022**, *2022*, 7366063. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Zhu, Z.; Shi, M.; Yu, Q.; Fei, J.; Song, B.; Qin, X.; Sun, L.; Zhang, Y. Burden and Risk Factors of Stroke Worldwide and in China: An Analysis from the Global Burden of Disease Study 2021. *Chin. Med. J.* **2025**, *138*, 2588–2595. [\[CrossRef\]](#)
23. Nazi, Z.A.; Peng, W. Large Language Models in Healthcare and Medical Domain: A Review. *Informatics* **2024**, *11*, 57. [\[CrossRef\]](#)
24. Tan, S.Y.; Sumner, J.; Wang, Y.; Wenjun Yip, A. A Systematic Review of the Impacts of Remote Patient Monitoring (RPM) Interventions on Safety, Adherence, Quality-of-Life and Cost-Related Outcomes. *npj Digit. Med.* **2024**, *7*, 192. [\[CrossRef\]](#)
25. Tharatipyakul, A.; Srikaewsiew, T.; Pongnumkul, S. Deep Learning-Based Human Body Pose Estimation in Providing Feedback for Physical Movement: A Review. *Heliyon* **2024**, *10*, e36589. [\[CrossRef\]](#)
26. Gu, B.; Kim, H.S.; Kim, H.; Yoo, J.-I. Advancements in Wearable Sensor Technologies for Health Monitoring in Terms of Clinical Applications, Rehabilitation, and Disease Risk Assessment: Systematic Review. *JMIR mHealth uHealth* **2026**, *14*, e76084. [\[CrossRef\]](#)
27. Alam, S.; Zhang, M.; Harris, K.; Fletcher, L.M.; Reneker, J.C. The Impact of Consumer Wearable Devices on Physical Activity and Adherence to Physical Activity in Patients with Cardiovascular Disease: A Systematic Review of Systematic Reviews and Meta-Analyses. *Telemed. e-Health* **2023**, *29*, 986–1000. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Sumner, J.; Lim, H.W.; Chong, L.S.; Bunde, A.; Mukhopadhyay, A.; Kayambu, G. Artificial Intelligence in Physical Rehabilitation: A Systematic Review. *Artif. Intell. Med.* **2023**, *146*, 102693. [\[CrossRef\]](#)
29. Sekhon, M.; Cartwright, M.; Francis, J.J. Acceptability of Healthcare Interventions: An Overview of Reviews and Development of a Theoretical Framework. *BMC Health Serv. Res.* **2017**, *17*, 88. [\[CrossRef\]](#)
30. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In *Proceedings of the Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30.
31. OpenAI; Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; Akkaya, I.; Aleman, F.L.; Almeida, D.; Altenschmidt, J.; Altman, S.; et al. GPT-4 Technical Report. *arXiv* **2023**, arXiv:2303.08774. [\[CrossRef\]](#)
32. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. Learning Transferable Visual Models from Natural Language Supervision. *Proc. Mach. Learn. Res.* **2021**, *139*, 8748–8763.
33. Hu, E.J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W. LoRA: Low-Rank Adaptation of Large Language Models. *arXiv* **2021**, arXiv:2106.09685v1.
34. Avogaro, A.; Cunico, F.; Rosenhahn, B.; Setti, F. Markerless Human Pose Estimation for Biomedical Applications: A Survey. *Front. Comput. Sci.* **2023**, *5*, 1153160. [\[CrossRef\]](#)
35. Scataglini, S.; Abts, E.; Van Bocxlaer, C.; Van Den Bussche, M.; Meletani, S.; Truijen, S. Accuracy, Validity, and Reliability of Markerless Camera-Based 3D Motion Capture Systems versus Marker-Based 3D Motion Capture Systems in Gait Analysis: A Systematic Review and Meta-Analysis. *Sensors* **2024**, *24*, 3686. [\[CrossRef\]](#)
36. Liu, Y.; Li, Y.; Zhang, Z.; Huo, B.; Dong, A. Quantitative Evaluation of Motion Compensation in Post-Stroke Rehabilitation Training Based on Muscle Synergy. *Front. Bioeng. Biotechnol.* **2024**, *12*, 1375277. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Ding, K.; Wang, J.; Wang, X.; Zhou, L.; Xiong, D.; Guo, L. System for Detection and Quantitative Evaluation of Compensatory Movement in Post-Stroke Patients Based on Wearable Sensor and Machine Learning Algorithm. *IEEE Sens. J.* **2024**, *24*, 22830–22842. [\[CrossRef\]](#)

38. Wei, S.; Wu, Z. The Application of Wearable Sensors and Machine Learning Algorithms in Rehabilitation Training: A Systematic Review. *Sensors* **2023**, *23*, 7667. [\[CrossRef\]](#)
39. Lam, W.W.T.; Tang, Y.M.; Fong, K.N.K. A Systematic Review of the Applications of Markerless Motion Capture (MMC) Technology for Clinical Measurement in Rehabilitation. *J. NeuroEng. Rehabil.* **2023**, *20*, 57. [\[CrossRef\]](#)
40. Moore, K.; O'Shea, E.; Kenny, L.; Barton, J.; Tedesco, S.; Sica, M.; Crowe, C.; Alamäki, A.; Condell, J.; Nordström, A.; et al. Older Adults' Experiences with Using Wearable Devices: Qualitative Systematic Review and Meta-Synthesis. *JMIR mHealth uHealth* **2021**, *9*, e23832. [\[CrossRef\]](#)
41. Muñoz Esquivel, K.; Gillespie, J.; Kelly, D.; Condell, J.; Davies, R.; McHugh, C.; Duffy, W.; Nevala, E.; Alamäki, A.; Jalovaara, J.; et al. Factors Influencing Continued Wearable Device Use in Older Adult Populations: Quantitative Study. *JMIR Aging* **2023**, *6*, e36807. [\[CrossRef\]](#)
42. Ding, H.; Ho, K.; Searls, E.; Low, S.; Li, Z.; Rahman, S.; Madan, S.; Igwe, A.; Popp, Z.; Burk, A.; et al. Assessment of Wearable Device Adherence for Monitoring Physical Activity in Older Adults: Pilot Cohort Study. *JMIR Aging* **2024**, *7*, e60209. [\[CrossRef\]](#)
43. Dobson, R.; Stowell, M.; Warren, J.; Tane, T.; Ni, L.; Gu, Y.; McCool, J.; Whittaker, R. Use of Consumer Wearables in Health Research: Issues and Considerations. *J. Med. Internet Res.* **2023**, *25*, e52444. [\[CrossRef\]](#)
44. Lederer, L.; Breton, A.; Jeong, H.; Master, H.; Roghanizad, A.R.; Dunn, J. The Importance of Data Quality Control in Using Fitbit Device Data from the Research Program. *JMIR mHealth uHealth* **2023**, *11*, e45103. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Daniore, P.; Nittas, V.; Haag, C.; Bernard, J.; Gonzenbach, R.; Von Wyl, V. From Wearable Sensor Data to Digital Biomarker Development: Ten Lessons Learned and a Framework Proposal. *npj Digit. Med.* **2024**, *7*, 161. [\[CrossRef\]](#)
46. Varcin, F.; Boocock, M.G. The Accuracy, Validity and Reliability of Theia3D Markerless Motion Capture for Studying the Biomechanics of Human Movement: A Systematic Review. *Artif. Intell. Med.* **2026**, *173*, 103332. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Yang, Z.; Zheng, X.; Hu, N.; Zhang, F.; Wang, A. "Challenges to Normalcy"—Perceived Barriers to Adherence to Home-Based Cardiac Rehabilitation Exercise in Patients with Chronic Heart Failure. *Patient Prefer. Adherence* **2023**, *17*, 3515–3524. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Doherty, C.; Lambe, R.; O'Grady, B.; O'Reilly-Morgan, D.; Smyth, B.; Lawlor, A.; Hurley, N.; Tragos, E. An Evaluation of the Effect of App-Based Exercise Prescription Using Reinforcement Learning on Satisfaction and Exercise Intensity: Randomized Crossover Trial. *JMIR mHealth uHealth* **2024**, *12*, e49443. [\[CrossRef\]](#)
49. Cavazzotto, T.G.; Dantas, D.B.; Queiroga, M.R. ChatGPT and Exercise Prescription: Human vs. Machine or Human plus Machine? *J. Sport Health Sci.* **2024**, *13*, 661–662. [\[CrossRef\]](#)
50. Puce, L.; Bragazzi, N.L.; Currà, A.; Trompetto, C. Harnessing Generative Artificial Intelligence for Exercise and Training Prescription: Applications and Implications in Sports and Physical Activity—A Systematic Literature Review. *Appl. Sci.* **2025**, *15*, 3497. [\[CrossRef\]](#)
51. Washif, J.A.; Pagaduan, J.; James, C.; Dergaa, I.; Beaven, C.M. Artificial intelligence in sport: Exploring the potential of using ChatGPT in resistance training prescription. *Biol. Sport* **2024**, *41*, 209–220. [\[CrossRef\]](#)
52. Wang, D.; Zhang, S. Large Language Models in Medical and Healthcare Fields: Applications, Advances, and Challenges. *Artif. Intell. Rev.* **2024**, *57*, 299. [\[CrossRef\]](#)
53. Zhou, H.; Liu, F.; Gu, B.; Zou, X.; Huang, J.; Wu, J.; Li, Y.; Chen, S.S.; Zhou, P.; Liu, J.; et al. A Survey of Large Language Models in Medicine: Progress, Application, and Challenge. *arXiv* **2023**, arXiv:2311.05112.
54. Shamshad, F.; Khan, S.; Zamir, S.W.; Khan, M.H.; Hayat, M.; Khan, F.S.; Fu, H. Transformers in Medical Imaging: A Survey. *Med. Image Anal.* **2023**, *88*, 102802. [\[CrossRef\]](#)
55. Yin, S.; Fu, C.; Zhao, S.; Li, K.; Sun, X.; Xu, T.; Chen, E. A Survey on Multimodal Large Language Models. *Natl. Sci. Rev.* **2024**, *11*, nwae403. [\[CrossRef\]](#)
56. Gao, Y.; Xiong, Y.; Gao, X.; Jia, K.; Pan, J.; Bi, Y.; Dai, Y.; Sun, J.; Wang, M.; Wang, H. Retrieval-Augmented Generation for Large Language Models: A Survey. *arXiv* **2023**, arXiv:2312.10997v1.
57. EQUATOR Network. TRIPOD + AI Statement: Updated Guidance for Reporting Clinical Prediction Models That Use Regression or Machine Learning Methods. Available online: <https://www.equator-network.org/reporting-guidelines/tripod-statement/> (accessed on 8 February 2026).
58. Moons, K.G.M.; Damen, J.A.A.; Kaul, T.; Hooft, L.; Andaur Navarro, C.; Dhiman, P.; Beam, A.L.; Van Calster, B.; Celi, L.A.; Denaxas, S.; et al. PROBAST+AI: An Updated Quality, Risk of Bias, and Applicability Assessment Tool for Prediction Models Using Regression or Artificial Intelligence Methods. *BMJ* **2025**, *388*, e082505. [\[CrossRef\]](#)
59. Ettetfagh, A.; Roshan Fekr, A. Technological Advances in Lower-Limb Tele-Rehabilitation: A Review of Literature. *J. Rehabil. Assist. Technol. Eng.* **2024**, *11*, 20556683241259256. [\[CrossRef\]](#)
60. Jubair, H.; Mehenaz, M. Smartwatch-Assisted Exercise Prescription: Utilizing Machine Learning Algorithms for Personalized Workout Recommendations and Monitoring: A Review. *Res. Sq.* **2024**, preprint. [\[CrossRef\]](#)
61. Banyai, A.D.; Brişan, C. Robotics in Physical Rehabilitation: Systematic Review. *Healthcare* **2024**, *12*, 1720. [\[CrossRef\]](#) [\[PubMed\]](#)

62. Tang, J.; Abedi, A.; Colella, T.J.F.; Khan, S.S. Rehabilitation Exercise Quality Assessment and Feedback Generation Using Large Language Models with Prompt Engineering. In *ArtifiAI for Aging Rehabilitation and Intelligent Assisted Living*; Khan, S.S., Romeo, L., Abedi, A., Eds.; Communications in Computer and Information Science; Springer Nature: Singapore, 2025; Volume 2620, pp. 60–75. ISBN 9789819505678.
63. Gazzarata, R.; Almeida, J.; Lindsköld, L.; Cangioli, G.; Gaeta, E.; Fico, G.; Chronaki, C.E. HL7 Fast Healthcare Interoperability Resources (HL7 FHIR) in Digital Healthcare Ecosystems for Chronic Disease Management: Scoping Review. *Int. J. Med. Inform.* **2024**, *189*, 105507. [\[CrossRef\]](#)
64. Saberi, M.A.; Mcheick, H.; Adda, M. From Data Silos to Health Records Without Borders: A Systematic Survey on Patient-Centered Data Interoperability. *Information* **2025**, *16*, 106. [\[CrossRef\]](#)
65. Baffert, S.; Hadouiri, N.; Fabron, C.; Burgy, F.; Cassany, A.; Kemoun, G. Economic Evaluation of Telerehabilitation: Systematic Literature Review of Cost-Utility Studies. *JMIR Rehabil. Assist. Technol.* **2023**, *10*, e47172. [\[CrossRef\]](#)
66. Su, Z.; Guo, Z.; Wang, W.; Liu, Y.; Liu, Y.; Chen, W.; Zheng, M.; Michael, N.; Lu, S.; Wang, W.; et al. The Effect of Telerehabilitation on Balance in Stroke Patients: Is It More Effective than the Traditional Rehabilitation Model? A Meta-Analysis of Randomized Controlled Trials Published during the COVID-19 Pandemic. *Front. Neurol.* **2023**, *14*, 1156473. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Liu, C.; Wang, Y.; Jiang, H.; Xie, Y. Internet-Based and Wearable-Device-Assisted Tele-Rehabilitation for Stroke Patients after Discharge: A Randomized Trial. *Can. J. Neurol. Sci.* **2025**, *FirstView*, 1–9. [\[CrossRef\]](#)
68. Boukhenoufa, I.; Zhai, X.; Utti, V.; Jackson, J.; McDonald-Maier, K.D. Wearable Sensors and Machine Learning in Post-Stroke Rehabilitation Assessment: A Systematic Review. *Biomed. Signal Process. Control.* **2022**, *71*, 103197. [\[CrossRef\]](#)
69. Wang, B.; Liu, Y.; Lu, A.; Wang, C. Application of Wearable Sensors in Constructing a Fall Risk Prediction Model for Community-Dwelling Older Adults: A Scoping Review. *Arch. Gerontol. Geriatr.* **2025**, *129*, 105689. [\[CrossRef\]](#)
70. Bonanno, M.; Ielo, A.; De Pasquale, P.; Celesti, A.; De Nunzio, A.M.; Quartarone, A.; Calabrò, R.S. Use of Wearable Sensors to Assess Fall Risk in Neurological Disorders: Systematic Review. *JMIR mHealth uHealth* **2025**, *13*, e67265. [\[CrossRef\]](#)
71. Chen, M.; Wang, H.; Yu, L.; Yeung, E.H.K.; Luo, J.; Tsui, K.-L.; Zhao, Y. A Systematic Review of Wearable Sensor-Based Technologies for Fall Risk Assessment in Older Adults. *Sensors* **2022**, *22*, 6752. [\[CrossRef\]](#)
72. Zhao, R.; Cheng, L.; Zheng, Q.; Lv, Y.; Wang, Y.-M.; Ni, M.; Ren, P.; Feng, Z.; Ji, Q.; Zhang, G. A Smartphone Application-Based Remote Rehabilitation System for Post-Total Knee Arthroplasty Rehabilitation: A Randomized Controlled Trial. *J. Arthroplast.* **2024**, *39*, 575–581.e8. [\[CrossRef\]](#)
73. Neo, J.R.E.; Ser, J.S.; Tay, S.S. Use of Large Language Model-Based Chatbots in Managing the Rehabilitation Concerns and Education Needs of Outpatient Stroke Survivors and Caregivers. *Front. Digit. Health* **2024**, *6*, 1395501. [\[CrossRef\]](#) [\[PubMed\]](#)
74. He, W.; Zhang, W.; Jin, Y.; Zhou, Q.; Zhang, H.; Xia, Q. Physician Versus Large Language Model Chatbot Responses to Web-Based Questions From Autistic Patients in Chinese: Cross-Sectional Comparative Analysis. *J. Med. Internet Res.* **2024**, *26*, e54706. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Megalla, M.; Hahn, A.K.; Bauer, J.A.; Windsor, J.T.; Grace, Z.T.; Gedman, M.A.; Arciero, R.A. ChatGPT and Google Provide Mostly Excellent or Satisfactory Responses to the Most Frequently Asked Patient Questions Related to Rotator Cuff Repair. *Arthrosc. Sports Med. Rehabil.* **2024**, *6*, 100963. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Naqvi, W.M.; Shaikh, S.Z.; Mishra, G.V. Large Language Models in Physical Therapy: Time to Adapt and Adept. *Front. Public Health* **2024**, *12*, 1364660. [\[CrossRef\]](#)
77. Lee, P.; Bubeck, S.; Petro, J. Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *N. Engl. J. Med.* **2023**, *388*, 1233–1239. [\[CrossRef\]](#)
78. Zakka, C.; Shad, R.; Chaurasia, A.; Dalal, A.R.; Kim, J.L.; Moor, M.; Fong, R.; Phillips, C.; Alexander, K.; Ashley, E.; et al. Almanac—Retrieval-Augmented Language Models for Clinical Medicine. *NEJM AI* **2024**, *1*, AIoa2300068. [\[CrossRef\]](#)
79. Tam, T.Y.C.; Sivarajkumar, S.; Kapoor, S.; Stolyar, A.V.; Polanska, K.; McCarthy, K.R.; Osterhoudt, H.; Wu, X.; Visweswaran, S.; Fu, S.; et al. A Framework for Human Evaluation of Large Language Models in Healthcare Derived from Literature Review. *npj Digit. Med.* **2024**, *7*, 258. [\[CrossRef\]](#)
80. Youssef, A.; Pencina, M.; Thakur, A.; Zhu, T.; Clifton, D.; Shah, N.H. External Validation of AI Models in Health Should Be Replaced with Recurring Local Validation. *Nat. Med.* **2023**, *29*, 2686–2687. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Denecke, K.; May, R.; LLM Health Group; Rivera Romero, O. Potential of Large Language Models in Health Care: Delphi Study. *J. Med. Internet Res.* **2024**, *26*, e52399. [\[CrossRef\]](#)
82. McDonald, P.L.; Foley, T.J.; Verheij, R.; Braithwaite, J.; Rubin, J.; Harwood, K.; Phillips, J.; Gilman, S.; Van Der Wees, P.J. Data to Knowledge to Improvement: Creating the Learning Health System. *BMJ* **2024**, *384*, e076175. [\[CrossRef\]](#)
83. Reid, R.J.; Wodchis, W.P.; Kuluski, K.; Lee-Foon, N.K.; Lavis, J.N.; Rosella, L.C.; Desveaux, L. Actioning the Learning Health System: An Applied Framework for Integrating Research into Health Systems. *SSM Health Syst.* **2024**, *2*, 100010. [\[CrossRef\]](#)
84. Luo, M.; Duan, Z.; Gao, J.; Sun, Y.; Chen, L.; Feng, X. Evaluating the Role of ChatGPT in Rehabilitation Medicine: A Narrative Review. *Front. Digit. Health* **2025**, *7*, 1618510. [\[CrossRef\]](#) [\[PubMed\]](#)

85. Hao, J.; Yao, Z.; Tang, Y.; Remis, A.; Wu, K.; Yu, X. Artificial Intelligence in Physical Therapy: Evaluating ChatGPT's Role in Clinical Decision Support for Musculoskeletal Care. *Ann. Biomed. Eng.* **2025**, *53*, 9–13. [\[CrossRef\]](#)
86. Dave, T.; Athaluri, S.A.; Singh, S. ChatGPT in Medicine: An Overview of Its Applications, Advantages, Limitations, Future Prospects, and Ethical Considerations. *Front. Artif. Intell.* **2023**, *6*, 1169595. [\[CrossRef\]](#)
87. Arbel, Y.; Gimmon, Y.; Shmueli, L. Evaluating the Potential of Large Language Models for Vestibular Rehabilitation Education: A Comparison of ChatGPT, Google Gemini, and Clinicians. *Phys. Ther.* **2025**, *105*, pzaf010. [\[CrossRef\]](#)
88. McBee, J.C.; Han, D.Y.; Liu, L.; Ma, L.; Adjero, D.A.; Xu, D.; Hu, G. Assessing ChatGPT's Competency in Addressing Interdisciplinary Inquiries on Chatbot Uses in Sports Rehabilitation: Simulation Study. *JMIR Med. Educ.* **2024**, *10*, e51157. [\[CrossRef\]](#)
89. Bilika, P.; Stefanouli, V.; Strimpakos, N.; Kapreli, E.V. Clinical Reasoning Using ChatGPT: Is It beyond Credibility for Physiotherapists Use? *Physiother. Theory Pract.* **2024**, *40*, 2943–2962. [\[CrossRef\]](#)
90. Sawamura, S.; Bito, T.; Ando, T.; Masuda, K.; Kameyama, S.; Ishida, H. Evaluation of the Accuracy of ChatGPT's Responses to and References for Clinical Questions in Physical Therapy. *J. Phys. Ther. Sci.* **2024**, *36*, 234–239. [\[CrossRef\]](#)
91. Hao, J.; Yao, Z.; Siu, K. Artificial Intelligence in Physical Therapy Education: Evaluating Clinical Reasoning Performance in Musculoskeletal Care Using ChatGPT. *Musculoskelet. Care* **2025**, *23*, e70177. [\[CrossRef\]](#)
92. Yang, Y.; Lin, J.; Zhang, J. The Effects of ChatGPT on Patient Education of Knee Osteoarthritis: A Preliminary Study of 60 Cases. *Int. J. Surg.* **2025**, *111*, 9753–9756. [\[CrossRef\]](#)
93. Yoo, M.; Jang, C.W. Presentation Suitability and Readability of ChatGPT's Medical Responses to Patient Questions about on Knee Osteoarthritis. *Health Inform. J.* **2025**, *31*, 14604582251315587. [\[CrossRef\]](#) [\[PubMed\]](#)
94. Huang, L.; Yu, W.; Ma, W.; Zhong, W.; Feng, Z.; Wang, H.; Chen, Q.; Peng, W.; Feng, X.; Qin, B.; et al. A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. *ACM Trans. Inf. Syst.* **2025**, *43*, 42. [\[CrossRef\]](#)
95. Xiong, L.; Zeng, Q.; Luo, W.; Liu, R. Nursing Retrieval-Augmented Generation: Retrieval Augmented Generation for Nursing Question Answering with Large Language Models. *Int. J. Nurs. Sci.* **2025**, *12*, 516–523. [\[CrossRef\]](#) [\[PubMed\]](#)
96. Kim, N.; Homer, M.; Jang, H. Clinical Application of Large Language Models for Intervention Plan Development in Speech-Language Pathology. *Am. J. Speech Lang. Pathol.* **2025**, *34*, 2098–2114. [\[CrossRef\]](#)
97. Sardari, S.; Sharifzadeh, S.; Daneshkhah, A.; Nakisa, B.; Loke, S.W.; Palade, V.; Duncan, M.J. Artificial Intelligence for Skeleton-Based Physical Rehabilitation Action Evaluation: A Systematic Review. *Comput. Biol. Med.* **2023**, *158*, 106835. [\[CrossRef\]](#)
98. Mourchid, Y.; Slama, R. D-STGCNT: A Dense Spatio-Temporal Graph Conv-GRU Network Based on Transformer for Assessment of Patient Physical Rehabilitation. *Comput. Biol. Med.* **2023**, *165*, 107420. [\[CrossRef\]](#)
99. Lim, Z.K.; Connie, T.; Goh, M.K.O.; Saedon, N.I.B. Fall Risk Prediction Using Temporal Gait Features and Machine Learning Approaches. *Front. Artif. Intell.* **2024**, *7*, 1425713. [\[CrossRef\]](#)
100. Cardenas, L.; Parajes, K.; Zhu, M.; Zhai, S. AutoHealth: Advanced LLM-Empowered Wearable Personalized Medical Butler for Parkinson's Disease Management. In Proceedings of the 2024 IEEE 14th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 8–10 January 2024; pp. 375–379.
101. Xiong, G.; Jin, Q.; Lu, Z.; Zhang, A. Benchmarking Retrieval-Augmented Generation for Medicine. In *Proceedings of the Findings of the Association for Computational Linguistics ACL*; Association for Computational Linguistics: Bangkok, Thailand, 2024; pp. 6233–6251.
102. Han, Z.; Gao, C.; Liu, J.; Zhang, J.; Zhang, S.Q. Parameter-Efficient Fine-Tuning for Large Models: A Comprehensive Survey. *arXiv* **2024**, arXiv:2403.14608.
103. Karlov, M.; Abedi, A.; Khan, S.S. Rehabilitation Exercise Quality Assessment through Supervised Contrastive Learning with Hard and Soft Negatives. *Med. Biol. Eng. Comput.* **2025**, *63*, 15–28. [\[CrossRef\]](#)
104. Retevoi, A.; Devittori, G.; Kowatsch, T.; Lambercy, O. Building Conversational Agents for Stroke Rehabilitation: An Evaluation of Large Language Models and Retrieval Augmented Generation. In Proceedings of the ACM International Conference on Intelligent Virtual Agents, Glasgow, UK, 16–19 September 2024; pp. 1–4.
105. Kim, J. Fine-Tuning the Llama2 Large Language Model Using Books on the Diagnosis and Treatment of Musculoskeletal System in Physical Therapy. *J. Musculoskelet. Sci. Technol.* **2024**, *8*, 65–73. [\[CrossRef\]](#)
106. Kourbane, I.; Papadakis, P.; Andries, M. SSL-Rehab: Assessment of Physical Rehabilitation Exercises through Self-Supervised Learning of 3D Skeleton Representations. *Comput. Vis. Image Underst.* **2025**, *251*, 104275. [\[CrossRef\]](#)
107. Gallifant, J.; Afshar, M.; Ameen, S.; Aphinyanaphongs, Y.; Chen, S.; Cacciamani, G.; Demner-Fushman, D.; Dligach, D.; Daneshjou, R.; Fernandes, C.; et al. The TRIPOD-LLM Reporting Guideline for Studies Using Large Language Models. *Nat. Med.* **2025**, *31*, 60–69. [\[CrossRef\]](#)
108. Tabassi, E. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*; National Institute of Standards and Technology (U.S.): Gaithersburg, MD, USA, 2023; NIST AI 100-1.
109. World Health Organization. Ethics and Governance of Artificial Intelligence for Health. 28 June 2021. Available online: <https://www.who.int/publications/i/item/9789240029200> (accessed on 29 January 2026).

110. Omiye, J.A.; Lester, J.C.; Spichak, S.; Rotemberg, V.; Daneshjou, R. Large Language Models Propagate Race-Based Medicine. *npj Digit. Med.* **2023**, *6*, 195. [CrossRef]
111. Yang, Y.; Liu, X.; Jin, Q.; Huang, F.; Lu, Z. Unmasking and Quantifying Racial Bias of Large Language Models in Medical Report Generation. *Commun. Med.* **2024**, *4*, 176. [CrossRef] [PubMed]
112. Fava, G.A.; Sonino, N.; Guidi, J. Measuring Clinical Findings: The Value of Clinimetrics. *Postgrad. Med. J.* **2025**, *102*, 88–94. [CrossRef] [PubMed]
113. Klukowska, A.M.; Vandertop, W.P.; Schröder, M.L.; Staartjes, V.E. Calculation of the Minimum Clinically Important Difference (MCID) Using Different Methodologies: Case Study and Practical Guide. *Eur. Spine J.* **2024**, *33*, 3388–3400. [CrossRef]
114. De Micco, F.; Di Palma, G.; Ferorelli, D.; De Benedictis, A.; Tomassini, L.; Tambone, V.; Cingolani, M.; Scendoni, R. Artificial Intelligence in Healthcare: Transforming Patient Safety with Intelligent Systems—A Systematic Review. *Front. Med.* **2025**, *11*, 1522554. [CrossRef]
115. Ten Hove, D.; Jorgensen, T.D.; Van Der Ark, L.A. Updated Guidelines on Selecting an Intraclass Correlation Coefficient for Interrater Reliability, with Applications to Incomplete Observational Designs. *Psychol. Methods* **2024**, *29*, 967–979. [CrossRef]
116. De Arruda, G.T.; Terwee, C.B.; Elsmann, E.B.M.; Avila, M.A.; Gagnier, J.J.; Mokkink, L.B.; PROM Reporting Group; Dibai-Filho, A.V.; Firth, A.D.; Mehdi-pour, A.; et al. Explanation & Elaboration Document of the COSMIN Reporting Guideline 2.0 for Studies on Measurement Properties of Patient-Reported Outcome Measures. *Qual. Life Res.* **2025**, *34*, 1891–1899. [CrossRef] [PubMed]
117. Caronni, A.; Picardi, M.; Scarano, S.; Rota, V.; Guidali, G.; Bolognini, N.; Corbo, M. Minimal Detectable Change of Gait and Balance Measures in Older Neurological Patients: Estimating the Standard Error of the Measurement from before-after Rehabilitation Data Thanks to the Linear Mixed-Effects Models. *J. NeuroEng. Rehabil.* **2024**, *21*, 44. [CrossRef]
118. Vach, W.; Saxer, F. Anchor-Based Minimal Important Difference Values Are Often Sensitive to the Distribution of the Change Score. *Qual. Life Res.* **2024**, *33*, 1223–1232. [CrossRef]
119. Dekker, J.; De Boer, M.; Ostelo, R. Minimal Important Change and Difference in Health Outcome: An Overview of Approaches, Concepts, and Methods. *Osteoarthr. Cartil.* **2024**, *32*, 8–17. [CrossRef]
120. Mishra, B.; Sudheer, P.; Agarwal, A.; Nilima, N.; Srivastava, M.V.P.; Vishnu, V.Y. Minimal Clinically Important Difference of Scales Reported in Stroke Trials: A Review. *Brain Sci.* **2024**, *14*, 80. [CrossRef] [PubMed]
121. Ayoub, N.F.; Balakrishnan, K.; Ayoub, M.S.; Barrett, T.F.; David, A.P.; Gray, S.T. Inherent Bias in Large Language Models: A Random Sampling Analysis. *Mayo Clin. Proc. Digit. Health* **2024**, *2*, 186–191. [CrossRef]
122. Omar, M.; Sorin, V.; Agbareia, R.; Apakama, D.U.; Soroush, A.; Sakhuja, A.; Freeman, R.; Horowitz, C.R.; Richardson, L.D.; Nadkarni, G.N.; et al. Evaluating and Addressing Demographic Disparities in Medical Large Language Models: A Systematic Review. *Int. J. Equity Health* **2025**, *24*, 57. [CrossRef]
123. U.S. Food and Drug Administration (FDA). Software as a Medical Device (SaMD). Content Current as of: 4 December 2018. Available online: <https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd> (accessed on 30 January 2026).
124. U.S. Food and Drug Administration (FDA). Artificial Intelligence in Software as a Medical Device. Content Current as of: 25 March 2025. Available online: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device> (accessed on 30 January 2026).
125. U.S. Food and Drug Administration (FDA). Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions: Guidance for Industry and Food and Drug Administration Staff (August 2025). Content Current as of: 18 August 2025. Available online: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence> (accessed on 30 January 2026).
126. Zhang, Z.; Chen, Y.; Zhao, X.; Fan, W.; Peng, D.; Li, T.; Zhao, L.; Fu, Y. A Review of Ethical Considerations for the Medical Applications of Brain-Computer Interfaces. *Cogn. Neurodyn.* **2024**, *18*, 3603–3614. [CrossRef] [PubMed]
127. Cheong, B.C. Transparency and Accountability in AI Systems: Safeguarding Wellbeing in the Age of Algorithmic Decision-Making. *Front. Hum. Dyn.* **2024**, *6*, 1421273. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.