



Artificial Intelligence—Based Gait Analysis in Neurological Rehabilitation: A Systematic Review Protocol

Dhirenbhai Kalal

Independent Researcher, USA

Abstract

Background: Gait impairment is common in neurological conditions, including stroke, Parkinson's disease, multiple sclerosis, cerebral palsy, and spinal cord injury. Reliable gait assessment can inform clinical evaluation, rehabilitation planning, and monitoring of functional change. Artificial intelligence (AI), including machine learning, deep learning, and computer vision, may enable automated analysis of gait data collected using wearable sensors, cameras, and instrumented systems. However, evidence varies across populations, technologies, outcomes, and validation designs.

Objective: This protocol describes a systematic review that will examine AI-based gait analysis in neurological rehabilitation, focusing on computational methods, data-acquisition technologies, assessment tasks, validation practices, reported performance, clinical relevance, and barriers to implementation.

Methods: The review will search relevant biomedical and engineering databases using a reproducible, database-adapted strategy. Eligible primary studies will evaluate AI methods applied to gait-related data in people with neurological disorders or in clearly defined neurological rehabilitation applications. Two reviewers will independently screen records and extract data where feasible. Risk of bias will be assessed using tools appropriate to study design and purpose. Findings will be synthesized narratively; meta-analysis will be considered only when studies are sufficiently comparable. Reporting will follow PRISMA 2020 guidance.

Conclusion: This protocol establishes a transparent plan for evaluating the evidence on AI-based gait analysis in neurological rehabilitation. The completed review will distinguish technical feasibility from measurement validity, generalizability, clinical utility, and patient-centered

* Corresponding author: **Dhirenbhai kalal, Independent Researcher, USA. kalaldhiren@gmail.com**

Received: 01-09-2026; **Accepted:** 25-09-2026; **Published:** 16-10-2026



Copyright: © The Author(s), 2026. Published by JAPMI. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

benefit. No study counts, performance estimates, or conclusions about the included evidence are reported here because the search, screening, and data extraction have not yet been verified.

Keywords:

Artificial intelligence; gait analysis; neurological rehabilitation; machine learning; deep learning; computer vision; wearable sensors; systematic review protocol.

1. Introduction

1.1 Background on Gait Analysis in Neurological Rehabilitation

Gait is a complex, coordinated motor activity involving the integration of the central nervous system, peripheral nervous system, musculoskeletal system, sensory feedback mechanisms, and postural control processes. Normal walking requires the coordinated regulation of muscle activation, joint movement, balance, and the timing of successive steps. Neurological disorders can disrupt these processes, resulting in abnormalities in walking speed, stride length, cadence, step symmetry, foot clearance, postural stability, and overall mobility. Such impairments may reduce functional independence, increase the risk of falls, restrict participation in daily activities, and negatively affect quality of life. Consequently, gait assessment is an essential component of neurological rehabilitation, supporting clinical evaluation, treatment planning, progress monitoring, and the measurement of rehabilitation outcomes.[1]

Gait impairment occurs across a broad range of neurological conditions, although its characteristics vary according to the underlying disease, severity, disease progression, and individual functional status. Following stroke, for example, motor weakness, impaired coordination, spasticity, and asymmetrical weight-bearing may contribute to reduced walking efficiency and uneven step patterns. Parkinson's disease is frequently associated with bradykinesia, reduced stride length, shuffling steps, freezing of gait, and impaired postural control. People with multiple sclerosis may experience variable walking capacity, fatigue-related deterioration, balance difficulties, and changes in gait performance over time. Cerebral palsy may involve persistent abnormalities in muscle tone, selective motor control, joint movement, and walking mechanics, whereas spinal cord injury can produce gait limitations associated with impaired motor and sensory pathways. These differences highlight the need for assessment approaches that can capture both shared gait abnormalities and condition-specific movement patterns [6, 11, 23, 25].

In clinical practice, gait assessment may involve observational examination, standardized functional tests, instrumented gait analysis, and patient-reported measures. Common functional assessments include walking-speed tests, the Timed Up and Go test, and the six-minute walk test. Instrumented approaches may use motion-capture systems, force plates, pressure-sensitive walkways, electromyography, and wearable inertial sensors to quantify movement characteristics. These methods provide complementary information: functional tests characterize overall mobility and endurance, whereas instrumented systems can provide more detailed measurements of gait mechanics and temporal-spatial parameters.

The purpose of gait assessment extends beyond identifying abnormal movement. Rehabilitation professionals also need to determine whether an intervention produces meaningful functional improvement, whether changes persist outside the clinical environment, and whether an individual's gait performance is improving in ways that support everyday mobility. Repeated measurements can help identify progress, guide adjustments to treatment intensity, and inform decisions concerning assistive devices or additional rehabilitation support. However, these objectives require assessment methods that are sufficiently reliable, sensitive to change, and feasible for repeated use. [3]

Recent developments in digital sensing, computational biomechanics, and artificial intelligence (AI) have created new opportunities to address these requirements. By analyzing movement data collected through wearable devices, cameras, pressure sensors, and laboratory-based motion-capture systems, AI methods can identify complex patterns that may be difficult to quantify through visual observation alone. Their potential relevance lies not only in automated classification of gait abnormalities but also in the possibility of supporting continuous monitoring, individualized assessment, and data-informed rehabilitation decisions. Nevertheless, the clinical value of these technologies depends on the quality of their measurements, the representativeness of the populations used for model development, and evidence that algorithmic outputs are meaningful for rehabilitation practice.

1.2 Limitations of Conventional Gait Assessment

Conventional gait assessment remains fundamental to neurological rehabilitation, but existing approaches have several limitations that may restrict the depth, frequency, and generalizability of gait evaluation. These limitations vary by assessment method and clinical setting; therefore, traditional clinical tests, observational examination, and instrumented gait analysis should not be considered interchangeable.

First, observational gait assessment can be influenced by examiner experience, visual judgment, assessment conditions, and the complexity of the movement abnormality. Experienced clinicians can identify clinically important patterns, but visual observation may be insufficient to quantify subtle changes in joint motion, step timing, variability, or movement symmetry. Differences in examiner interpretation can also affect consistency, particularly when abnormalities are complex or when changes over time are small. Instrumented measurements can complement clinical observation by providing quantitative data, although measurement quality still depends on the equipment, protocol, and interpretation used.

Second, sophisticated instrumented gait-analysis systems may require specialized facilities, trained personnel, equipment maintenance, and substantial financial investment. Optical motion-capture laboratories and force-plate systems can provide detailed biomechanical measurements, but their availability may be limited in smaller rehabilitation facilities and resource-constrained healthcare systems. Laboratory testing also commonly occurs under controlled conditions, which may not fully represent the movement patterns individuals demonstrate at home, in the community, or during everyday activities.

Third, many conventional assessments provide information about performance during a specific test or at a particular time. Standardized walking tests are useful for evaluating functional capacity, but they may not capture fluctuations in gait across different environments, fatigue states, medication cycles, or daily activities. A single assessment may therefore fail to represent the variability of an individual's walking performance. Wearable sensing and other remote-monitoring technologies offer potential ways to address this limitation, although they introduce their own concerns regarding sensor placement, adherence, data completeness, and measurement accuracy.

Fourth, conventional gait measures may not fully capture the multidimensional nature of neurological movement impairment. Walking speed, cadence, stride length, and symmetry are clinically relevant, but no single parameter adequately describes every aspect of gait quality, stability, adaptability, and functional independence. Two individuals may demonstrate similar walking speeds while differing substantially in balance, compensatory movement strategies, fatigue, or the effort required to walk. A comprehensive assessment therefore requires consideration of multiple parameters and their relationships to functional outcomes.

Finally, translating gait measurements into individualized treatment decisions remains challenging. Clinicians must interpret measurements in the context of diagnosis, disease severity, cognitive and sensory function, comorbidities, rehabilitation goals, and the individual's everyday environment. Conventional assessments can support these decisions, but their ability to combine large volumes of heterogeneous movement data is limited by practical constraints [11, 23].

AI-based methods may help address some of these challenges by automating feature extraction, identifying complex relationships, and integrating information from multiple sensors. However, AI does not automatically eliminate measurement error, subjectivity, or accessibility barriers. Algorithmic predictions can be unreliable when input data are noisy, patient populations differ from training datasets, or models have not been externally validated. AI should therefore be evaluated as a potential complement to established clinical assessment rather than as an assumed replacement for professional judgment.

1.3 Role of Artificial Intelligence in Gait Analysis

Artificial intelligence encompasses computational methods that perform tasks such as pattern recognition, prediction, classification, and decision support by learning from data or applying structured computational rules. In gait analysis, AI can be used to process movement signals, identify gait events, extract meaningful features, classify movement abnormalities, estimate functional outcomes, and examine changes in gait over time. Its application is particularly relevant to neurological rehabilitation because gait impairment often involves complex, nonlinear, and interacting movement characteristics that may not be adequately represented by a small set of manually selected measurements.

1.3.1 Machine learning for gait classification and assessment

Traditional machine-learning techniques, including support vector machines, random forests, decision trees, and other supervised-learning algorithms, can use extracted gait features to distinguish between predefined classes or estimate clinically relevant outcomes. Input features may include step duration, stride length, cadence, acceleration patterns, angular velocity, joint trajectories, and measures of movement variability.

In neurological populations, these techniques may be investigated for identifying gait patterns associated with particular disorders, differentiating levels of functional impairment, detecting freezing-of-gait episodes, or classifying changes in walking performance. Their performance depends on the representativeness of the training data, the reliability of the reference labels, the choice of features, and the evaluation design. High classification accuracy in a small or highly controlled dataset does not necessarily demonstrate clinical usefulness in a different hospital, population, or real-world environment.

1.3.2 Deep learning and automated feature extraction

Deep-learning approaches, including convolutional neural networks and recurrent neural networks, can learn representations from complex signals and sequential data. Depending on the task and dataset, these models may be applied to wearable sensor recordings, video sequences, or joint-coordinate trajectories. Temporal models and other sequence-learning architectures may help capture changes in movement across successive steps, while convolutional architectures may identify patterns in spatial or time-series representations.

These methods could reduce reliance on manually engineered features and support analysis of subtle or temporally distributed gait characteristics. However, deep-learning models often require substantial amounts of representative data and careful validation. They may be difficult to interpret, sensitive to

differences in sensor configuration, and vulnerable to overfitting when the number of participants is small relative to model complexity. For these reasons, model performance should be assessed using appropriate data-partitioning strategies and, where feasible, independent external datasets.

1.3.3 Computer vision and markerless gait analysis

Computer vision enables computational analysis of movement recorded by cameras. Human-pose estimation methods can identify body landmarks, estimate joint coordinates, and derive movement features from video. These capabilities may support gait assessment in clinical rooms, rehabilitation facilities, or other environments where specialized motion-capture equipment is unavailable.

Markerless systems may reduce the need for physical markers and facilitate more accessible movement assessment. They may also provide opportunities for remote assessment when video acquisition is feasible and appropriate. Nevertheless, estimated body landmarks can be affected by camera angle, lighting, occlusion, clothing, movement speed, and the characteristics of the pose-estimation model. Two-dimensional video measurements may not adequately capture three-dimensional biomechanics, and estimates of clinically important parameters require validation against suitable reference measurements.

1.3.4 Wearable sensors and continuous monitoring

Wearable devices, including inertial measurement units, accelerometers, and gyroscopes, can capture movement signals during walking. AI algorithms may process these signals to identify gait events, estimate temporal-spatial characteristics, detect abnormal episodes, and characterize variability across repeated walking bouts.

Wearable-based assessment is particularly relevant to rehabilitation monitoring because it may allow measurements outside specialized laboratories. Repeated or continuous data collection can help investigate fluctuations in walking performance and changes associated with treatment, fatigue, or environmental context. However, real-world monitoring introduces technical and practical challenges, including sensor-placement variability, battery life, missing observations, user adherence, and differences between devices. The validity of estimates derived from wearable data must therefore be established for the specific devices, populations, and intended uses [14, 15, 19, 20].

1.3.5 AI-supported rehabilitation planning and outcome prediction

Beyond gait classification, AI may support rehabilitation by identifying patterns associated with functional recovery, estimating future performance, stratifying patients according to movement characteristics, or helping clinicians monitor responses to intervention. Models combining gait features with relevant clinical information may eventually contribute to individualized rehabilitation planning.

These applications require particular methodological care. A model that predicts an outcome is not necessarily capable of identifying which treatment will produce the greatest benefit for an individual. Similarly, an association between gait features and recovery does not establish that changing those features will improve clinical outcomes. Predictive performance, calibration, clinical utility, and the potential consequences of erroneous predictions should be evaluated before such models are used to inform care.

1.3.6 Explainability, reliability, and responsible implementation

The usefulness of AI-based gait analysis depends on more than predictive accuracy. Rehabilitation professionals need to understand what a model measures, which patients it has been evaluated on, how uncertain its predictions are, and whether its output can be interpreted in a clinically meaningful way.

Explainable AI techniques may help clarify the contribution of particular features, but an explanation generated by a model should not automatically be interpreted as evidence of a causal mechanism.

Additional considerations include patient privacy, secure handling of movement and video data, informed consent, algorithmic bias, accessibility, interoperability, and integration into clinical workflows. Models developed using narrow or unrepresentative populations may perform less reliably for people with different ages, body characteristics, disability severity, assistive-device use, or comorbidities. Accordingly, responsible implementation requires transparent reporting, appropriate validation, clinician oversight, and evaluation of patient-centered outcomes.

Overall, AI offers a range of complementary approaches to gait assessment, from automated signal processing to predictive modelling. The extent to which these technologies improve neurological rehabilitation must be established through systematic evaluation of their accuracy, reliability, generalizability, feasibility, and clinical relevance.

1.4 Research Gap and Study Rationale

Research on AI-based gait analysis has expanded across biomedical engineering, movement science, neurology, and rehabilitation. Studies have investigated a range of data-acquisition technologies, including wearable inertial sensors, pressure-sensitive systems, optical motion capture, and video-based pose estimation. They have also examined diverse computational approaches, ranging from conventional machine-learning algorithms to deep-learning models. Although this research provides a foundation for technological development, the diversity of methods, populations, and assessment objectives makes it difficult to determine how consistently the available evidence supports clinical use [7, 18, 25, 28].

One important challenge is methodological heterogeneity. Studies may differ in participant characteristics, neurological diagnoses, sample sizes, sensor locations, data-preprocessing procedures, feature definitions, model architectures, and outcome measures. Some investigations focus on distinguishing clinical groups, whereas others examine gait-event detection, episode recognition, functional prediction, or monitoring of rehabilitation-related changes. Consequently, performance estimates from different studies may not be directly comparable. A model designed to detect a specific gait event, for example, cannot be evaluated on the same terms as a model intended to predict walking independence or measure treatment response.

A second concern is the distinction between technical performance and clinical relevance. AI studies frequently report measures such as accuracy, sensitivity, specificity, or F1-score, depending on the task. These metrics are useful, but they do not independently establish whether a system improves rehabilitation decisions, detects meaningful functional change, reduces assessment burden, or leads to better patient outcomes. For predictive models, calibration and external validation may be as important as discrimination. For measurement systems, agreement with an appropriate reference standard and reliability across repeated assessments are also important [2, 17, 26, 27].

A third research gap concerns generalizability. Models trained using small, single-site, or highly controlled datasets may not perform similarly in routine rehabilitation settings. Variations in neurological presentation, disease severity, age, assistive-device use, walking environment, and sensor configuration can affect model performance. Limited external validation may therefore restrict confidence in the transportability of published findings. In addition, class imbalance, participant overlap between training and testing datasets, and inadequate separation of repeated recordings from the same individual can lead to overly optimistic estimates if not handled appropriately.

A fourth challenge is the integration of AI outputs into clinical practice. Automated classification or prediction may be technically feasible without providing information that is actionable for a rehabilitation professional. Questions remain regarding interpretability, uncertainty communication, workflow integration, accessibility, privacy, and the practical costs of implementation. The value of AI-based gait analysis should therefore be assessed in relation to its intended clinical purpose, not solely in terms of computational novelty.

Finally, neurological rehabilitation encompasses diverse conditions and treatment objectives. Evidence obtained from one diagnostic population or gait abnormality may not be transferable to another. A review that examines the field broadly must account for these differences while identifying shared methodological patterns and opportunities for cross-condition learning. It must also distinguish evidence supporting measurement or classification from evidence demonstrating improved rehabilitation outcomes.

These issues provide the rationale for a systematic review of AI-based gait analysis in neurological rehabilitation. By systematically identifying and synthesizing relevant studies, the review can characterize the technologies being investigated, compare their intended applications, summarize reported performance and validation practices, and identify limitations affecting interpretation and clinical translation. The review may also help establish priorities for future research, including more representative datasets, participant-level data separation, external validation, transparent reporting, clinically meaningful outcomes, and prospective evaluation in rehabilitation settings.

The purpose is not to assume that AI is superior to conventional assessment, but to determine what the available evidence demonstrates, where findings remain uncertain, and which developments are most likely to support safe, reliable, and clinically meaningful neurological rehabilitation.

1.5 Review Objectives and Research Questions

1.5.1 General objective

The general objective of this systematic review is to evaluate the current evidence on AI-based gait analysis in neurological rehabilitation, with particular attention to computational methods, data-acquisition technologies, gait-assessment applications, reported performance, validation practices, and barriers to clinical implementation.

1.5.2 Specific objectives

The review will aim to:

1. Identify and classify AI techniques used for gait analysis in people with neurological disorders.
2. Examine the sensing technologies and data sources used to acquire gait-related information.
3. Summarize the gait parameters, movement abnormalities, and clinical outcomes assessed by AI-based systems.
4. Evaluate the reported performance, reliability, and validation of these systems using task-appropriate measures.
5. Investigate applications in functional assessment, rehabilitation monitoring, outcome prediction, and individualized rehabilitation support.
6. Identify methodological weaknesses, including dataset limitations, risk of bias, insufficient external validation, and limited interpretability.

7. Describe ethical, technical, and practical barriers to clinical implementation and propose priorities for future research.

1.5.3 Research questions

The review will address the following research questions:

- RQ1: Which AI techniques are used in gait analysis for neurological rehabilitation, and what tasks are they designed to perform?
- RQ2: Which data-acquisition technologies and gait features are most frequently used in the included studies?
- RQ3: What performance, reliability, and validation results have been reported for AI-based gait-analysis systems?
- RQ4: How are these systems applied to clinical assessment, rehabilitation monitoring, functional prediction, and individualized care?
- RQ5: What methodological, ethical, technical, and clinical barriers limit the generalizability and implementation of AI-based gait analysis?
- RQ6: What research priorities could improve the reliability, transparency, and clinical usefulness of AI-based gait assessment?

By addressing these questions, the review will seek to provide a structured account of the current evidence while distinguishing established findings from promising but insufficiently validated applications.

2. Materials and Methods

2.1 Review Design and Reporting Guidelines

This study will be conducted as a systematic review of the scientific literature examining the application of artificial intelligence in gait analysis for neurological rehabilitation. The review will adopt a structured and reproducible methodology to identify relevant studies, evaluate their methodological characteristics, and synthesize evidence concerning AI techniques, gait-assessment technologies, clinical applications, and reported performance. A systematic review approach is appropriate because the research area encompasses multiple neurological populations, data-acquisition methods, computational approaches, and clinical assessment objectives.

The review will follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) reporting guideline. PRISMA 2020 provides recommendations for transparent reporting of study identification, screening, eligibility assessment, inclusion, and synthesis. Its application will help ensure that the review process is documented clearly and that the reasons for excluding studies are reported systematically [12, 13].

Protocol registration: [Insert the registry name and registration identifier if the protocol has been registered. If it has not been registered, state this accurately.]

Before the literature search begins, the review protocol will define the research questions, eligibility criteria, information sources, search strategy, screening procedures, data-extraction fields, risk-of-bias assessment, and synthesis plan. Any substantive amendments made during the review will be documented with

explanations. Where eligible, the protocol may be registered in PROSPERO or made publicly available through an appropriate repository [12].

The review will focus on primary research studies that investigate AI-based gait analysis in people with neurological disorders or evaluate technologies with a clearly defined application to neurological rehabilitation. Relevant study designs may include observational studies, diagnostic-accuracy studies, prospective or retrospective investigations, experimental studies, and clinical validation studies. Methodological or technical studies will also be considered when they provide empirical evidence relevant to the review questions.

The review will distinguish between studies demonstrating technical feasibility and those evaluating clinical validity or rehabilitation outcomes. This distinction is important because successful algorithm development does not necessarily establish reliability in routine clinical practice or demonstrate improvements in patient outcomes [16, 17, 26].

2.2 Eligibility Criteria

Eligibility criteria will be established before screening to promote consistency and minimize selection bias. The criteria will be organized around the study population, AI-based intervention or analytical approach, gait-related outcomes, study design, and publication characteristics.

2.2.1 Inclusion criteria

Studies will be eligible for inclusion when they satisfy the following criteria:

1. Population: Studies involving individuals with neurological disorders associated with gait impairment, including stroke, Parkinson's disease, multiple sclerosis, cerebral palsy, spinal cord injury, and other relevant neurological conditions. Studies involving mixed populations will be considered when neurological participants or findings can be identified separately, or when the study directly addresses a technology intended for neurological rehabilitation.
2. Technology and analytical approach: Studies applying AI techniques, including machine learning, deep learning, neural networks, computer vision, or other data-driven learning methods, to gait-related data. Relevant applications may include gait-event detection, abnormal gait classification, movement-pattern recognition, functional prediction, and rehabilitation monitoring.
3. Data sources: Studies using wearable sensors, inertial measurement units, accelerometers, gyroscopes, pressure-sensitive devices, force plates, optical motion capture, video recordings, skeletal pose estimation, or other appropriate sources of gait-related information.
4. Outcomes: Studies reporting at least one relevant gait-analysis outcome, such as gait speed, cadence, stride length, step symmetry, gait-event timing, gait variability, freezing-of-gait detection, classification performance, measurement agreement, reliability, functional mobility, or rehabilitation-related outcomes.
5. Study design: Original empirical research, including observational, experimental, diagnostic, predictive, and clinical validation studies, as well as relevant technical studies with reported experimental results.
6. Publication characteristics: Full-text publications in English, from database inception to the final search date. Any language or date restrictions beyond these criteria must be specified and justified in the protocol before screening.

2.2.2 Exclusion criteria

Studies will be excluded if they meet any of the following conditions:

1. They do not investigate gait, walking performance, or a directly relevant locomotor assessment.
2. They discuss AI conceptually without presenting relevant empirical evidence.
3. They focus exclusively on non-neurological populations without a relevant neurological subgroup or applicable validation evidence.
4. They describe general rehabilitation technologies without evaluating an AI-based gait-analysis component.
5. They are editorials, opinion pieces, narrative reviews, conference abstracts without sufficient methodological information, or other publications that do not provide eligible primary research data.
6. They are duplicate publications or overlapping reports that do not contribute distinct relevant data.
7. Their full text or essential methodological information cannot be obtained after reasonable retrieval attempts.

Review articles may be examined to identify additional primary studies, but they will not be treated as primary empirical evidence in the main synthesis unless the protocol explicitly includes a separate evidence-mapping component.

The publication period, language restrictions, and treatment of conference papers will be finalized before screening. These decisions should reflect the scope of the review and the indexing practices of the selected databases rather than being chosen after examining the results.

2.3 Information Sources and Databases

A comprehensive literature search will be conducted across biomedical, rehabilitation, and engineering databases to capture the interdisciplinary nature of AI-based gait analysis. The selected sources will provide coverage of clinical research, biomechanics, sensor technologies, computational methods, and rehabilitation science.

The proposed databases are:

- PubMed/MEDLINE: For neurological disorders, clinical gait assessment, and rehabilitation research.
- Scopus: For interdisciplinary literature spanning medicine, engineering, computer science, and health sciences.
- Web of Science Core Collection: For multidisciplinary research and citation tracking.
- IEEE Xplore: For machine learning, wearable sensing, signal processing, computer vision, and biomedical engineering.
- Embase: For biomedical and clinical research, subject to institutional access.
- Cochrane Library: For relevant clinical evidence and references to primary studies.

The final database selection will depend on availability and the scope defined in the protocol. The search will also consider backward citation searching of included studies and relevant reviews, together with

forward citation searching where available. Supplementary searches may be conducted through appropriate scholarly platforms to identify eligible studies missed by database indexing.

For each database, the search date, platform, complete search string, filters, and number of retrieved records will be documented. Database-specific syntax and controlled vocabulary will be adapted where necessary. For example, MeSH terms may be used in PubMed, while equivalent subject headings or free-text terms will be applied in other databases [12].

The search will cover the period from the earliest available records to the final search date specified in the protocol. The final date must be recorded precisely so that the review can be reproduced and updated.

2.4 Search Strategy

The search strategy will combine terms relating to artificial intelligence, gait analysis, neurological disorders, and rehabilitation. Both controlled vocabulary and free-text keywords will be used to capture differences in terminology across disciplines and publications.

The search will be developed iteratively, beginning with a core strategy and then adapting it to each database. Boolean operators will combine related concepts, while truncation and phrase searching will be applied where supported.

The principal search concepts will include:

- Artificial intelligence: artificial intelligence, machine learning, deep learning, neural network, computer vision, pattern recognition, and predictive modelling.
- Gait analysis: gait analysis, gait assessment, walking, locomotion, gait pattern, gait event, spatiotemporal parameters, and movement analysis.
- Neurological conditions: stroke, Parkinson's disease, multiple sclerosis, cerebral palsy, spinal cord injury, neurological disorders, and neuromotor impairment.
- Rehabilitation: neurological rehabilitation, physical rehabilitation, gait rehabilitation, mobility assessment, and rehabilitation monitoring.

2.5 Study Selection Process

Study selection will be conducted in two principal stages: title-and-abstract screening and full-text assessment. The process will follow predefined eligibility criteria and will be documented using a PRISMA 2020 flow diagram.

First, all database records will be imported into reference-management or review software. Duplicate records will be identified using automated matching and, where necessary, manual verification. The number of records retrieved, duplicates removed, and records remaining for screening will be recorded.

Second, titles and abstracts will be screened to identify potentially relevant studies. Records will be excluded at this stage when they clearly fail to meet the eligibility criteria, such as studies unrelated to gait analysis, studies without an AI component, or studies involving exclusively irrelevant populations.

Third, full texts of potentially eligible records will be retrieved and assessed against all inclusion and exclusion criteria. Full-text exclusions will be assigned explicit reasons, such as an ineligible population, absence of an eligible AI-based gait-analysis method, an irrelevant outcome, or an ineligible publication type.

Where feasible, two reviewers will independently screen records at both stages. Disagreements will be resolved through discussion, with consultation from a third reviewer if consensus cannot be reached. If the review is conducted by a single reviewer, this will be reported transparently, together with any verification or checking procedures used to reduce selection errors.

The selection process will record the number of studies included in the final synthesis. Where multiple publications report the same participants or dataset, their relationship will be assessed to avoid double-counting participants or outcomes. Distinct analyses from the same dataset may be retained when they contribute separate relevant findings, provided their dependence is acknowledged.

The final report will include a PRISMA flow diagram showing the number of records identified, duplicates removed, records screened, full-text reports assessed, reports excluded with reasons, and studies included in the review.

2.6 Data Extraction

A standardized data-extraction form will be developed and pilot-tested on a small sample of eligible studies. The form will be designed to capture study characteristics, participant information, data-acquisition technologies, AI methods, assessment objectives, evaluation procedures, reported findings, and methodological limitations.

The following categories will be extracted where available:

Table 1. Data items to be extracted from each included study

Data category	Information to be collected
Bibliographic information	First author, publication year, title, journal, country, and DOI
Study design	Observational, experimental, diagnostic, predictive, or validation design
Population	Neurological diagnosis, sample size, age, sex distribution, disease severity, and relevant inclusion criteria
Data acquisition	Sensor type, device model, placement, sampling frequency, recording environment, and data modality
Gait features	Gait speed, cadence, stride length, step timing, symmetry, variability, joint kinematics, or other reported features
AI approach	Algorithm family, model architecture, input features, preprocessing, and classification or prediction task
Reference standard	Clinical diagnosis, clinician annotation, validated instrument, motion-capture reference, or other comparison method
Model development	Training, validation, and test-set design; participant-level separation; feature selection; and handling of class imbalance
Performance measures	Accuracy, sensitivity, specificity, precision, F1-score, area under the curve, error measures, calibration, or agreement metrics, as applicable
Validation	Internal validation, cross-validation, independent test set, external validation, and prospective clinical evaluation
Clinical relevance	Functional assessment, rehabilitation monitoring, outcome prediction, treatment-response assessment, or decision support
Limitations	Small samples, missing data, risk of bias, restricted generalizability, interpretability concerns, and implementation barriers

Numerical results will be extracted with their relevant context, including the evaluation task, test population, reference standard, and validation method. Performance metrics will not be interpreted in isolation because their meaning depends on the intended task and the way the evaluation was conducted.

When information is missing or unclear, the reviewers will distinguish between data that are not reported and data that are explicitly unavailable. Where feasible, study authors may be contacted for clarification of important missing information.

If two reviewers are available, data extraction will be conducted independently or through independent verification of the extracted data. Any discrepancies will be resolved through discussion and, where necessary, consultation with another reviewer.

When multiple reports describe the same study or participant cohort, data will be linked to the underlying study to minimize duplicate counting. Extraction decisions and any assumptions made during synthesis will be documented to support transparency and reproducibility.

2.7 Risk-of-Bias and Quality Assessment

The methodological quality of the included studies will be evaluated using assessment tools appropriate to their design and intended purpose. Because the review may include diagnostic, predictive, observational, experimental, and technical validation studies, a single assessment instrument may not adequately capture all relevant sources of bias.

For diagnostic-accuracy studies, QUADAS-2 may be considered to assess patient selection, the conduct and interpretation of index tests, reference standards, and participant flow. For studies developing or evaluating prediction models, PROBAST may be appropriate for examining participants, predictors, outcomes, and analysis. Other study designs will be assessed using suitable established tools selected in advance and justified in the protocol.

For AI-based prediction studies, additional attention will be given to methodological features that can affect the reliability of reported performance. These may include:

- Sample size relative to model complexity.
- Potential data leakage between training and testing datasets.
- Separation of participants rather than merely individual recordings.
- Appropriate handling of missing values and class imbalance.
- Transparency of preprocessing and feature-selection procedures.
- Adequacy of internal and external validation.
- Reporting of uncertainty, calibration, and relevant performance metrics.
- Representativeness of the study population.
- Availability of sufficient methodological detail to reproduce the analysis.

For gait-measurement systems, the assessment will also consider agreement with an appropriate reference method, measurement reliability, repeatability, and evidence that the system performs adequately in the intended environment.

Where applicable, the review may use a reporting guideline such as the relevant TRIPOD guidance for prediction-model studies, including applicable extensions for machine-learning prediction models. Reporting completeness and risk of bias will be treated as related but distinct concepts: a study may report its methods incompletely without necessarily demonstrating that the method itself is biased.

Quality assessments will be performed independently by two reviewers where feasible. Disagreements will be resolved through discussion or consultation with a third reviewer. If a single reviewer performs the assessment, this limitation will be acknowledged [16, 17, 27].

The findings will be summarized in tables and, where useful, visual risk-of-bias summaries. Studies will not automatically be excluded solely because they have methodological limitations; instead, the assessment will inform the interpretation of their findings and the confidence placed in the overall evidence. [11]

2.8 Data Synthesis and Analysis

The extracted data will be synthesized using a structured narrative approach, with quantitative synthesis considered when the available studies are sufficiently comparable. This approach is appropriate because AI-based gait-analysis research may differ substantially in neurological populations, sensor technologies, model architectures, outcome definitions, reference standards, and validation procedures.

2.8.1 Descriptive synthesis

The included studies will first be summarized descriptively. The synthesis will report the number of studies, publication years, neurological populations, study designs, data-acquisition technologies, AI approaches, and clinical applications. Study characteristics will be presented in structured tables to facilitate comparison across research settings.

Where meaningful, the review will describe the distribution of studies by neurological condition, sensing modality, AI model family, and intended assessment task. Counts and percentages may be reported for these characteristics, with denominators stated clearly when information is missing or categories overlap.

2.8.2 Classification of AI methods and applications

AI methods will be grouped into broad categories, such as traditional machine learning, deep learning, computer vision, and other data-driven approaches. These categories will be used to organize the evidence rather than to assume that one model family is inherently superior to another.

Applications will be classified according to their intended purpose, including:

1. Detection of gait events and movement abnormalities.
2. Classification of neurological gait patterns.
3. Estimation of temporal-spatial and biomechanical gait parameters.
4. Monitoring of gait variability and functional mobility.
5. Prediction of clinical or rehabilitation-related outcomes.
6. Support for individualized assessment and rehabilitation decision-making.

The synthesis will distinguish studies focused on algorithm development from studies that evaluate measurement validity, clinical performance, or patient-centered outcomes.

2.8.3 Analysis of reported model performance

Performance results will be summarized according to the specific task and evaluation design. Classification studies may report accuracy, sensitivity, specificity, precision, F1-score, or area under the receiver operating characteristic curve. Regression and gait-parameter estimation studies may report mean absolute error, root

mean square error, correlation, or agreement statistics. The suitability of each metric will be considered in relation to the study objective.

Results will not be pooled simply because studies report the same metric. Comparability of the neurological population, target outcome, reference standard, validation method, and test conditions will first be assessed. For example, high accuracy obtained from a participant-independent test set will be distinguished from performance obtained using randomly divided observations that may contain recordings from the same individuals in both training and test sets.

Where available, confidence intervals and measures of variability will be extracted. Findings from external validation studies will be considered separately from internally validated model-development studies because they provide different levels of evidence regarding generalizability.

2.8.4 Assessment of heterogeneity

Potential sources of heterogeneity will be examined qualitatively, including differences in diagnosis, disease severity, participant characteristics, sensor placement, sampling frequency, preprocessing, feature selection, model architecture, outcome definitions, and validation strategies.

The review will also examine whether studies use controlled laboratory recordings or data collected during everyday activities. Performance in a controlled setting may not directly predict performance in routine clinical practice or community environments.

Where a sufficiently comparable group of studies is identified, subgroup analyses may be considered according to neurological condition, sensing modality, model family, assessment task, or validation approach. Such analyses will be conducted only when supported by the available data and specified or justified in the protocol.

2.8.5 Meta-analysis

A meta-analysis will be performed only if there are sufficient studies with comparable populations, target outcomes, evaluation methods, and statistical information. If appropriate, effect estimates and uncertainty intervals will be calculated using methods suited to the outcome and study design. Between-study heterogeneity will be assessed using suitable statistical methods, and a random-effects model may be considered when variation in underlying effects is expected and the available data support such analysis.

Diagnostic-accuracy studies require particular care because sensitivity and specificity are often correlated and depend on thresholds and study populations. If a sufficient number of comparable studies is available, an appropriate diagnostic-accuracy meta-analysis may be considered rather than simply averaging reported accuracy percentages.

For studies reporting multiple models, outcomes, or datasets, the analysis will account for the dependence between estimates wherever possible. The review will avoid treating multiple results from the same participants as independent observations.

If the data are too heterogeneous or insufficient for meaningful pooling, the findings will be synthesized narratively and presented in structured tables. The absence of a meta-analysis will not be interpreted as evidence that the reviewed technologies are ineffective.

2.8.6 Interpretation of evidence and clinical implications

The final synthesis will integrate findings on technical performance, measurement reliability, validation, clinical relevance, and implementation barriers. Particular attention will be given to whether studies demonstrate practical utility beyond algorithmic classification or prediction.

The review will distinguish evidence supporting technical feasibility from evidence supporting clinical validity and evidence demonstrating improved rehabilitation outcomes. It will also identify recurring limitations, such as small samples, insufficient external validation, lack of representative populations, limited interpretability, and inadequate evaluation in real-world settings.

The findings will be used to identify methodological priorities for future research, including standardized outcome definitions, transparent reporting, robust validation, representative datasets, and prospective evaluation of clinical utility. Conclusions will be proportionate to the strength and consistency of the available evidence.

No new patient-level data will be collected for this review. Ethical approval and informed consent are generally not required for a review of previously published literature, although institutional requirements and the journal's policies will be checked before submission.

3. Planned evidence synthesis and reporting

The review findings will be reported only after the database searches, deduplication, screening, full-text assessment, and data extraction have been completed and documented. No illustrative or hypothetical study counts will be presented as empirical results.

The PRISMA 2020 flow diagram will report the actual numbers of records identified, duplicates removed, records screened, full-text reports assessed, exclusions with reasons, and studies included. Each number will be traceable to the search and screening records.

A study-characteristics table will summarize each included study's bibliographic details, neurological population, sample size, design, data-acquisition technology, AI approach, target task, reference standard, validation strategy, performance metrics, and relevant limitations. Where information is not reported, it will be labelled "not reported" rather than inferred.

The narrative synthesis will group studies by neurological condition, data modality, AI method, and intended use. Classification, event detection, gait-parameter estimation, and outcome-prediction tasks will be analyzed separately because their metrics and clinical interpretations differ. Any quantitative pooling will be justified by clinical and methodological comparability and will account for dependence between estimates from the same participants or dataset.

The final report will distinguish internal validation from participant-independent testing and external validation. It will also separate technical performance from measurement agreement, clinical validity, feasibility, and evidence of improved rehabilitation outcomes. The discussion and conclusions will be revised to reflect the actual findings and the certainty of the evidence.

4. Discussion and anticipated contribution

4.1 Rationale and anticipated contribution

AI-based gait analysis is an evolving approach to the quantitative assessment of walking abnormalities in neurological rehabilitation. By combining movement data with computational techniques, these systems may help identify gait patterns, estimate movement parameters, detect abnormal events, and monitor changes in functional performance. The review is designed to assess these applications across neurological conditions without assuming that any method is clinically effective solely because it demonstrates promising technical performance.

Once the review is completed, findings will be interpreted in relation to the populations studied, data-acquisition technologies, AI approaches, outcome definitions, and methodological quality. The completed manuscript should state clearly whether the evidence supports reliable measurement, accurate classification, generalizable prediction, or improvements in rehabilitation practice; these are distinct claims and require different forms of evidence.

An important distinction exists between technical performance and clinical effectiveness. A model may demonstrate promising performance in identifying gait abnormalities within a research dataset without necessarily improving clinical decision-making or patient outcomes. Consequently, the usefulness of AI-based gait analysis should be evaluated not only through computational metrics but also through its reliability, reproducibility, clinical relevance, and applicability to routine rehabilitation.

The interpretation of the evidence must also account for differences in study populations, recording environments, sensor technologies, outcome definitions, and validation procedures. These differences may limit direct comparisons between studies and influence the extent to which findings can be generalized across neurological diagnoses and clinical settings.

The completed review will identify which applications are supported by consistent evidence, which remain at the proof-of-concept stage, and which require further validation. These conclusions cannot be drawn until the search and synthesis are complete.

4.2 Comparison of AI Approaches

AI-based gait analysis encompasses several methodological approaches, including traditional machine learning, deep learning, computer vision, time-series analysis, and predictive modelling. Each approach offers different advantages and limitations depending on the task, data characteristics, and intended clinical application.

Traditional machine-learning algorithms, such as support vector machines, random forests, and decision trees, commonly use predefined gait features extracted from sensor recordings. These features may include cadence, stride duration, acceleration statistics, gait variability, and measures of bilateral asymmetry. Such approaches can be practical when datasets are relatively small and the relevant gait characteristics are well defined. Their performance, however, depends on appropriate feature selection, preprocessing, and validation [7, 25, 28].

Deep-learning methods can learn complex spatial and temporal representations from sensor signals, video sequences, and body-joint trajectories. These methods may be useful when gait abnormalities involve subtle patterns that are difficult to describe using manually selected features. Nevertheless, deep-learning models often require substantial training data, careful tuning, and robust evaluation to reduce overfitting. Their complexity does not automatically guarantee better performance than traditional machine learning [7, 25].

Computer-vision approaches provide another important direction by enabling the estimation of body landmarks and movement trajectories from video. Markerless methods may reduce the need for specialized equipment and facilitate repeated assessments. However, their performance can be affected by lighting, camera position, occlusion, clothing, and limitations in two-dimensional movement estimation. Validation against appropriate reference systems is therefore important when these methods are used to estimate clinically relevant gait parameters. [11, 23]

Time-series methods focus on the temporal structure of gait signals and may be applied to identify heel strikes, toe-offs, freezing-of-gait episodes, or other abnormal movement events. Their evaluation requires task-specific metrics because the detection of short events differs from the classification of entire participants.

Table 2. Comparison of AI approaches in neurological gait analysis

Approach	Potential strengths	Principal limitations
Traditional machine learning	Can work with engineered gait features; some models are relatively interpretable	Depends on feature quality and preprocessing
Deep learning	Can learn complex spatial and temporal patterns	Data requirements, overfitting, and limited interpretability
Computer vision	Supports video-based and potentially markerless assessment	Sensitive to recording conditions and pose-estimation errors
Time-series analysis	Suitable for temporal events and movement patterns	Performance depends on signal quality and event definitions
Multimodal AI	Combines complementary sources of movement information	Greater integration, synchronization, and computational complexity
Predictive modelling	May estimate functional status or future outcomes	Requires appropriate validation, calibration, and clinically meaningful targets[25]

A fair comparison of these approaches requires studies to be evaluated according to their intended tasks, datasets, reference standards, and validation designs. Reported accuracy values should not be compared directly when the studies investigate different neurological conditions or use different outcome definitions. Where sufficiently comparable evidence is available, the review should examine whether any approach demonstrates a consistent advantage in performance, reliability, generalizability, or clinical utility. [7, 18, 25, 26]

4.3 Clinical Implications for Rehabilitation

AI-based gait analysis may support neurological rehabilitation by providing objective and repeatable information about walking performance. Conventional clinical assessment often relies on observation, standardized functional tests, and clinician judgment. Although these methods remain essential, automated measurement may provide additional quantitative detail about movement characteristics that are difficult to assess consistently through visual observation alone.

One potential application is the automated assessment of gait abnormalities. AI systems may help quantify changes in step timing, walking speed, gait symmetry, movement trajectories, and other features relevant to neurological impairment. Such information could complement clinical examinations and support the identification of movement characteristics requiring further assessment [18, 19, 20].

A second application is the monitoring of rehabilitation progress. Repeated measurements may help clinicians examine whether gait characteristics change over time and whether these changes correspond with functional improvement. Wearable sensors and video-based systems could facilitate assessments

outside specialized gait laboratories, provided that their measurements are sufficiently reliable under the intended conditions [19, 20].

AI may also support the assessment of specific neurological gait patterns, including freezing of gait in Parkinson's disease and asymmetric walking after stroke. However, the clinical value of such applications depends on appropriate validation in the relevant patient population. Performance demonstrated in healthy participants or highly controlled experimental settings may not transfer directly to patients with complex neurological impairments. [28]

Importantly, AI should be regarded as a potential decision-support tool rather than a replacement for clinical expertise. Rehabilitation decisions require consideration of multiple factors, including strength, balance, cognition, fatigue, pain, patient goals, comorbidities, and environmental barriers. Gait-analysis outputs should therefore be interpreted alongside other clinical information.

The final review should distinguish systems that have been evaluated only for technical feasibility from those that have been tested during actual rehabilitation practice. Evidence of improved measurement accuracy alone is insufficient to establish that an AI system improves patient care.

4.4 Personalized and Predictive Rehabilitation

Personalized rehabilitation aims to adapt assessment and treatment to the functional characteristics, needs, and progress of individual patients. AI-based gait analysis may contribute to this objective by identifying movement patterns, quantifying functional differences, and tracking changes across repeated assessments.

For example, gait measurements may help characterize differences in walking speed, stride length, symmetry, and movement variability among patients with similar diagnoses. Such information could contribute to identifying individualized rehabilitation targets and monitoring whether particular aspects of gait change during treatment. [29]

Predictive modelling may extend these applications by estimating functional status, identifying patterns associated with poor mobility, or predicting rehabilitation-related outcomes. However, the validity of these predictions depends on the quality and representativeness of the training data, the definition of the predicted outcome, and the evaluation of the model in independent populations.

A model developed using a single hospital dataset may reflect local patient characteristics, treatment practices, equipment, or data-collection procedures. Its predictions may therefore be less reliable when applied to different clinical settings. External validation is particularly important before predictive models are used to guide high-impact rehabilitation decisions.

Personalized rehabilitation also requires a distinction between prediction and treatment effectiveness. A model that accurately predicts recovery does not necessarily identify which intervention will produce the best outcome for an individual patient. Demonstrating treatment-selection benefits requires appropriate clinical evaluation, ideally including comparisons of AI-informed rehabilitation with usual care or other relevant strategies.

Future studies should investigate whether AI-generated assessments can support individualized treatment planning, identify meaningful changes in function, and improve patient-centered outcomes. Such evaluations should include functional independence, walking ability, participation, quality of life, and other outcomes relevant to patients and rehabilitation professionals.

4.5 Explainability and Clinical Interpretability

Explainability is an important consideration in the clinical application of AI-based gait analysis. Rehabilitation professionals need to understand what a model measures, what its output represents, and how much confidence should be placed in its predictions. Systems that provide only a classification or risk score may be difficult to interpret when the clinical reasoning behind the output is unclear.

Some traditional machine-learning approaches offer relatively straightforward explanations when their features and decision rules are transparent. Deep-learning systems, by contrast, may use complex internal representations that are difficult to interpret directly. Explainability methods, including feature attribution and visual explanations, may help investigate which inputs contribute to a model's output.

However, an explanation of a model's prediction should not automatically be interpreted as evidence of a causal relationship. A feature identified as important may reflect a genuine gait characteristic, a participant-specific pattern, a recording artifact, or a bias in the dataset. Explanations should therefore be assessed for stability, plausibility, and consistency with relevant clinical knowledge. [30]

Clinical interpretability also depends on how the results are presented. Outputs expressed in familiar gait parameters may be easier to integrate into rehabilitation workflows than unexplained numerical scores. Where appropriate, systems should communicate measurement uncertainty, limitations, and conditions under which their predictions may be unreliable.

Future research should evaluate not only whether AI explanations can be generated but also whether they help clinicians understand model behavior, identify errors, and make better-informed decisions. Human-factors studies and clinical evaluations will be important for determining whether explainable AI provides practical benefits in neurological rehabilitation.

4.6 Technical, Ethical, and Implementation Challenges

Despite its potential, AI-based gait analysis faces technical, ethical, and practical challenges that may limit its adoption in clinical rehabilitation.

Technical challenges

Data quality is a fundamental concern. Wearable sensors may be affected by inconsistent placement, device calibration, motion artifacts, and missing recordings. Video-based methods may be sensitive to lighting, occlusion, camera angle, and variations in movement. Differences between devices and acquisition protocols can also make it difficult to combine datasets or compare findings across studies.

Dataset size and representativeness are additional concerns. Neurological gait abnormalities can vary considerably according to disease severity, age, functional ability, medication status, assistive-device use, and other patient characteristics. Models developed from small or homogeneous datasets may not perform reliably across the broader clinical population.

Data leakage can further inflate reported performance. If recordings from the same participant are distributed across training and test datasets, the evaluation may not accurately represent performance in previously unseen patients. Participant-independent data partitioning and appropriate external validation are therefore important.

Ethical and privacy considerations

Gait recordings, video images, and sensor signals may contain sensitive personal information. Systems intended for remote or continuous monitoring should incorporate appropriate data protection, access controls, informed consent, and secure data management.

Potential algorithmic bias also requires attention. If training datasets underrepresent particular age groups, disability levels, neurological diagnoses, or demographic populations, model performance may differ across groups. Evaluations should examine subgroup performance where sample sizes permit reliable assessment.

The use of AI should not reduce patient autonomy or remove appropriate human oversight. Patients and clinicians should understand the intended purpose of the technology, its limitations, and the role of its outputs in clinical decision-making [26, 27].

Implementation challenges

Clinical adoption depends on more than technical performance. Costs, equipment availability, staff training, workflow integration, interoperability, maintenance, and technical support may influence whether a system can be used consistently.

Systems designed for specialized laboratories may be unsuitable for resource-limited rehabilitation facilities. Conversely, lower-cost wearable or video-based systems may require additional validation before being used for decisions that depend on precise measurement [29, 30].

Implementation studies should examine usability, clinician acceptance, patient adherence, cost-effectiveness, and the consequences of false-positive and false-negative outputs. Clear procedures for human review and handling uncertain predictions are also important.

Overall, successful implementation requires collaboration among rehabilitation clinicians, engineers, data scientists, patients, and healthcare organizations. Technical development should proceed alongside evaluation of safety, usability, accessibility, and clinical value.

4.7 Research Gaps and Future Directions

Several research priorities should be addressed to strengthen the evidence for AI-based gait analysis in neurological rehabilitation.

First, future studies should recruit larger and more representative patient populations. Multicenter studies involving different neurological diagnoses, disease severities, age groups, and functional abilities could help establish whether models generalize beyond narrowly selected research samples.

Second, independent external validation should become a central component of model evaluation. Studies should report participant-level data separation, preprocessing procedures, missing-data management, and the methods used to select and tune models. Transparent reporting would improve reproducibility and help readers assess whether reported performance is likely to translate into clinical practice.

Third, greater standardization of gait parameters and evaluation procedures is needed. Differences in outcome definitions, sensor specifications, walking protocols, and performance metrics can prevent meaningful comparisons between studies. Researchers should clearly define their target variables, reference standards, and evaluation units.

Fourth, the relationship between technical performance and clinical utility requires further investigation. Prospective clinical studies should assess whether AI-supported gait analysis improves rehabilitation planning, monitoring, decision-making, or patient outcomes compared with existing approaches. Where

feasible, these studies should evaluate clinically meaningful changes rather than relying exclusively on computational performance metrics.

Fifth, research should expand beyond controlled laboratory environments. Real-world monitoring could provide information about walking patterns during daily activities, but such applications require assessment of data quality, adherence, privacy, device reliability, and feasibility for people with substantial mobility limitations.

Sixth, explainability and uncertainty estimation should receive greater attention. Models should communicate when their predictions are unreliable and should be evaluated for performance across relevant patient subgroups. Explainability methods should be tested for practical value rather than treated as sufficient evidence of clinical transparency.

Finally, research should examine the cost-effectiveness and accessibility of AI-based gait-analysis systems. Lower-cost solutions may have particular relevance where access to specialized gait laboratories is limited, but affordability should not come at the expense of measurement validity, safety, or patient privacy.

These priorities should guide the interpretation of the final systematic-review findings. The most important research gaps should be selected according to the limitations actually observed in the included studies rather than assumed in advance.

5. Limitations of the Protocol

Several limitations may affect the planned review. The actual limitations will be reported transparently after the search and synthesis have been completed.

First, the findings may be influenced by the search strategy and the databases selected. Relevant studies may be missed if they use different terminology for artificial intelligence, gait analysis, movement assessment, or neurological rehabilitation. The omission of databases, unpublished research, conference proceedings, or non-English publications may also affect the completeness of the evidence. The final manuscript should specify the databases searched, the search dates, the language restrictions applied, and any other limits imposed.

Second, substantial heterogeneity may exist among the included studies. Differences in neurological diagnoses, participant characteristics, sample sizes, acquisition technologies, AI architectures, and assessment environments may limit direct comparison. Studies involving stroke, Parkinson's disease, cerebral palsy, and spinal cord injury may address distinct clinical questions and should not automatically be treated as evidence for a single common application.

Third, outcome measures and reporting practices may differ considerably. Classification metrics, gait-event detection measures, parameter-estimation errors, and clinical outcome measures represent different types of evidence. Even when studies report the same metric, differences in class distribution, reference standards, or validation methods may prevent valid comparisons. A pooled quantitative analysis should only be conducted when sufficient methodological and clinical comparability is established.

Fourth, the quality of the available evidence may be limited by small datasets, restricted participant recruitment, inadequate external validation, incomplete reporting, or potential data leakage. Such limitations may reduce confidence in reported model performance and constrain generalizability. The actual review should document these concerns through a structured risk-of-bias or methodological-quality assessment.

Fifth, publication and reporting biases may affect the evidence base. Studies with favorable results may be more likely to be published or to emphasize their strongest performance metrics. The extent of this problem will depend on the available evidence and whether appropriate assessment methods can be applied.

Sixth, the rapid development of AI technologies may limit the currency of the review. New algorithms, sensing devices, and validation studies may become available after the final search date. The manuscript should report its last search date so that readers can assess the period covered.

At the protocol stage, the number and characteristics of eligible studies and their findings are not yet known. Accordingly, this document does not present empirical review results. Any future manuscript that reports completed results must replace this protocol-stage text with findings derived from a documented search, screening process, verified data extraction, and structured appraisal.

6. Conclusion

This protocol sets out a reproducible approach for reviewing AI-based gait analysis in neurological rehabilitation. The planned review will examine wearable sensors, video-based computer vision, motion-capture systems, and machine-learning methods used for gait classification, event detection, movement-parameter estimation, and rehabilitation-related assessment.

The review will assess whether reported performance is supported by reliable measurement, representative data, participant-independent evaluation, external validation, and clinically relevant outcomes. Technical performance alone will not be treated as evidence that a system improves clinical decisions or patient outcomes.

The synthesis will identify research priorities based on the evidence actually included, with attention to representative multicenter datasets, transparent reporting, appropriate validation, standardized outcomes, uncertainty estimation, and prospective evaluation of clinical utility.

The completed systematic review will provide an evidence-based account of the current applications, strengths, limitations, and clinical readiness of AI-based gait analysis. Until the planned searches and analyses are completed, conclusions about the distribution or effectiveness of these technologies should remain provisional.

References

1. Bächlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J. M., Giladi, N., & Tröster, G. (2010). A wearable system to assist walking of Parkinson's disease patients. *IEEE Transactions on Information Technology in Biomedicine*, 14(2), 436–446. <https://doi.org/10.1109/TITB.2009.2036165>
2. Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine*, 162(1), 55–63. <https://doi.org/10.7326/M14-0697>
3. Del Din, S., Godfrey, A., Galna, B., Lord, S., & Rochester, L. (2016). Free-living gait characteristics in ageing and Parkinson's disease: Impact of environment and ambulatory bout length. *Journal of NeuroEngineering and Rehabilitation*, 13, Article 46. <https://doi.org/10.1186/s12984-016-0154-5>
4. Hausdorff, J. M. (2007). Gait dynamics, fractals and falls: Finding meaning in the stride-to-stride fluctuations of human walking. *Human Movement Science*, 26(4), 555–589. <https://doi.org/10.1016/j.humov.2007.05.003>

5. Hausdorff, J. M., Rios, D. A., & Edelberg, H. K. (2001). Gait variability and fall risk in community-living older adults: A 1-year prospective study. *Archives of Physical Medicine and Rehabilitation*, 82(8), 1050–1056. <https://doi.org/10.1053/apmr.2001.24893>
6. Iosa, M., Picerno, P., Paolucci, S., & Morone, G. (2016). Wearable inertial sensors for human movement analysis. *Expert Review of Medical Devices*, 13(7), 641–659. <https://doi.org/10.1080/17434440.2016.1198694>
7. Kohneshahri, F. S., Merlo, A., Mazzoli, D., Bò, M. C., & Stagni, R. (2024). Machine learning applied to gait analysis data in cerebral palsy and stroke: A systematic review. *Gait & Posture*, 111, 105–121. <https://doi.org/10.1016/j.gaitpost.2024.04.007>
8. Lima, A. L. S., Evers, L. J. W., Hahn, T., Bataille, L., Hamilton, J. L., Little, M. A., Okuma, Y., Bloem, B. R., & Faber, M. J. (2017). Freezing of gait and fall detection in Parkinson's disease using wearable sensors: A systematic review. *Journal of Neurology*, 264(8), 1642–1654. <https://doi.org/10.1007/s00415-017-8424-0>
9. Lin, S., Gao, C., Li, H., Huang, P., Ling, Y., Chen, Z., Ren, K., & Chen, S. (2023). Wearable sensor-based gait analysis to discriminate early Parkinson's disease from essential tremor. *Journal of Neurology*, 270(4), 2283–2301. <https://doi.org/10.1007/s00415-023-11577-6>
10. Lord, S., Galna, B., & Rochester, L. (2013). Moving forward on gait measurement: Toward a more refined approach. *Movement Disorders*, 28(11), 1534–1543. <https://doi.org/10.1002/mds.25545>
11. Muro-de-la-Herran, A., García-Zapirain, B., & Méndez-Zorrilla, A. (2014). Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications. *Sensors*, 14(2), 3362–3394. <https://doi.org/10.3390/s140203362>
12. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>
13. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372, n160. <https://doi.org/10.1136/bmj.n160>
14. Pardoel, S., Kofman, J., Nantel, J., & Lemaire, E. D. (2019). Wearable-sensor-based detection and prediction of freezing of gait in Parkinson's disease: A review. *Sensors*, 19(23), Article 5141. <https://doi.org/10.3390/s19235141>
15. Peraza, L. R., Kinnunen, K. M., McNaney, R., Craddock, I. J., Whone, A. L., Morgan, C., Joules, R., & Wolz, R. (2021). An automatic gait analysis pipeline for wearable sensors: A pilot study in Parkinson's disease. *Sensors*, 21(24), Article 8286. <https://doi.org/10.3390/s21248286>
16. Whiting, P. F., Rutjes, A. W. S., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., Leeflang, M. M. G., Sterne, J. A. C., & Bossuyt, P. M. M. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*, 155(8), 529–536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
17. Wolff, R. F., Moons, K. G. M., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., et al. (2019). PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine*, 170(1), 51–58. <https://doi.org/10.7326/M18-1376>

18. Boukhenoufa, I., Zhai, X., Utti, V., Jackson, J., & McDonald-Maier, K. D. (2022). Wearable sensors and machine learning in post-stroke rehabilitation assessment: A systematic review. *Biomedical Signal Processing and Control*, 71(Part B), 103197. <https://doi.org/10.1016/j.bspc.2021.103197>
19. Michaud, M., de l'Escalopier, N., Voisard, C., Oudre, L., et al. (2026). Real-world gait study of Parkinson's disease using wearable sensors: A systematic review. *Journal of NeuroEngineering and Rehabilitation*, 23, Article 251. <https://doi.org/10.1186/s12984-026-01957-0>
20. Neumann, S., Albites-Sanabria, J., Naef, A. C., Luft, A. R., Chiari, L., & Easthope Awai, C. (2026). Wearable sensors in gait assessment for Parkinson disease and stroke in real-world environments: Systematic review and meta-analysis. *Journal of Medical Internet Research*, 28, Article e93413. <https://doi.org/10.2196/93413>
21. Huang, T. H., Li, M., & Huang, J. (2023). Recent trends in wearable device used to detect freezing of gait and falls in people with Parkinson's disease: A systematic review. *Frontiers in Aging Neuroscience*, 15, Article 1119956. <https://doi.org/10.3389/fnagi.2023.1119956>
22. Salarian, A., Russmann, H., Vingerhoets, F. J. G., Dehollain, C., Blanc, Y., Burkhard, P. R., & Aminian, K. (2004). Gait assessment in Parkinson's disease: Toward an ambulatory system for long-term monitoring. *IEEE Transactions on Biomedical Engineering*, 51(8), 1434–1443. <https://doi.org/10.1109/TBME.2004.827933>
23. Tao, W., Liu, T., Zheng, R., & Feng, H. (2012). Gait analysis using wearable sensors. *Sensors*, 12(2), 2255–2283. <https://doi.org/10.3390/s120202255>
24. Zhang, W., Sun, H., Huang, D., Zhang, Z., Li, J., Wu, C., Sun, Y., Gong, M., Wang, Z., Sun, C., Cui, G., Guo, Y., & Chan, P. (2024). Detection and prediction of freezing of gait with wearable sensors in Parkinson's disease. *Neurological Sciences*, 45(2), 431–453. <https://doi.org/10.1007/s10072-023-07017-y>
25. Lim, A. C. Y., Natarajan, P., Fonseka, R. D., Maharaj, M., & Mobbs, R. J. (2022). The application of artificial intelligence and custom algorithms with inertial wearable devices for gait analysis and detection of gait-altering pathologies in adults: A scoping review of literature. *Digital Health*, 8, 20552076221074128. <https://doi.org/10.1177/20552076221074128>
26. Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., et al. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*, 385, Article e078378. <https://doi.org/10.1136/bmj-2023-078378>
27. Moons, K. G. M., Wolff, R. F., Riley, R. D., Whiting, P. F., Westwood, M., Collins, G. S., et al. (2019). PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. *Annals of Internal Medicine*, 170(1), W1–W33. <https://doi.org/10.7326/M18-1377>
28. McGinnis, R. S., Mahadevan, N., Moon, Y., Seagers, K., Sheth, N., Wright, J. A., DiCristofaro, S., Jortberg, E., Ceruolo, M., & Silva, I. (2017). A machine learning approach for gait speed estimation using skin-mounted wearable sensors: From healthy controls to individuals with multiple sclerosis. *PLOS ONE*, 12(6), e0178366. <https://doi.org/10.1371/journal.pone.0178366>
29. Kobsar, D., Charlton, J. M., Tse, C. T. F., Esculier, J.-F., Graffos, A., Krowchuk, N. M., Thatcher, D., & Hunt, M. A. (2020). Validity and reliability of wearable inertial sensors in healthy adult walking: A systematic review and meta-analysis. *Journal of NeuroEngineering and Rehabilitation*, 17, Article 62. <https://doi.org/10.1186/s12984-020-00685-3>

30. Godfrey, A., Del Din, S., Barry, G., Mathers, J. C., & Rochester, L. (2015). Instrumenting gait with an accelerometer: A system and algorithm examination. *Medical Engineering & Physics*, 37(4), 400–407. <https://doi.org/10.1016/j.medengphy.2015.02.003>