Research Proposal: Development and Validation of the Distributional Gait Assessment Index: A Novel Probabilistic Approach to Quantifying Gait Abnormality

Stimuloop Research Consortium

November 4, 2025

Abstract

We propose the **Distributional Gait Assessment Index (DGAI)**, a novel data-driven metric that quantifies patient improvement in neurological rehabilitation by measuring deviation from normative gait patterns. Unlike the existing Gait Deviation Index which relies on discrete kinematic parameters, the DGAI models complete gait cycles as probability distributions, learns normative patterns from healthy controls, and measures patient deviation using optimal distributional distance metrics. By processing 4D spatio-temporal motion capture data and constructing individual and population-level distributions, the DGAI provides objective, comprehensive assessment of gait quality that captures natural movement variability. This distributional approach offers a sensitive tool for tracking rehabilitation progress across diverse neurological conditions, addressing critical gaps in current gait assessment methodologies.

Contents

1 Introduction **3**

2 Research Objectives **3**

3 Literature Review **4**

3.1 Current Approaches 4

3.1.1 Existing Gait Deviation Index 4

3.2 Probabilistic Approaches 4

3.3 Generative Models 4

3.4 Physics-Constrained Models 4

3.5 Distributional Assessment Framework 5

3.5.1 Beyond the Traditional Gait Deviation Index 5

3.6 Distance Measures 5

3.7 Clinical Context 5

4 Methodology **6**

4.1 Framework Overview 6

4.2 Data Representation 6

4.3 Gait Cycle Segmentation and Normalization 6

4.4 Individual Gait Distribution Estimation 6

4.5 Normative Distribution Construction 7

4.5.1 Distribution Aggregation 7

4.5.2 Demographic Conditioning 7

4.6 The Distributional Gait Assessment Index (DGAI) 7

4.6.1 DGAI Definition 7

4.6.2 DGAI Interpretation 8

4.6.3 Distance Measure Selection 8

4.6.4 DGAI Validation Framework 9

4.7 Study Design and Participants 9

4.7.1 Phase 1: Normative Database Construction 9

4.7.2 Phase 2: Clinical Validation 9

4.8 Validation 9

5 Expected Outcomes **10**

5.1 Primary Outcome: The Validated Distributional Gait Assessment Index . 10

5.2 Supporting Deliverables 10

5.3 Clinical Impact 10

1 Introduction

Current gait assessment methods rely on discrete parameters or subjective scales, failing to capture the holistic nature of human locomotion (Baker, 2013). Traditional approaches examine isolated parameters, missing complex multidimensional changes in neurological conditions (Whittle, 2014).

We propose the **Distributional Gait Assessment Index (DGAI)**, which models gait as probability distributions over complete cycles to capture natural variability and movement quality (Hausdorff et al., 2001). While the existing Gait Deviation Index provides valuable kinematic assessment, the DGAI addresses the critical gap in comprehensive distributional gait assessment tools by quantifying deviation from healthy patterns and providing sensitive detection of rehabilitation improvements (Horst et al., 2019). This novel metric represents a paradigm shift from discrete parameter analysis to holistic distributional assessment of human locomotion.

2 Research Objectives

Primary Objective:

- Develop and validate the **Distributional Gait Assessment Index (DGAI)** as a comprehensive distributional metric for quantifying deviation from healthy gait patterns

Secondary Objectives:

1. Establish robust normative gait distributions with demographic conditioning (age, sex, height, weight)
2. Validate DGAI sensitivity for detecting rehabilitation improvements in stroke and Parkinson's patients
3. Optimize distributional distance measures for maximal clinical sensitivity and interpretability
4. Demonstrate DGAI reliability and responsiveness across diverse rehabilitation scenarios
5. Create a standardized assessment protocol for clinical implementation of the DGAI

3 Literature Review

3.1 Current Approaches

Traditional gait analysis uses discrete measurements (stride length, cadence, joint angles) that fail to capture coordinated locomotion dynamics (Baker, 2013; Whittle, 2014).

3.1.1 Existing Gait Deviation Index

The Gait Deviation Index (GDI), developed by **?**, represents a significant advancement in gait assessment by providing a single summary measure of overall gait pathology. The GDI uses principal component analysis of nine key kinematic variables from 3D gait analysis to create a multivariate index, where 100 represents typical gait and lower scores indicate greater deviation. While the GDI has proven valuable for clinical assessment and research (**?**), it is limited to discrete kinematic parameters and does not capture the full distributional characteristics of gait patterns or the natural variability inherent in human locomotion (**?**).

3.2 Probabilistic Approaches

Probabilistic models show promise for comprehensive gait representation. Gholami et al. (2019) demonstrated hidden Markov models for temporal dependencies. Stihi et al. (2025) used Hierarchical Variational Sparse Heteroscedastic Gaussian Processes for multiple sclerosis analysis, capturing variability patterns and neurological impairment hallmarks.

3.3 Generative Models

Recent AI-driven generative models address data scarcity in clinical populations. Peppes et al. (2023) developed FoGGAN for Parkinson's episodes. Adeli et al. (2025) introduced GAITGen for conditional 3D gait generation. Yamada et al. (2025) demonstrated physics-based approaches achieving superior performance through synthetic data diversity. Rezvani et al. (2025) applied diffusion models for Parkinson's severity assessment.

3.4 Physics-Constrained Models

Takeishi and Kalousis (2021) developed physics-integrated VAEs embedding biomechanical models for realistic gait generation. Ghosh et al. (2025) validated biomechanical feasibility requirements, ensuring plausible and interpretable synthetic patterns.

3.5 Distributional Assessment Framework

Our approach leverages distributional perspectives for assessment rather than generation. We focus on measuring deviation from normative patterns using probabilistic representations (Adeli et al., 2025; Takeishi and Kalousis, 2021). This provides direct quantitative measures of gait abnormality while capturing variability inherent in neurological conditions. Zhang et al. (2025) demonstrated 65% improvement in estimation accuracy through distributional modeling, supporting our hypothesis that comprehensive distributional representations provide more sensitive assessment tools.

3.5.1 Beyond the Traditional Gait Deviation Index

While the established GDI provides valuable clinical insights through kinematic parameter analysis, it has inherent limitations. The GDI reduces complex gait patterns to a single score based on nine discrete variables, potentially missing subtle distributional changes and natural movement variability that characterize neurological conditions. Our proposed Distributional Gait Assessment Index (DGAI) addresses these limitations by:

- Modeling complete gait cycles as probability distributions rather than discrete parameters
- Capturing natural movement variability and temporal dynamics
- Providing sensitivity to distributional changes that may precede kinematic abnormalities
- Offering complementary information to the existing GDI for comprehensive assessment

The DGAI is designed to work alongside, not replace, the established GDI, providing a more comprehensive view of gait quality through distributional analysis.

3.6 Distance Measures

Distance metric selection critically impacts assessment sensitivity and interpretability. We will evaluate information-theoretic measures, geometric distances respecting gait structure, and optimal transport distances. Selection criteria include computational efficiency, clinical sensitivity, noise robustness, and practitioner interpretability.

3.7 Clinical Context

Post-stroke gait shows asymmetry and compensatory patterns (Dobkin, 2005). Parkinson's disease presents reduced step length and increased variability (Keus et al., 2014).

Traditional metrics miss complex interactions between movement aspects. Rehabilitation requires sensitive measures detecting gradual improvements (Mehrholz et al., 2017).

4 Methodology

4.1 Framework Overview

We model gait as probability distributions over complete cycles with three components: (1) normative distribution learning, (2) patient-specific estimation, and (3) distributional distance computation.

4.2 Data Representation

Gait data is represented as 4-dimensional spatio-temporal graphs, where each node corresponds to an anatomical landmark with 3D spatial coordinates (x, y, z) and temporal information t. For a single gait cycle, we have:

$$G_i = \{(x_i^j(t), y_i^j(t), z_i^j(t), t) : j \in \mathcal{J}, t \in [0, T_i]\} \tag{1}$$

where i indexes the gait cycle, j indexes anatomical landmarks in the set $\mathcal{J}$, and T_i is the duration of cycle i.

4.3 Gait Cycle Segmentation and Normalization

Raw motion capture recordings undergo automatic segmentation into individual gait cycles using heel-strike detection algorithms. Each cycle is temporally normalized to a standard duration to enable comparison across different walking speeds:

$$\tilde{G}_i = \{(x_i^j(\tau), y_i^j(\tau), z_i^j(\tau), \tau) : j \in \mathcal{J}, \tau \in [0, 1]\} \tag{2}$$

where $\tau = t/T_i$ represents normalized time within the gait cycle.

4.4 Individual Gait Distribution Estimation

For each person p, let $\mathcal{G}_p = \{\tilde{G}_{p,1}, \tilde{G}_{p,2}, \ldots, \tilde{G}_{p,n_p}\}$ represent normalized gait cycles. We estimate $P_p(\tilde{G})$ using density modeling techniques selected based on dimensionality, sample size, and distributional characteristics. Approaches include parametric, non-parametric, or hybrid methods optimized for variability capture and computational tractability.

4.5 Normative Distribution Construction

For healthy controls $\mathcal{H} = \{P_1, P_2, \ldots, P_N\}$, we construct the normative distribution $P_{norm}(\tilde{G}|\mathbf{z})$.

4.5.1 Distribution Aggregation

We evaluate:

- Gaussian mixture models: $P_{norm} = \sum_{k=1}^{K} \pi_k \mathcal{N}(\mu_k, \Sigma_k)$
- Kernel density estimation with adaptive bandwidth
- Variational autoencoders for latent space modeling
- Weighted ensemble averaging

4.5.2 Demographic Conditioning

We will also explore conditioning on demographic variables $\mathbf{z} = \{age, sex, height, weight\}$. Methods include similarity-based weighting schemes and parametric conditioning approaches to ensure appropriate normative comparisons across different population characteristics.

4.6 The Distributional Gait Assessment Index (DGAI)

The **Distributional Gait Assessment Index (DGAI)** is the core innovation of this research, providing a unified metric for quantifying gait abnormality through distributional analysis. Unlike the existing Gait Deviation Index which focuses on discrete kinematic variables, the DGAI captures the comprehensive deviation of a patient's gait pattern from normative healthy patterns while accounting for natural demographic variation and movement variability.

4.6.1 DGAI Definition

The DGAI for a patient p is defined as:

$$DGAI_p = \min\{D(P_p, P_{norm}(\cdot|\mathbf{z}_p)), D(P_p, P_{norm})\} \tag{3}$$

where:

- P_p is the patient's estimated gait distribution
- $P_{norm}(\cdot|\mathbf{z}_p)$ is the demographic-conditioned normative distribution
- P_{norm} is the general population normative distribution

- $D(\cdot, \cdot)$ is the optimal distributional distance measure
- $\mathbf{z}_p$ represents patient demographic characteristics

The minimum operation ensures robustness by selecting the most appropriate normative comparison based on demographic similarity and distributional characteristics.

4.6.2 DGAI Interpretation

The DGAI provides an intuitive scale where:

- **DGAI $\approx$ 0**: Gait pattern indistinguishable from healthy controls
- **DGAI $>$ 0**: Increasing deviation from normative patterns
- **Higher values**: Greater gait abnormality requiring intervention

4.6.3 Distance Measure Selection

We systematically evaluate multiple distributional distance measures for optimal DGAI performance:

Information-Theoretic Measures:

- Kullback-Leibler divergence: $D_{KL}(P_p \| P_{norm}) = \int P_p(x) \log \frac{P_p(x)}{P_{norm}(x)} dx$
- Jensen-Shannon divergence for symmetric comparison
- Mutual information for dependency structure

Geometric Distances:

- Wasserstein distance respecting gait manifold structure
- Hellinger distance for probabilistic similarity
- Maximum Mean Discrepancy for kernel-based comparison

Statistical Measures:

- Energy distance for distribution comparison
- Cramér-von Mises test statistics
- Anderson-Darling measures for tail sensitivity

Selection criteria prioritize clinical sensitivity, computational efficiency, noise robustness, practitioner interpretability, and stability across diverse populations.

4.6.4 DGAI Validation Framework

We establish comprehensive validation protocols for the DGAI:

Construct Validity: Correlation with established clinical measures (Berg Balance Scale, Fugl-Meyer Assessment, UPDRS-III)

Discriminant Validity: Ability to distinguish healthy controls from patients across neurological conditions

Sensitivity: Detection of subtle changes during rehabilitation progress

Specificity: Minimal false positives in healthy population variation

Responsiveness: Tracking rehabilitation improvements over time

Reliability: Test-retest consistency and inter-session stability

4.7 Study Design and Participants

4.7.1 Phase 1: Normative Database Construction

- **Healthy controls**: $N = 30$ participants
- **Age range**: 40-70 years, stratified by decade
- **Gender**: Balanced representation
- **Exclusion criteria**: No history of neurological, orthopedic, or cardiovascular conditions affecting gait

4.7.2 Phase 2: Clinical Validation

- **Stroke patients**: $n = 15$, recruited from rehabilitation centers
- **Parkinson's patients**: $n = 15$, recruited from movement disorder clinics
- **Stimuloop participants**: $n = 30$, undergoing specific rehabilitation protocol
- **Assessment timeline**: Baseline, 4 weeks, 8 weeks, 12 weeks post-intervention

4.8 Validation

We assess construct validity (correlation with clinical measures), sensitivity (healthy vs. patient differences), responsiveness (intervention detection), and test-retest reliability. Additional validation includes synthetic data testing (Yamada et al., 2025), generative augmentation robustness (Zhang et al., 2025), and cross-modal consistency with generative model representations.

5 Expected Outcomes

5.1 Primary Outcome: The Validated Distributional Gait Assessment Index

The primary deliverable of this research is a fully validated **Distributional Gait Assessment Index (DGAI)** that provides:

- **Quantitative Assessment**: Objective measurement of gait deviation from normative patterns with established clinical thresholds
- **Clinical Sensitivity**: Demonstrated ability to detect subtle rehabilitation improvements in stroke and Parkinson's patients
- **Demographic Robustness**: Age, sex, and anthropometric conditioning for appropriate normative comparisons
- **Standardized Protocol**: Ready-to-implement assessment methodology for clinical settings

5.2 Supporting Deliverables

1. **Normative Database**: Comprehensive normative gait distributions with demographic conditioning capabilities, serving as the foundation for DGAI computation
2. **Validation Evidence**: Clinical validation demonstrating DGAI sensitivity, specificity, and responsiveness across neurological conditions
3. **Implementation Guidelines**: Standardized protocols for DGAI assessment, interpretation, and clinical decision-making
4. **Software Framework**: Open-source implementation enabling widespread clinical adoption of the DGAI

5.3 Clinical Impact

The DGAI will transform gait assessment by providing:

- Enhanced objective tracking of rehabilitation progress
- Sensitive detection of subtle therapeutic improvements
- Standardized outcome measurement across rehabilitation centers
- Evidence-based optimization of intervention protocols
- Improved patient engagement through quantifiable progress metrics

References

Adeli, V., Ehsani, K., Rose, C., Kapoor, H., and Adeli, E. (2025). Gaitgen: Disentangled motion-pathology impaired gait generative model – bringing motion generation to the clinical domain. *arXiv preprint arXiv:2503.22397*.

Baker, R. (2013). *Measuring Walking: A Handbook of Clinical Gait Analysis*. Mac Keith Press, London.

Dobkin, B. H. (2005). Rehabilitation after stroke. *New England Journal of Medicine*, 352(16):1677–1684.

Gholami, M., Rezaei, S., and Khodayari, A. (2019). A probabilistic approach to gait analysis using hidden markov models. *Medical Engineering & Physics*, 66:54–63.

Ghosh, M., Dhara, B. C., and Mukherjee, J. (2025). Gait data generation using lightweight generative deep learning framework. *Journal of Biomechanics*, 192:112951.

Hausdorff, J. M., Rios, D. A., and 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.

Horst, F., Lapuschkin, S., Samek, W., Müller, K.-R., and Schöllhorn, W. I. (2019). Explaining the unique nature of individual gait patterns with deep learning. *Scientific Reports*, 9(1):1–13.

Keus, S. H., Bloem, B. R., Hendriks, E. J., Bredero-Cohen, A. B., and Munneke, M. (2014). European physiotherapy guideline for parkinson's disease. *KNGF/ParkinsonNet, The Netherlands*.

Mehrholz, J., Thomas, S., Kugler, J., Pohl, M., and Elsner, B. (2017). Treadmill training and body weight support for walking after stroke. *Cochrane Database of Systematic Reviews*, (8).

Peppes, N., Daskalakis, E., Alexakis, T., Adamopoulou, E., and Demestichas, K. (2023). Foggan: Generating realistic parkinson's disease freezing of gait data using gans. *Sensors*, 23(19):8158.

Rezvani, A., Xu, R., Little, M. A., Pontil, M., and Ciccarelli, O. (2025). Diffusegait-net: Improving parkinson's disease gait severity assessment with a diffusion model framework. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 33:2858–2869.

Stihi, A., Kaski, S., and Solin, A. (2025). Hierarchical gaussian processes for characterizing gait variability in multiple sclerosis. *Data-Centric Engineering*, 6:e36.

Takeishi, N. and Kalousis, A. (2021). Physics-integrated variational autoencoders for robust and interpretable generative modeling. In *Advances in Neural Information Processing Systems*, volume 34.

Whittle, M. W. (2014). *Gait Analysis: An Introduction*. Butterworth-Heinemann, Oxford, 5th edition.

Yamada, Y., Dembia, C. L., Ong, C. F., Uhlrich, S. D., Hicks, J. L., and Delp, S. L. (2025). Utility of synthetic musculoskeletal gaits for generalizable healthcare applications. *Nature Communications*, 16:6188.

Zhang, W., Jiao, Y., Liu, X., Wang, S., and Li, Q. (2025). Real-time forecasting of pathological gait via imu navigation: a few-shot and generative learning framework for wearable devices. *Discover Electronics*, 2:51.