Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Foundation model for gait kinetics estimation across neurological and developmental cohorts from flexible kinematic inputs

Wenqi Liang
University of Auckland
Zhi-Qiang Zhang
University of Leeds
Xingye Cheng
University of Auckland
Xi Gao
University of Auckland
Jiping Jin
ShanghaiTech University
Shuyun Jiang
Shanghai University of Traditional Chinese Medicine
Silmara Gusso
University of Auckland
Yanxin Zhang
yanxin.zhang@auckland.ac.nz
University of Auckland

Article

Keywords:

Posted Date: September 30th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11197962/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

In neurological disorders, gait kinetics can inform clinical assessment and treatment planning. Conventional gait-kinetic measurement remains confined to specialised gait laboratories, and the alternative modelling approaches are limited to a specific population cohort, measurement protocol or input configuration, limiting their clinical utility. PathoGaitFM is a foundation model estimating vertical ground reaction force and joint moments across neurological and developmental cohorts from flexible kinematic inputs. Pretrained on healthy participants, PathoGaitFM reached a mean Pearson correlation of 0.811 in cross-validation across four cohorts of participants with cerebral palsy, typical development, stroke or Parkinson's disease, showing that one shared model tracked kinetic waveforms across heterogeneous cohorts. In five unseen external datasets it reached 0.898 without retraining, showing consistently high accuracy across laboratories and populations. Incomplete inputs lowered correlation by at most 0.041, supporting reuse across joint-angle sets. These findings motivate prospective validation with flexible kinematic inputs in clinical populations.

Introduction

Gait impairments are common across neurological disorders, including cerebral palsy (CP), stroke, and Parkinson's disease (PD). Quantitative gait analysis provides objective information that can support treatment planning[1], guide surgery[2], target rehabilitation and monitor recovery or treatment response[3,4]. Among the biomechanical measures obtained from gait analysis, joint moments are an important component because they provide information about the mechanical function and loading of individual joints. Their clinical relevance has been demonstrated across neurological populations. In CP, joint moments inform treatment planning, and instrumented gait analysis can influence surgical and non-surgical interventions[2,5]. After stroke, altered lower-limb kinetics and impaired paretic-limb propulsion are associated with walking performance and motor recovery[3]. In PD, joint kinetics can reveal joint-specific responses to dopaminergic treatment[6], such as levodopa, which may improve hip kinetics while producing relatively little change at the ankle[4]. Despite their clinical value, joint moments are difficult to obtain routinely because they are conventionally calculated using inverse dynamics, which requires measured ground reaction forces (GRFs) and kinematic data[7]. These measurements typically rely on force plates, marker-based motion-capture systems and trained personnel in specialised laboratories, making data collection and processing time-consuming. This substantially limits the use of joint-moment analysis in routine clinical care, home-based rehabilitation and remote monitoring.

Recent advances in motion-sensing technologies have substantially reduced the practical burden of collecting gait kinematics. Video-based markerless systems and inertial measurement units (IMUs) require less specialised infrastructure than conventional marker-based motion capture and allow data to be collected across a wider range of clinical, community and home environments[8–10]. However, greater access to kinematic data does not by itself provide the external forces required to calculate gait kinetics using conventional inverse dynamics. As a result, GRFs and joint moments remain difficult to obtain in these settings.

Deep-learning and model-based approaches have sought to bridge this gap by estimating GRFs and joint moments directly from kinematic inputs, without requiring force-plate measurements as model input[11–16]. This creates an opportunity to extend kinetic gait assessment to settings where kinematic data can be collected but direct force measurements are unavailable. For such an approach to be broadly useful, however, the benefit of removing force plates should not be offset by restrictive requirements on how the kinematic input is obtained. Ideally, a model should operate on kinematics acquired from different measurement systems, including marker-based motion capture, video-based methods and IMUs, and should tolerate differences in the available joint-angle inputs across datasets and clinical protocols[17,18]. However, most existing deep-learning models are trained and evaluated within a single cohort and protocol, using fixed kinematic inputs and kinetic targets[11–13,19]. Changes in acquisition systems or available joint angles may reduce accuracy or require retraining, limiting generalisation to new patient cohorts and measurement settings.

Pretraining offers one potential approach to improving transfer across tasks and datasets by learning reusable representations that can be adapted with less task-specific retraining and fewer labelled data[20–23]. Building on this idea, recent work has explored more general gait representations that can support different input configurations and biomechanical tasks within a shared modelling framework. A notable example is GaitDynamics[24], which uses a generative model to jointly represent kinematics and GRFs, thereby supporting incomplete inputs and multiple biomechanical tasks. This shared representation has been applied to force estimation, prediction of knee loading following gait modifications and gait generation across different running speeds. Related work has also explored representation learning for pathological gait[25,26], estimation of clinical measures from wearable sensors[27] and generation of skeletal motion conditioned on pathology labels[28], although these approaches do not directly model GRF or joint-moment waveforms. However, a generalisable framework that can estimate GRFs and joint moments across different neurological cohorts and datasets, while accommodating complete or incomplete joint-angle inputs from different sensing modalities, remains lacking. This limits the translation of kinetic estimation methods to routine clinical use, where patient populations, measurement systems and available gait variables can vary considerably.

To improve generalisability across these settings, we developed PathoGaitFM, a foundation model for gait-kinetics estimation across neurological and developmental cohorts. It accepts flexible joint-angle inputs, whether complete or incomplete and whether acquired using marker-based motion capture, video-based methods or IMUs, and reconstructs unobserved kinetic signals from the available joint angles by masked inpainting[29]. After pretraining on healthy gait, the model was fine-tuned in four development cohorts totalling 90 participants with CP, stroke, PD or typical development. We evaluated PathoGaitFM at two levels. First, five-fold cross-validation across these four cohorts assessed vertical GRF (vGRF) and available sagittal joint moments. Within the cross-validation, we also tested whether the same model could estimate kinetics from fewer joint angles. Second, PathoGaitFM was evaluated without retraining in five external datasets of 94 participants, none of which had been used for training. In two paired-modality datasets, we compared vGRF estimates from video- and IMU-based kinematics with those from marker-based kinematics recorded during the same trials.

Results

Overview of PathoGaitFM

PathoGaitFM estimates vGRF and sagittal hip, knee and ankle moments from flexible joint-angle inputs (Fig. 1). The available joint angles guide reconstruction of the masked kinetic channels. PathoGaitFM was pretrained on healthy gait[30,31] and fine-tuned in four development cohorts of children with CP, typically developing (TD) children, adults after stroke and adults with PD[31,32]. Within these cohorts we assessed estimation with all available and with incomplete joint-angle inputs, and then transfer to five external datasets without retraining[8,33–36]. We also compared vGRF estimates from marker-based and video- or IMU-based kinematics recorded during the same trials[8,37]. Table 1 summarises the datasets and their roles.

PathoGaitFM estimates gait kinetics across four neurological and developmental cohorts

Within the four development cohorts, PathoGaitFM estimated bilateral vGRF and, where reference data were available, sagittal hip, knee and ankle moments using joint angles. Example ankle-moment waveforms illustrate differences between PathoGaitFM and the same architecture without pretraining in TD, CP and stroke (Fig. 2a). The mean Pearson correlation coefficient (PCC) across the available kinetic targets and the four cohorts was 0.811 (95% confidence interval 0.787 to 0.834). The estimates captured the main features of the reference kinetic waveforms across these cohorts.

Among the three published models[11,13,38,39] retrained under the same protocol, the highest mean PCC was 0.799 (Fig. 2b). The paired difference between masked inpainting and this baseline was 0.012 (95% confidence interval − 0.001 to 0.026). Mean PCC was 0.892 in TD children, 0.834 in PD, 0.801 in CP and 0.717 after stroke. Range-normalised root-mean-square error (nRMSE), averaged across the available target variables, was 11.4% in TD children, 19.6% in CP, 22.4% in PD and 34.3% after stroke (Fig. 2c). Joint-moment root-mean-square error (RMSE) was 0.10–0.12, 0.14–0.18 and 0.22–0.27 N·m/kg in TD, CP and stroke, respectively. vGRF RMSE ranged from 9.86 to 23.44% of body weight (BW) across the four cohorts (Fig. 2d,e).

Pretraining increased the equally weighted four-cohort mean PCC from 0.745 to 0.811, a difference of 0.066 (95% confidence interval 0.056 to 0.076), compared with the same architecture trained without pretraining. The largest PCC gains were observed for joint moments, reaching 0.203 for the knee moment after stroke and 0.147 for the knee moment in CP, compared with 0.015 to 0.045 for vGRF. Knee-moment RMSE was lower with pretraining in 19 of 22 children with CP and 28 of 36 adults after stroke. These gains support the value of healthy-gait pretraining for pathological joint-moment estimation, particularly at the knee in CP and stroke.

We compared three methods for estimating kinetic waveforms: a direct regressor from joint angles, masked inpainting and a supervised head that uses the model's features. The direct regressor achieved a PCC of 0.753, compared with 0.811 for masked inpainting and 0.821 for the supervised head (Fig. 2b). Masked inpainting had a lower nRMSE than the supervised head, 21.9% compared with 22.7% (Fig. 2c). Mean PCC for the supervised head exceeded that of the best published comparison model by 0.022 (95% confidence interval 0.009 to 0.035, Fig. 2b). Subsequent evaluations used masked inpainting, which accepts complete or incomplete joint-angle inputs, whereas the supervised head was trained for one fixed input configuration.

PathoGaitFM transfers to external datasets without retraining

Without retraining, PathoGaitFM estimated vGRF in 84 participants from four external datasets[33–36]. Mean participant-level vGRF PCC (± SD) was 0.847 ± 0.038 in 7 children with CP in the Meyns (2017) dataset and 0.906 ± 0.046 in 10 children with CP in the Steele (2010) dataset (Fig. 3a). nRMSE was 23.2% in the Meyns CP group and 17.9% in the Steele dataset (Fig. 3b). For comparison, vGRF PCC in the development cohort with CP was 0.914. For the five TD children in the same Meyns dataset, the PCC was 0.927.

Mean participant-level vGRF PCC (± SD) was 0.979 ± 0.007 in 42 healthy adults in the Fukuchi (2018) dataset and 0.952 ± 0.031 in 20 adults with PD in the Viglialoro (2025) dataset, with mean participant-level nRMSE values of 8.3% and 16.2%, respectively. Participant-level vGRF PCC ranged from 0.967 to 0.989 in the Fukuchi dataset and from 0.865 to 0.993 in the Viglialoro dataset.

Sagittal joint moments were evaluated in 52 healthy adults: 42 in the Fukuchi dataset and 10 in the Uhlrich (2023) dataset[8]. Evaluation covered the full gait cycle in the Fukuchi dataset and left stance in the Uhlrich dataset. In the Fukuchi dataset, hip-, knee- and ankle-moment PCCs were 0.908, 0.891 and 0.978, and the corresponding values in the Uhlrich dataset were 0.886, 0.688 and 0.947 (Fig. 3c). Ankle-moment PCC was highest and knee-moment PCC was lowest in both datasets. nRMSE ranged from 14.9% to 15.6% in the Fukuchi dataset and from 19.4% to 25.4% in the Uhlrich dataset (Fig. 3d). Mean joint-moment RMSE ranged from 0.13 to 0.23 N·m/kg in the Fukuchi dataset and from 0.22 to 0.29 N·m/kg in the Uhlrich dataset (Fig. 3e). Averaged with equal weight over the ten combinations of dataset and kinetic target, the mean PCC across the five external datasets was 0.898.

Gait kinetics estimation with incomplete and portable kinematic inputs

Example ankle-moment waveforms show the effects of all available inputs, withholding hip angles and using hip angles alone in TD, CP and stroke (Fig. 4a). Within the same five-fold cross-validation, we changed only which joint angles were provided and ran the same model under seven input configurations without retraining (Fig. 4b). Starting from all joint angles available in each cycle, we withheld the pelvis, hip, knee or ankle angles one group at a time, always on every available side. Two further configurations withheld the ankle and pelvis angles together, or all angles except the hip. Withholding the hip angles lowered the overall mean PCC by 0.041, from 0.811 to 0.770. Withholding the knee angles lowered it by 0.022 and withholding the pelvis or the ankle angles by no more than 0.004. Among the single-group omissions, withholding hip angles also caused the largest increase in nRMSE. It was 21.9% with all available inputs and with the pelvis withheld, 23.3% with the hip withheld, 22.8% with the knee and 22.3% with the ankle. Withholding the ankle and pelvis angles together lowered the mean PCC by 0.009. Withholding the ankle, pelvis and knee angles, leaving only hip angles, lowered mean PCC by 0.036, a smaller drop than withholding the hip angles, and raised nRMSE to 23.4%. Overall waveform correspondence was largely preserved across the tested configurations, although effects varied by cohort. Withholding hip angles reduced PCC after stroke from 0.717 to 0.609; using hip angles alone reduced PCC in CP from 0.801 to 0.723.

We further examined input flexibility by estimating vGRF from video- and IMU-based kinematics. The paired-modality evaluation used 20 healthy adults, 10 with video-based kinematics in the Uhlrich dataset[8] and 10 with IMU-based kinematics in the Grouvel (2023) dataset[37] (Fig. 4c). For each participant, the marker-based and alternative-modality estimates used recordings from the same walking trials. In the Uhlrich dataset the vGRF PCC was 0.901 for video-based kinematics and 0.901 for marker-based kinematics. In the Grouvel dataset it was 0.889 for IMU-based kinematics and 0.862 for marker-based kinematics. In the Uhlrich dataset nRMSE was 16.9% for video-based and 16.8% for marker-based kinematics, and in the Grouvel dataset 19.1% for IMU-based and 21.1% for marker-based kinematics.

Discussion

To our knowledge, PathoGaitFM is the first generative foundation model for pathological gait-kinetics estimation across multiple neurological cohorts. It further transfers across laboratories, sensing modalities and kinematic input configurations without dataset-specific retraining. PathoGaitFM achieved an overall mean PCC of 0.811 under cross-validation across the four development cohorts of 90 participants and 0.898 in five external datasets never used for training. Joint-moment RMSE was 0.10–0.12 N·m/kg in TD, 0.14–0.18 N·m/kg in CP and 0.22–0.27 N·m/kg after stroke. In direct comparison with prior work, three published kinetic-estimation models were retrained on our development cohorts under the same protocol, and the best-performing among them achieved an overall mean PCC of 0.799. PathoGaitFM (PCC = 0.811) marginally exceeded the strongest published model, but without requiring dataset-specific retraining or task-specific architecture design. This suggests that the representational capacity gained through large-scale pretraining and masked inpainting confers a level of transfer that task-specific architectures trained without pretraining do not readily achieve.

The external datasets used for validation differed from the development cohorts not only in acquisition systems (Vicon, Motion Analysis and Qualisys systems) but also in population (children with CP and TD, adults with PD and healthy adults) and anthropometric characteristics. Joint-angle estimates can differ substantially between computational models[17] and acquisition systems. The development and external datasets span both direct kinematic models and inverse kinematic models (Table 1), so the evaluation crosses that difference rather than holding it fixed. Despite these sources of variation, the model maintained a mean PCC of 0.898 across all five external datasets. These results support transfer of the learned kinematics-to-kinetics mapping across independent datasets with different acquisition and processing procedures. This level of performance suggests that the features PathoGaitFM extracts do not overfit to cohort-specific or laboratory-specific idiosyncrasies; rather, they appear to capture the underlying physical structure of walking that is common across participants and measurement configurations. The model, through large-scale pretraining

and exposure to diverse input masks (masked inpainting), appears to have approximated this constraint directly from data.

Within the cross-validation, withholding joint-angle groups lowered the overall mean PCC by at most 0.041, and this small reduction supports reuse of the same model across assessment protocols that provide different joint-angle sets. Taken together, the cross-cohort, cross-laboratory and incomplete-input results provide converging evidence that PathoGaitFM has learned a shared representation of gait mechanics. That representation appears to be robust to sensor-specific differences, to span population-level variability and to preserve the mapping from kinematics to kinetics with different sets of joint-angle inputs. This capability may distinguish the present work from prior models that, although accurate within their training domain, may lose accuracy when applied to new acquisition protocols or cohorts, often requiring costly and time-consuming retraining. The model also estimated vGRF from video-based and IMU-based kinematics with accuracy comparable to that obtained from marker-based motion capture. By demonstrating that one model can transfer across these axes of variation, our findings provide a promising basis for deploying deep-learning-based kinetic estimation outside the confines of the specialised gait laboratory.

Test–retest variability in laboratory gait analysis provides a reference for evaluating emerging gait-assessment technologies[40]. A widely used quantification of this variability is the minimal detectable change (MDC)[41]. Published MDC values in peak joint moments are broadly on the scale of 0.1–0.2 N·m/kg in healthy adults and adults with chronic low back pain[42,43]. The joint-moment waveform RMSEs of PathoGaitFM in TD and CP (0.10–0.18 N·m/kg) are of a similar magnitude. MDC values are also larger for CP joint kinetics and post-stroke GRFs than in their respective control groups[44,45], highlighting the value of interpreting estimation errors in relation to the population and biomechanical measure. These comparisons provide a clinical reference for interpreting the observed errors and support further evaluation of PathoGaitFM's potential clinical utility.

Given this level of prediction fidelity, the most direct clinical implication is the model's potential to estimate joint moments and vGRF from joint angles derived from video or wearable IMU recordings. In routine clinical practice, instrumented gait analysis remains confined to specialised gait laboratories due to the cost of force plates, the expertise required for marker placement, and the time burden of processing. PathoGaitFM could reduce this dependency by using joint angles derived from video and wearable IMU recordings[10,46,47], making quantitative kinetic estimates available in outpatient rehabilitation clinics, community health centres, and potentially even home-based tele-rehabilitation settings. This does not imply that the model replaces force plates for research-grade measurement, but rather that it offers a promising alternative where force plates are logistically inaccessible. In addition, the model's ability to operate with incomplete kinematic inputs is a valuable feature that allows the model to be deployed in routine clinical settings, where video occlusion, sensor dropouts, and heterogeneous acquisition protocols are the norm.

This study has two main limitations. First, external validation in pathological gait covered two CP datasets and one PD dataset, and the pathological cohorts were relatively small. As relevant datasets become available, the same fixed model can undergo further external validation across a broader range of clinical cohorts. Larger cohorts, including more severe presentations, would also allow finer stratification by severity and more precise cohort-specific accuracy estimates. Second, comparable MDC values for joint moments are not yet well established for stroke and PD, and the model's ability to detect within-person change over time requires further evaluation. Repeated assessments across sessions, comparing both reference and model-estimated joint moments, would help establish this ability and inform the model's use in longitudinal clinical follow-up.

In summary, one model, without dataset-specific retraining, estimated gait kinetics across four neurological and developmental cohorts, across independent laboratories, and across sensing modalities and kinematic input configurations. These findings support prospective evaluation of out-of-lab gait-kinetics estimation from portable kinematic measurements.

Methods

Study design and data preparation

PathoGaitFM was pretrained by self-supervised denoising using the training partition of a healthy-adult dataset comprising 259 participants and 290,408 gait cycles from eight public studies[30,31] and fine-tuned using four development cohorts comprising 90 participants with CP, stroke, PD or typical development. Evaluation comprised five-fold cross-validation in the development cohorts (90 participants, 4,653 gait cycles, three laboratories) and evaluation without retraining in five external datasets[8,33–36] of 94 participants (612 gait cycles, five laboratories) not used for training. We also assessed incomplete joint-angle configurations under the same cross-validation and compared paired input modalities in two datasets[8,37] of 20 participants. Uhlrich contributed to both external joint-moment evaluation and paired marker/video vGRF comparisons, whereas Grouvel contributed to paired marker/IMU vGRF comparisons. Table 1 summarises the pretraining sources, development cohorts and external datasets, with participant characteristics, evaluated targets and analysis units. Available joint angles and kinetic targets differed across datasets. We therefore aligned the datasets to one shared representation with a fixed set of kinematic and kinetic variables, using masks to identify unavailable variables in each cycle.

Data from children with CP and TD children had previously been collected at the Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine[14]. The present analysis used de-identified data and was approved by the Auckland Health Research Ethics Committee, University of Auckland (reference AH26772). The study followed the Declaration of Helsinki. Written informed consent was obtained from each child's parent or legal guardian before data collection, and age-appropriate assent was obtained from the children. All other datasets are public.

Table 1 | Data sources, participant characteristics and analysis units. A, the included healthy-adult pretraining dataset across all partitions; B, development cohorts; C, external and paired-modality evaluation. Participant counts in B and C refer to evaluated participants. Ages are mean ± SD (range), in years; M, male; F, female. Gait cycles (n), number of gait cycles after preprocessing for the listed participants.

A Healthy-adult pretraining dataset (all partitions)

Study	Internal source ID	Reference pipeline	Participants (n)	Gait cycles (n)
Camargo et al. (2021)	addbio_Camargo2021	OpenSim Rajagopal, inverse kinematics[48,49]	21	37,642
Carter et al. (2024; data release)	addbio_Carter2023	OpenSim Rajagopal, inverse kinematics[48,49]	49	131,311
Moore et al. (2015)	addbio_Moore2015	OpenSim Rajagopal, inverse kinematics[48,49]	11	32,037
Tan et al. (2022; running)	addbio_Tan2021	OpenSim Rajagopal, inverse kinematics[48,49]	7	22,392
Tan et al. (2023; walking)	addbio_Tan2022	OpenSim Rajagopal, inverse kinematics[48,49]	16	27,565
Wang et al. (2023)	addbio_Wang2023	OpenSim Rajagopal, inverse kinematics[48,49]	8	8,128
van der Zee et al. (2022)	addbio_vanderZee2022	OpenSim Rajagopal, inverse kinematics[48,49]	9	29,682
Van Criekinge et al. (2023)	vdk_healthy	Plug-in Gait, direct kinematics[50]	138	1,651

B Development cohorts

Dataset	Participants and demographics	Reference pipeline	Evaluated targets	Gait cycles (n)
Yueyang CP	22 children with CP; 7.6 ± 2.1 (5–12); 15 M / 7 F	Adjusted Helen Hayes and Visual3D, direct kinematics[51]	vGRF and sagittal moments	370
Van Criekinge (2023)[31]	36 adults after stroke; 62.4 ± 15.1 (19–84); 26 M / 10 F	Plug-in Gait, direct kinematics[50]	vGRF and sagittal moments	559
Shida (2023; BMClab)[32]	25 adults with PD; 46 medication-state recordings; 65.0 ± 9.0 (46–81); 20 M / 5 F	OpenSim Rajagopal, inverse kinematics[48,49]	vGRF; evaluated per medication-state recording	3,589
Yueyang TD	7 TD children; 8.9 ± 1.6 (7–11); 4 M / 3 F	Adjusted Helen Hayes and Visual3D, direct kinematics[51]	vGRF and sagittal moments	135

C External and paired-modality datasets

Dataset	Participants and demographics	Reference pipeline	Evaluated targets	Gait cycles (n)	Analysis
Meyns (2017; Leuven)[33]	7 children with CP; 5 TD children. CP: 9.8 ± 2.4 (6.3–12.5); 6 M / 1 F	OpenSim Rajagopal, inverse kinematics[48,49]	vGRF	36	External vGRF; CP and TD groups shown separately

Steele (2010; Gillette)[34]	10 children with spastic diplegic CP and mild crouch gait; 8.1 ± 1.7	OpenSim gait2392, inverse kinematics[52]	vGRF	21	External vGRF
Fukuchi (2018; WBDS)[35]	42 healthy adults; 42.6 ± 18.6 (21–84); 24 M / 18 F	Leardini marker set and Visual3D, direct kinematics[53]	vGRF and sagittal moments	126	External vGRF and moments
Viglialoro (2025; PD-Nordic)[36]	20 adults with PD	PyCGM2.5, direct kinematics[54]	vGRF	339	External vGRF
Uhlrich (2023; OpenCap LabValidation)[8]	10 healthy adults; 27.7 ± 3.8; 4 M / 6 F	OpenSim Lai, inverse kinematics[55]	vGRF and sagittal moments	90	Paired marker/video vGRF; external moments during left stance
Grouvel (2023)[37]	10 healthy adults; 29.7 ± 6.8 (22–42); 6 M / 4 F	OpenSim Rajagopal, inverse kinematics[48,49]; IMU angles from OpenSense	vGRF	173	Paired marker/IMU vGRF

Signal preprocessing. Joint angles came from direct kinematic models in some datasets and from inverse kinematics in others, and Table 1 gives the reference pipeline for each dataset. For Shida, Meyns and marker-based Grouvel recordings, we calculated joint angles using OpenSim inverse kinematics without additional filtering of marker trajectories. Grouvel IMU-based joint angles were calculated using OpenSense[9]. For the other datasets, we used released joint angles with their source-specific preprocessing, including Uhlrich video-based kinematics from OpenCap[8]. Yueyang joint angles had been low-pass filtered at 6 Hz during the original Visual3D processing. We low-pass filtered Grouvel force-plate signals at 15 Hz with a fourth-order Butterworth filter before decimation from 1,000 to 100 Hz.

Kinetic references. Reference vGRF was measured with force plates. Reference joint moments were obtained from the source datasets and had been calculated by inverse dynamics using kinematics, force-plate data and body-segment parameters in each source's processing pipeline. We aligned sign conventions and units across datasets before use. Segmented gait cycles or released mean waveforms were normalised to 101 points. Kinematic and kinetic signals were arranged in the shared representation in physical units. The first 100 points were used as model input. External evaluation included only samples with valid and technically reliable reference kinetics for the evaluated targets.

PathoGaitFM architecture and training

Model inputs and masking. PathoGaitFM estimates bilateral vGRF and sagittal hip, knee and ankle moments from pelvis segment angles and bilateral hip, knee and ankle joint angles (Fig. 1). The kinematic input consists of joint-angle waveforms only, so marker-based, video-based and IMU-based kinematics share the same input representation. The available joint-angle channels are visible to the model; the kinetic targets are hidden channels reconstructed by denoising. This approach extends diffusion inpainting[29] and conditional diffusion models for multivariate time-series imputation and forecasting[56,57] to inputs with an entire variable or one limb missing. The input mask identifies the variables supplied to the model. The validity mask excludes unavailable values from the training loss. Missing or deliberately withheld variables are represented by a learned token. Changing the input mask allows the model to use all available joint-angle inputs or selected subsets without retraining.

Model architecture. The network is a one-dimensional diffusion Transformer[58] with 33.5 million parameters (Fig. 5). Each waveform comprises 40 channels and 100 time points and is divided into patches of four time points, giving 25 tokens of 384 features each. A learned positional embedding is added to the tokens, which then pass through twelve Transformer blocks, each with six attention heads and a token-wise feed-forward network of 1,536 hidden units. The diffusion process[59] uses 1,000 steps, a cosine noise schedule[60] and v-prediction parameterisation[61].

Model conditioning. Dataset source, cohort and severity embeddings are added to the diffusion-step embedding, and the resulting condition is applied through adaptive layer normalisation (AdaLN). Conditioning uses the AdaLN-Zero variant, which produces normalisation shifts, scales and residual gates from the condition. One source identifier is reserved for datasets not seen during training. Severity conditions are based on the clinical scales reported by each dataset[62,63]. During training, conditions are dropped with probability 0.1 to enable classifier-free guidance[64].

Model training. Both training stages used AdamW to minimise the mean squared error of the v-prediction over valid entries. Stage 1 pretrained the model by self-supervised denoising on healthy-adult gait only, with uniform sampling across the eight sources. No task-specific input mask was applied. All variables in the shared representation were modelled, with unavailable values excluded from the loss by the validity mask. Stage 2 fine-tuned the pretrained weights on the four development cohorts to estimate vGRF and sagittal joint moments. The eight auxiliary waveform channels and non-sagittal hip moments were represented by a learned token throughout this stage. Only the evaluated kinetic targets contributed to the loss. For all other variables, the input mask followed each cycle's validity mask. Cycles with a vGRF label were sampled uniformly. Stage 1 ran for 25 epochs at a learning rate of 1×10^{-4}, with 5,000 warm-up steps followed by cosine decay. Validation loss was monitored, and Stage 2 used the final Stage 1 checkpoint. Stage 2 ran for 200 epochs at a constant learning rate of 1×10^{-5} after 1,000 warm-up steps. Both stages used a batch size of 64, weight decay of 1×10^{-4}, gradient clipping at 1.0 and mixed precision.

Model inference. Masked inpainting was the main estimation method (Fig. 5b). The waveform was initialised as Gaussian noise and denoised over 50 steps using a denoising diffusion implicit model sampler[65] with a classifier-free guidance scale of 1.0. At each step, the hidden channels were updated and the visible channels were re-noised from the observed kinematics to the same noise level. The two sets of channels were then recombined using the input mask. Estimates were saved at steps 5, 10, 15, 25, 35 and 45 and at the end, giving seven estimates from a single denoising trajectory. At each channel and time point, the highest and lowest values were discarded and the remaining five were averaged. This procedure was repeated with three random seeds, and the three waveforms were averaged. Only the hidden channels were evaluated.

Evaluation design

Waveform agreement and error were assessed by five-fold cross-validation in the development cohorts and without retraining in external datasets not used for training. We also evaluated incomplete kinematic inputs, video- and IMU-based inputs, and a supervised head on PathoGaitFM features as an alternative to masked inpainting.

Internal cross-validation, incomplete-input evaluation and comparisons between estimation methods used the model from each training fold to evaluate that fold's held-out evaluation units. External-dataset and paired-modality evaluation used one model trained on all four development cohorts, without further retraining.

Internal validation. Accuracy was reported as the unweighted mean PCC across the kinetic targets available in each cohort, using the prespecified evaluation units in Table 1. Evaluation used a participant-level design in TD, CP and stroke and a recording-level design in PD. Because levodopa materially changes gait, medication ON and OFF recordings from the same participant were treated as two gait conditions and as separate evaluation units, and PD folds were split at the

recording level. The other three cohorts were split by participants. Medication-state recordings from the same participant were not grouped when assigning folds.

External validation. Each external dataset was assigned the reserved source identifier and a single cohort and severity condition determined by its release population. vGRF was evaluated in four external datasets and sagittal joint moments in Fukuchi and Uhlrich. The pelvis rotation angle in the Meyns dataset includes walking direction in the laboratory frame.

Incomplete-input evaluation. We changed only the kinematic variables supplied to the model; weights and all other evaluation settings remained unchanged. Seven configurations were evaluated. The reference configuration used all joint angles available in each cycle. Six reduced configurations withheld pelvis, hip, knee or ankle angles separately, withheld ankle and pelvis angles together, or retained hip angles alone. Additional masking removed only channels available in the source cycle. Each withheld variable was represented by the learned token.

Paired-modality evaluation. Comparisons were made within each paired-modality dataset. For each participant, two estimates were obtained from the same walking recording using the same force-plate reference: one from marker-based kinematics and the other from video- or IMU-based kinematics. We compared participant-level PCC and nRMSE for vGRF between the paired input modalities.

Model comparisons. To test whether PathoGaitFM features improved kinetic estimation, we compared two three-layer convolutional networks that differed only in their inputs. Each network estimated the kinetic targets and was trained within each fold using only the training evaluation units. The supervised head used features from the penultimate layer of PathoGaitFM; the direct regressor used joint angles without passing them through PathoGaitFM. Both were compared with masked inpainting. The supervised head was deterministic, used one fixed input configuration, and was evaluated within cohorts using all available inputs.

To measure the effect of pretraining, we trained a second model from random initialisation. It used the same folds, masks, optimiser, fine-tuning schedule, batch size and three-seed sampling.

We retrained GroundLinkNet[38], the Sugai long short-term memory (LSTM) network[39] and the Ozates convolutional neural network[11,13] under a common protocol, using the same folds, valid cycles, inputs, targets and validity masks. Each architecture used one hyperparameter set, and predictions were averaged across three independently trained networks. Five joint-angle channels from one limb were unavailable in every post-stroke and PD cycle, so a validity indicator for each kinematic input was appended to the model inputs. Published-model comparisons used the same internal cross-validation protocol.

Example waveforms. Figs. 2a and 4a show example ankle-moment waveforms from the same participant and limb in each cohort, using all eligible cycles for every comparison. Means and band limits in Figs. 2a and 4a used seven-point Gaussian filtering (σ = 1 sample) followed by shape-preserving cubic interpolation. Waveform metrics were calculated before display processing.

Statistical analysis and reproducibility

Evaluation metrics. Waveform agreement was quantified by the PCC, and error by RMSE and nRMSE. For each limb-cycle, nRMSE was calculated by dividing RMSE by the peak-to-peak range of its reference waveform and expressing the result as a percentage. RMSE was reported in physical units, in N·m/kg for joint moments and as a percentage of body weight (%BW) for vGRF. Errors in physical units were not averaged across variables. PCC, nRMSE and RMSE were calculated for each valid limb-cycle. Limb-cycles with non-finite PCC or error values were excluded. Limb-cycles whose reference range was 1×10^{-9} or less, in N·m/kg or BW, were excluded from nRMSE averaging.

Within each evaluation unit and variable, cycle-level metric values were averaged over valid cycles from both limbs. Physical-unit RMSE was then averaged equally across evaluation units for each variable separately.

Internal PCC and nRMSE composites were calculated by averaging available variable means equally within each evaluation unit, then averaging units within each cohort, and finally averaging the four cohorts with equal weight.

For each external dataset, cycle-level metric values from both limbs were averaged to obtain one value per participant for each kinetic variable. The overall external value was the unweighted mean across datasets and targets, using the CP children only in Meyns.

Statistical analysis. For internal performance estimates and model comparisons, 95% confidence intervals were calculated from 5,000 percentile bootstrap resamples of participants within each cohort. All cycles and medication-state recordings from the same participant were kept together. Paired differences used the same resampled participants for both methods.

Declarations

Code availability

The code and trained model weights are available at https://github.com/wlia728-sketch/PathoGaitFM.

Competing interests

All authors declare no financial or non-financial competing interests.

Author Contribution

Conceptualisation: W.L. and Y.Z.; Methodology: W.L., Z.-Q.Z. and J.J.; Software: W.L.; Formal analysis: W.L.; Investigation: W.L. and S.J.; Data curation: W.L. and S.J.; Resources: S.J. and Y.Z.; Visualisation: W.L.; Supervision: Z.-Q.Z., S.G. and Y.Z.; Project administration: Y.Z.; Writing (original draft): W.L.; Writing (review and editing): Z.-Q.Z., X.C., X.G., J.J., S.G., S.J. and Y.Z. All authors read and approved the final manuscript.

Acknowledgement

We thank Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, for the Yueyang CP and Yueyang TD paediatric gait cohorts, and the maintainers of the public AddBiomechanics 1.0 Dataset and the Van Criekinge, Shida, Meyns, Steele, Fukuchi, Viglialoro, Uhlrich and Grouvel datasets. This work was supported in part by the China Scholarship Council (Grant No. 202406290062). The funder had no role in the study design, analysis or writing.

Data Availability

The Yueyang CP and TD data are not publicly available due to the ethics restrictions on this secondary analysis (AH26772) but are available from the corresponding author on reasonable request. All other datasets are publicly available from their original releases, namely the AddBiomechanics 1.0 Dataset (https://addbiomechanics.org/download_data.html), the Van Criekinge dataset (doi:10.1038/s41597-023-02767-y), the Shida dataset (figshare doi:10.6084/m9.figshare.14896881), the Meyns dataset (https://simtk.org/projects/cp-child-

gait), the Steele dataset (https://simtk.org/projects/crouchgait), the Fukuchi dataset (doi:10.7717/peerj.4640), the Viglialoro dataset (figshare doi:10.6084/m9.figshare.29371769), the Uhlrich dataset (doi:10.1371/journal.pcbi.1011462) and the Grouvel dataset (doi:10.1038/s41597-023-02077-3).

References

1. Hulleck, A. A., Menoth Mohan, D., Abdallah, N., El Rich, M. & Khalaf, K. Present and future of gait assessment in clinical practice: towards the application of novel trends and technologies. Front. Med. Technol. 4, 901331 (2022).
2. Kadhim, M. & Miller, F. Crouch gait changes after planovalgus foot deformity correction in ambulatory children with cerebral palsy. Gait Posture 39, 793–798 (2014).
3. Roelker, S. A., Bowden, M. G., Kautz, S. A. & Neptune, R. R. Paretic propulsion as a measure of walking performance and functional motor recovery post-stroke: a review. Gait Posture 68, 6–14 (2019).
4. Baudendistel, S. T., Schmitt, A. C., Roemmich, R. T., Harrison, I. L. & Hass, C. J. Levodopa facilitates improvements in gait kinetics at the hip, not the ankle, in individuals with Parkinson's disease. J. Biomech. 121, 110366 (2021).
5. States, R. A. et al. Three-dimensional instrumented gait analysis for children with cerebral palsy: an evidence-based clinical practice guideline. Pediatr. Phys. Ther. 36, 182–206 (2024).
6. Mirelman, A. et al. Gait impairments in Parkinson's disease. Lancet Neurol. 18, 697–708 (2019).
7. Werling, K. et al. AddBiomechanics: automating model scaling, inverse kinematics, and inverse dynamics from human motion data through sequential optimization. PLoS ONE 18, e0295152 (2023).
8. Uhlrich, S. D. et al. OpenCap: human movement dynamics from smartphone videos. PLoS Comput. Biol. 19, e1011462 (2023).
9. Al Borno, M. et al. OpenSense: an open-source toolbox for inertial-measurement-unit-based measurement of lower extremity kinematics over long durations. J. Neuroeng. Rehabil. 19, 22 (2022).
10. Phan, V. et al. Biomechanics-informed inertial tracking achieves the accuracy of marker-based kinematics. Nat. Commun. 17, 7927 (2026).
11. Ozates, M. E., Karabulut, D., Salami, F., Wolf, S. I. & Arslan, Y. Z. Machine learning-based prediction of joint moments based on kinematics in patients with cerebral palsy. J. Biomech. 155, 111668 (2023).
12. Mundt, M. et al. Estimation of gait mechanics based on simulated and measured IMU data using an artificial neural network. Front. Bioeng. Biotechnol. 8, 41 (2020).
13. Ozates, M. E., Salami, F., Wolf, S. I. & Arslan, Y. Z. Estimating ground reaction forces from gait kinematics in cerebral palsy: a convolutional neural network approach. Ann. Biomed. Eng. 53, 634–643 (2025).
14. Liang, W. et al. Biomechanics-embedded deep learning for joint moment estimation in clinical gait analysis for cerebral palsy: a learnable Hill-type pathway. IEEE Trans. Neural Syst. Rehabil. Eng. 34, 3736–3747 (2026).
15. Lin, R. et al. Deep learning-based estimation of ground reaction forces in parkinsonian gait using an optimized set of IMU data. IEEE Trans. Neural Syst. Rehabil. Eng. 34, 2729–2740 (2026).
16. Meng, Q. et al. Spatio-temporal graph convolutional networks with transfer learning for continuous ground reaction force estimation in hemiparetic gait. Front. Bioeng. Biotechnol. 14, 1871692 (2026).
17. Kainz, H. et al. Joint kinematic calculation based on clinical direct kinematic versus inverse kinematic gait models. J. Biomech. 49, 1658–1669 (2016).
18. Gorton, G. E., Hebert, D. A. & Gannotti, M. E. Assessment of the kinematic variability among 12 motion analysis laboratories. Gait Posture 29, 398–402 (2009).
19. Museck, I. J., Brinton, D. L. & Dean, J. C. The use of wearable sensors and machine learning methods to estimate biomechanical characteristics during standing posture or locomotion: a systematic review. Sensors 24, 7280 (2024).

20. Moor, M. et al. Foundation models for generalist medical artificial intelligence. Nature 616, 259–265 (2023).
21. Yuan, H. et al. Self-supervised learning for human activity recognition using 700,000 person-days of wearable data. npj Digit. Med. 7, 91 (2024).
22. Brand, Y. E. et al. Continuous assessment of daily-living gait using self-supervised learning of wrist-worn accelerometer data. npj Digit. Med. 9, 338 (2026).
23. Li, Y. et al. Self-supervised learning enhances accuracy and data efficiency in lower-limb joint moment estimation from gait kinematics. Front. Bioeng. Biotechnol. 13, 1633513 (2025).
24. Tan, T. et al. GaitDynamics: a generative foundation model for analyzing human walking and running. Nat. Biomed. Eng. 10, 1659–1671 (2026).
25. Magruder, R. D., Gilon, S., Falisse, A. & Uhlrich, S. D. GaitEncoder: a foundation model of gait kinematics for diverse clinical applications and pathologies. Preprint at https://doi.org/10.64898/2026.07.07.26357479 (2026).
26. Dey, S. & Paez-Granados, D. PathoFM: toward a foundation model for pathological gait. in NeurIPS 2025 Workshop on Learning from Time Series for Health https://openreview.net/forum?id=Hq9Afq7EMk (2025).
27. Wu, P. et al. Hierarchical multi-task learning for comprehensive gait assessment using wearable inertial sensors. npj Digit. Med. https://doi.org/10.1038/s41746-026-02988-6 (2026).
28. Adeli, V. et al. GAITGen: disentangled motion-pathology impaired gait generative model, bringing motion generation to the clinical domain. in Proc. IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) 3150–3161 (2026).
29. Lugmayr, A. et al. RePaint: inpainting using denoising diffusion probabilistic models. in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 11451–11461 (2022).
30. Werling, K. et al. AddBiomechanics Dataset: capturing the physics of human motion at scale. in Computer Vision – ECCV 2024 490–508 (Springer, 2024).
31. Van Criekinge, T. et al. A full-body motion capture gait dataset of 138 able-bodied adults across the life span and 50 stroke survivors. Sci. Data 10, 852 (2023).
32. Shida, T. K. F. et al. A public data set of walking full-body kinematics and kinetics in individuals with Parkinson's disease. Front. Neurosci. 17, 992585 (2023).
33. Meyns, P., Molenaers, G., Duysens, J. & Jonkers, I. The differential effect of arm movements during gait on the forward acceleration of the centre of mass in children with cerebral palsy and typically developing children. Front. Hum. Neurosci. 11, 96 (2017).
34. Steele, K. M., Seth, A., Hicks, J. L., Schwartz, M. S. & Delp, S. L. Muscle contributions to support and progression during single-limb stance in crouch gait. J. Biomech. 43, 2099–2105 (2010).
35. Fukuchi, C. A., Fukuchi, R. K. & Duarte, M. A public dataset of overground and treadmill walking kinematics and kinetics in healthy individuals. PeerJ 6, e4640 (2018).
36. Viglialoro, R. M. et al. Dataset on gait analysis of parkinsonian subjects: effect of Nordic walking. Sci. Data 12, 1937 (2025).
37. Grouvel, G., Carcreff, L., Moissenet, F. & Armand, S. A dataset of asymptomatic human gait and movements obtained from markers, IMUs, insoles and force plates. Sci. Data 10, 180 (2023).
38. Han, X. et al. GroundLink: a dataset unifying human body movement and ground reaction dynamics. in SIGGRAPH Asia 2023 Conference Papers Article 48 (ACM, 2023).
39. Sugai, R. et al. LSTM network-based estimation of ground reaction forces during walking in stroke patients using markerless motion capture system. IEEE Trans. Med. Robot. Bionics 5, 1016–1024 (2023).
40. Horsak, B., Kranzl, A., Slijepcevic, D., Svehlik, M. & Dumphart, B. Test-retest reliability of spatiotemporal, kinematic, and kinetic measures in marker-based 3D gait analysis: a systematic review. Gait Posture 130, 110270 (2026).

41. Haley, S. M. & Fragala-Pinkham, M. A. Interpreting change scores of tests and measures used in physical therapy. Phys. Ther. 86, 735–743 (2006).
42. Hawkins, J. L. & Milner, C. E. Within-session reliability and minimum detectable differences for discrete lower-extremity angles and moments during walking. J. Appl. Biomech. 37, 477–480 (2021).
43. Fernandes, R., Armada-da-Silva, P., Pool-Goudzwaard, A. L., Moniz-Pereira, V. & Veloso, A. P. Test-retest reliability and minimal detectable change of three-dimensional gait analysis in chronic low back pain patients. Gait Posture 42, 491–497 (2015).
44. Kainz, H. et al. Reliability of four models for clinical gait analysis. Gait Posture 54, 325–331 (2017).
45. Aguirre Ramirez, A., Kuch, A., Johnson, R. T. & Sanchez, N. Minimal detectable change in gait biomechanics post-stroke: disentangling the effects of walking speed and stroke-related variability. Preprint at https://doi.org/10.64898/2026.08.25.26361347 (2026).
46. Peiffer, J. D. et al. Portable biomechanics laboratory enables clinically accessible movement analysis from a handheld smartphone. npj Digit. Med. 9, 692 (2026).
47. Miller, E. Y., Tan, T., Falisse, A. & Uhlrich, S. D. Integrating machine learning with musculoskeletal simulation improves OpenCap video-based dynamics estimation. IEEE Trans. Biomed. Eng. https://doi.org/10.1109/TBME.2026.3691975 (2026).
48. Delp, S. L. et al. OpenSim: open-source software to create and analyze dynamic simulations of movement. IEEE Trans. Biomed. Eng. 54, 1940–1950 (2007).
49. Rajagopal, A. et al. Full-body musculoskeletal model for muscle-driven simulation of human gait. IEEE Trans. Biomed. Eng. 63, 2068–2079 (2016).
50. Davis, R. B., Õunpuu, S., Tyburski, D. & Gage, J. R. A gait analysis data collection and reduction technique. Hum. Mov. Sci. 10, 575–587 (1991).
51. Kadaba, M. P., Ramakrishnan, H. K. & Wootten, M. E. Measurement of lower extremity kinematics during level walking. J. Orthop. Res. 8, 383–392 (1990).
52. Delp, S. L. et al. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. IEEE Trans. Biomed. Eng. 37, 757–767 (1990).
53. Leardini, A. et al. A new anatomically based protocol for gait analysis in children. Gait Posture 26, 560–571 (2007).
54. Leboeuf, F. et al. The conventional gait model, an open-source implementation that reproduces the past but prepares for the future. Gait Posture 69, 235–241 (2019).
55. Lai, A. K. M., Arnold, A. S. & Wakeling, J. M. Why are antagonist muscles co-activated in my simulation? A musculoskeletal model for analysing human locomotor tasks. Ann. Biomed. Eng. 45, 2762–2774 (2017).
56. Tashiro, Y., Song, J., Song, Y. & Ermon, S. CSDI: conditional score-based diffusion models for probabilistic time-series imputation. in Advances in Neural Information Processing Systems 34 (NeurIPS) 24804–24816 (2021).
57. Lopez Alcaraz, J. M. & Strodthoff, N. Diffusion-based time series imputation and forecasting with structured state space models. Trans. Mach. Learn. Res. https://openreview.net/forum?id=hHiIbk7ApW (2022).
58. Peebles, W. & Xie, S. Scalable diffusion models with transformers. in Proc. IEEE/CVF International Conference on Computer Vision (ICCV) 4172–4182 (2023).
59. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. in Advances in Neural Information Processing Systems 33 (NeurIPS) 6840–6851 (2020).
60. Nichol, A. Q. & Dhariwal, P. Improved denoising diffusion probabilistic models. in Proc. International Conference on Machine Learning (ICML) 8162–8171 (2021).
61. Salimans, T. & Ho, J. Progressive distillation for fast sampling of diffusion models. in Proc. International Conference on Learning Representations (ICLR) (2022).

62. Holden, M. K., Gill, K. M., Magliozzi, M. R., Nathan, J. & Piehl-Baker, L. Clinical gait assessment in the neurologically impaired: reliability and meaningfulness. Phys. Ther. 64, 35–40 (1984).
63. Hoehn, M. M. & Yahr, M. D. Parkinsonism: onset, progression and mortality. Neurology 17, 427–442 (1967).
64. Ho, J. & Salimans, T. Classifier-free diffusion guidance. Preprint at https://arxiv.org/abs/2207.12598 (2022).
65. Song, J., Meng, C. & Ermon, S. Denoising diffusion implicit models. in Proc. International Conference on Learning Representations (ICLR) (2021).

Figures

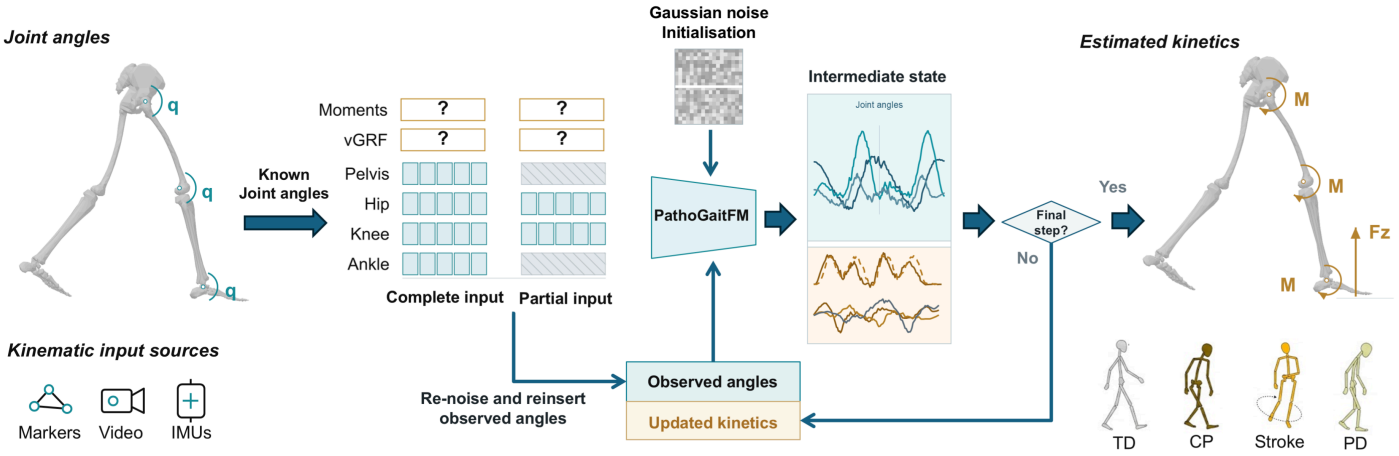


Figure 1

Estimating gait kinetics from available joint angles. Complete and partial inputs supply observed joint angles. Hatching marks unavailable angles; question marks denote unobserved kinetics. No measured kinetics are provided at inference. All channels are initialised as Gaussian noise. After each model update, observed angles are re-noised to the next sampling level and reinserted, while updated kinetic channels are retained. Sampling continues for 50 steps. Saved estimates are aggregatedacross steps and random seeds to obtain vGRF and sagittal joint moments. Icons indicate input sources and development cohorts. Waveforms are schematic. PD joint moments were not evaluated. vGRF, vertical ground reaction force; TD, typically developing; CP, cerebral palsy; PD, Parkinson's disease; IMU, inertial measurement unit; q, joint angle; M, joint moment; Fz, vGRF.

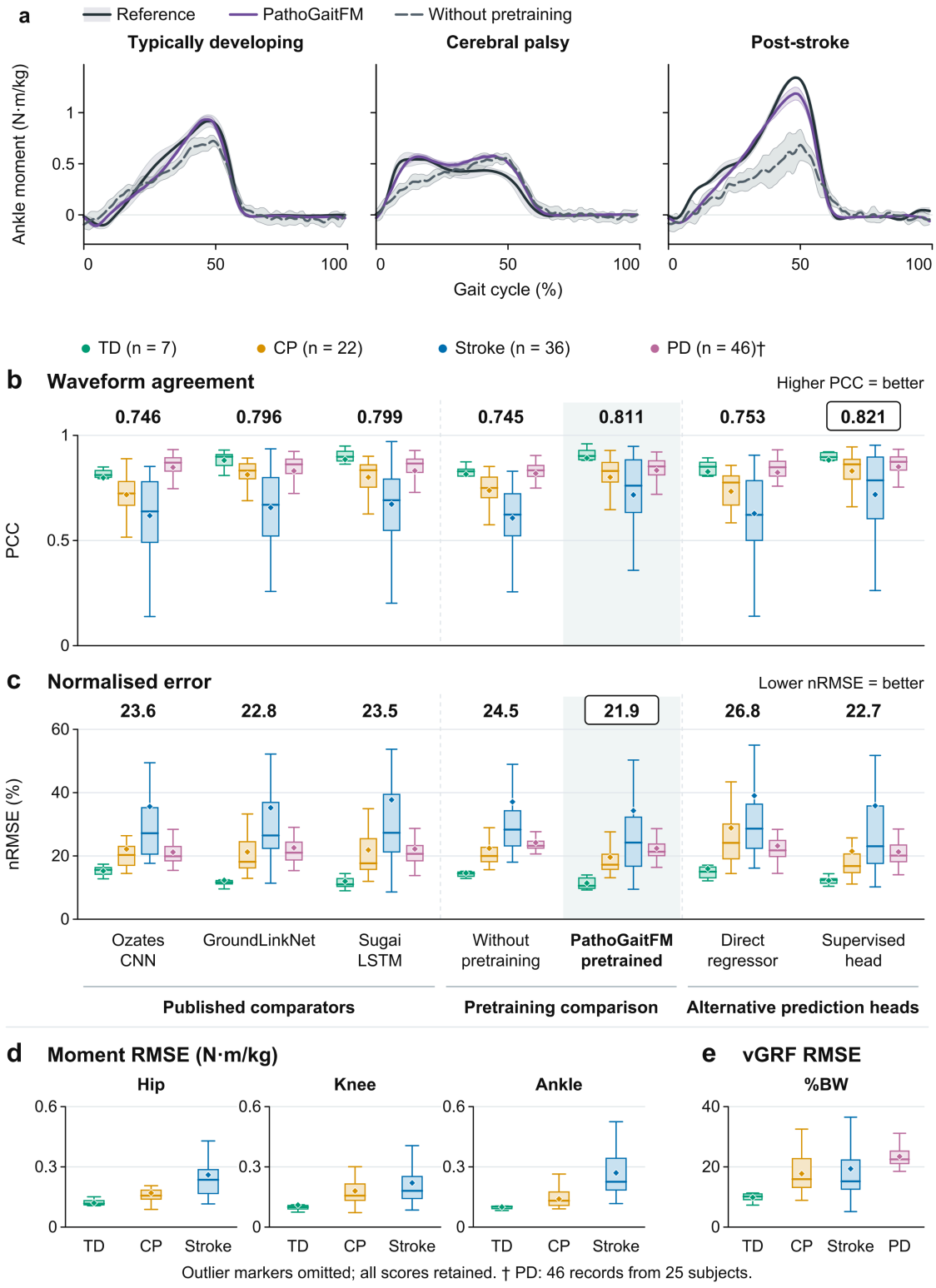


Figure 2

Within-cohort estimation and pretraining. (a) Example ankle-moment waveforms from one participant per cohort. Lines and bands show the mean ± 1 standard deviation (SD) across each participant's own cycles, on one shared vertical scale. (b, c) PCC and nRMSE for seven model configurations. Diamonds show cohort means, numbers the equally weighted four-cohort mean and outlines the best mean. (d, e) Joint-moment RMSE (N·m/kg) and vGRF RMSE (%BW) for pretrained PathoGaitFM. Boxes show medians and interquartile ranges and whiskers the most extreme observations within 1.5 interquartile ranges.

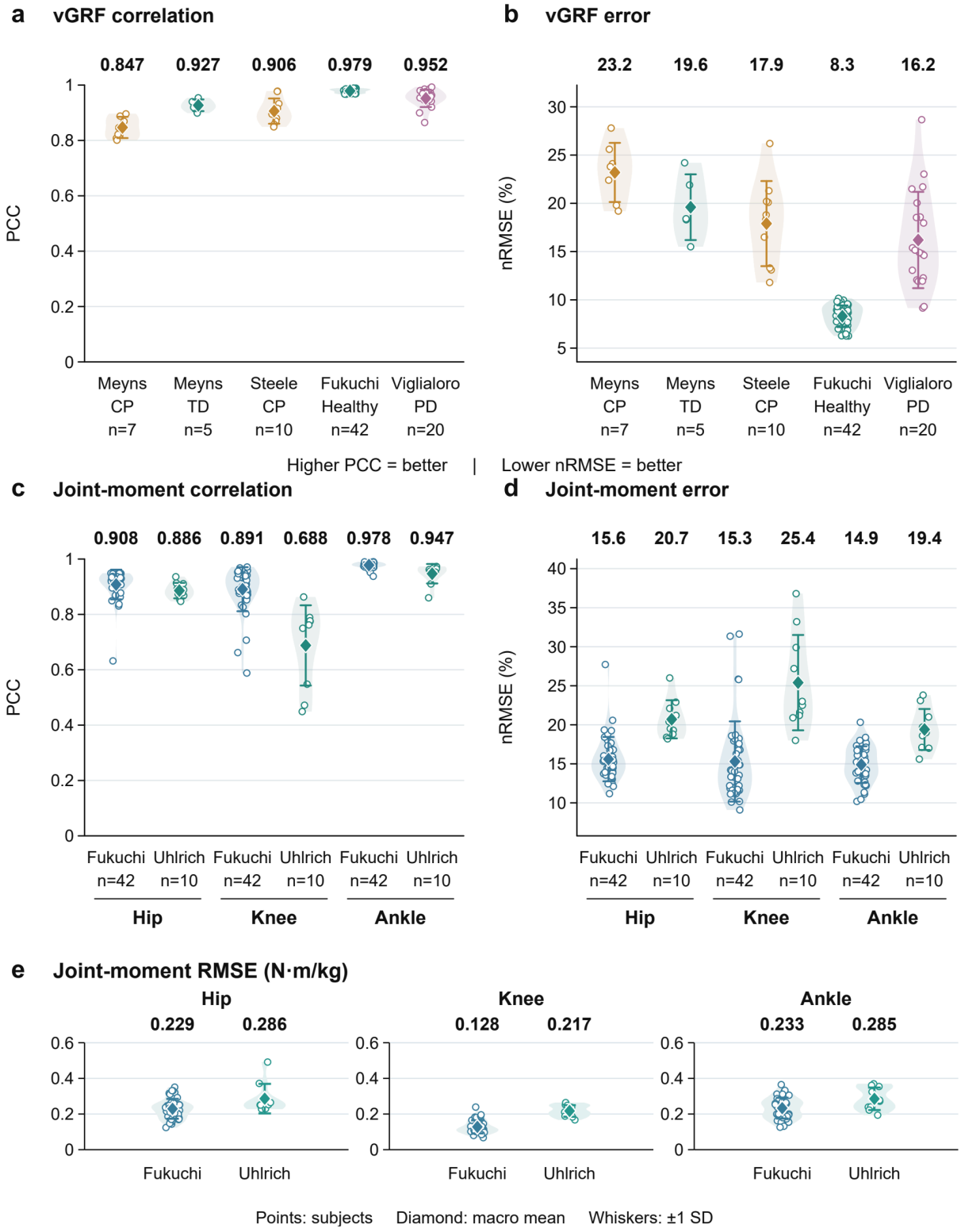


Figure 3

External validation without retraining. (a, b) vGRF estimation in four datasets, with Meyns CP and TD groups shown separately. (c–e) Sagittal joint-moment PCC, nRMSE and RMSE (N·m/kg) in the Fukuchi and Uhlrich datasets, evaluated over the full gait cycle and during left stance, respectively. Points show participant scores; violins show their distributions. Diamonds and numbers show means; whiskers indicate ±1 SD; n denotes the number of participants.

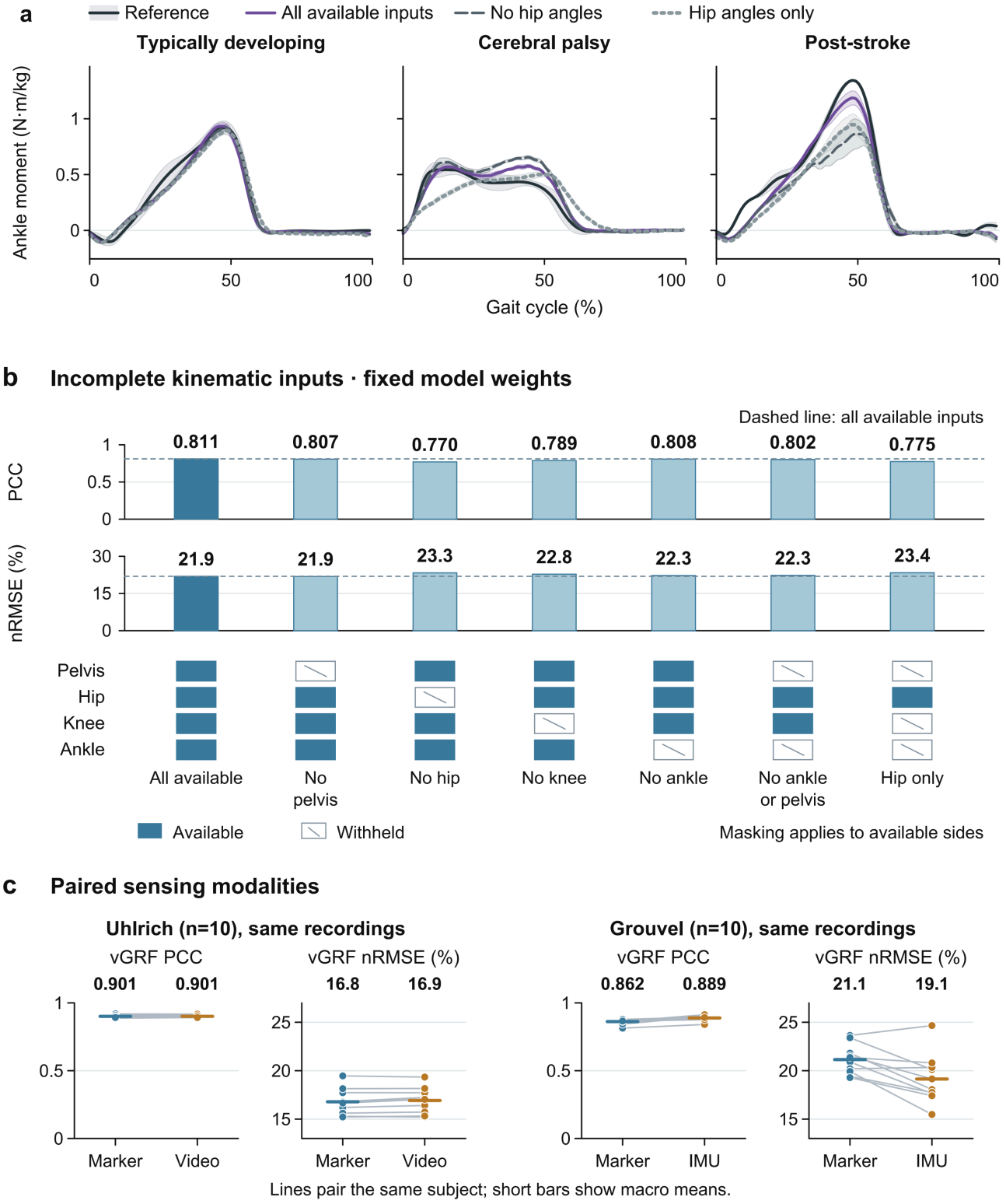


Figure 4

Incomplete and portable kinematic inputs with fixed model weights.(a) Example ankle-moment waveforms under three input configurations, from one participant per cohort. Lines and bands show the mean ± 1 SD across each participant's own cycles, on one shared vertical scale. (b) Mean scores with all available inputs and six reduced configurations. Cells indicate retained angle groups; masking applies to available sides. Dashed lines mark performance with all available inputs. (c) Paired vGRF scores using marker- and video-based kinematics (Uhlrich) or marker- and IMU-based kinematics (Grouvel), each in 10 healthy participants. Lines connect paired scores and bars show means.

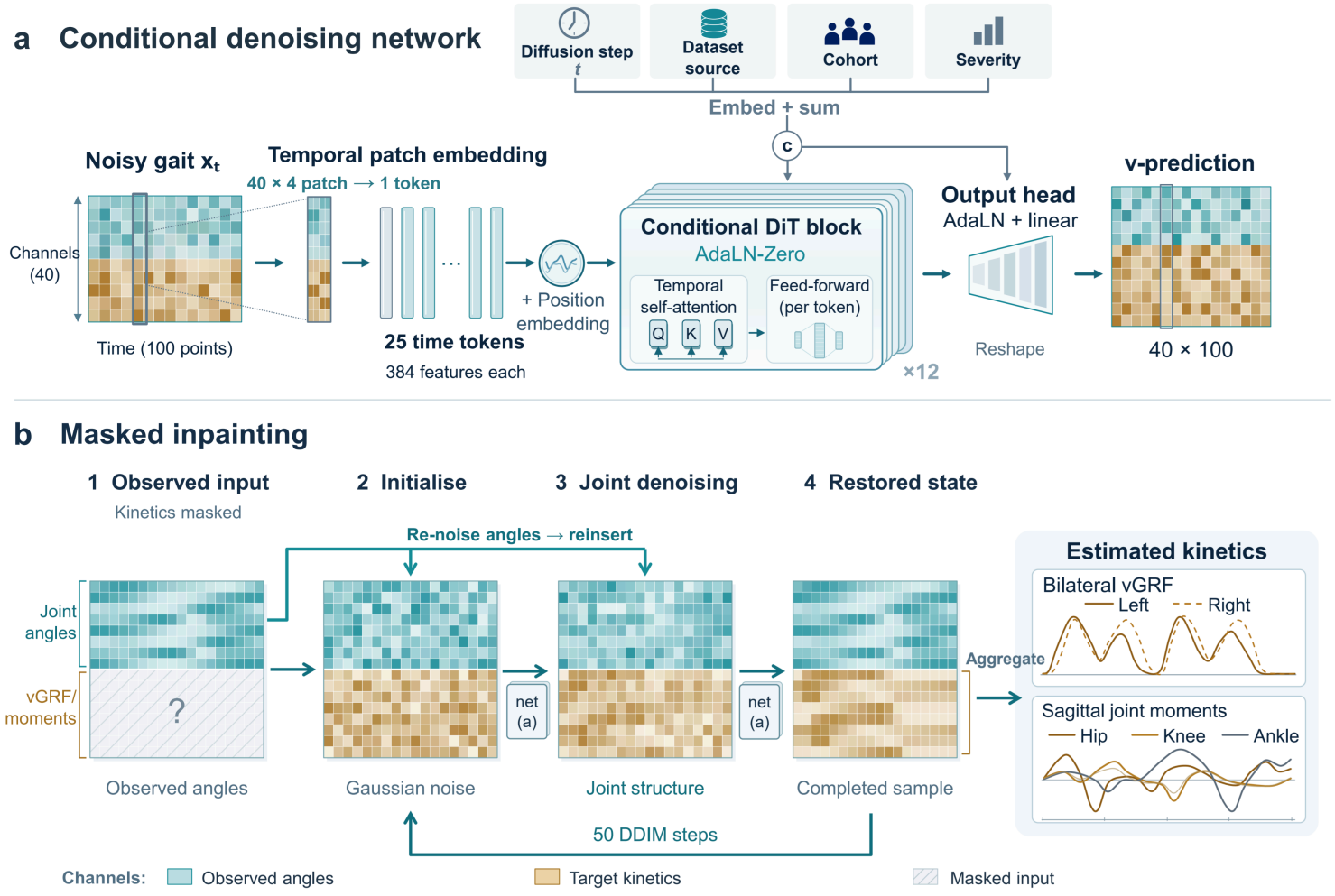


Figure 5

Conditional denoising network and masked inpainting. (a) One denoising pass. Patches of four time points across all channels are embedded as tokens and passed through the conditional Transformer blocks. Summed diffusion-step, source, cohort and severity embeddings form c, which modulates every block and the output head. The output head restores the channel-by-time layout and predicts v, the diffusion target. (b) Kinetic estimation. From Gaussian noise, the network in (a) denoises the hidden kinetic channels over 50 sampling steps, while the observed joint angles are re-noised to the current noise level and reinserted at each step. Saved estimates are aggregated to obtain vGRF and joint moments. Matrices and waveforms are schematic.