Gait parameters measured from wearable sensors reliably detect freezing of gait in a stepping in place task

Cameron Diep1, Johanna O’Day1,2, Yasmine Kehnemouyi1, Gary Burnett1,3, and Helen Bronte-Stewart1,4*

1 Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA 94305, United States
2 Department of Bioengineering, Stanford University, Stanford, CA 94305, United States
3 Department of Biomedical Data Science, Stanford University, Stanford, CA 94305, United States
4 Department of Neurosurgery, Stanford University School of Medicine, Stanford, CA 94305, United States
* Correspondence: hbs@stanford.edu

Abstract: Freezing of gait (FOG), a debilitating symptom of Parkinson’s disease (PD), can be safely studied using the stepping in place (SIP) task. However, clinical, visual identification of FOG during SIP is subjective and time consuming, and automatic FOG detection during SIP currently requires measuring center of pressure on dual force plates. This study examines whether FOG elicited during SIP in 10 individuals with PD could be reliably detected using kinematic data measured from wearable inertial measurement unit sensors (IMUs). A general, logistic regression model (AUC = 0.81) determined that three gait parameters together were overall the most robust predictors of FOG during SIP: arrhythmicity, swing time coefficient of variation, and swing angular range. Participant-specific models revealed varying sets of gait parameters that best predicted FOG for each participant, highlighting variable FOG behaviors, and demonstrated equal or better performance for 6 out of the 10 participants, suggesting the opportunity for model personalization. The results of this study demonstrated that gait parameters measured from wearable IMUs reliably detect FOG during SIP, and the general and participant-specific gait parameters allude to variable FOG behaviors that could inform more personalized approaches for treatment of FOG and gait impairment in PD.

Keywords: Parkinson’s disease; wearables; inertial measurement unit; sensors; freezing of gait

1. Introduction

Freezing of gait (FOG), often defined as a feeling of one’s feet being “glued” to the floor [1, 2], is a debilitating phenomenon in Parkinson’s disease (PD) that negatively impacts quality of life and can lead to falls, serious injury, or even death [3-6]. FOG is a challenging phenomenon to objectively measure in the clinic and laboratory [7], but several tasks, such as 360-degree turning in place [8], the turning and barrier course (TBC) [9], and stepping in place (SIP) [10], have been developed to safely and reliably elicit FOG.

The SIP task is a safe and powerful tool for studying FOG. However, clinical, visual identification of FOG during SIP is subjective and time consuming, and automatic FOG detection during SIP currently requires measuring center of pressure on dual force plates [10]. These force plates are large, immobile, and only used in laboratories and clinics. Instead, wearable inertial measurement unit sensors (IMUs) are small and portable, allowing for use in daily life. IMUs have been used to characterize and predict FOG during several tasks [9, 11-37], and many of these IMU-based FOG detection schemes rely on statistical and machine learning methods [9, 11-22]. For example, IMUs measured kinematic data during straight walking and the TBC [9, 37]. A statistical model developed...
using this data revealed that certain gait parameters—stride time, swing angular range, asymmetry, and arrhythmicity—reliably predicted FOG during the TBC [9, 37]. No study has investigated if IMUs can detect FOG during SIP, and no statistical model has been developed to determine which gait parameters most reliably predict FOG during SIP.

In this study, we examined whether FOG elicited during SIP could be reliably detected using kinematic data measured from two IMUs worn on the shanks. We aimed to develop a statistical model using multiple gait features to determine which gait parameters best predicted FOG during SIP and gain insight into potential personalized approaches for treatment of FOG and gait impairment in PD.

2. Materials and Methods

2.1. Participant

Data were obtained from 10 participants (5 female) with clinically established PD at the Stanford Movement Disorders Center. All participants gave their written informed consent to participate in this study, which was approved by the Food and Drug Administration and the Stanford University School of Medicine Institutional Review Board. Participants were tested off therapy (medication and/or deep brain stimulation). Long-acting dopaminergic medication was withdrawn over 24 hours prior to testing (48-72 hours for extended-release dopamine agonists), and short-acting dopaminergic medication was withdrawn over 12 hours prior to testing. Deep brain stimulation was turned off at least 2 minutes before testing. A certified rater performed the Unified Parkinson’s Disease Rating Scale Part III (UPDRS III) [38] and the Freezing of Gait Questionnaire (FOG-Q) [39].

2.2. Experimental Protocol

The SIP task is a validated task that elicits FOG and consists of repetitive alternating stepping in place at a self-selected pace on dual force plates [10]. Participants begin by standing at rest. At a “go” cue, participants step in place for 100 seconds, ending with a “stop” cue. All participants are harnessed during the task as a safety measure. Data from one trial of SIP per participant were analyzed in this work.

2.3. Data Acquisition

External videos of the SIP task were recorded on an encrypted clinical iPad (Apple Inc., Cupertino, CA) for offline video review. Ground reaction forces were captured at 100 Hz with two force plates (Bertec, Columbus, OH). Participants were instrumented with IMUs (APDM, Inc., Portland, OR) on the lateral side of the shanks so that one axis of each sensor was aligned with the sagittal plane of the participant. Triaxial gyroscope and accelerometer signals from the IMUs were sampled at 128 Hz. The data were filtered using a zero-phase 8th order low pass Butterworth filter with a 9 Hz cut-off frequency, and principal component analysis was used to align the angular velocity with the sagittal plane.

Data gathered from the IMUs were used to measure and calculate a total of eight gait parameters: peak shank angular velocity, stride time, swing angular range, swing time, swing time coefficient of variation (CV), asymmetry, arrhythmicity, and freeze index. Using the angular velocity measured by the IMUs (Figure 1), individual steps were identified as positive peaks in the left and right sagittal shank angular velocity plot. These peaks represent the absolute maxima shank angular velocity in a step cycle and were marked as steps only if they exceeded a minimum threshold of 10 deg/s [9]. Stride time was defined as the time between two successive positive peaks on the angular velocity plot. Swing angular range was calculated as the area under a peak on the sagittal angular velocity plot. Swing time was computed as the time between swing phase initiation and end, as
determined by zero-crossings on the angular velocity plot. Swing time CV was defined as the standard deviation of swing time divided by average swing time over a window of the previous 6 steps. Asymmetry was defined as $100 \times |\ln(\text{SSWT}/\text{LSWT})|$, where SSWT and LSWT correspond to the leg with the shortest and longest mean swing time over a window of the previous six steps. Arrhythmicity was calculated as the average stride time CV of the previous three stride times of the left and right leg. Freeze index was defined as the power in the freezing band (3-8 Hz) divided by the power in the gait band (0.5-3 Hz) [40]. Analysis of these gait parameters was performed in MATLAB (version 9.8, The MathWorks Inc., Natick, MA). All gait parameters were scaled and normalized to have zero mean and unit variance.

2.4. FOG Detection

All participants exhibited at least one freezing episode during the SIP task, as identified by offline video review.

Ground-truth labels of freezing episodes were automatically identified from force plate data by a previously validated, computerized algorithm [10]. These ground-truth labels of freezing and non-freezing episodes were used to create and balance the training and testing sets used to build the models, as discussed in the following section.

Data from the force plates and IMUs were synchronized by inducing a force or rapid acceleration that was detected by both systems, allowing for the validation of kinematic, sensor-based FOG detection with kinetic, force plate-based FOG detection.

2.5. Logistic Regression Models of FOG during SIP

Binomial logistic regression models were developed to calculate the probability that a given step during SIP was considered part of a freezing episode. A threshold of 0.5 was set such that a given step was categorized as part of a freezing episode when the probability that it was part of a freezing episode exceeded 50%. Ground-truth, binary labels of freeze steps (non-freeze = 0, freeze = 1) from the previously validated algorithm [10] and eight gait parameters (peak shank angular velocity, stride time, swing angular range, swing time, swing time CV, asymmetry, arrhythmicity, and freeze index) were used to build the binomial logistic regression models in R (version 4.0.2, R Core Team (2020)).
To develop the general model, the entire data set of eight gait parameters from all 10 participants was scaled to have zero mean and unit variance, balanced so that there was an equal number of freeze steps as non-freeze steps, and shuffled. 75% of each participant’s data were pooled into the training data set, while the remaining 25% of each participant’s data were pooled into the testing data set. The general model was then trained and tested on the respective data sets using 10-fold cross validation.

To develop the participant-specific models, the data set of eight gait parameters was first separated by participant. Each participant’s data set was then scaled to have zero mean and unit variance, balanced so that there was an equal number of freeze steps as non-freeze steps, and shuffled. The participant-specific models were trained on 75% of the respective participant’s data and tested on the remaining 25% of the participant’s data using 10-fold cross validation.

Model performance was evaluated by calculating the area under the receiver operating curve (AUC), accuracy (number of correct predictions divided by the total number of predictions), sensitivity (number of freezes identified by the model divided by the total number of true freezes), and specificity (number of non-freezes identified by the model divided by the total number of true non-freezes).

3. Results

3.1. Participants

Among the 10 participants (5 female), average age was 62.5 ± 8.8 years, average disease duration was 10.3 ± 3.5 years, average off therapy UPDRS III score was 38.4 ± 11.2, and average FOG-Q question 3 (FOG-Q3) score was 2.1 ± 1.5 (Table 1).

Table 1. Participant demographics (n = 10).

<table>
<thead>
<tr>
<th>Participant</th>
<th>Sex</th>
<th>Age (years)</th>
<th>Disease Duration (years)</th>
<th>UPDRS III Score</th>
<th>FOG-Q3 Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
<td>65</td>
<td>6</td>
<td>25*</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>60</td>
<td>6</td>
<td>48</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>M</td>
<td>44</td>
<td>7</td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>57</td>
<td>11</td>
<td>55</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>61</td>
<td>12</td>
<td>36</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>61</td>
<td>16</td>
<td>38**</td>
<td>3**</td>
</tr>
<tr>
<td>7</td>
<td>F</td>
<td>75</td>
<td>8</td>
<td>27</td>
<td>N/A***</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>56</td>
<td>11</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>69</td>
<td>14</td>
<td>38</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>F</td>
<td>71</td>
<td>8</td>
<td>32</td>
<td>4</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td></td>
<td>62.5</td>
<td>10.3</td>
<td>38.4</td>
<td>2.1</td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td></td>
<td>8.8</td>
<td>3.5</td>
<td>11.2</td>
<td>1.5</td>
</tr>
</tbody>
</table>

* and ** report scores from previous visit (* 1 month prior and ** 3 months prior), since no UPDRS III and/or FOG-Q3 score was recorded at the time of visit.

*** no FOG-Q3 score was recorded at the time of visit and no other visits occurred.

3.2. Gait parameters measured by wearable IMUs detected FOG during SIP
Gait parameters measured from wearable IMUs reliably detected FOG during SIP, capturing changes in gait metrics, such as swing angular range and arrhythmicity, that distinguish freezing from non-freezing episodes (Figure 2).

Figure 2. Detection of freezing of gait (FOG) during the stepping in place (SIP) task. FOG detection during SIP from kinetics derived from vertical force measured by dual force plates (A) and from kinematics derived from shank angular velocity measured by wearable inertial measurement unit sensors (B). Blue and red traces correspond to the left and right legs, respectively. Swing angular range (C) and arrhythmicity (D) calculated from shank angular velocity show visually detectable differences between freezing and non-freezing episodes.

3.3. A general model revealed that three gait parameters best predicted FOG during SIP overall

Logistic regression models based on kinematic data measured from wearable IMUs detected FOG during SIP on a step-by-step basis. The general model, tested on all participants’ data, achieved an AUC value of 0.81, accuracy of 0.84, sensitivity of 0.86, and specificity of 0.81 (Table 2).

Table 2. General logistic regression model performance.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Area Under the Curve (AUC)</td>
<td>0.81</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.84</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.86</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.81</td>
</tr>
</tbody>
</table>

The general model determined that the three most robust predictors of FOG during SIP were arrhythmicity (coefficient = 1.076), swing time CV (coefficient = 0.894), and swing angular range (coefficient = -0.06), with an intercept of 0.017 (Figure 3). Peak shank angular velocity (coefficient = -0.006) also distinguished freeze from non-freeze steps during SIP with a smaller coefficient. The probability that a step during SIP was...
considered part of a freezing episode was calculated using the aforementioned, normalized gait parameters, denoted as $X_{AR}$, $X_{SWCV}$, $X_{SA}$, and $X_{AV}$, respectively:

$$P(FOG) = \frac{1}{1 + e^{-(0.017+1.076\times X_{AR}+0.894\times X_{SWCV}-0.06\times X_{SA}-0.006\times X_{AV})}}$$

(1)

Figure 3. All steps from all participants mapped according to the top three general predictors of freezing of gait (FOG) during the stepping in place (SIP) task. The general model revealed that arrhythmicity, swing time coefficient of variation (CV), and swing angular range together were the most robust predictors of FOG during SIP. There is a clear separation between freeze (red) and non-freeze (green) steps during SIP.

3.4. Participant-specific models revealed varying sets of gait parameters that best predicted FOG during SIP and could outperform the general model for some participants

Although arrhythmicity, swing time CV, and swing angular range together were the most robust predictors of FOG in general (Figure 3), the participant-specific models revealed that varying sets of gait parameters best predicted FOG for each participant (Table 3). For example, the most robust predictors of FOG for participant 1 were a combination of peak shank angular velocity, swing time, swing angular range, and arrhythmicity, while the most robust predictors of FOG for participant 2 were a combination of arrhythmicity, swing time, swing time CV, and asymmetry (Figure 4). External video recordings of participants’ SIP trials showed that each participant exhibited one, some, or all of the following FOG behaviors: complete akinesia (i.e. no observable motion of the legs), small stepping (i.e. festination or shuffling), trembling in place (i.e. small movements of the leg or knee with no effective stepping motion), and slow stepping (i.e. sticky feet).

Although the general model achieved an overall AUC of 0.81 (Table 2), it performed with variable accuracy among the participants, ranging from 0.60 to 1.00. Participant-specific models, on the other hand, achieved equal or higher accuracies than the general model for 6 out of the 10 participants. Sensitivity was equal or higher for 8 out of the 10 participants, and specificity was equal or higher for 7 out of the 10 participants (Table 4). Values of 1.00 for accuracy, sensitivity, and specificity were likely the result of small training and testing sets used to build the models.

Table 3. Coefficients of general and participant-specific gait parameters that best predicted freezing of gait during the stepping in place task.
Figure 4. All steps from participants 1 and 2 mapped according to their top three participant-specific predictors of freezing of gait (FOG) during the stepping in place (SIP) task. Participant-specific models revealed that varying sets of gait parameters best predicted FOG during SIP for each participant. For example, participant 1’s most robust predictors of FOG were a combination of peak shank angular velocity, swing time, and swing angular range (left), while participant 2’s most robust predictors of FOG were a combination of arrhythmicity, swing time, and swing time coefficient of variation (CV) (right). For both participants, there is a clear separation between freeze (red) and non-freeze (green) steps during SIP.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.578</td>
<td>-0.479</td>
<td>0.318</td>
</tr>
<tr>
<td>2</td>
<td>0.627</td>
<td>0.544</td>
<td>0.181</td>
</tr>
<tr>
<td>3*</td>
<td>0.065</td>
<td>0.466</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.619</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.505</td>
<td>0.399</td>
<td>0.341</td>
</tr>
<tr>
<td>8*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-0.542</td>
<td>-0.118</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-1.165</td>
<td>-2.474</td>
<td>1.293</td>
</tr>
</tbody>
</table>

* participant-specific models calculated unstable coefficients due to small training and testing sets, thus no coefficients were reported.

Table 4. Participant-specific model performance.
4. Discussion

This study determined that gait parameters measured from wearable IMUs reliably detected FOG during the SIP task. A general, logistic regression model (AUC = 0.81) determined that three gait parameters together were overall the most robust predictors of FOG during SIP: arrhythmicity, swing time CV, and swing angular range. Participant-specific models revealed varying sets of gait parameters that best predicted FOG for each participant and demonstrated equal or better performance for 6 out of the 10 participants.

4.1. Gait parameters measured from wearable inertial sensors detected FOG in PD

Automatic FOG detection during SIP on dual force plates has been previously validated [10], but these force plates are large, immobile, and only used in laboratories and clinics. Instead, wearable IMUs are small and portable, allowing for use in daily life. This study determined that IMUs worn on the shanks reliably detected FOG elicited during SIP by capturing changes in gait parameters, such as swing angular range and arrhythmicity, that distinguish freezing from non-freezing episodes. These results extend the accessibility of the SIP task for FOG assessment to the home environment.

4.2. A general model for FOG detection during SIP determined that three gait parameters were overall the most robust predictors of FOG

Logistic regression models based on kinematic data measured from wearable IMUs detected FOG during SIP on a step-by-step basis. The general model achieved an AUC value of 0.81, accuracy of 0.84, sensitivity of 0.86, and specificity of 0.81, which are similar to other IMU-based FOG detection algorithms [8-9, 11-34, 41].

The general model revealed that three gait parameters together were overall the most robust predictors of FOG during SIP: arrhythmicity, swing time CV, and swing angular range. The probability that a step was considered part of a freezing episode increased as arrhythmicity and swing time CV increased and as swing angular range decreased. This is consistent with the abnormal gait pattern generation thought to cause FOG. The positive correlation between arrhythmicity and FOG supports previous findings that people who experience FOG exhibit greater arrhythmicity and bilateral dyscoordination of left-right stepping than people who do not experience FOG [10, 42-44]. This relationship can
also be explained by the impaired temporal control of gait cycles during FOG [42, 45-46], further reinforcing the rhythmicity and bilateral coordination required while performing the SIP task [10]. In addition, increased step time variability and cadence often precede and accompany FOG [42, 45-46]. Our model also captured increased swing time CV during FOG in SIP, consistent with the finding that gait variability increases in PD [47] and prior to FOG [48-49]. Finally, gait impairment preceding freezing episodes has been characterized by reduced joint angle ranges in the hip, knee, and ankle [45]. Similarly, our model revealed that FOG during SIP was often characterized by decreased swing angular range.

4.3. Participant-specific models highlight the opportunity for personalized approaches for the treatment of FOG and gait impairment in PD

In general, arrhythmicity, swing time CV, and swing angular range were the three most robust predictors of FOG during SIP. However, the participant-specific models revealed that varying sets of gait parameters best predicted FOG for each participant. For example, the most robust predictors of FOG for participant 1 were a combination of peak shank angular velocity, swing time, swing angular range, and arrhythmicity, while the most robust predictors of FOG for participant 2 were a combination of arrhythmicity, swing time, swing time CV, and asymmetry. The differences between sets of predictive gait parameters suggest that each participant’s FOG may be caused by differing levels of gait impairment, possibly due to variable disease durations and severity [45, 50]. PD subtype (i.e. akinetic rigid or tremor dominant), or FOG behaviors, such as complete akinesia, small stepping, and trembling in place [51] as well as slow stepping. External video recordings of participants’ SIP trials showed that some participants exhibited one, some, or all of the four types of FOG behaviors. This observation combined with the improved performance of participant-specific models compared to the general model in some participants support the use of personalized approaches for the treatment of FOG and gait impairment in PD.

4.4. Limitations

This study focused on sensors worn on the shanks, the most common body location for sensors aimed to detect FOG [41]. Future studies could look into sensor placements on other body locations. Although our logistic regression models achieved performance results similar to other IMU-based FOG algorithms [8-9, 11-34, 41], the training and testing sets were built using data from one trial of SIP from a small cohort of participants and were further reduced after balancing freeze and non-freeze steps. To increase the sizes of the training and testing sets, future models could include more trials of SIP from people who experience FOG in addition to data from people who do not experience FOG. The models’ freeze classifications were also binary (non-freeze and freeze). Future models could further discriminate freezes based on the variable FOG behaviors (i.e. complete akinesia, small stepping, trembling in place, and slow stepping).

5. Conclusions

The SIP task is a useful tool for eliciting FOG in the clinic and laboratory setting to study and treat FOG and gait impairment in PD. This study has demonstrated that gait parameters measured from wearable sensors reliably detected FOG during SIP, extending the accessibility of the SIP task for FOG assessment to the home environment. A general, logistic regression model built from the kinematic data revealed that three gait parameters together were overall the most robust predictors of FOG during SIP: arrhythmicity, swing time CV, and swing angular range. Participant-specific models revealed varying sets of gait parameters that best predicted FOG for each participant, highlighting variable FOG behaviors, and demonstrated equal or better performance than the general model for 6
out of the 10 participants. These results support the use of personalized approaches for the treatment of FOG and gait impairment in PD both inside and outside the clinical environment.


**Funding:** This study was supported in part by the National Institute of Neurological Disorders and Stroke (NINDS UH3NS107709 and R21 NS096398-02), the Michael J. Fox Foundation (9605), the Robert and Ruth Halperin Foundation, the John A. Blume Foundation, the John E. Cahill Foundation, and the National Institutes of Health Big Data to Knowledge (BD2K) Center of Excellence Grant (U54EB020405).

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and was approved by the Institutional Review Board of Stanford University (Protocol 52548 approved September 3, 2019).

**Informed Consent Statement:** Informed consent was obtained from all participants involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** We thank Sudeep Aditham for his assistance in analyzing kinematic data and developing the participant-specific models, as well as Matthew Petrucci, Kevin Wilkins, Shannon Hoffman, and Gerrit Orthlieb for their support in collecting data during the experiments and for their helpful feedback. We also thank the participants who dedicated their time to participating in our study.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


47. Wilkins, K.B.; Parker, J.E.; Bronte-Stewart, H.M. Gait variability is linked to the atrophy of the Nucleus Basalis of Meynert and is resistant to STN DBS in Parkinson’s disease. Neurobiology of Disease, 2020, 146. doi: 10.1016/j.nbd.2020.105134


