

RESEARCH

Open Access



A feasibility study of objective outcome measures used in clinical trials of freezing of gait

Gonzalo J. Revuelta^{1,2*}, Aaron Embry^{2,3}, Jordan J. Elm⁴, Shonna Jenkins¹, Philip Lee^{1,5} and Steve Kautz^{2,3}

Abstract

Background: Freezing of gait (FOG) is notoriously difficult to quantify, which has led to the use of multiple markers as outcomes for clinical trials. The instrumented timed up and go (TUG) and the many parameters that can be derived from it are commonly used as objective markers of FOG severity in clinical trials; however, it is unknown if they represent actual FOG severity.

Objective: To determine the specificity and responsiveness of objective surrogate markers of FOG severity commonly utilized in FOG studies.

Methods: Study design: We compared the specificity and responsiveness of commonly used markers in FOG clinical trials. Markers compared included velocity, step/stride length, step/stride length variability, TUG, and turn duration. Data was collected in four conditions (ON and OFF dopaminergic drugs, with and without a dual task). Unified Parkinson's Disease Rating Scale (UPDRS) was administered in the ON and OFF states.

Results: Thirty-three subjects were recruited (17 PD subjects without FOG (PD-control) and 16 subjects with PD and dopa-responsive FOG (PD-FOG)). The UPDRS motor scores were 24.9 for the PD-control group in the ON state, 24.8 for the FOG group in the ON state, and 42.4 for the FOG group in the OFF state. Significant mean differences between the ON and OFF conditions were observed with all surrogate markers ($p < 0.01$). However, only dual task turn duration and step variability showed trends toward significance when comparing PD-control and ON-FOG ($p = 0.08$). Test-retest reliability was high (ICC > 0.90) for all markers except standard deviations. Step length variability was the only marker to show an area under the ROC curve analysis > 0.70 comparing ON-FOG vs. PD-control.

Conclusions: Multiple candidate surrogate markers for FOG severity showed responsiveness to levodopa challenge; however, most were not specific for FOG severity.

Keywords: Parkinson's disease, Freezing of gait, Severity, Surrogate markers, Outcome measures

Background

Freezing of gait (FOG) is a debilitating condition occurring in the majority of patients with Parkinson's disease (PD) [1–3], for which there is no effective therapy. It is defined as the episodic inability to walk, often triggered

by environmental factors [4]. A major barrier toward therapeutic development in FOG is the lack of validated, objective outcome measures of FOG severity [5]. Measures like the Freezing of Gait Questionnaire [6] (FOG-Q) are limited by their subjective nature and cannot be repeated in one session since they are meant to be retrospective over a period of 1 month. The new FOG-Q has recently been found to be unreliable and not responsive to small effect sizes [7]. Measures that rely on capturing a FOG episode in the laboratory (direct measures) [8–12] are limited by the inherent variability of each episode;

*Correspondence: revuelta@musc.edu

¹ Movement Disorders Division, Department of Neurology, College of Medicine, Medical University of South Carolina, 208B Rutledge Avenue, MSC 108, Charleston, SC 29425, USA
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

therefore, a captured episode may not be representative of overall FOG severity. Furthermore, approaches to reliably trigger an episode have not been established. Long-term continuous monitoring approaches [12, 13] are ideal since they capture FOG severity over a period of days or weeks accounting for variability of individual episodes; however, they cannot be repeated in one session (since they must be administered over a long term) and require the analysis of large amounts of data.

Surrogate markers of FOG severity present an option for therapeutic trials, since they are objective assessments, easy to administer, can be administered multiple times in one session, and do not depend on triggering an episode of FOG; however, their specificity for FOG severity has not been determined. This type of marker is particularly useful for dose-finding studies and determining immediate effects of therapeutic interventions, e.g., neuromodulation therapies which require testing of multiple variables to optimize. For these reasons, clinical trials of therapies for FOG have utilized multiple candidate surrogate markers including instrumented timed up and go (TUG), turn duration, velocity, dual task interference, and step length variability [14–17]. However, it is not clear which (if any) of these markers best represent FOG severity, and if they are responsive to the interventions that are being tested, impairing our ability to interpret these studies and providing little guidance for future study design.

We selected markers that have been commonly utilized as surrogate markers of FOG severity in previous studies (velocity, step length, step length variability, dual task interference [13, 14, 17, 18], and turn duration [19–21]). To determine the feasibility of objective markers as outcomes in future clinical trials aiming to improve FOG severity, we defined two feasibility objectives: (1) *specificity for FOG* and (2) *responsiveness to intervention*. Given that these markers are indirect measures of gait, specificity is of particular interest. To determine the specificity of each marker for FOG, we selected a group of patients with FOG, and a control group of PD patients without FOG (that had otherwise similar motor severity as the PD-FOG group) and tested the ability of each marker to differentiate between groups. To determine responsiveness, we ensured each of the FOG patients selected had a clear response to dopaminergic medications (dopa-response) and compared the ability of each marker to differentiate between the OFF and ON medication state.

Methods

Subjects

Subjects (ages 18–80) who met UK Brain Bank criteria for idiopathic PD (Hoehn and Yahr stage 2–4 [22]) were recruited from the Medical University of South Carolina

Movement Disorder Clinics. Referrals were made by clinic providers and patients were then contacted by research staff and provided with information of the study. Interested potential subjects were brought into the Murray Center for Research in Parkinson's Disease and Related Diseases for a screening visit and were consented by study staff.

Inclusion and exclusion criteria

Subjects with a score of 0 in question 1 of the new freezing of gait questionnaire [23] (nFOGQ) and item 14 of the UPDRS part 2 were enrolled into the PD-control group. Subjects with a score of 1 in question 1 of the nFOGQ were enrolled in the FOG group. To ensure subjects in the FOG group had dopa-responsive FOG, an improvement of at least 1 point on item 14 of the Unified Parkinson's Disease Rating Scale (UPDRS) from the OFF to the ON state was required. In addition, each subject was observed to have FOG at screening and confirmed through multiple comprehensive clinical evaluations by a movement disorder neurologist (GJR) in the ON and OFF states. Subjects who exhibited FOG due to any trigger (initiation, turning, upon reaching destination, or on straightaway walking) or any phenomenological subtype of FOG (akinetetic, knee trembling) were included in the FOG group; therefore, all FOG subgroups [24] were included in this study. Subjects with a Mini-Mental Status Examination score of < 26, or who were unable to walk 30 feet unassisted in the OFF state, or had any other significant gait impairment (festination, or major orthopedic disturbance affecting gait) were also excluded from the study.

The Institutional Review Board of the Medical University of South Carolina approved the study (Pro00037836). All participants provided written informed consent to take part in this study. The datasets generated during the current study are available from the corresponding author upon request.

Assessments

Clinical

All patients had full UPDRS (parts 1–4) in the practically defined ON and OFF states (OFF: 12 h off all dopaminergic agents, and ON: at least 30 min after taking dopaminergic agents), and nFOGQ, performed by a movement disorder neurologist (GJR).

Description of gait assessment

Spatiotemporal parameters were obtained by a research physical therapist (AE) from the GAITRite® (CIR Systems, Franklin NY) electronic walkway in the ON and OFF states with and without a dual cognitive task. The dual tasks alternated between serial 7's and every other

letter of the alphabet. Performance on the dual task was monitored to ensure adequate effort as reported in prior studies [25]. Spatiotemporal data was collected and averaged from four passes over the GAITRite[®] walkway (two trials). Specifically, they were asked to stand up, walk over the GAITRite[®] walkway, step off the GAITRite[®] onto the M² walkway, turn around a cone set at the center of the M² (54 inches to the center of the cone from the leading edge of the M2/GAITRite[®] interface), and walk back to the chair. The M² walkway is a square digital walkway placed at the end of the GaitRite[®] walkway designed to capture the turn. The instructions for the walking task were identical to what is commonplace during the TUG [26]. The turn was 180° and the diameter of the turn was only limited to the 48" width or lateral boundaries of the M² walkway. The turn was performed by each subject in their preferred direction. Participants were not required to pre-select their direction of turn and were not mandated to turn in either or both directions.

Quantitative data generated from gait assessment

This protocol yielded two walking periods or passes on the GAITRite[®] per trial and one turn duration per trial. Two trials (4 passes) were completed in each condition (two trials ON dopaminergic medication, single task; two trials ON dopaminergic medication, dual task; two trials OFF dopaminergic medication, single task; and two trials OFF dopaminergic medication, dual task). The average and standard deviation (SD) was estimated for each side (left and right) for a total of four passes in each condition (each trial producing two passes on the GAITRite[®], one departing and another returning to the chair). Step length was not corrected for leg length. Step and stride length coefficient of variability (CV) were calculated from the standard deviation of each parameter (again a total of 4 trials on the GAITRite[®] were used to calculate CV) as previously described [27]. The turn task (mean time to turn) was calculated as the difference between the moment the individual stepped off the end of the GAITRite[®] and onto the M2 walkway to the time of the end of the final foot fall leaving the M2 and returning to the GAITRite[®]. The distance from the end of the GAITRite[®] to the cone was kept constant for all participants. The difference for each spatiotemporal parameter with and without a concurrent cognitive task was calculated and labeled dual task interference (e.g., the measured step length without a dual task was subtracted from the measured step length with a dual task to generate step length dual task interference). If subjects experienced a freezing episode during a walking trial, accurate spatiotemporal data could not always be obtained. For those trials, manual step identification was attempted to include as many steps as possible in each trial.

Timed data (TUG and turn duration) included the occurrence of FOG episodes when they occurred. This protocol is not designed to precipitate FOG episodes or to directly measure the duration or severity of an individual episode, but rather to describe a marker's properties to *indirectly* function as a *surrogate* of FOG severity.

Statistical analysis

Raw spatiotemporal data was removed from the GAITRite[®], into a spreadsheet for calculation of the surrogate markers of interest. Clinical data was entered into paper forms. All data was then uploaded into a RedCap database for statistical analysis. Turn duration under the dual task condition was pre-specified as the primary parameter of interest as it had been utilized effectively in a previous clinical trial for FOG (without the dual task component) [14].

Sample size estimation

Sample size was estimated based on the ability for turn duration to distinguish between severity groups. A prior study [14] found the mean duration turn task was 31s for the PD patients with FOG versus 2.7s for PD patients without FOG (SD = 25). Assuming a similar difference in groups and standard deviation when assessed under dual task, a two-sample *t*-test has 85% power when there are *n*=15 patients in each group with two-sided alpha=0.05.

Test-retest reliability was calculated for each spatiotemporal parameter for the 3–4 trials on a single visit using the intraclass correlation coefficient (ICC) reliability for the mean of *k* ratings (SAS %INTRACC macro). For each spatiotemporal parameter, the Wilcoxon rank sum test was used to compare group differences in FOG patients to PD-control. Similarly, the Wilcoxon signed-rank test was used to determine whether there were differences in response to dopaminergic medications (dopa-response) within FOG patients (tested under the ON and OFF condition, respectively). The statistical significance level was set at alpha=0.05 for all comparisons. These analyses are purely to demonstrate the measurement properties of the spatiotemporal parameters by examining the extent to which the means differ in the expected fashion using groups that are known to be different (ON-FOG, OFF-FOG, and PD-control). Results will be presented with 95% confidence intervals.

Area under the receiver operating characteristic curve (AUC) analysis was performed as a measure of responsiveness (or the ability to distinguish one group from another) for each spatiotemporal parameter. This was done by fitting a series of logistic models of PD-control versus PD-FOG as the response modeled with a separate model for each dopa-response condition (ON/OFF). Similarly, a logistic model with a random effect

for subject was fit with the ON/OFF condition as the response (PROC GLIMMIX). AUC values of 0.70 or higher are generally considered adequate to demonstrate that a measure is able to distinguish one group from another [28].

Results

Demographic and clinical descriptive data

The mean (SD) PD-control (no FOG) group ($n = 17$) was 67.3 (5.4) years of age, 5.2 (3.7) years of disease duration, with 6 females, 15 whites, one African American, and one of other race/ethnicity. The mean (SD) PD-FOG group ($n = 16$) was 64.3 (5.7) years of age, 10.2 (4.6) years of disease duration, with 5 females, all whites.

The mean UPDRS, part III (motor) scores were 24.8 (10.4) for the PD-control group in the ON condition, 24.2 (9.1) for the FOG group in the ON condition, and 42.4 (8.6) for the FOG group in the OFF condition. The UPDRS part II, item 14 FOG scores (a subjective measure of FOG severity) were 0 (0) for the PD-control group, 0.8 (0.7) for the FOG group in the ON condition, and 2.6 (0.6) for the FOG group in the OFF condition (severe FOG severity level). The mean nFOGQ score was 17.8 (5.5) in the FOG group and 0 in the PD-control group.

Test-retest reliability

Given that an AUC of > 0.70 is generally considered to be adequate [28], test-retest reliability of the spatiotemporal parameter under a single type of condition (i.e., SINGLE or DUAL) was high ($ICC > 0.90$) for all measures, except the standard deviation (SD) measures (e.g., step length standard deviation left, etc.). ICC was poor (< 0.50) for the SD measures under the SINGLE condition and fair under the dual task condition for the PD patients with

FOG in the ON state, PD patients with FOG in the OFF state, and the PD-control subjects. See Table 1.

Comparison of surrogate markers

The group means (or medians) were different for all spatiotemporal measures, with and without a dual task, between the PD-control versus PD-FOG-OFF groups and for the PD-FOG-ON versus PD-FOG-OFF condition within the FOG group. However, no differences in the means/medians were detected between the PD-FOG-ON and PD-control groups, with only trends for dual task step CV and dual task turn duration. See Table 2. The dual task interference for average step length and average stride length was significantly different between the PD-control versus PD-FOG-OFF groups, but no other group differences in the dual task interference metrics were detected.

For the area under the ROC curve (AUC) analysis, all dual task and single task spatiotemporal metrics had $AUC > 0.70$ when discriminating between PD-control vs. PD-FOG-OFF. Likewise, all dual task metrics and single task metrics had $AUC > 0.70$ when discriminating between PD patients with FOG in the ON vs. OFF condition. However, only one metric, step CV under dual task, had $AUC > 0.70$ when discriminating between PD-control vs. PD-FOG-ON. See Fig. 1. For the dual task interference metrics, very few had AUC greater than 0.70, namely average step length ($AUC = 0.76$) and average stride length ($AUC = 0.79$) when comparing control versus off and step CV ($AUC = 0.73$) when comparing control versus ON. In sum, although many markers are capable of differentiating very different groups, particularly showing responsiveness, only step CV in the dual task condition was able to differentiate groups that were

Table 1 Test-retest reliability for the gait parameters in PD-controls and PD-FOG

	PD-control		PD-FOG			
	Single	Dual	ON		OFF	
			Single	Dual	Single	Dual
ICC						
Velocity (m/s)	0.98	0.98	0.97	0.96	0.97	0.98
Step length, R (cm)	0.99	0.99	0.99	0.98	0.99	0.98
Step length, L (cm)	0.99	0.99	0.99	0.98	0.99	0.97
Step length SD, R (cm)	0.33	0.57	0.25	0.59	0.62	0.95
Step length SD, L (cm)	0.23	0.66	0.68	0.68	0.68	0.86
Stride length, R (cm)	0.99	0.99	0.99	0.98	0.99	0.99
Stride length, L (cm)	0.996	0.99	0.99	0.98	0.99	0.98
Stride length SD, R (cm)	0.29	0.53	0.16	0.77	0.68	0.97
Stride length SD, L (cm)	0.36	0.71	0.53	0.69	0.70	0.92

Table 2 Spatiotemporal data across groups

	PD-control (no FOG) (n = 17)				Freezers (FOG) (n = 16)				p-value			
	OFF				ON				Control vs OFF	Control vs ON	ON vs OFF (paired)	
	Median (IQR)	Mean difference control-OFF 95% CI	Effect size	Median (IQR)	Mean difference control-ON 95% CI	Effect size	Median (IQR)	Mean difference OFF-ON 95% CI	Effect size			
Dual task												
Velocity (m/s)	108.9 (95.7, 120.3)	39.9 (21.0, 58.8)	1.53	74.0 (52.8, 89.9)	12.1 (-3.1, 27.2)	0.57	95.0 (85.2, 114.3)	-28.2 (-38.9, -17.5)	1.46	<0.001	0.19	<0.001
Step length (R and L) (cm)	59.0 (52.4, 68.5)	20.9 (10.8, 31.1)	1.49	39.5 (29.6, 50.1)	5.6 (-3.6, 14.9)	0.43	51.9 (48.8, 58.4)	-15.6 (-21.2, -10.0)	1.55	<0.01	0.14	<0.001
Stride length (R and L) (cm)	117.9 (102.5, 136.7)	41.3 (21.0, 61.7)	1.47	79.2 (59.2, 100.2)	10.9 (-7.7, 29.6)	0.42	104.1 (96.1, 117.7)	-31.3 (-42.6, -20.0)	1.54	<0.01	0.20	<0.001
Step CV	4.5 (3.3, 5.9)	-13.1 (-26.6, 0.3)	0.77	7.4 (5.4, 15.0)	-2.1 (-4.1, 0.0)	0.81	6.0 (5.1, 8.7)	11.0 (-2.3, 24.3)	0.50	<0.01	0.08	<0.01
Stride CV	3.2 (2.4, 5.0)	-10.8 (-21.8, 0.3)	0.77	5.3 (4.2, 13.1)	-1.4 (-3.2, 0.5)	0.59	3.9 (3.2, 5.6)	9.4 (-1.3, 20.2)	0.53	0.01	0.25	<0.01
Time to turn (s)	5.0 (4.3, 5.7)	-13.6 (-25.4, -1.8)	0.83	8.2 (6.2, 19.8)	-1.2 (-2.4, 0.0)	0.71	5.5 (5.0, 6.5)	12.4 (-0.3, 25.0)	0.54	<0.01	0.08	<0.01
Single task												
Velocity (m/s)	128.2 (114.0, 137.7)	36.3 (16.3, 56.4)	1.29	94.5 (67.4, 106.9)	12.8 (-2.8, 28.5)	0.58	111.2 (100.7, 131.8)	-23.5 (-35.0, -12.0)	1.09	<0.01	0.10	<0.001
Step length (R and L) (cm)	67.4 (58.6, 73.9)	16.2 (5.3, 27.1)	1.06	51.3 (39.1, 57.9)	4.6 (-4.9, 14.0)	0.34	61.2 (54.6, 64.8)	-11.6 (-16.0, -7.3)	1.43	<0.01	0.40	<0.001
Stride length R and L (cm)	135.3 (117.5, 148.1)	31.3 (9.6, 53.1)	1.02	103.0 (78.4, 116.1)	8.5 (-10.2, 27.1)	0.32	122.7 (109.4, 130.3)	-22.9 (-31.8, -14.0)	1.37	<0.01	0.35	<0.001
Step CV	3.9 (2.8, 4.2)	-8.1 (-16.5, 0.4)	0.76	6.4 (5.4, 8.6)	-0.8 (-2.1, 0.5)	0.46	4.0 (3.7, 4.7)	7.3 (-0.9, 15.5)	0.54	<0.01	0.33	<0.001
Stride CV	2.8 (2.1, 3.3)	-7.2 (-15.1, 0.7)	0.73	5.0 (3.7, 7.8)	-0.6 (-1.6, 0.4)	0.50	3.1 (2.7, 3.6)	6.6 (-1.3, 14.5)	0.51	<0.01	0.26	<0.001
Time to turn (s)	4.4 (3.6, 4.8)	-6.9 (-12.5, -1.3)	0.90	6.1 (4.8, 9.9)	-0.9 (-2.0, 0.3)	0.54	4.9 (4.2, 5.4)	6.0 (0.6, 11.5)	0.61	<0.01	0.10	<0.01
Interference												
Velocity (m/s)	-15.3 (-20.1, -8.9)	5.2 (-6.4, 16.9)	0.32	-14.8 (-36.9, -10.1)	-0.8 (-8.1, 6.6)	0.07	-13.2 (-24.5, -8.1)	-5.3 (-14.6, 4.1)	0.31	0.74	0.84	0.45
Step length (R and L) (cm)	-5.4 (-7.1, -3.8)	5.9 (1.3, 10.6)	0.93	-8.6 (-21.5, -6.3)	1.0 (-2.1, 4.2)	0.24	-7.2 (-10.9, -4.1)	-4.4 (-8.4, -0.3)	0.60	<0.05	0.26	0.06

Table 2 (continued)

	PD-control (no FOG) (n = 17)				Freezers (FOG) (n = 16)				p-value			
	OFF		ON		OFF		ON		Control vs OFF	Control vs ON		
	Median (IQR)	Mean difference control-OFF 95% CI	Effect size	Median (IQR)	Mean difference control-ON 95% CI	Effect size	Median (IQR)	Mean difference OFF-ON 95% CI	Effect size	ON vs OFF (paired)		
Stride length Rand L (cm)	-11.0 (-14.1, -8.9)	12.7 (3.8, 21.7)	1.03	-17.3 (-43.6, -12.8)	2.5 (-3.3, 8.3)	0.30	-14.0 (-21.4, -7.1)	-9.2 (-17.6, -0.9)	0.61	<0.01	0.28	0.06
Step CV	0.7 (-0.6, 2.7)	-5.0 (-14.1, 4.0)	0.44	0.6 (-0.1, 1.6)	-1.3 (-2.7, 0.0)	0.79	2.1 (1.2, 3.2)	3.7 (-5.8, 13.2)	0.24	0.65	0.06	0.74
Stride CV	0.9 (-0.9, 1.6)	-3.6 (-10.5, 3.4)	0.41	0.7 (-0.3, 1.2)	-0.7 (-1.9, 0.5)	0.48	1.1 (0.4, 1.9)	2.8 (-4.5, 10.2)	0.23	0.94	0.35	0.95
Time to turn (s)	0.6 (0.2, 0.9)	-6.7 (-13.6, 0.2)	0.70	2.1 (0.3, 5.6)	-0.3 (-1.0, 0.3)	0.36	0.5 (0.2, 1.3)	6.3 (-1.6, 14.2)	0.44	0.08	0.58	0.17

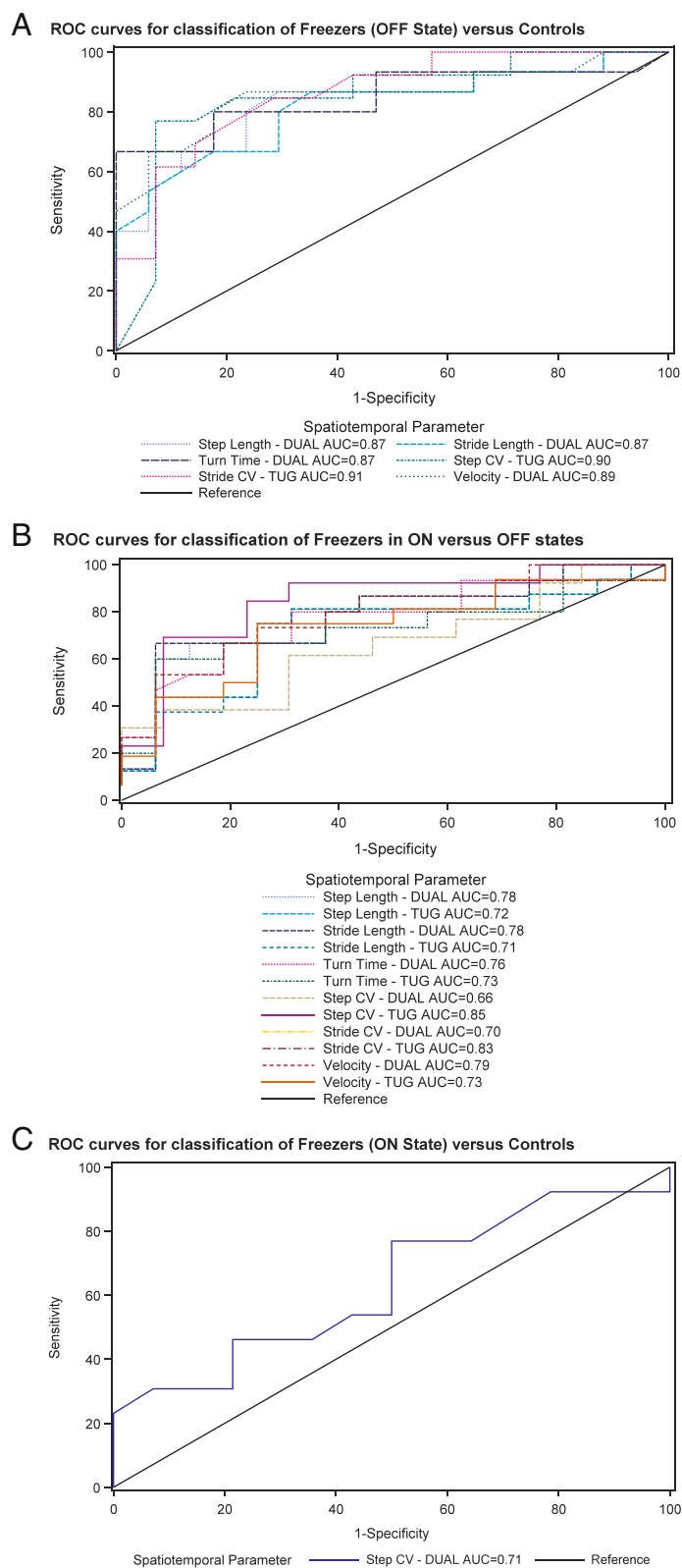


Fig. 1 Graphical representation of the area under the receiving operating characteristics (AUC) analysis, showing sensitivity and specificity for each surrogate marker comparing PD-control (no FOG) vs. ON-FOG (A), PD-control vs. OFF-FOG (B), and ON-FOG vs. OFF-FOG (C)

very similar except for the presence of FOG, demonstrating specificity for FOG.

Discussion

We report our findings on direct comparisons of commonly used outcome measures in FOG clinical trials. The study was designed to determine the feasibility of the use of objective markers of FOG severity in clinical trials of interventions for the treatment of FOG. The feasibility objectives include (1) the specificity of each marker for FOG and (2) the responsiveness of each marker to an intervention. In addition, we investigated whether adding a dual task or calculating dual task interference changed the biometric properties of each marker or should be considered as a separate marker. The goal of our study was to provide objective data regarding the utility of each of these markers for clinical trials or behavioral association studies in order to assist investigators in choosing the appropriate marker for the scientific question being asked. The findings of our study can inform future clinical trials investigating the effectiveness of novel interventions for FOG and can help interpret previous trials that have reported changes in these surrogate markers.

All of the surrogate markers studied were able to differentiate between ON and OFF indicating the responsiveness to levodopa challenge with and without a dual task. However, only step CV in the dual task was able to distinguish between the PD-control group and the FOG group when ON medications ($AUC > 0.70$). These were two very similar groups (with very similar UPDRS scores) who only differed by the fact that the FOG group had the underlying propensity for FOG behavior when in the OFF state. These findings imply that the remaining markers are not specific for FOG; however, the rigorous design of this study comparing very similar groups should be taken into account when interpreting this finding. Therefore, most of these markers may be used in clinical trials to study the magnitude of response to an intervention, however, may not to represent a change in FOG severity independently of other gait factors. Turn duration in the dual task condition showed a strong trend toward significance when comparing the ON-FOG group and the PD-control. Therefore, dual task turn duration should not be ruled out as a proxy for FOG severity in cross-sectional studies or imaging-behavioral associations investigating the relationship of a specific finding to FOG, or as an outcome in clinical trials of a therapeutic intervention. Similar markers like stride time variability have been shown to correlate with overall disease severity [27] and have also been shown to be greater in patients with PD and FOG as compared to PD alone [29, 30].

Patients with PD have been shown to have decreased automaticity of motor tasks [31] and of gait [32, 33].

Automatic gait generates effective stepping with little or no variability; therefore, it is not surprising that there is increased variability of step length in patients with FOG compared to those without FOG. Increasing cognitive load challenges automatic gait by diverting cortical control from gait. If there is a deficit in gait automaticity, greater declines in gait would be expected as a cognitive load is added. In fact, dual task interference is considered a marker for gait automaticity [34]. We interpret the finding that step CV in the dual task was the most specific marker of FOG severity to be related to a loss of automaticity of gait in FOG.

Turn duration is a very simple metric to obtain and has been utilized effectively in clinical trials for FOG in the past [35]. Our finding that adding the dual task to multiple surrogate markers improves the biometric properties of the marker informs this and future studies when selecting markers of this condition. Curtze et al. found that turning measurements were the strongest correlates of disease severity as measured by the UPDRS, in a large PD cohort with similar disease duration, although this study did not look at FOG [36]. It is important to note that although some patients may experience a FOG episode during turning (particularly the severe FOG level), this setup (using a large turning space and a cone) is designed to minimize — not precipitate — a FOG episode, and each parameter's value is an average of at least two trials in each condition. Therefore, these results are independent of whether or not a FOG episode is triggered and differ from studies of the turn condition designed to trigger a freezing episode and then quantify each episode individually. By understanding the biometric properties of markers of FOG severity that do not depend on eliciting a FOG episode, we can remove the inherent variability of the episode, presumably allowing a more consistent and representative assessment of FOG severity. Furthermore, such a marker is inherently simple to capture and can be repeated in one session, making it ideal for same day dose-finding studies or early-stage neuromodulation clinical trials. However, this comes at the cost of specificity for FOG, for most of the parameters derived from this approach.

Study limitations include our inability to determine which condition (ON or OFF) best indicates severity, since we were comparing each marker in the ON and OFF states. However, other studies have assessed turn measurements and have found the OFF condition to be superior [36]. We were powered to determine a difference between PD-controls and PD-FOG, but not between ON and OFF FOG, or ON FOG and PD-controls. Small sample size is also a limitation and should be taken into consideration when interpreting p -values, especially trends. Therefore, non-significant differences

or strong trends should not be discarded. Also, due to the design, we could not compare each marker's ability to differentiate between severity levels with the nFOGQ. This is due to the fact that retrospective subjective questionnaires, when administered, provide an overall assessment of severity over a period of time (usually weeks) and cannot be administered reliably to predict severity in the ON and OFF states. There was a small difference in age between the control and FOG groups (67.2 years for the control and 64.3 years for the FOG group) and a significant difference in disease duration (5.2 years control, 10.2 years FOG group). The disease duration difference is to be expected as FOG occurs later in the disease course. Finally, this is not a validation study of any one surrogate marker, but our findings help identify the most appropriate markers to answer future scientific questions or to be used in clinical trials and should lead to future validation studies of such.

Based on the findings of this comparative study of surrogate markers of FOG severity, we conclude that (1) objective gait assessment is a feasible outcome measure in clinical trials and behavioral association studies for FOG, (2) dual task turn duration and dual task step CV are most specific for FOG of the markers compared, and (3) velocity, step/stride length, and dual task turn duration are responsive to levodopa challenge. Further validation studies of these surrogate markers are warranted for their use as outcome measures in clinical trials. As more markers become available such as continuous monitoring, novel approaches to capture FOG, or more nuanced calculations of dual task interference (e.g., dual task effect [37]), studies to validate and compare them can help guide their use in future clinical trials.

Abbreviations

PD: Parkinson's disease; FOG: Freezing of gait; TUG: Timed up and go; UPDRS: Unified Parkinson's Disease Rating Scale; ROC: Receiver operating characteristic; SD: Standard deviation; CV: Coefficient of variability; AUC: Area under the curve.

Acknowledgements

Not applicable.

Authors' contributions

Dr. Revuelta: research project: conception, organization, execution; statistical analysis: design, review, and critique; manuscript: execution, review, and critique. Dr. Embry: research project: organization, execution. Dr. Elm: statistical analysis: execution, review, and critique; manuscript: review and critique. Ms. Jenkins: research project: organization, execution. Dr. Lee: research project: organization, execution. Dr. Kautz: research project: conception, organization; statistical analysis: design, review, and critique; manuscript: review and critique. The authors read and approved the final manuscript.

Funding

This study was supported by Dr. Revuelta's funding which included Barmore Fund for Parkinson's Research, the MUSC Departments of Neurology and Neurosurgery, by the South Carolina Clinical & Translational Research (SCTR) Institute, with an academic home at the Medical University of South

Carolina, supported by NIH/NCATS Grant Number UL1TR000062, and by NIH NINDS Grant Number 1K23NS091391-01A1. Dr. Kautz was also supported by P20GM109040. Dr. Elm and Dr. Embry also had NIH support. Mrs. Jenkins and Dr. Lee had nothing to report. None of the authors have any competing interests in the manuscript.

Availability of data and materials

The authors believe the data necessary for analysis and interpretation is provided in the manuscript; however, any further data found to be necessary can be made available by request.

Declarations

Ethics approval and consent to participate

The institutional review board of the Medical University of South Carolina approved the study.

Consent for publication

Written patient consent was obtained and documented on all patients. This manuscript does not report individual data.

Competing interests

Dr. Revuelta serves on the advisory board for Boston Scientific. The remaining authors declare that there are no additional disclosures to report. This material was presented at the 4th International Workshop on Freezing of Gait, Belgium, 2018.

Author details

¹Movement Disorders Division, Department of Neurology, College of Medicine, Medical University of South Carolina, 208B Rutledge Avenue, MSC 108, Charleston, SC 29425, USA. ²Ralph H. Johnson VA Medical Center, Charleston, SC, USA. ³Center for Rehabilitation Research in Neurological Conditions, Department of Health Sciences and Research, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA. ⁴Department of Public Health Sciences, College of Medicine, Medical University of South Carolina, Charleston, SC, USA. ⁵Department of Neurology, Virginia Commonwealth University, Richmond, VA, USA.

Received: 19 January 2021 Accepted: 15 June 2022

Published online: 04 July 2022

References

1. Forsaa EB, Larsen JP, Wentzel-Larsen T, Alves G. A 12-year population-based study of freezing of gait in Parkinson's disease. *Parkinsonism Relat Disord.* 2015;21(3):254–8. <https://doi.org/10.1016/j.parkreldis.2014.12.020>.
2. Perez-Lloret S, Negre-Pages L, Damier P, Delval A, Derkinderen P, Destee A, et al. Prevalence, determinants, and effect on quality of life of freezing of gait in Parkinson disease. *JAMA Neurol.* 2014;71(7):884–90. <https://doi.org/10.1001/jamaneurol.2014.753>.
3. Nieuwboer A, Giladi N. Characterizing freezing of gait in Parkinson's disease: models of an episodic phenomenon. *Mov Disord.* 2013;28(11):1509–19. <https://doi.org/10.1002/mds.25683>.
4. Nutt JG, Bloem BR, Giladi N, Hallett M, Horak FB, Nieuwboer A. Freezing of gait: moving forward on a mysterious clinical phenomenon. *Lancet Neurol.* 2011;10(8):734–44. [https://doi.org/10.1016/S1474-4422\(11\)70143-0](https://doi.org/10.1016/S1474-4422(11)70143-0).
5. Thevathasan W, Debu B, Aziz T, Bloem BR, Blahak C, Butson C, et al. Pedunculopontine nucleus deep brain stimulation in Parkinson's disease: a clinical review. *Mov Disord.* 2018;33(1):10–20. <https://doi.org/10.1002/mds.27098>.
6. Abnoosian A, Maguire G. Case report of an interaction of a vagal nerve stimulation system with a microwave current from a body fat analyzer. *Ann Clin Psychiatry.* 2008;20(4):229–30. <https://doi.org/10.1080/10401230802435542>.
7. Hulzinga F, Nieuwboer A, Dijkstra BW, Mancini M, Strouwen C, Bloem BR, et al. The new freezing of gait questionnaire: unsuitable as an outcome in clinical trials? *Mov Disord Clin Pract.* 2020;7(2):199–205. <https://doi.org/10.1002/mdc3.12893>.

8. Pham TT, Moore ST, Lewis SJG, Nguyen DN, Dutkiewicz E, Fuglevand AJ, et al. Freezing of gait detection in Parkinson's disease: a subject-independent detector using anomaly scores. *IEEE Trans Biomed Eng.* 2017;64(11):2719–28. <https://doi.org/10.1109/TBME.2017.2665438>.
9. Mancini M, Smulders K, Cohen RG, Horak FB, Giladi N, Nutt JG. The clinical significance of freezing while turning in Parkinson's disease. *Neuroscience.* 2017;343:222–8. <https://doi.org/10.1016/j.neuroscience.2016.11.045>.
10. Delval A, Snijders AH, Weerdesteyn V, Duysens JE, Defebvre L, Giladi N, et al. Objective detection of subtle freezing of gait episodes in Parkinson's disease. *Mov Disord.* 2010;25(11):1684–93. <https://doi.org/10.1002/mds.23159>.
11. Ziegler K, Schroeteler F, Ceballos-Baumann AO, Fietzek UM. A new rating instrument to assess festination and freezing gait in Parkinsonian patients. *Mov Disord.* 2010;25(8):1012–8. <https://doi.org/10.1002/mds.22993>.
12. Rodriguez-Martin D, Sama A, Perez-Lopez C, Catala A, Moreno Arostegui JM, Cabestany J, et al. Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer. *PLoS ONE.* 2017;12(2):e0171764. <https://doi.org/10.1371/journal.pone.0171764>.
13. Bachlin M, Plotnik M, Roggen D, Maidan I, Hausdorff JM, Giladi N, et al. Wearable assistant for Parkinson's disease patients with the freezing of gait symptom. *IEEE Trans Inf Technol Biomed.* 2010;14(2):436–46. <https://doi.org/10.1109/ITTB.2009.2036165>.
14. Thevathasan W, Cole MH, Graepel CL, Hyam JA, Jenkinson N, Brittain JS, et al. A spatiotemporal analysis of gait freezing and the impact of pedunculopontine nucleus stimulation. *Brain.* 2012;135(Pt 5):1446–54. <https://doi.org/10.1093/brain/awt039>.
15. Espay AJ, Dwivedi AK, Payne M, Gaines L, Vaughan JE, Maddux BN, et al. Methylphenidate for gait impairment in Parkinson disease: a randomized clinical trial. *Neurology.* 2011;76(14):1256–62. <https://doi.org/10.1212/WNL.0b013e3182143537>.
16. Moro E, Hamani C, Poon YY, Al-Khairallah T, Dostrovsky JO, Hutchison WD, et al. Unilateral pedunculopontine stimulation improves falls in Parkinson's disease. *Brain.* 2010;133(Pt 1):215–24. <https://doi.org/10.1093/brain/awp261>.
17. Mi TM, Garg S, Ba F, Liu AP, Wu T, Gao LL, et al. High-frequency rTMS over the supplementary motor area improves freezing of gait in Parkinson's disease: a randomized controlled trial. *Parkinsonism Relat Disord.* 2019;68:85–90. <https://doi.org/10.1016/j.parkreldis.2019.10.009>.
18. Nanhoe-Mahabier W, Snijders AH, Delval A, Weerdesteyn V, Duysens J, Overeem S, et al. Walking patterns in Parkinson's disease with and without freezing of gait. *Neuroscience.* 2011;182:217–24. <https://doi.org/10.1016/j.neuroscience.2011.02.061>.
19. Bertoli M, Croce UD, Cereatti A, Mancini M. Objective measures to investigate turning impairments and freezing of gait in people with Parkinson's disease. *Gait Posture.* 2019;74:187–93. <https://doi.org/10.1016/j.gaitpost.2019.09.001>.
20. Spildooren J, Vinken C, Van Baekel L, Nieuwboer A. Turning problems and freezing of gait in Parkinson's disease: a systematic review and meta-analysis. *Disabil Rehabil.* 2018;1-11. <https://doi.org/10.1080/09638288.2018.1483429>.
21. Mancini M, Weiss A, Herman T, Hausdorff JM. Turn around freezing: community-living turning behavior in people with Parkinson's disease. *Front Neurol.* 2018;9:18. <https://doi.org/10.3389/fneur.2018.00018>.
22. Goetz CG, Poewe W, Rascol O, Sampaio C, Stebbins GT, Counsell C, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord.* 2004;19(9):1020–8. <https://doi.org/10.1002/mds.20213>.
23. Nieuwboer A, Rochester L, Herman T, Vandenberghe W, Emil GE, Thomaes T, et al. Reliability of the new freezing of gait questionnaire: agreement between patients with Parkinson's disease and their carers. *Gait Posture.* 2009;30(4):459–63. <https://doi.org/10.1016/j.gaitpost.2009.07.108>.
24. Ehoetz Martens KA, Shine JM, Walton CC, Georgiades MJ, Gilat M, Hall JM, et al. Evidence for subtypes of freezing of gait in Parkinson's disease. *Mov Disord.* 2018;33(7):1174–8. <https://doi.org/10.1002/mds.27417>.
25. Bloem BR, Grimbergen YA, van Dijk JG, Munneke M. The "posture second" strategy: a review of wrong priorities in Parkinson's disease. *J Neurol Sci.* 2006;248(1–2):196–204. <https://doi.org/10.1016/j.jns.2006.05.010>.
26. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc.* 1991;39(2):142–8. <https://doi.org/10.1111/j.1532-5415.1991.tb01616.x>.
27. Hausdorff JM, Cudkowicz ME, Firtion R, Wei JY, Goldberger AL. Gait variability and basal ganglia disorders: stride-to-stride variations of gait cycle timing in Parkinson's disease and Huntington's disease. *Mov Disord.* 1998;13(3):428–37. <https://doi.org/10.1002/mds.870130310>.
28. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>.
29. Hausdorff JM, Schaafsma JD, Balash Y, Bartels AL, Gurevich T, Giladi N. Impaired regulation of stride variability in Parkinson's disease subjects with freezing of gait. *Exp Brain Res.* 2003;149(2):187–94. <https://doi.org/10.1007/s00221-002-1354-8>.
30. Barbe MT, Amarell M, Snijders AH, Florin E, Quatuor EL, Schonau E, et al. Gait and upper limb variability in Parkinson's disease patients with and without freezing of gait. *J Neurol.* 2014;261(2):330–42. <https://doi.org/10.1007/s00415-013-7199-1>.
31. Wu T, Hallett M, Chan P. Motor automaticity in Parkinson's disease. *Neurobiol Dis.* 2015;82:226–34. <https://doi.org/10.1016/j.nbd.2015.06.014>.
32. Hallett M. The intrinsic and extrinsic aspects of freezing of gait. *Mov Disord.* 2008;23(Suppl 2):S439–43. <https://doi.org/10.1002/mds.21836>.
33. Snijders AH, Takakusaki K, Debu B, Lozano AM, Krishna V, Fasano A, et al. Physiology of freezing of gait. *Ann Neurol.* 2016;80(5):644–59. <https://doi.org/10.1002/ana.24778>.
34. Clark DJ. Automaticity of walking: functional significance, mechanisms, measurement and rehabilitation strategies. *Front Hum Neurosci.* 2015;9:246. <https://doi.org/10.3389/fnhum.2015.00246>.
35. Thevathasan W, Coyne TJ, Hyam JA, Kerr G, Jenkinson N, Aziz TZ, et al. Pedunculopontine nucleus stimulation improves gait freezing in Parkinson's disease. *Neurosurgery.* 2011. <https://doi.org/10.1227/NEU.0b013e31822b6f71>.
36. Curtze C, Nutt JG, Carlson-Kuhta P, Mancini M, Horak FB. Objective gait and balance impairments relate to balance confidence and perceived mobility in people with Parkinson disease. *Phys Ther.* 2016;96(11):1734–43. <https://doi.org/10.2522/ptj.20150662>.
37. Plummer P, Eskes G. Measuring treatment effects on dual-task performance: a framework for research and clinical practice. *Front Hum Neurosci.* 2015;9:225. <https://doi.org/10.3389/fnhum.2015.00225>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

