Exploring Responsiveness to Highly Challenging Balance and Gait Training in Parkinson's Disease

Franziska Albrecht, PhD,[1,2,*] Hanna Johansson, RPT, PhD,[1,2,3] Konstantinos Poulakis, PhD,[4] Eric Westman, PhD,[4] Maria Hagströmer, RPT, PhD,[1,5] and Erika Franzén, RPT, PhD[1,2,3]

Abstract: Background: Exercise potentially improves gait, balance, and habitual physical activity in Parkinson's disease (PD). However, given the heterogeneous nature of the disease, it is likely that people respond differently to exercise interventions. Factors determining responsiveness to exercise interventions remain unclear.

Objectives: To address this uncertainty, we explored the responsiveness to our highly challenging balance and gait intervention (HiBalance) in people with PD.

Methods: Thirty-nine participants with mild–moderate PD who underwent the HiBalance intervention from our randomized controlled trial were included. We defined response in three domains: (1) balance based on Mini-BESTest, (2) gait based on gait velocity, and (3) physical activity based on accelerometry-derived steps per day. In each domain, we explored three responsiveness levels: high, low, or non-responders according to the change from pre- to post-intervention. Separate Random Forests for each responder domain classified these responsiveness levels and identified variable importance.

Results: Only the Random Forest for the balance domain classified all responsiveness levels above the chance level indicated by a Cohen's kappa of "slight" agreement. Variable importance differed among the responsiveness levels. Slow gait velocity indicated high responders in the balance domain but showed low probabilities for low and non-responders. For low and non-responders, fall history or no falls, respectively, were more important.

Conclusions: Among three responder domains and responsiveness levels, we could moderately classify responders in the balance domain, but not for the gait or physical activity domain. This can guide inclusion criteria for balance-targeted, personalized intervention studies in people with PD.

Parkinson's disease (PD) is a complex and diverse neurodegenerative disease that adversely affects balance, gait, and cognition. Evidence is growing that exercise interventions potentially promote improvements in gait and balance performance in people with PD.[1–3] Exercise drives changes within the neural system, which may contribute to improved functional outcomes.[4] Recent meta-analyses concluded that exercise interventions—resistance training, aerobic exercise, treadmill training, and dance—improved motor function, balance, and physical capacity in people with PD.[2,5]

Our group has earlier shown that highly challenging balance training—the HiBalance intervention—is effective in improving balance, gait, and habitual physical activity in people with PD.[1,6,7] The HiBalance intervention is based on motor learning principles (ie, intensity, specificity, and practice variation), and

incorporates both motor and cognitive dual-tasks. Dual-tasking involves the simultaneous performance of two tasks which can be either motor, cognitive or a combination of both.[8] The intervention was also designed to facilitate effects through exercise-induced neuroplasticity and by progressively increasing the exercise's difficulty. However, in our most recent randomized controlled trial (RCT) (EXercise in Parkinson's disease and Neuroplasticity [EXPANd] trial) with blinded assessors, an active control group, and where we reduced the group training sessions to twice a week, we were unable to find significant changes in gait, balance, or physical activity on a group level.[9] Reasons for this discrepancy might be the added magnetic resonance imaging exclusion criteria and extended measurement battery that led to the inclusion of individuals with milder motor symptoms and fewer balance problems. Further, due to the heterogeneity of PD it is unlikely that all individuals respond in the same way to exercise interventions. Thus, we might not see beneficial effects on a group level of the HiBalance intervention, but this does not mean that no one benefited on an individual level. Responder analyses after exercise interventions are scarce and could contribute to interventions that are customized for individual needs. Moreover, the variables used, responder cut-offs, and analysis methods differ between studies.[10–13]

In one of our former RCTs, we identified by linear regression that low self-perceived general health 36-Item Short Form Health Survey (SF-36),[14] less functional mobility (Timed Up and Go, TUG), and greater number of errors on a cognitive task predicted higher responsiveness in balance performance mini balance evaluation systems test (Mini-BESTest).[10] On the other hand, in a following implementation study, we employed logistic regression and found out that lower balance confidence (Activities-specific Balance Confidence scale, ABC) and higher attention rate led to increased balance performance (Mini-BESTest).[12]

Besides those studies, there is a considerable gap in the literature about what determines responsiveness to balance and gait interventions for people with PD.

To fill this gap, we identified variables that indicate HiBalance intervention response in people with PD. We explored three response domains (balance, gait velocity, and physical activity) and three responsiveness levels (high, low, and non-responders) to find variables that characterize specific responses. We aimed at describing even low responders to capture people that at least did not worsen in their progressive disease. The selected variables may be important when screening people with PD for future balance and gait interventions.

Methods

Participants

We included 39 people with PD of our RCT (EXPANd trial) that underwent the balance and gait intervention–HiBalance.[9,15,16] Inclusion criteria were mild–moderate idiopathic PD, Hoehn &

Yahr 2–3, age ≥60 years, Montreal Cognitive Assessment (MoCA)[17] score ≥21, and ≤27 Mini-BESTest.[18] The HiBalance intervention is delivered in a group setting, consisting of two 60-min sessions per week for 10 weeks. Additionally, participants are instructed to perform a once-weekly home exercise program during the training period.

Standard Protocol Approvals, Registrations, and Patient Consents

The RCT was approved by the Regional Ethical Review Board Stockholm (2016/1264–31/4, 2017/1258–32, 2017/2445–32). Participants received written and oral information and provided written informed consent.

Data

We selected 15 variables of the RCT to be relevant for gait and balance intervention responsiveness with the reasoning that these are easy to collect in a clinical setting before an intervention and yield clinical relevance. The analysis included pre-intervention variables of (details, see study protocol[16]): Age, Levodopa equivalent daily dosage (LEDD), balance performance (Mini-BESTest), dual-tasking ability (Timed Up and Go test with a cognitive dual-task, TUG Cog), anxiety and depression (Hospital Anxiety and Depression scales, HADS), balance confidence (ABC), self-reported walking ability (Walking Impact Scale, Walk 12), motor severity (Movement Disorder Society-sponsored Revision of the Unified Parkinson's Disease Rating Scale, MDS-UPDRS III), number of falls within 6 months, and habitual physical activity measured as steps per day (GT3X+, ActiGraph, worn on the hip for 7 days, see Supplement). Gait velocity was measured using an electronic walkway system (GAITRite®, active zone: 8.3 m), with a three-meter acceleration and deceleration distance on each side. Participants were instructed to walk at their usual speed. A total of eight trials on the walkway were performed, whereof the first two were considered practice runs and removed from analysis. The TUG Cog is the time in seconds that is needed to rise from a chair, walk 3 m at normal gait velocity, turn around, walk back, and sit down while performing a serial subtraction task (item 14 of the Mini-BESTest).

Moreover, three calculated variables were added to the Random Forest classifiers (see Section Random Forest Supervised Classification Analysis): cognitive subtype (cognitively normal and mild cognitive impairment[19]), motor subtype (tremor-dominant, postural instability/gait difficulty, and indeterminate[20]), and physical activity subtype (Sedentary, Light Movers, and Steady Movers[21]). The Sedentary subtype spent most time in sedentary behavior and the minority in light or moderate or higher intensity physical activity. The Light Movers engaged equally in sedentary and physical activity behavior. The Steady Movers were mostly active, either with light or moderate to high intensity, while spending less time in sedentary behavior. Those subtypes could provide clinically meaningful knowledge regarding how

physical activity levels are related to gait and balance interventions.

Ten participants had missing values (1.1% out of all variables), which we imputed with Random Forest regression (R package "missForest"[22]) (see Supplement).

Responder Classification

We calculated HiBalance intervention response by subtracting post- from pre-intervention values, ie, delta values. We defined responders in three different domains: (1) balance performance as measured by the Mini-BESTest, (2) gait velocity as measured by an electronic walkway, and (3) physical activity as measured by steps per day during 7 days using an accelerometer.

We further defined three levels of intervention responsiveness: High responders based on literature cut-off values,[23–25] exploratory low responders, and non-responders. Our reasoning for a low responder level was to explore the factors for people who could not much improve since they already had high values before the intervention and did not have much range to improve left. Regarding balance domain, out of 39 people with PD, 11 were classified as high responders (increase ≥3 points) in balance performance (Fig. 1, left panel). Of the remaining, 12 were low responders (increase of 1–2 points), and 16 were non-responders (no response or decrease). In the gait domain, 11 were high responders (increase ≥0.14 m/s), 11 were low responders (0.139–0.04 m/s), and 17 were non-responders (≤0.039 m/s) (Fig. 1, middle panel). Concerning the physical activity domain, 13 were high responders (increase ≥500 steps per day), three low responders (increased between 499 and 1), and 23 non-responders (no increase or decrease) (Fig. 1, right panel). We excluded low responders for the physical activity domain due to the low cohort size.

We compared demographic, clinical, and motor variables between the responsiveness levels using Kruskal-Wallis or chi-squared tests (R package "arsenal"[26]).

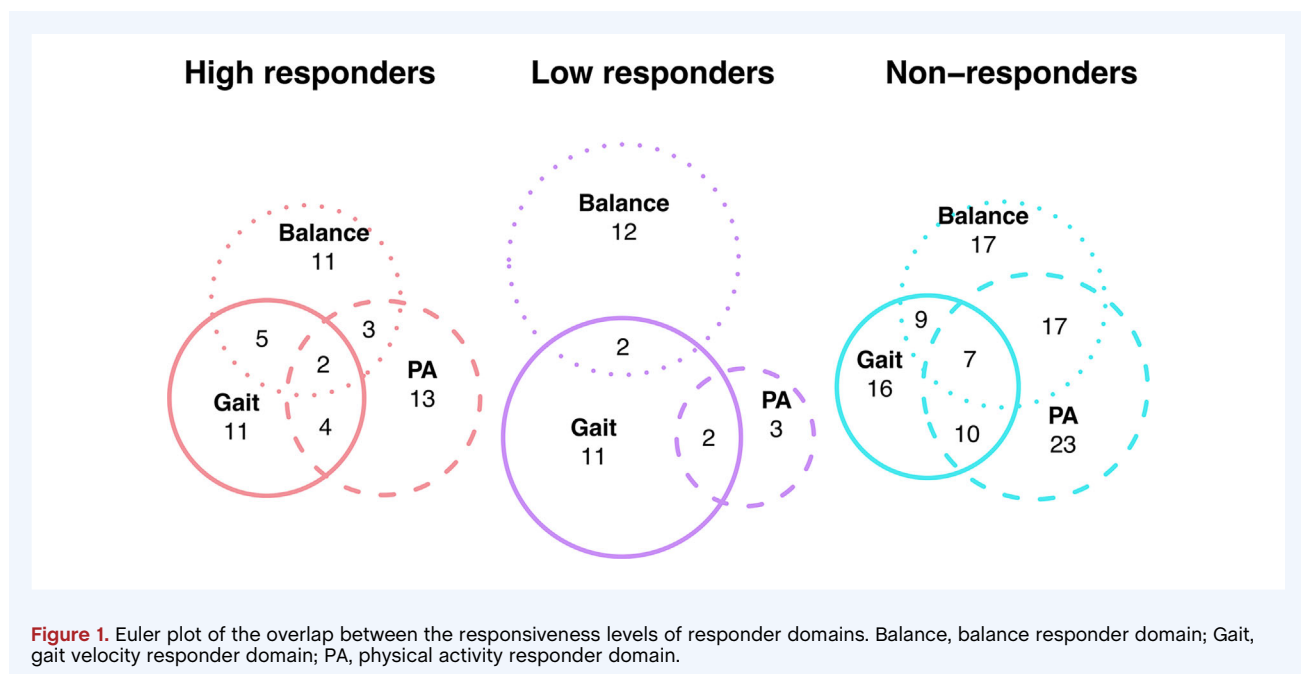## Random Forest Supervised Classification Analysis

Random Forest

The Random Forest aggregates many decision trees to predict a classification.[27] Each decision tree is a supervised classifier built on a randomly selected training set of the original dataset, ie, bagged sample. The data that is not used to build the decision tree will be used as a classification test set, ie, Out-of-bag (OOB) sample. At each node-split in the decision tree, only a restricted number of randomly selected variables are considered as possible splitting variables. The final classification is found by aggregating the classifications provided by each tree as a majority vote. The Random Forest can find a classification rule in a supervised setting and find variables that are important for the classification. Random Forests can cope with high dimensional data (ie, more variables than observations) and complex interaction structures between variables.

Analysis

Each Random Forest was built using 20,000 trees and four possible variables for each split. A node size of seven was chosen for gait responders and eight for balance and step responders Random Forests. Otherwise, default settings were used (R package "randomForest", version 4.7-1.1[28]).

We first included the pre-intervention variable of the responders' domain variable in the Random Forests, eg, Mini-

Figure 1. Euler plot of the overlap between the responsiveness levels of responder domains. Balance, balance responder domain; Gait, gait velocity responder domain; PA, physical activity responder domain.

BESTest in the balance responders classification. The respective pre-intervention variables were always ranked as the most important according to mean accuracy decrease. As this seemed circular and yielded no clinical value, we excluded the pre-values when building the respective Random Forests. For example, the balance responders' Random Forest was built without Mini-BESTest. Each Random Forest classifier was built on 17 variables.

Variable Importance

Random Forest produces measures of ranking the variables in their classification relevance. The mean decrease in accuracy—a permutation-based measure—was chosen[29] to extract interpretable information since it is widely accepted[30] (more details, see Supplement). A decrease in accuracy underlines a strong influence of the variable on the classification, whereas a permutation of a non-influential variable leads to no or less accuracy decrease.[31] Negative values in variable importance mean that the variables do not contribute to the classification.

Partial Dependence

To inspect the influence of these important variables on responsiveness classification, we used partial dependence plots. The plots visualize the variable influence with respect to all the other variables included in the Random Forest.[29] The variable of interest is plotted as the predicted probability of the observation belonging to a responsiveness level.

Classification Quality Assessment

To understand if the responsiveness classification results are interpretable, we calculated the analysis-specific chance levels for each responder domain and each responsiveness level (ie, the number of responders divided by the number of the complete cohort = 39). For the balance domain responders, the chance level was non-responders = 16/39 = 41%, low responders = 31%, and high responders = 28%. For the gait domain responders, chance levels were non-responders = 44%, low responders = 28%, and high responders = 28%. Regarding the step domain responders, chance levels were non-responders = 64% and high responders = 36%. We compared these chance levels with the classification error of the respective Random Forests.

The OOB estimate of the error rate is derived from predictions on the OOB samples, when comparing it to the real class label. A lower OOB error indicates a better prediction score.[32]

Additionally, we calculated Cohen's Kappa using the confusion matrix of the Random Forest (R package "psych"[33]). The unweighted Kappa only considers agreements, ie, a match between the Random Forest classification and the true responsiveness label (diagonal of the confusion matrix, Table 2). Disagreements were considered for the weighted kappa (off-diagonal of the confusion matrix, Table 2). A Kappa of −1 indicates total disagreement, while 0 indicates a random classification.[34] Values between 0.01–0.20 indicate none to slight

agreement, 0.21–0.40 fair agreement, 0.41–0.6 moderate agreement, 0.61–0.80 substantial agreement, and 0.81–1.00 nearly perfect agreement.

Results

Balance Responders

Responsiveness levels differed significantly in physical activity subtype ($P = 0.045$), balance performance (Mini-BESTest, $P = 0.007$), fall history ($P = 0.023$), and marginally in gait velocity ($P = 0.053$) (Kruskal-Wallis test, Table 1). High responders had the worst balance, most falls, and slowest gait velocity, and three out of five of the Sedentary subtype were high responders. Non-responders had the best balance, fewest falls, a gait velocity faster than high responders but slower than low responders, and most of the Steady Movers subtype were non-responders.

The OOB estimate of the error rate for the Random Forest was 59%. A Cohen's weighted kappa of 0.15 summarized only a slight agreement between the true responsiveness levels and the Random Forest classification (Table S1). The confusion matrix indicated that the classification of responsiveness levels exceeded the calculated chance level (Table 2).

Feature importance derived from the Random Forest showed that the number of falls, gait velocity, and dual-task ability (TUG Cog) had the highest importance for responsiveness classification (Table S2, Fig. 2A). Gait velocity and self-perceived walking ability (Walk 12) could classify high responders. Fall history followed by gait velocity could classify low responsiveness. Non-responder classification accuracy was the highest for fall history followed by dual-task ability (TUG Cog).

The partial dependence plots visualized that slow gait velocity (<1.09 m/s) indicated high responsiveness (Fig. 2B). A faster gait velocity (>1.09 m/s) was indicative of low or non-responsiveness. The partial dependence on the fall history showed that the main criterion for classification was if someone had fallen or not, ie, people with PD that fell were likely to be low while people that did not fall were non-responders (Fig. 2C). The partial dependence on dual-task ability (TUG Cog) showed that faster completion time (<15.6 s) indicated non- and slower completion time indicated low responsiveness.

Gait Responders

Responsiveness levels differed significantly in global cognition (MoCA, $P = 0.027$) and gait velocity ($P = 0.017$) (Kruskal-Wallis test, Table S3). High responders had the slowest gait velocity of the three groups, and performed worse on global cognition compared to non-responders, but higher than low responders. Non-responders had the highest global cognition and fastest gait velocity, while low responders had the lowest global cognition.

The OOB estimate of the error rate was 59%. A Cohen's weighted Kappa of −0.05 showed that the model and data were not good at classifying gait domain responders (Table S1). Only

TABLE 1 *Demographics of the participants in the HiBalance intervention group divided into balance responders according to Mini-BESTest differences (post-pre), non-responders ≤0, low responders = 1−2 points, and high responders ≥3 points*

Variable	Non-responder (N = 16)	Low responder (N = 12)	High responder (N = 11)	Total (N = 39)	P-value
Age, yrs					0.722[a]
Mean	70.81	70.83	69.00	70.31	
Range	64.00–83.00	61.00–83.00	62.00–81.00	61.00–83.00	
Sex, n (%)					0.296[b]
Male	11 (68.8%)	9 (75.0%)	5 (45.5%)	25 (64.1%)	
Female	5 (31.2%)	3 (25.0%)	6 (54.5%)	14 (35.9%)	
Cognitive subtype, n (%)					0.304[b]
Non-MCI	13 (81.2%)	9 (75.0%)	6 (54.5%)	28 (71.8%)	
MCI	3 (18.8%)	3 (25.0%)	5 (45.5%)	11 (28.2%)	
Motor subtype, n (%)					0.650[b]
TD	3 (18.8%)	1 (8.3%)	3 (27.3%)	7 (17.9%)	
PIGD	9 (56.2%)	9 (75.0%)	7 (63.6%)	25 (64.1%)	
IND	4 (25.0%)	2 (16.7%)	1 (9.1%)	7 (17.9%)	
Physical activity subtype, n (%)					0.045[b]
Sedentary	1 (6.2%)	1 (8.3%)	3 (27.3%)	5 (12.8%)	
Light movers	4 (25.0%)	8 (66.7%)	2 (18.2%)	14 (35.9%)	
Steady movers	11 (68.8%)	3 (25.0%)	6 (54.5%)	20 (51.3%)	
LEDD, mg					0.488[a]
Mean	696.12	623.25	511.50	621.63	
Range	78.00–1385.00	0.00–1324.00	100.00–992.50	0.00–1385.00	
MDS-UPDRS III					0.306[a]
Mean	28.38	30.92	36.36	31.41	
Range	11.00–46.00	10.00–70.00	20.00–60.00	10.00–70.00	
MDS-UPDRS Total					0.900[a]
Mean	49.94	49.92	54.36	51.18	
Range	24.00–76.00	22.00–102.00	23.00–110.00	22.00–110.00	
MoCA					0.316[a]
Mean	26.25	26.75	25.18	26.10	
Range	22.00–29.00	22.00–30.00	21.00–29.00	21.00–30.00	
Mini-BESTest					0.007[a]
Mean	22.75	21.33	18.64	21.15	
Range	14.00–27.00	15.00–26.00	15.00–24.00	14.00–27.00	
Mini-BESTest post-pre					<0.001[a]
Mean	−1.31	1.67	3.64	1.00	
Range	−3.00–0.00	1.00–2.00	3.00–8.00	−3.00–8.00	
Gait velocity, cm/s					0.053[a]
Mean	126.58	128.82	111.79	123.10	
Range	100.10–150.60	109.70–143.80	82.20–150.90	82.20–150.90	

(Continues)

TABLE 1 *Continued*

Variable	Non-responder (N = 16)	Low responder (N = 12)	High responder (N = 11)	Total (N = 39)	*P*-value
Gait velocity post-pre, cm/s					0.056[a]
Mean	1.80	6.09	12.68	6.19	
Range	−20.10–20.60	−9.70–29.00	0.90–30.80	−20.10–30.80	
Steps per day					0.389[a]
Mean	5800	4690	5987	5511	
Range	1858–8792	1689–9048	2334–11,482	1689–11,482	
Steps per day post-pre					0.168[a]
Mean	−965	−72	579	−255	
Range	−2997–1749	−2637–2572	−3561–5830	−3561–5830	
TUG					0.211[a]
Mean	10.04	9.96	12.22	10.63	
Range	7.09–13.09	6.09–14.53	7.13–18.12	6.09–18.12	
TUG Cog					0.069[a]
Mean	12.50	15.58	17.15	14.76	
Range	7.12–17.84	10.44–23.00	9.16–40.72	7.12–40.72	
Presence intervention, %					0.484[a]
Mean	81.56	85.42	87.27	84.36	
Range	60.00–100.00	65.00–100.00	70.00–100.00	60.00–100.00	
HADS Anxiety					0.390[a]
Mean	5.31	4.42	3.64	4.56	
Range	0.00–11.00	1.00–7.00	1.00–10.00	0.00–11.00	
HADS Depression					0.502[a]
Mean	3.31	3.67	2.73	3.26	
Range	0.00–8.00	1.00–7.00	0.00–8.00	0.00–8.00	
ABC					0.974[a]
Mean	82.60	81.20	81.36	81.82	
Range	48.75–96.88	46.25–96.88	43.75–100.00	43.75–100.00	
Walk 12					0.890[a]
Mean	9.94	10.92	11.09	10.56	
Range	4.00–22.00	5.00–23.00	0.00–25.00	0.00–25.00	
Falls last 6 months					0.023[a]
Mean	0.19	0.92	1.27	0.72	
Range	0.00–1.00	0.00–3.00	0.00–5.00	0.00–5.00	

[a]Kruskal-Wallis test.
[b]χ^2-test.
Abbreviations: ABC, Activities-specific Balance Confidence scale; HADS, Hospital Anxiety and Depression scale; IND, indetermined; LEDD, levodopa-equivalent daily dose; MCI, mild cognitive impairment; MDS-UPDRS, the Movement Disorder Society-sponsored Revision of the Unified Parkinson's disease Rating Scale; MDS-UPDRS-III, motor scale of the MDS-UPDRS; Mini-BESTest, Mini Balance Evaluation Systems Test; MoCA, Montreal Cognitive Assessment; PIGD, postural instability gait difficulty; TD, tremor dominant; TUG Cog, Timed Up and Go test with a serial subtraction task; TUG, Timed Up and Go; Walk 12, Walking Impact Scale.

TABLE 2 *Confusion matrices of the random forest classification analyses*

Responder group	Non-responder	Low responder	High responder	Classification error
Balance domain				
Non-responder	7	8	1	0.56
Low responder	6	5	1	0.58
High responder	5	2	4	0.64
Gait domain				
Non-responder	10	3	4	0.41
Low responder	6	4	1	0.64
High responder	7	2	2	0.81
Physical activity domain				
Non-responder	19		4	0.17
High responder	9		4	0.69

low and non-responder classifications exceeded the calculated chance level (Table 2). A more detailed description regarding variable importance and accuracy results can be found in the supplement (Fig. S1 and Table S4).

Physical Activity Responders

The responsiveness levels differed significantly in the motor subtype ($P = 0.026$), steps per day ($P = 0.046$), balance confidence (ABC, $P = 0.038$), and self-perceived walking ability (Walk 12, $P = 0.047$) (Kruskal-Wallis test, Table S5). High responders had the lowest steps per day, highest balance confidence, and lowest self-perceived walking ability. Five out of six people with PD that were tremor-dominant were high responders. Non-responders had the most steps per day, lowest balance confidence, highest self-perceived walking ability, and the majority were of the postural instability/gait difficulty subtype.

The Random Forest had the lowest OOB estimate of error rate with 36.11%. A Cohen's weighted Kappa of 0.15 indicated only a slight agreement (Table S1). Only the classification of non-responders exceeded the chance level (Table 2). A more detailed description regarding variable importance and accuracy results can be found in the supplement (Fig. S2 and Table S6).

Discussion

We investigated responsiveness to a highly challenging balance intervention (HiBalance) in 39 people with PD. High, low, or non-responsiveness were explored according to three responder domains relating to balance, gait, and physical activity. All Random Forests reached low Cohen's Kappa values. Only the balance responder domain model exceeded the chance level in classifying all responsiveness levels and will thus be the discussion focus. Given that the HiBalance intervention primarily targets balance impairments, it is perhaps not surprising that the balance

model was a better fit compared to gait and physical activity. The gait model exceeded the chance level for low responders and non-responders, and the physical activity model for non-responders only. Hence, the physical activity domain model seemed not to be a good model to characterize responsiveness to the HiBalance intervention. Regarding the gait domain model, high balance confidence (ABC) indicated low responsiveness and high anxiety (HADS) indicated non-responsiveness. Slow gait velocity indicated high responsiveness in the balance domain but falls better characterized low responsiveness and no falls non-responsiveness. Thus, both performance-based and self-reported gait difficulties, dual-task performance, and previous falls yield potential as screening variables. This may ultimately increase the accuracy in finding responders to HiBalance than by random chance inclusion.

Gait Velocity and Fall History Indicate Balance Responsiveness

Slow gait velocity and high self-perceived walking difficulties (Walk 12) were indicative of high balance responsiveness. This finding that both performance-based as well as self-reported decrease in walking ability indicated better balance performance after the HiBalance intervention may be of clinical value. Gait and balance are said to represent different mobility constructs in people with PD, which makes our finding that slow gait velocity was indicative of balance response interesting.[35] However, decreased gait velocity often interacts with poor balance and results in higher fall risk, which in turn reduces independence and quality of life.[36] Thus, balance is an important factor for gait stability, making them interdependent constructs.

Dual-task ability (TUG Cog) was indicative of low or non-responsiveness in the balance domain ie, a fast completion time classified non-responders while a slower completion time
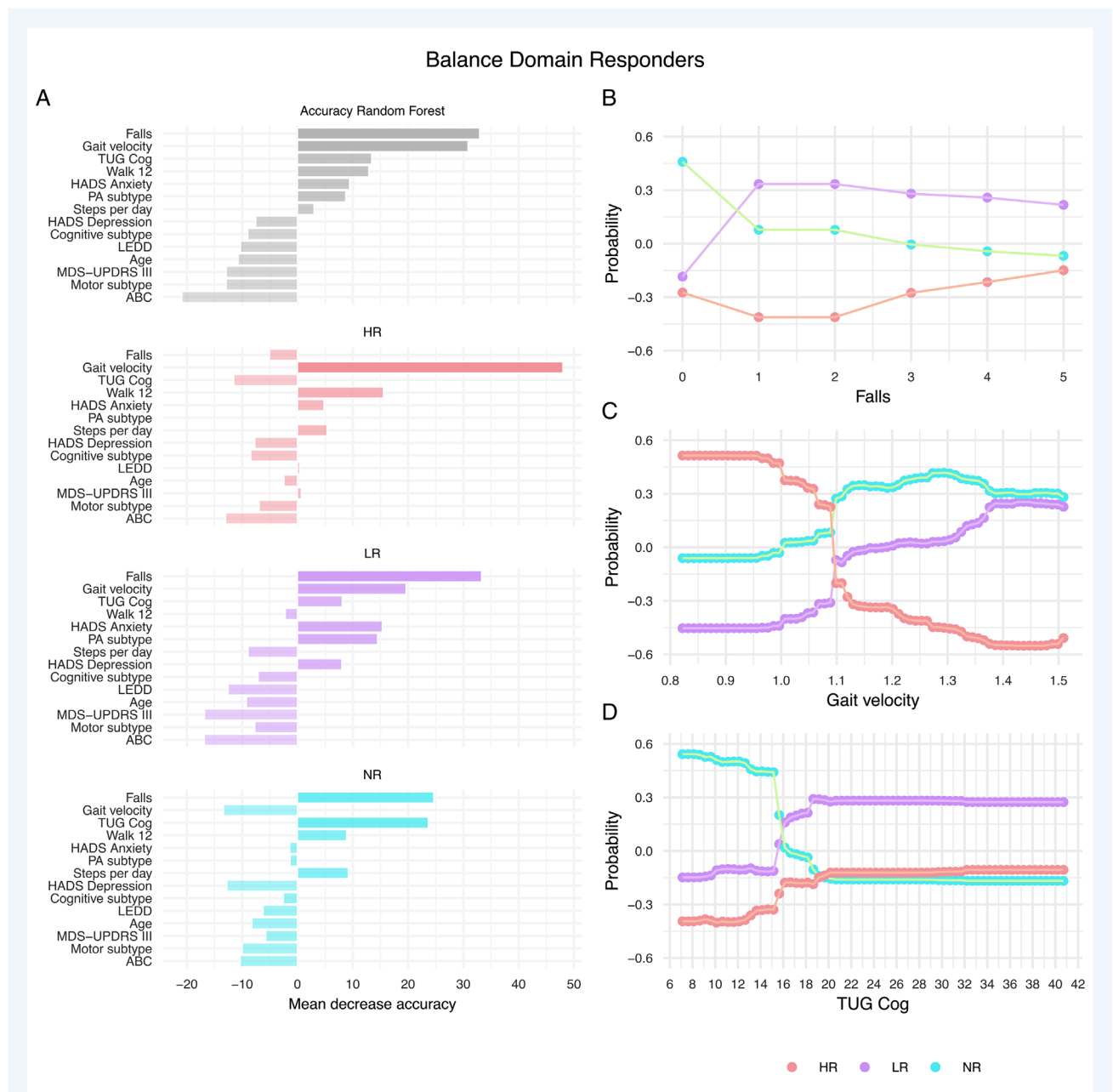
Figure 2. Variable importance. (A) Mean decrease in the Random Forest prediction accuracy for each variable used for the decision trees. The decreased accuracy is shown as a percent increase in the misclassification rate as compared to the out-of-bag rate. Classification probability for each responsiveness level is shown as a function of predicted values of classification variables: (B) Partial dependence on the number of falls within 6 months before the HiBalance intervention, (C) Partial dependence on gait velocity in cm/s, and (D) Partial dependence on Timed Up and Go test with a serial subtraction task (TUG Cog) in s. ABC, Activities-specific Balance Confidence scale; HADS, Hospital Anxiety and Depression scale; HR, high responder; LR, low responder; LEDD, levodopa-equivalent daily dose; MDS-UPDRS-III, Movement Disorder Society-sponsored Revision of the Unified Parkinson's disease Rating Scale–motor severity; NR, non-responder; PA, physical activity; Walk 12, Walking Impact Scale.

classified low responders. The TUG Cog does not just impose a demand on attentional resources, but it also includes functional mobility aspects such as transfers and turns. Thus, balance plays an essential role in a fast TUG Cog completion time. Being able to successfully find those individuals who may not respond to a certain type of intervention can be of equal importance as to finding those who will. By doing so, clinicians can better help

their patients allocate time and commitment to the type of training that will best suit their specific symptom profiles and needs. Further, the TUG Cog can be readily performed in most settings, it is quick and only requires a chair, tape, and a stopwatch.[37]

People with a fall history 6 months prior to the HiBalance intervention indicated low balance responsiveness while

non-fallers were likely to be non-responders. Of note, we did not dichotomize falls since the Random Forest can be biased by including too many categorical or short-range variables. Interestingly, the partial dependence plots show that the actual number of falls was unimportant. Falls in general might be a sign of more progressed motor or cognitive decline which is in line with our findings of slower gait velocity being indicative of high balance responsiveness. Fall history and even subjective fear of falling impact motor learning negatively by dampening the effect of exercise in people with PD, older adults, and people with depressive symptoms.[38–40] Noteworthy, our participants had no clinical depression with values between 2 and 3 on the HADS Depression items. But even though there was no significant difference between the responsiveness levels, descriptively, high responsiveness had the lowest anxiety.

Balance Response in Other Studies and Different Interventions

In a previous study on the HiBalance intervention responsiveness, factors characterizing improved balance and gait were functional mobility (TUG) and cognitive dual-task difficulties.[10] Remarkably, only self-perceived health (SF-36) additionally indicated balance performance responsiveness. There could be various reasons why only the finding of the TUG is in line with the present study. For example, the previous study predicted continuous outcomes while we used responder classifications. In another responder analysis study on effectiveness-implementation of our HiBalance intervention, low balance confidence and high attendance rate (>80%) indicated balance responsiveness.[12] Noteworthy, in the present study, the attendance rate was 85% and we found no significant difference between responsiveness levels. Both aforementioned studies investigated response to the HiBalance intervention and found different variables to be indicative of responsiveness in the balance domain but have also significant methodological differences such as the analysis method, included variables, and participant inclusion criteria. In summary, two out of three HiBalance responder studies identified TUG as important and thus maybe indicative for balance response. Still, more research using a bigger data set and the same analysis method is needed to draw robust conclusions.

Another study about training response to an agility boot camp with cognitive training (ABC-C) systematically compared objective and clinical measures to the predictive value of a training response.[11] Among those variables indicating response were also dual-task gait velocity and anticipatory postural adjustment. Higher effect sizes were found for objectively measured balance aspect changes, using accelerometers, than for clinical measures such as the Mini-BESTest. Interestingly, the gait domain showed the greatest intervention response. Even though the study protocol and analysis method differed, gait velocity was, as in our study, the variable that classified response.

Non-responsiveness to HiBalance

The identification of responders to the HiBalance intervention is imperative to the development of future studies and implementation protocols. However, of equal importance is finding those participants who did not respond to this type of exercise. There are various reasons why people respond differently to the same type of intervention. Although extensive in nature regarding both motor and non-motor symptoms, the set of outcome variables included in this study may have prevented those participants performing at the higher end at baseline from showing any range of improvement post-intervention. Given that non-responders had higher balance function and were more physically active in everyday life, it follows that exercise intervention targeting these participants should be adapted to a level where they can be challenged to the same extent as the high responders presumably were. With an ongoing shift in rehabilitation research towards a precision-based approach, the intention is to be able to tailor exercise to each individual and adapt it according to their symptom profile.

Limitations

We acknowledge the explorative nature of this study. Thus, conclusions need to be interpreted with caution and need further replication. A limitation is the small sample size due to the splitting into three responsiveness levels. The intervention in the current study primarily targets balance performance, and it is thus not surprising that the balance model showed a better fit than both the gait and the physical activity model. Our models' relatively high misclassification rate showed that we do not explain all factors that indicate responsiveness with our set of included variables. Random Forests (built with the CART method) tend to be biased towards variables with more categories/ wider ranges.[41] Variable importance could be biased towards questionnaire scales with a broader range, but we could not see that in our data.

Conclusions

Balance domain response could be classified with moderate Cohen's Kappa and above chance level, while gait and physical activity domains could not. Slow baseline gait velocity and low self-perceived walking ability indicated a high response in balance performance after HiBalance participation. Recent falls and slow baseline gait velocity predicted a low response in balance performance. This suggests that people with slower gait velocity, subjective diminished walking ability, and a history of falls should perhaps be prioritized in participating in highly challenging balance interventions. We provide valuable insights into the potential screening variables for exercise interventions to be used as a promising therapeutic intervention for the management of PD-related motor symptoms which cannot yet be treated medically.

Author Roles

(1) Research project: A. Conception, B. Organization, C. Execution; (2) Statistical Analysis: A. Design, B. Execution, C. Review and Critique; (3) Manuscript: A. Writing of the first draft, B. Review and Critique.

F.A.: 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B
H.J.: 1A, 1B, 2C, 3B
E.W.: 1A, 3B
K.P.: 2A, 2C, 3B
M.H.: 1A, 2C, 3B
E.F.: 1A, 1B, 2C, 3B

Acknowledgments

Disclosures

Data Availability Statement

The data are part of the EXPANd trial.[16] The analysis scripts and meta-data are available on the OSF (osf.io/8truq). The original data are not publicly available due to Swedish/EU law. Data sharing will be regulated via a data transfer and user agreement upon a reasonable request. ■

References

1. Conradsson D, Löfgren N, Nero H, Hagströmer M, Ståhle A, Lökk J, Franzén E. The effects of highly challenging balance training in elderly with Parkinson's disease: a randomized controlled trial. *Neurorehabil Neural Repair* 2015;29:827–836.

2. Mak MKY, Wong-Yu ISK. Exercise for Parkinson's disease. *Int Rev Neurobiol* 2019;147:1–44.

3. Radder DLM, Silva L, de Lima A, Domingos J, et al. Physiotherapy in Parkinson's disease: a meta-analysis of present treatment modalities. *Neurorehabil Neural Repair* 2020;34:871–880.

4. Petzinger GM, Fisher BE, McEwen S, Beeler JA, Walsh JP, Jakowec MW. Exercise-enhanced neuroplasticity targeting motor and cognitive circuitry in Parkinson's disease. *Lancet Neurol* 2013;12:716–726.

5. Allen NE, Sherrington C, Paul SS, Canning CG. Balance and falls in Parkinson's disease: a meta-analysis of the effect of exercise and motor training. *Mov Disord* 2011;26:1605–1615.

6. Leavy B, Joseph C, Lofgren N, Johansson H, Hagstromer M, Franzen E. Outcome evaluation of highly challenging balance training for people with Parkinson disease: a multicenter effectiveness-implementation study. *J Neurol Phys Ther* 2020;44:15–22.

7. Rennie L, Opheim A, Dietrichs E, Löfgren N, Franzén E. Highly challenging balance and gait training for individuals with Parkinson's disease improves pace, rhythm and variability domains of gait–a secondary analysis from a randomized controlled trial. *Clin Rehabil* 2021;35:200–212.

8. McIsaac TL, Lamberg EM, Muratori LM. Building a framework for a dual task taxonomy. *Biomed Res Int* 2015;2015:591475.

9. Freidle M, Johansson H, Ekman U, et al. Behavioural and neuroplastic effects of a double-blind randomised controlled balance exercise trial in people with Parkinson's disease. *Npj Parkinsons Dis* 2022;8:1–10.

10. Löfgren N, Conradsson D, Joseph C, Leavy B, Hagstromer M, Franzen E. Factors associated with responsiveness to gait and balance training in people with Parkinson disease. *J Neurol Phys Ther* 2019;43:42–49.

11. Hasegawa N, Shah VV, Harker G, et al. Responsiveness of objective vs. clinical balance domain outcomes for exercise intervention in Parkinson's disease. *Front Neurol* 2020;11:940.

12. Joseph C, Leavy B, Franzen E. Predictors of improved balance performance in persons with Parkinson's disease following a training intervention: analysis of data from an effectiveness-implementation trial. *Clin Rehabil* 2020;34:837–844.

13. Strouwen C, Molenaar E, Münks L, et al. Determinants of dual-task training effect size in Parkinson disease: who will benefit Most? *J Neurol Phys Ther* 2019;43:3–11.

14. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–483.

15. Albrecht F, Pereira JB, Mijalkov M, et al. Effects of a highly challenging balance training program on motor function and brain structure in Parkinson's disease. *J Parkinsons Dis* 2021;11:2057–2071.

16. Franzén E, Johansson H, Freidle M, et al. The EXPANd trial: effects of exercise and exploring neuroplastic changes in people with Parkinson's disease: a study protocol for a double-blinded randomized controlled trial. *BMC Neurol* 2019;19:280.

17. Borland E, Nagga K, Nilsson PM, Minthon L, Nilsson ED, Palmqvist S. The Montreal cognitive assessment: normative data from a large Swedish population-based cohort. *J Alzheimers Dis* 2017;59:893–901.

18. Franchignoni F, Horak F, Godi M, Nardone A, Giordano A. Using psychometric techniques to improve the balance evaluation systems test: the mini-BESTest. *J Rehabil Med* 2010;42:323–331.

19. Litvan I, Goldman JG, Troster AI, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society task force guidelines. *Mov Disord* 2012;27:349–356.

20. Stebbins GT, Goetz CG, Burn DJ, Jankovic J, Khoo TK, Tilley BC. How to identify tremor dominant and postural instability/gait difficulty groups with the movement disorder society unified Parkinson's disease rating scale: comparison with the unified Parkinson's disease rating scale. *Mov Disord* 2013;28:668–670.

21. von Rosen P, Hagströmer M, Franzén E, Leavy B. Physical activity profiles in Parkinson's disease. *BMC Neurol* 2021;21:71.

22. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2012;28:112–118.

23. Hass CJ, Bishop M, Moscovich M, et al. Defining the clinically meaningful difference in gait speed in persons with Parkinson disease. *J Neurol Phys Ther* 2014;38:233–238.

24. Löfgren N, Lenholm E, Conradsson D, Ståhle A, Franzén E. The mini-BESTest - a clinically reproducible tool for balance evaluations in mild to moderate Parkinson's disease? *BMC Neurol* 2014;14:235.

25. Rowlands A, Davies M, Dempsey P, Edwardson C, Razieh C, Yates T. Wrist-worn accelerometers: recommending ∼1.0 mg as the minimum clinically important difference (MCID) in daily average acceleration for inactive adults. *Br J Sports Med* 2021;55: 814–815.

26. Heinzen E. arsenal: An Arsenal of 'R' Functions for Large-Scale Statistical Summaries. R package version 3.6.3; https://cran.r-project.org/web/packages/arsenal/index.html.

27. Breiman L. Random forests. *Mach Learn* 2001;45:5–32.

28. Liaw A, Wiener M. Classification and regression by randomForest. *R News* 2002;2:18–22.

29. Couronné R, Probst P, Boulesteix A-L. Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinformatics* 2018;19:270.

30. Boulesteix AL, Janitza S, Kruppa J, König IR. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdiscip Rev Data Min Knowl Discov* 2012;2:493–507.

31. Bénard C, Da Veiga S, Scornet E. Mean decrease accuracy for random forests: inconsistency, and a practical solution via the Sobol-MDA. *Biometrika* 2022;109:881–900.

32. Janitza S, Hornung R. On the overestimation of random forest's out-of-bag error. *PLoS One* 2018;13:e0201904.

33. Revelle W. *Psych: Procedures for Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois. R Package Version 1.9. 12. Comprehensive R Archive Network (CRAN)*. Evanston, IL, USA: Northwestern University; 2019.

34. Ben-David A. Comparison of classification accuracy using Cohen's weighted kappa. *Expert Syst Appl* 2008;34:825–832.

35. Horak FB, Mancini M, Carlson-Kuhta P, Nutt JG, Salarian A. Balance and gait represent independent domains of mobility in Parkinson disease. *Phys Ther* 2016;96:1364–1371.

36. Mirelman A, Bonato P, Camicioli R, et al. Gait impairments in Parkinson's disease. *Lancet Neurol* 2019;18:697–708.

37. Conradsson D, Leavy B, Hagstromer M, Nilsson MH, Franzen E. Physiotherapy for Parkinson's disease in Sweden: provision, expertise, and multi-professional collaborations. *Mov Disord Clin Pract* 2017;4(6):843–851.

38. Ellis T, Boudreau JK, DeAngelis TR, et al. Barriers to exercise in people with Parkinson disease. *Phys Ther* 2013;93:628–636.

39. Afshari M, Yang A, Bega D. Motivators and barriers to exercise in Parkinson's disease. *J Parkinsons Dis* 2017;7:703–711.

40. Boyd LA, Vidoni ED, Siengsukon CF. Multidimensional motor sequence learning is impaired in older but not younger or middle-aged adults. *Phys Ther* 2008;88:351–362.

41. Strobl C, Boulesteix A-L, Zeileis A, Hothorn T. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 2007;8:25.

Supporting Information

Supporting information may be found in the online version of this article.

Figure S1. Variable importance of the gait responders analysis.

Figure S2. Variable importance of the step responders analysis.

TABLE S1. Cohen's Kappa for the responder domain prediction.

TABLE S2. Mean decrease accuracy for balance domain responders.

TABLE S3. Demographics of the participants in the HiBalance intervention group divided into responders according to gait velocity differences.

TABLE S4. Mean decrease accuracy for gait domain responders.

TABLE S5. Demographics of the participants in the HiBalance intervention group divided into physical activity responders according to steps per day differences.

TABLE S6. Mean decrease accuracy for step domain responders.