



Published in final edited form as:

*IEEE Trans Neural Syst Rehabil Eng.* 2021 ; 29: 576–586. doi:10.1109/TNSRE.2021.3056908.

## Robot-Based Assessment of HIV-Related Motor and Cognitive Impairment for Neurorehabilitation

**Kevin D. Bui [Graduate Student Member, IEEE],**

Rehabilitation Robotics Laboratory, University of Pennsylvania, Philadelphia, PA 19104 USA, and also with the Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104 USA

**Carol A. Wamsley,**

Penn Institute for Rehabilitation Medicine, Philadelphia, PA 19146 USA

**Frances S. Shofer,**

Department of Epidemiology and Biostatistics in Emergency Medicine, University of Pennsylvania, Philadelphia, PA 19104 USA

**Dennis L. Kolson,**

Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104 USA

**Michelle J. Johnson [Member, IEEE]**

Rehabilitation Robotics Laboratory, University of Pennsylvania, Philadelphia, PA 19104 USA, and also with the Department of Physical Medicine and Rehabilitation and Bioengineering, University of Pennsylvania, Philadelphia, PA 19104 USA

### Abstract

There is a pressing need for strategies to slow or treat the progression of functional decline in people living with HIV. This paper explores a novel rehabilitation robotics approach to measuring cognitive and motor impairment in adults living with HIV, including a subset with stroke. We conducted a cross-sectional study with 21 subjects exhibiting varying levels of cognitive and motor impairment. We tested three robot-based tasks – trajectory tracking, N-back, and spatial span – to assess if metrics derived from these tasks were sensitive to differences in subjects with varying levels of executive function and upper limb motor impairments. We also examined how well these metrics could estimate clinical cognitive and motor scores. The results showed that the average sequence length on the robot-based spatial span task was the most sensitive to differences between various cognitive and motor impairment levels. We observed strong correlations between robot-based measures and clinical cognitive and motor assessments relevant to the HIV population, such as the Color Trails 1 ( $\rho = 0.83$ ), Color Trails 2 ( $\rho = 0.71$ ), Digit Symbol – Coding ( $\rho = 0.81$ ), Montreal Cognitive Assessment – Executive Function subscore ( $\rho = 0.70$ ), and Box and Block Test ( $\rho = 0.74$ ). Importantly, our results highlight that gross motor impairment may be overlooked in the assessment of HIV-related disability. This study shows that

---

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see <https://creativecommons.org/licenses/by/4.0/>

Corresponding author: Michelle J. Johnson. johnmic@penmedicine.upenn.edu.

rehabilitation robotics can be expanded to new populations beyond stroke, namely to people living with HIV and those with cognitive impairments.

## Keywords

Rehabilitation robotics; HIV; stroke; cognitive impairment; motor impairment; neurorehabilitation; upper limb

---

## I. Introduction

TODAY, there are nearly 37 million persons living with human immunodeficiency virus (PLWH) worldwide [1]. As PLWH age due to the success of antiretroviral therapy (ART), the challenges have shifted to managing the chronic effects of living with HIV. Many of these challenges can be attributed to neurological complications caused by HIV-associated neurocognitive disorders (HAND), accelerated aging, drug abuse, and HIV-related comorbidities [2]. Together, the broad range of impairments experienced by PLWH has been shown to impact instrumental activities of daily living (IADLs), such as medication management, telephone communication, cooking, and financial management [3]. In one study, upwards of 80% of PLWH reported dealing with at least one impairment, activity limitation or disability, or social participation restriction [4]. These deficits are often tied to impairments in executive function, memory, and visuospatial domains [5]. PLWH also experience motor impairments in gait, coordination, upper limb fine motor skills, and strength, with 69% of PLWH in one study demonstrating at least one motor impairment [6]–[9]. As such, there is a pressing need for effective neurorehabilitation strategies to slow or treat the progression of functional decline in PLWH.

The gold standard for diagnosing neurocognitive impairment has been established by the Frascati criteria, an extensive neuropsychological battery that classifies HAND subtypes as asymptomatic neurocognitive impairment, mild neurocognitive disorder, or HIV-associated dementia [10]. However, the assessments used to diagnose HAND often test domains in isolation, which is not reflective of the dual involvement of cognitive and motor demands in most IADL tasks. Differences between HIV and non-HIV populations are also seen in more nuanced tasks. Kronemer *et al.* demonstrated that even when there was no motor impairment detected on clinical assessments, PLWH demonstrated upper limb motor impairment while multitasking compared to a non-HIV control group that did not relate to HAND stage [11]. Assessments of multitasking have been shown to be more reflective of IADL performance in PLWH compared to standard clinical assessments [12]. These results demonstrate that current clinical assessments and biomarkers of HIV do not necessarily correspond well to more subtle impairments in cognition and motor performance [11].

HIV-associated non-communicable diseases, such as cerebrovascular disease (CVD), are a secondary effect of HIV infection that can further exacerbate existing cognitive and motor impairments. HIV is an independent risk factor for CVD such as stroke [13]. With an incidence rate of 3.87 per 1000 years lived, CVDs occur at an average age of 48 years in the HIV population [13]. These numbers are 1.5 times higher and 22 years younger than the general U.S. population [14]. Augustyn *et al.* recently showed that stroke survivors with HIV

experienced a decline in ADL functions one month after discharge compared to stroke survivors without HIV who continued to show improvement, highlighting how HIV can impact stroke recovery [15].

Efforts to develop neurorehabilitation strategies have been made in the stroke population, but there is a paucity of established solutions for PLWH despite evidence that rehabilitation can positively address HIV-related challenges in physical, social, and psychological well-being [16], [17]. The rehabilitation robotics field provides a potential solution to address these challenges [18]. Robot-assisted stroke therapy has been shown to be as effective as high-intensity physical therapy for chronic stroke patients [19]. Additionally, robotic systems allow for a variety of kinematic metrics to be observed that relate to clinical measures of motor impairment [20]–[24].

While the primary focus to date has been on motor impairment, recent studies have started to look at robot-based measures of cognitive impairment in stroke and traumatic brain injury populations [21], [22]. Both of these studies have demonstrated a relationship between robot-based metrics and overall cognitive scores. However, given that cognition is broadly defined, more work needs to be done to establish robot-based metrics relating to specific domains.

The strengths of a rehabilitation robotics-based approach include the ability to standardize assessments with a greater range of objective measures, collect a vast amount of data, and develop personalized neurorehabilitation strategies based on the patient's presenting characteristics. Our prior work has also shown the feasibility of deploying cost-effective rehabilitation robotics systems in lower-resource contexts [25]. Cost-effective rehabilitation robotics systems can bridge healthcare gaps in countries with low-to-middle income economies that are dealing with large populations of patients with impairments and a shortage of rehabilitation professionals. This approach has the potential to positively impact PLWH by building upon the body of work that has been done in the stroke population.

This preliminary cross-sectional study aims to establish objective, robot-based measures of executive function and upper limb motor impairment in PLWH – including a subset with stroke – and assess the strength of the relationship between these robot-based and clinical assessment scores. This study tests three hypotheses to demonstrate the utility of a robotic approach in assessing impairments in PLWH. Given the heterogeneous nature of impairments in this population, the first part of this study tests the hypothesis that robot-based metrics can differentiate subjects with and without moderate executive function or upper-limb motor impairments (H1). The second hypothesis measures the relationship between robot-based metrics and clinical assessments used in PLWH by testing whether robot-based metrics are good predictors of clinical cognitive assessment scores (H2) as well as clinical motor assessment scores (H3). This work lays the foundation for the development of novel neurorehabilitation strategies for PLWH.

## II. Methods

### A. Subject Population and Procedure

Individuals over the age of 18 years old were recruited from the community through flyers posted at local HIV clinics and organizations. Inclusion criteria for the HIV group consisted of documented HIV status that was ART-treated and virally-suppressed, the ability to ambulate, the ability to comprehend study procedures, and the ability to provide written informed consent. Individuals with neuropathy (i.e. distal symmetric polyneuropathy) were excluded.

Subjects were included in the HIV-stroke subgroup if they met the inclusion criteria for the HIV group and were at least three months removed from a stroke event. HIV-stroke subjects with severe aphasia, visual neglect, or basal ganglia stroke were excluded. Subjects were excluded if they were more than mildly depressed as assessed by the Beck's Depression Inventory – Fast Screen (score  $\geq 4$ ) [26]. Subjects were compensated for time and travel. This protocol was approved by the Internal Review Board of the University of Pennsylvania (Protocol no. 823511).

Subjects underwent a preliminary phone screen to screen for study eligibility. They were then sent a copy of the informed consent to review prior to coming in for their scheduled in-person appointment. After written informed consent was obtained in-person, cognitive and motor assessments were performed. Participants then completed three robot-based tasks in a randomized order with the dominant and non-dominant upper-extremity limb.

### B. Cognitive Assessments

The cognitive assessments consisted of the Color Trails, Digit Symbol–Coding (WAIS-III ®), Montreal Cognitive Assessment (MoCA), and International HIV Dementia Scale (IHDS) [27]–[30]. These tests have all been administered in PLWH previously to measure neurocognitive impairment [8], [30]–[33]. These tests were chosen to reflect the cognitive domains commonly affected by HIV.

**1) Color Trails:** The Color Trails is a set of two cognitive pencil and paper tests based on the Trail Making Test but does not require knowledge of the alphabet, thus reducing potential bias [27]. Color Trails 1 tests for sustained visual attention and simple sequencing, while Color Trails 2 assesses frontal systems such as selective attention, mental flexibility, visual spatial skills, and motor speed. Performance was measured by the time to complete the task, with a higher time indicating worse performance. These scores were normalized by age, gender, and education [27].

**2) Digit Symbol – Coding (WAIS-III ®):** The Digit Symbol–Coding (WAIS-III ®) test is another neuropsychological test assessing processing speed [28]. Subjects use a number-symbol key to copy symbols under a sequence of numbers. Performance was measured by the number of symbols coded in the span of two minutes, with a higher number of symbols copied in the time span representing better performance. Scores were normalized by age, gender, and education.

**3) Montreal Cognitive Assessment (MoCA):** The MoCA is a screening tool to detect impairment in a number of cognitive domains – visuospatial/executive, naming, memory, attention, language, abstraction, delayed recall, and orientation – and reflects the degree of cognitive impairment in a subject [29]. A score above 25 out of 30 generally indicates normal cognitive function, while a score below 19 indicates likely moderate cognitive impairment.

An executive function subscore (MoCA-EF) was calculated to serve as a proxy in place of a more extensive neuropsychological assessment of executive function, based on work by Lam *et al.* demonstrating good convergent validity between this subscore and standardized neuropsychological tests of executive function [34]. This subscore, scored out of five points, was calculated from summing the scores from the backward digit span, trail making, word similarities, and ‘F’-word list generation tasks [34]. Lam *et al.* demonstrated that a cutoff score of 4 had a sensitivity of 0.79 to executive function impairment [34].

**4) International HIV Dementia Scale (IHDS):** The IHDS is a screening test for cognitive impairment designed to screen for HAND, with a score below 10 out of 12 indicating potential cognitive impairment [30]. It was developed as a culturally appropriate adaptation of the HIV Dementia Scale. However, the IHDS has not been validated in the stroke population.

### C. Motor Assessments

The motor assessments tested gross motor function, fine motor function, and strength. They consisted of the Box and Blocks Test (BBT), Grooved Pegboard (GP), and grip strength.

**1) Box and Blocks (BBT):** The BBT is a test of gross motor function measuring how many blocks subjects are able to transfer across a partition in one minute, with a higher number of transferred blocks indicating better motor function [35]. Scores were normalized by age, gender, and limb. It is typically used to measure reach and grasp function in the stroke population.

**2) Grooved Pegboard (GP):** GP is a common motor assessment in PLWH. It tests fine motor function and dexterity, measuring the amount of time a subject takes to insert all of the grooved pegs into matched holes on a board. Performance was measured by the time to complete the task with longer times indicating worse fine motor function [36]. GP data for subjects unable to complete the task were not included in the analysis (one subject).

**3) Grip Strength:** Grip strength is measured with a dynamometer. Three trials were taken with each hand, with the average and standard deviation being recorded. Accelerated grip strength decline has been shown in a study of HIV-infected men, which may contribute to decreased life expectancy and lower quality of life with aging [37].

### D. Robot Assessment

**1) Rehabilitation Robot System:** The rehabilitation robot used in this study, the Haptic TheraDrive, is a one degree-of-freedom robot for upper limb stroke rehabilitation

(Fig. 1) [25]. The user operates the TheraDrive by manipulating a vertically-mounted crank handle equipped with force sensors and an optical encoder. For assessment purposes, it is run in a gravity-compensation mode, which uses force sensors as an input to a proportional-integral-derivative (PID) controller to calculate the necessary response by the motor to give the sensation that there is no resistance or assistance while the user manipulates the handle.

**2) Trajectory Tracking Motor Task:** The trajectory tracking task is designed to assess upper limb motor performance. A single trial consists of the user moving the crank arm forward and backward to follow a vertically scrolling sinusoidal path for 15 seconds. This task is repeated 15 times after one training trial. The outcome measures include performance error, movement smoothness, and the normalized distance traversed. Performance error was calculated as the root mean square error (RMSE) of the position relative to the displayed trajectory and normalized by the RMSE assuming no movement. A lower performance error indicates better tracking performance. Spectral arc length was used as the measure of smoothness, which has the benefit of being less sensitive to noise compared to other measures of smoothness [39]. More negative values of smoothness indicate less smooth movements. Normalized distance traversed was calculated from dividing the total angular distance that the subject traversed by the expected angular distance of the displayed trajectory path. A value closer to 1 reflects that the actual distance traversed matched the expected distance. A lower value could reflect moderate motor impairment, while a higher value could reflect inefficient movement.

**3) N-Back Cognitive Task:** The N-back test is commonly used in the cognitive neuroscience field as a test of working memory and working memory capacity [40]. In this version, the subject is presented with a sequence of numerical digits (1-4) with three different conditions. For the 0-back condition, the easiest condition, the subject indicates when the current stimulus shown on the screen is the number '2.' For the more cognitively-involved 1-back and 2-back conditions, the subject indicates when the current stimulus matches the stimulus shown one stimulus or two stimuli prior, respectively. The subject indicates a match by pressing a button on the TheraDrive. The number then flashes green or red for a correct or incorrect response, respectively. Each subject performed the task with each limb, cycling through the 0-back, 1-back, and 2-back conditions four times for a total of 12 trials, all with different numerical sequences. The first set of trials is used as a training set and not included in the analysis. Ten responses are recorded per trial. Each subject was shown the same set of 12 sequences, with each sequence having a minimum of three button press responses. N-back performance was measured as the total number of correct responses divided by the total number of responses across the trials, resulting in a score ranging from 0 to 1, with a score closer to 1 representing better performance.

**4) Spatial Span Cognitive-Motor Task:** The Spatial Span is a test of visuospatial working memory based on the Corsi block-tapping task used in neuropsychological assessments [41]. While computerized versions of the Spatial Span exist [42], this version incorporates an added motor component to concurrently test for arm coordination, visuospatial ability, and working memory. A 3-by-3 grid of tiles is displayed to the user on a computer screen, and a sequence of tiles is shown one tile at a time. The user must operate

the TheraDrive to select the tiles in the order shown. If the user successfully repeats the sequence by selecting the correct tiles in order, the next displayed sequence increases in length by one to make the task more difficult. If the user is unsuccessful, the sequence decreases in length by one. The metrics of interest for the task include the normalized distance traversed, movement smoothness, mean sequence length across all the trials, and performance. Normalized distance traversed and movement smoothness were calculated the same way as in the trajectory tracking task. Mean sequence length is the average number of tiles displayed to the user per trial and reflects the capacity of the subject. Spatial span performance was measured as the total number correct tile matches divided by the total number of tiles shown across the trials. Thus, spatial span performance is a score ranging from 0 to 1, with 1 representing perfect performance.

### E. Data Processing

A one-sample Kolmogorov-Smirnov test for normal distribution was run on the raw continuous demographic, clinical, and robot metrics. Given that the data were not normally distributed, non-parametric Wilcoxon rank-sum tests were conducted to test for differences between HIV and HIV-stroke groups. To adjust for multiple comparisons, separate Bonferroni corrections were applied for the clinical (adjusted  $p = 0.004$ ) and robot-based (adjusted  $p = 0.006$ ) scores.

All robot metrics were Z-score normalized by the entire subject population in this study, resulting in a distribution with a mean of zero and standard deviation of one. This was done to ensure metrics were evenly weighted in the regression analysis.

### F. Functional Subgroup Comparison Analysis

To investigate the first hypothesis that robot-based metrics can differentiate between subjects with and without moderate executive function impairments or upper-limb motor impairment, all study subjects were categorized by their motor and cognitive status based on clinical score cutoffs. The subject population demonstrated motor impairment on both the BBT and GP based on healthy population norms, but BBT was chosen to avoid excluding individual subjects who did not complete the GP. To categorize subjects by motor status, raw BBT scores were normalized by published gender, age, and limb side norms and converted into a Z-score. A BBT Z-score of  $-2$  and below was used to indicate moderate motor impairment. To categorize subjects by cognitive status, a MoCA-EF score of 3.5 and below was used as a cutoff for likely moderate executive function impairment [34]. Subjects were then categorized into one of four functional subgroups based on the possible combinations of motor and cognitive status. Because this was done for both dominant and non-dominant limb motor status, subjects could be classified into two different functional subgroup classifications based on differing motor performance between dominant and non-dominant limbs.

For each robot-based metric, a two-way analysis of variance (ANOVA) was conducted where the factors were functional group and limb performance side. To adjust for all pairwise comparisons between functional groups, a Tukey-Kramer honest significance test

was applied if the ANOVA was significant. An alpha level of 0.05 was used to establish the significance on all statistical tests.

### G. Multiple Linear Regression Analysis

To investigate whether the robot-based metrics were significant predictors of clinical assessment scores, a multiple linear regression approach was used. Bosecker *et al.* previously used a backward multiple linear approach to identify a set of robot-based metrics reflective of various stroke outcome measures [23]. Rather than start with all of the robot-based metrics and remove terms, a forward stepwise approach was implemented here. This consisted of individually testing each robot-based metric and subsequently adding it to the model only if it was a statistically significant predictor individually. Given the sample size of the subject population, the model was limited to two terms. In order to adjust for the number of predictors used in the model and to compare performance between models with different numbers of predictors, the adjusted  $R^2$  is reported. A power analysis revealed that the linear regression models were powered to detect a minimum  $R^2$  of 0.40 with one predictor and 0.43 with two predictors ( $n = 21$ , power = 0.80, alpha = 0.05). A small, medium, and large effect size were defined as an  $R^2$  value of 0.01, 0.25, and 0.50, respectively. The non-parametric Spearman's rho was also calculated to measure the correlation between predicted and actual clinical scores. All analysis was conducted in Matlab 2019A.

## III. Results

### A. Subject Population Breakdown

The descriptive statistics for demographic and clinical information for the subject groups (HIV, HIV-stroke, and combined) are presented in Table 1. Twenty-one subjects in total – thirteen male and eight female – participated in the study. Six subjects had a history of stroke. The average age of the HIV and HIV-stroke groups were  $56.2 \pm 5.4$  years old and  $54.2 \pm 8.1$  years old, respectively, while the average age of the entire subject population was  $55.5 \pm 6.3$  years old. Fifteen subjects had 12 or more years of education. Fourteen subjects had MoCA-EF scores below 3.5 and sixteen subjects displayed moderate motor impairment in at least one limb based on BBT scores. There were no statistically significant differences – even at the unadjusted alpha level of 0.05 – between HIV and HIV-stroke groups or between limbs on the clinical motor assessments.

### B. Robot-Based Performance for Example Subjects

Performance data from two sample subjects (Subjects 12 and 18) on the trajectory tracking and spatial span tasks are presented, highlighting the wide variety of impairments seen in the subject population (Fig. 2). Subject 12 is a 56-year-old male HIV subject with moderate cognitive and moderate motor impairment, scoring a 13 on the MoCA and more than two standard deviations below Box and Block population norms on both the dominant and non-dominant limb. Subject 18 is a 49-year-old male HIV-stroke subject with low cognitive and low motor impairment, scoring a 25 on the MoCA and less than two standard deviations below BBT populations norms on both the dominant and non-dominant limb. Qualitatively, Subject 12 demonstrates poorer performance compared to Subject 18 (Fig. 2; left). This can be seen in comparing the average trajectory of each subject to the desired trajectory and the



larger variance across the trials as seen in the shaded regions. On the robotic spatial span task, the histogram of sequence lengths across the trials shows a distinct difference between the two subjects (Fig. 2; right).

### C Raw Robot Performance Metrics

Table II shows the mean and standard deviations for the raw robot-based metrics across the HIV-only group, HIV-stroke group, and the entire subject population. The scores for both the dominant and non-dominant limb are reported. There were no statistically significant differences – even at the unadjusted alpha level of 0.05 – in any robot metrics between dominant and non-dominant limbs or between HIV and HIV-stroke groups. However, some qualitative differences are notable. For example, while trajectory tracking performance was similar on both limbs in the HIV-only group, it was noticeably different for the HIV-stroke group, reflecting the presence of motor impairments in the non-dominant limb likely caused by stroke. The spatial span mean sequence length in each group was lower than the reported average span of 4.8 in a study that developed a computer-based version of the Corsi block-tapping task [42]. Given that moderate cognitive impairment may mask motor performance, the study subjects were further stratified by their cognitive and motor function.

### D. Stratification by Functional Subgroups (Hypothesis 1)

Figure 3 shows the distribution of the subject population by their functional groups using MoCA-EF subscores and BBT Z-scores to separate subjects by cognitive and motor function, respectively. The number of subjects in each of the four functional groups were the same when using dominant versus non-dominant BBT z-scores. There were two subjects in the low cognitive and low motor impairment group, five subjects in the low cognitive and moderate motor impairment group, six subjects in the moderate cognitive and low motor impairment group, and eight subjects in the moderate cognitive and moderate motor impairment group. Five HIV subjects and one stroke subject had different functional group classifications based on their dominant and non-dominant motor scores.

There was a statistically significant main effect of functional group on N-back performance ( $F(3,34) = 6.64, p = 0.001$ ). There was no main effect of limb side or interaction effect. Subjects with low cognitive and low motor impairments performed better on the N-back task compared to subjects with moderate cognitive and moderate motor impairments ( $0.96 \pm 0.01$  vs.  $0.83 \pm 0.05, p = 0.001$ ) and subjects with moderate cognitive and low motor impairments ( $0.96 \pm 0.01$  vs.  $0.85 \pm 0.06, p = 0.01$ ). Fig. 4 (top) shows the N-back performance scores for each of the functional subgroups.

There was a statistically significant main effect of functional group on trajectory tracking performance error ( $F(3,34) = 7.78, p = 0.0004$ ). There was no main effect of limb side or interaction effect. Subjects with moderate cognitive and moderate motor impairment had significantly higher performance error scores compared to subjects with low cognitive and moderate motor impairment ( $0.50 \pm 0.25$  vs.  $0.28 \pm 0.08, p = 0.04$ ) and subjects with low cognitive and low motor impairment ( $0.50 \pm 0.25$  vs.  $0.20 \pm 0.04, p = 0.04$ ). Fig. 4 (middle) shows the trajectory tracking performance for each of the functional groups.

There was a statistically significant main effect of functional group on spatial span mean sequence length ( $F(3, 34) = 8.23, p = 0.0004$ ). There was no main effect of limb side or interaction effect. Subjects with low cognitive and low motor impairment had longer average sequence lengths compared to subjects with moderate cognitive and moderate motor impairments ( $4.48 \pm 0.56$  vs.  $2.39 \pm 0.74, p = 0.0004$ ) and subjects with moderate cognitive and low motor impairments ( $4.48 \pm 0.56$  vs.  $3.12 \pm 1.00, p = 0.04$ ). Subjects with low cognitive and moderate motor impairment also had longer average sequence lengths compared to subjects with moderate cognitive and moderate motor impairment ( $3.31 \pm 0.70$  vs.  $2.39 \pm 0.74, p = 0.04$ ). Fig. 4 (bottom) shows the spatial span mean sequence length for each of the functional subgroups.

There was a statistically significant main effect of functional group on spatial span performance, but there were no significant differences between any of the functional subgroups after correcting for multiple comparisons.

There were no statistically significant main or interaction effects for trajectory tracking normalized distance traversed, trajectory tracking smoothness, spatial span normalized distance traversed, or spatial span smoothness scores.

## E. Estimating Clinical Cognitive Scores (Hypothesis 2)

**1) Dominant Limb Predictors:** Fig. 5 shows the multiple linear regression models for each of the clinical cognitive assessments using dominant limb robot-based metrics as the predictors.

Color Trails 1 was predicted by a combination of trajectory tracking normalized distance traversed and spatial span mean sequence length ( $p = 0.03$  and  $0.001$ , respectively). The robot-based predictors accounted for 60% of the variance in the model, and the predicted scores strongly correlated with actual Color Trails 1 scores ( $\rho = 0.83, p = 3.33 \times 10^{-6}$ ; adjusted  $R^2 = 0.60, p = 1.13 \times 10^{-4}$ ).

Color Trails 2 was predicted by spatial span mean sequence length ( $p = 0.002$ ). The robot-based predictor accounted for 36% of the variance in the model, and the predicted scores strongly correlated with actual Color Trails 2 scores ( $\rho = 0.71, p = 3.34 \times 10^{-4}$ ; adjusted  $R^2 = 0.36, p = 0.002$ ).

Digit Symbol Coding was predicted by spatial span mean sequence length ( $p = 1.83 \times 10^{-5}$ ). The robot-based predictor accounted for 61% of the variance in the model, and the predicted scores strongly correlated with actual Digit Symbol Coding scores ( $\rho = 0.81, p = 7.06 \times 10^{-6}$ ; adjusted  $R^2 = 0.61, p = 1.83 \times 10^{-5}$ ).

MoCA was predicted by spatial span mean sequence length ( $p = 0.003$ ). The robot-based predictor accounted for 34% of the variance in the model, and the predicted scores moderately correlated with actual MoCA scores ( $\rho = 0.64, p = 0.002$ ; adjusted  $R^2 = 0.34, p = 0.003$ ).

MoCA-EF was predicted by spatial span mean sequence length ( $p = 5.30 \times 10^{-4}$ ). The robot-based predictor accounted for 45% of the variance in the model, and the predicted

scores strongly correlated with actual MoCA-EF scores ( $\rho = 0.70$ ,  $p = 4.07 \times 10^{-4}$ ; adjusted  $R^2 = 0.45$ ,  $p = 5.30 \times 10^{-4}$ ).

IHDS was predicted by spatial span performance ( $p = 5.31 \times 10^{-4}$ ). The robot-based predictor accounted for 45% of the variance in the model, and the predicted scores moderately correlated with actual IHDS scores ( $\rho = 0.52$ ,  $p = 0.02$ ; adjusted  $R^2 = 0.45$ ,  $p = 5.30 \times 10^{-4}$ ).

**2) Non-Dominant Limb Predictors:** Fig. 6 shows the linear regression models for each of the clinical cognitive assessments using non-dominant limb robot-based metrics as the predictors.

Color Trails 1 was predicted by spatial span mean sequence length ( $p = 0.001$ ). The robot-based predictor accounted for 39% of the variance in the model, and the predicted scores strongly correlated with actual Color Trails 1 scores ( $\rho = 0.70$ ,  $p = 3.73 \times 10^{-4}$ ; adjusted  $R^2 = 0.39$ ,  $p = 0.001$ ).

Color Trails 2 was predicted by N-back performance ( $p = 0.003$ ). The robot-based predictor accounted for 35% of the variance in the model, and the predicted scores moderately correlated with actual Color Trails 2 scores ( $\rho = 0.68$ ,  $p = 7.79 \times 10^{-4}$ ; adjusted  $R^2 = 0.35$ ,  $p = 0.003$ ).

Digit Symbol Coding was predicted by spatial span mean sequence length ( $p = 1.51 \times 10^{-4}$ ). The robot-based predictor accounted for 51% of the variance in the model, and the predicted scores strongly correlated with actual Digit Symbol Coding scores ( $\rho = 0.76$ ,  $p = 6.38 \times 10^{-5}$ ; adjusted  $R^2 = 0.51$ ,  $p = 1.51 \times 10^{-4}$ ).

MoCA was predicted by N-back performance ( $p = 0.007$ ). The robot-based predictor accounted for 28% of the variance in the model, and the predicted scores weakly correlated with actual MoCA scores ( $\rho = 0.48$ ,  $p = 4.07 \times 10^{-4}$ ; adjusted  $R^2 = 0.28$ ,  $p = 0.007$ ).

MoCA-EF was predicted by spatial span mean sequence length ( $p = 0.001$ ). The robot-based predictor accounted for 41% of the variance in the model, and the predicted scores moderately correlated with actual MoCA-EF scores ( $\rho = 0.68$ ,  $p = 7.00 \times 10^{-4}$ ; adjusted  $R^2 = 0.41$ ,  $p = 0.001$ ).

IHDS was predicted by a combination of trajectory tracking smoothness and spatial span smoothness ( $p = 9.76 \times 10^{-5}$  and  $0.02$ , respectively). The robot-based predictors accounted for 53% of the variance in the model, and the predicted scores strongly correlated with actual IHDS scores ( $\rho = 0.80$ ,  $p = 1.46 \times 10^{-5}$ ; adjusted  $R^2 = 0.53$ ,  $p = 4.12 \times 10^{-4}$ ).

## F. Estimating Clinical Motor Scores (Hypothesis 3)

**1) Dominant Limb Predictors:** Fig. 5 shows the linear regression models for each of the clinical motor assessments using dominant limb robot-based metrics as the predictors.

Dominant limb BBT was predicted by a combination of trajectory tracking normalized distance traversed and spatial span normalized distance traversed ( $p = 0.003$  and  $0.02$ ,

respectively). The robot-based predictors accounted for 53% of the variance in the model, and the predicted scores strongly correlated with actual BBT scores ( $\rho = 0.74$ ,  $p = 1.46 \times 10^{-4}$ ; adjusted  $R^2 = 0.53$ ,  $p = 4.72 \times 10^{-4}$ ).

Dominant limb GP was predicted by trajectory tracking performance ( $p = 0.002$ ). The robot-based predictor accounted for 38% of the variance in the model, and the predicted scores moderately correlated with actual GP scores ( $\rho = 0.58$ ,  $p = 0.006$ ; adjusted  $R^2 = 0.38$ ,  $p = 0.002$ ).

Dominant limb grip strength was predicted by spatial span normalized distance traversed, but it was neither a significant predictor nor correlated to actual grip strength scores ( $\rho = 0.29$ ,  $p = 0.20$ ; adjusted  $R^2 = 0.10$ ,  $p = 0.09$ ).

**2) Non-Dominant Limb Predictors:** Fig. 6 shows the linear regression models for each of the clinical motor assessments using non-dominant limb robot-based metrics as the predictors.

Non-dominant limb BBT was predicted by a combination of trajectory tracking normalized distance traversed and spatial span mean sequence length ( $p = 0.002$  and  $0.01$ , respectively). The robot-based predictors accounted for 64% of the variance in the model, and the predicted scores strongly correlated with actual BBT scores ( $\rho = 0.71$ ,  $p = 3.41 \times 10^{-4}$ ; adjusted  $R^2 = 0.64$ ,  $p = 4.44 \times 10^{-5}$ ).

Non-dominant limb GP was predicted by trajectory tracking normalized distance traversed ( $p = 0.005$ ). The robot-based predictor accounted for 31% of the variance in the model while the predicted scores were not significantly correlated with actual GP scores ( $\rho = 0.29$ ,  $p = 0.21$ ; adjusted  $R^2 = 0.31$ ,  $p = 0.005$ ).

Non-dominant limb grip strength was predicted by spatial span performance ( $p = 0.03$ ). The robot-based predictor accounted for 19% of the variance in the model, and the predicted scores weakly correlated with actual grip strength scores ( $\rho = 0.44$ ,  $p = 0.04$ ; adjusted  $R^2 = 0.19$ ,  $p = 0.03$ ).

## IV. Discussion

### A. Gross Motor Impairments Are Prevalent in PLWH

This study aimed to use a robot-based approach to explore objective measures of cognitive and motor impairment in HIV and HIV-stroke populations. The HIV and HIV-stroke groups displayed no significant differences in clinical or robot-based scores. Subjects in both the HIV and HIV-stroke groups demonstrated mild to moderate impairment in executive function, information processing, and upper limb fine and gross motor domains relative to published population normal performance values in uninfected populations. These results are consistent with previous research demonstrating impairments in these domains in PLWH [6], [10], [43].

We found it notable that the HIV-only group demonstrated not only fine motor impairment as previously reported in the literature [7]–[9], [11], but also gross upper limb motor

impairment. Gross motor impairment has generally been considered a pre-ART era manifestation of HIV infection, and studies since then have focused on the fine motor deficits that result from HIV [7]. Moderate bilateral gross motor impairment, as measured by the BBT and adjusted to healthy population norms, was present in 7 of 15 subjects in the HIV group. The prevalence of moderate bilateral fine motor impairment in the HIV-only subjects in this study (5 out of 15), as measured by the GP, is higher than what was reported in Wilson *et al.* (2 out of 12) in a group of PLWH with a similar average age of 57.9 years old [7]. These results suggest that gross upper limb motor impairments may be an overlooked effect of chronic HIV and that the BBT can be used to identify these impairments as an alternative to the GP. This approach could be useful when examining patients with both HIV and stroke in particular, when motor impairments may be more prevalent [44].

## B. Robot-Based Metrics Capture Differences in Functional Subgroups

A wide range of impairments was observed in the subject population and there was no clear separation between the HIV and HIV-stroke groups on either the clinical assessments or robot-based metrics. As such, subjects were classified into one of four functional groups by their cognitive and motor performance. The results provide evidence in support of the study's first hypothesis that robot-based metrics can differentiate subjects with and without moderate executive function or upper-limb motor impairments.

Subjects with moderate executive function impairment, regardless of motor status, performed worse on the N-back compared to subjects with low cognitive and low motor impairment. These results suggest the robot-based N-back can be used to isolate executive function deficits. This is consistent with previous findings that the paper-based N-back test, although specifically a test for working memory, engages executive function domains impacted by HIV [45].

Subjects with moderate executive function and moderate gross motor impairments performed worse on the robot-based trajectory tracking task compared to subjects with low cognitive impairment, regardless of motor status. This suggests that there might be a cognitive component to the trajectory tracking task that exacerbates performance error in the presence of executive function impairments.

Similarly to the robot-based N-back, subjects with moderate executive function impairment, regardless of motor status, had shorter sequences on the robot-based spatial span task compared to subjects with low cognitive and low motor impairment. Additionally, subjects with low cognitive and moderate motor impairment performed better than subjects with moderate cognitive and moderate motor impairment. These results suggest that a robot-based spatial span task can be used to detect executive function impairment, even in the presence of moderate motor impairment.

Together, these robot-based metrics provide a set of measures that are able distinguish between certain functional groups. Going forward, these represent a potential set of objective metrics that can be used to track longitudinal performance that relate to functional

status in PLWH, stroke, and other conditions presenting with both motor and cognitive impairments.

### C. Robot-Based Metrics Relate to HIV-Related Clinical Assessments

To our knowledge, we are the first group to explore objective robot-based measures of both motor and cognitive impairments in PLWH. This study is a first step in developing more targeted neurorehabilitation strategies for PLWH exhibiting both motor and cognitive decline. The results support the study's second hypothesis and show that both individual and linear combinations of robot-based metrics can successfully estimate clinical cognitive scores. The regression models for Color Trails 1, Digit Symbol–Coding, MoCA–EF and IHDS (adjusted  $R^2 = 0.41–0.60$ ) — excluding the non-dominant limb model for Color Trails 1 — performed the best, exceeding the effect size for which the study was powered. The robot-based measures also demonstrated statistically significant relationships with Color Trails 2 and MoCA.

This is one of the first studies to establish objective robot-based measures that relate to Digit Symbol–Coding, MoCA-EF subscores, or IHDS. Given that the Digit Symbol–Coding, MoCA-EF, and IHDS look at more specific cognitive domains related to executive function, this suggests the potential of robot-based metrics to identify more specific impairments going forward that are relevant to PLWH. Notably, the robot-based metrics that best predicted these clinical scores were consistent with the robot-based metrics that showed differences between functional groups.

Two other studies that examine the relationship between robotic metrics and MoCA scores in stroke and traumatic brain injury populations reported correlation coefficients ranging between 0.49 and 0.65 that are similar to the values observed in this study ( $\rho = 0.48–0.64$ ) [21], [22].

The results provide evidence that robot-based metrics can successfully estimate clinical motor scores in PLWH. The dominant and non-dominant limb models for BBT scores (adjusted  $R^2 = 0.53$  and  $0.64$ , respectively) performed the best, exceeding the effect size for which the study was powered and demonstrating strong correlations between predicted and actual scores. Using a multiple linear regression with eight robotic predictors derived from three tasks, Bosecker *et al.* demonstrated correlation coefficients between estimated and actual scores for the Fugl-Meyer, Motor Status Score, Motor Power Scale, and Modified Ashworth Scale of  $0.42–0.80$  on training models [23]. While the clinical motor metrics differed from those used in this study, these values were similar for the dominant and non-dominant BBT and GP models ( $\rho = 0.31–0.74$ , respectively) with fewer predictors.

While computerized versions of the spatial span exist [42], the robotic aspect implemented in this study allows for kinematic measures to be observed that are reflective of motor function. This enables more detailed study of the interactions between cognitive and motor domains. The utility of this task can be seen by the high prevalence of metrics from this task demonstrating strong relationships with both cognitive and motor clinical scores.

#### D. Relevance to HAND Assessment, Neurorehabilitation, Global Health, and Robotics

Taken together, these results show the potential clinical utility of a robotics-based approach to assess motor and cognitive function in PLWH. Due to the involved nature of performing a complete HAND assessment, other alternatives have been explored to capture HIV-related neurocognitive impairments. For example, Fogel *et al.* used a stepwise multiple linear regression approach to predict a global deficit score (GDS) from a set of 24 metrics extracted from basic medical history in an older HIV population with an average age of  $61.1 \pm 4.6$  years, which was similar to the average of the HIV-only population in this study ( $56.2 \pm 5.4$  years old) [46]. The GDS was calculated from a set of neuropsychological tests encompassing working memory and memory, motor, information processing, and learning domains that overlapped with some of the assessments in this study – specifically the GP, Trail Making A (equivalent to the Color Trails 1), and Digit Symbol–Coding. The ultimate three-term model from the Fogel *et al.* study had a  $R^2$  of 0.29, which is weaker compared to the  $R^2$  values for the Color Trails 1, Digit Symbol–Coding, and GP models in this study ( $R^2 = 0.31–0.60$ ) [46]. In Botswana, a lower-resource setting, a six-part neurocognitive battery, which also utilizes many of the same assessments as this study, was used to identify impairments in cognitive-motor areas in PLWH [8].

From a clinical rehabilitation perspective, increasing access to effective rehabilitation interventions and enhancing outcome measurement have been identified as research priorities in HIV, disability, and rehabilitation [47]. There is a need to develop interventions addressing the rapid aging and frailty associated with HIV to reduce disparities in health outcomes that can compound in the presence of other comorbidities or complications. No gold standard exists to capture the relationship between cognitive impairment and physical frailty as it relates to HIV [48]. While a limited number of studies have shown that physical exercise can induce improvements in physical, cognitive, and emotional wellbeing in both HIV and non-HIV populations, there is a need for further work to understand what impact exercise – including robot-based exercise – might have on the aging immune system in PLWH. A benefit to the objective quantification used in this study is the ability to track changes during the course of rehabilitation with specific metrics. This approach can be practical within a neurorehabilitation context because the metrics are reflective of clinically-relevant tests and can be administered in a less time-intensive way.

From a global health perspective, this technology-based approach provides a possible scalable strategy that is sensitive to subtle signs of functional decline. With more affordable rehabilitation robot systems becoming increasingly available, this approach has the potential to meet a huge rehabilitation need in lower resource settings where the capacity to supply additional rehabilitation professionals is lacking but the prevalence of non-communicable diseases necessitating rehabilitation is increasing [25]. This would be valuable particularly when medical history may be lacking or harder to assess. This preliminary work lays the groundwork for identifying specific impairments and developing HIV-specific neurorehabilitation strategies to address the various cognitive and motor impairments associated with aging with HIV. Our group is currently exploring this in Botswana.

From a robotics perspective, this study expands the application of rehabilitation robotics beyond stroke to PLWH and those living with cognitive impairments. Given that

neurocognitive impairment is associated with instrumental ADL function [49], assessments and treatments should reflect the integration of both motor and cognitive domains that are often assessed in isolation. Like other robotic studies, large effect sizes were observed in this study, which can significantly reduce the sample size needed for clinical trials going forward [24]. This study also shows that clinical measures can be estimated from both limbs, which can be helpful in avoiding confounding factors, such as the presence of unilateral motor impairment that could result from stroke. Although these results do not provide enough information to generalize to other neurological conditions, this approach allows for future studies on other neurological conditions because it is rooted in standard clinical assessments used in other populations beyond HIV and stroke.

## E. Study Limitations

Given the small sample size, lack of control group (either non-HIV healthy control or non-HIV stroke group), and predominance of Black persons within the HIV group, we may not be able to fully generalize these results. Although, the sample population is small, we were able to see significant differences and the population was reflective of the aging HIV population in the U.S. While we observed strong correlations between robot-based measures and clinical cognitive and motor assessments relevant to the HIV population, correlation studies are susceptible to the distribution of the data across the span of the predictor variables. While we had adequate distribution across many variables, we were not able to get an even distribution across functional groups, which could have biased the analysis. Despite these limitations, further studies with a larger sample size and a longitudinal evaluation of this approach is warranted.

## Acknowledgment

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors would like to acknowledge Kristine Lima and Christine Kurian for their contributions to the study.

This work was supported by the National Institute of Neurological Disorders and Stroke, National Institutes of Health under Grant T32NS091006; the University of Pennsylvania's Center for AIDS Research under Grant P30AI 045008; by the Global E3 GIRE program (NSF grant number 1829436); and by the University of Pennsylvania's Departments of Bioengineering and Physical Medicine and Rehabilitation.

## References

- [1]. The Gap Report 2014, UNAIDS, Geneva, Switzerland, 2014.
- [2]. Gill AJ and Kolson DL, "Chronic inflammation and the role for cofactors (hepatitis C, drug abuse, antiretroviral drug toxicity, aging) in HAND persistence," *Current HIV/AIDS Rep*, vol. 11, no. 3, pp. 325–335, 9. 2014.
- [3]. Heaton RK et al., "The impact of HIV-associated neuropsychological impairment on everyday functioning," *J. Int. Neuropsychol. Soc.*, vol. 10, no. 3, pp. 317–331, 2004, doi: 10.1017/S1355617704102130. [PubMed: 15147590]
- [4]. Rusch M, Nixon S, Schilder A, Braitstein P, Chan K, and Hogg RS, "Impairments, activity limitations and participation restrictions: Prevalence and associations among persons living with HIV/AIDS in British Columbia," *Health Qual. Life Outcomes*, vol. 2, no. 1, p. 46, 2004. [PubMed: 15350202]



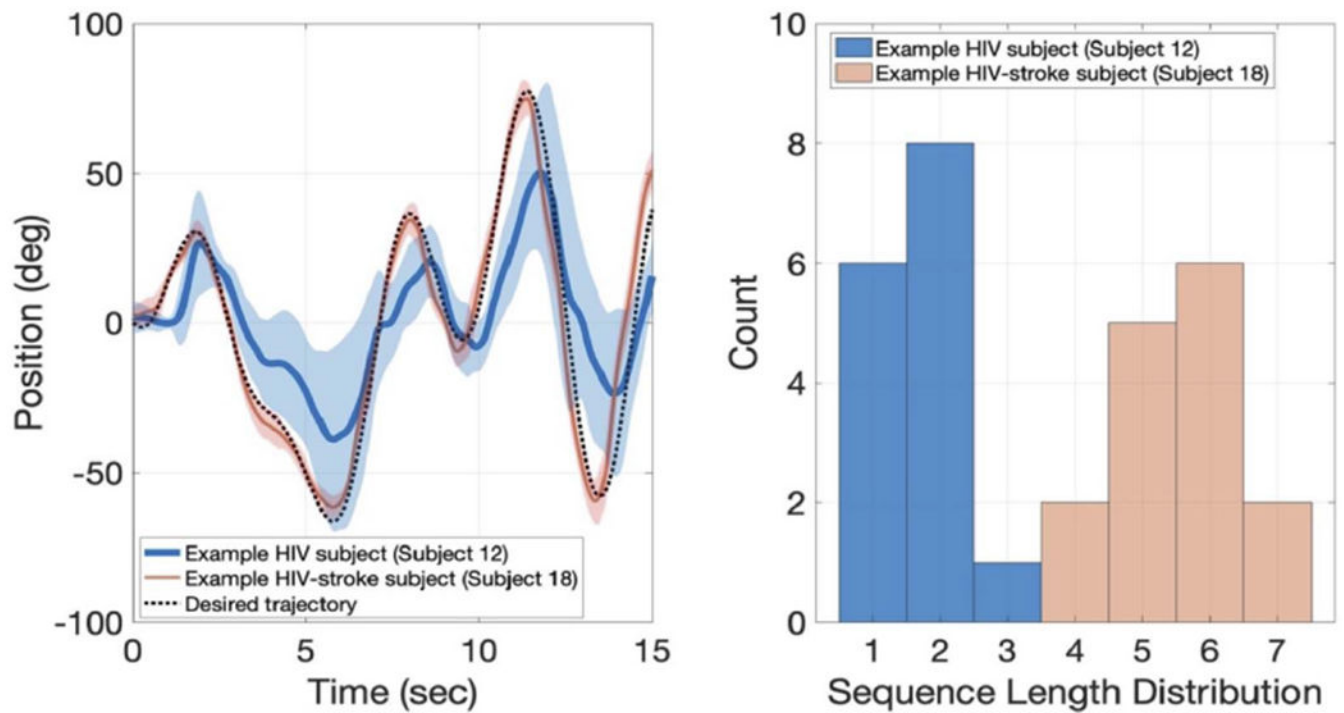
- [5]. Heaton RK et al., “HIV-associated neurocognitive disorders before and during the era of combination antiretroviral therapy: Differences in rates, nature, and predictors,” *J. Neurovirol*, vol. 17, no. 1, pp. 3–16, 2011. [PubMed: 21174240]
- [6]. Pullen S, Chigbo N, Chukwudi NE, Chim AC, Chinwe C, and Chukwunoso IS, “Physiotherapy intervention as a complementary treatment for people living with HIV/AIDS,” *HIV/AIDS Res. Palliative Care*, vol. 6, p. 99, 6. 2014.
- [7]. Wilson TW et al., “Functional brain abnormalities during finger-tapping in HIV-infected older adults: A magnetoencephalography study,” *J. Neuroimmune Pharmacol*, vol. 8, no. 4, pp. 965–974, 9. 2013. [PubMed: 23749418]
- [8]. Lawler K et al., “Neurobehavioral effects in HIV-positive individuals receiving highly active antiretroviral therapy (HAART) in Gaborone, Botswana,” *PLoS ONE*, vol. 6, no. 2, 2. 2011, Art. no. e17233.
- [9]. Robinson-Papp J et al., “Characteristics of motor dysfunction in longstanding human immunodeficiency virus,” *Clin. Infectious Diseases*, vol. 71, no. 6, pp. 1532–1538, 9. 2020.
- [10]. Antinori A et al., “Updated research nosology for HIV-associated neurocognitive disorders,” *Neurology*, vol. 69, no. 18, pp. 1789–1799, 10. 2007. [PubMed: 17914061]
- [11]. Kronemer SI, Mandel JA, Sacktor NC, and Marvel CL, “Impairments of motor function while multitasking in HIV,” *Frontiers Hum. Neurosci*, vol. 11, p. 212, 4. 2017.
- [12]. Scott JC et al., “A neuropsychological investigation of multitasking in HIV infection: Implications for everyday functioning,” *Neuropsychology*, vol. 25, no. 4, p. 511, 2011. [PubMed: 21401259]
- [13]. Chow FC, Regan S, Feske S, Meigs JB, Grinspoon SK, and Triant VA, “Comparison of ischemic stroke incidence in HIV-infected and non-HIV-infected patients in a US health care system,” *J. Acquired Immune Deficiency Syndromes*, vol. 60, no. 4, p. 351, 2012.
- [14]. Vinikoor MJ, Napravnik S, Floris-Moore M, Wilson S, Huang DY, and Eron JJ, “Incidence and clinical features of cerebrovascular disease among HIV-infected adults in the southeastern united states,” *AIDS Res. Hum. Retroviruses*, vol. 29, no. 7, pp. 1068–1074, 7. 2013. [PubMed: 23565888]
- [15]. Augustyn H, de Witt P, and Franzsen D, “The effect of HIV status on post-stroke outcomes in personal activities of daily living,” *Brit. J. Occupational Therapy*, vol. 83, no. 11, 2020, Art. no. 0308022620902683.
- [16]. Weber E, Blackstone K, and Woods SP, “Cognitive neurorehabilitation of HIV-associated neurocognitive disorders: A qualitative review and call to action,” *Neuropsychol. Rev*, vol. 23, no. 1, pp. 81–98, 3. 2013. [PubMed: 23417497]
- [17]. deBoer H et al., “Where and how does physical therapy fit? Integrating physical therapy into interprofessional HIV care,” *Disability Rehabil*, vol. 41, no. 15, pp. 1768–1777, 7. 2019.
- [18]. Bui KD and Johnson MJ, “Designing robot-assisted neurorehabilitation strategies for people with both HIV and stroke,” *J. NeuroEng. Rehabil*, vol. 15, no. 1, p. 75, 12. 2018. [PubMed: 30107849]
- [19]. Lo AC et al., “Robot-assisted therapy for long-term upper-limb impairment after stroke,” *New England J. Med*, vol. 362, no. 19, pp. 1772–1783, 5 2010. [PubMed: 20400552]
- [20]. Tran VD, Dario P, and Mazzoleni S, “Kinematic measures for upper limb robot-assisted therapy following stroke and correlations with clinical outcome measures: A review,” *Med. Eng. Phys*, vol. 53, pp. 13–31, 3. 2018. [PubMed: 29361407]
- [21]. Bourke TC, Lowrey CR, Dukelow SP, Bagg SD, Norman KE, and Scott SH, “A robot-based behavioural task to quantify impairments in rapid motor decisions and actions after stroke,” *J. NeuroEng. Rehabil*, vol. 13, no. 1, p. 91, 12. 2016. [PubMed: 27724945]
- [22]. Logan LM, Semrau JA, Debert CT, Kenzie JM, Scott SH, and Dukelow SP, “Using robotics to quantify impairments in sensorimotor ability, visuospatial attention, working memory, and executive function after traumatic brain injury,” *J. Head Trauma Rehabil*, vol. 33, no. 4, pp. E61–E73, 2018.
- [23]. Bosecker C, Dipietro L, Volpe B, and Krebs HI, “Kinematic robot-based evaluation scales and clinical counterparts to measure upper limb motor performance in patients with chronic stroke,” *Neurorehabilitation Neural Repair*, vol. 24, no. 1, pp. 62–69, 1. 2010. [PubMed: 19684304]

- [24]. Krebs HI et al., “Robotic measurement of arm movements after stroke establishes biomarkers of motor recovery,” *Stroke*, vol. 45, no. 1, pp. 200–204, 1. 2014. [PubMed: 24335224]
- [25]. Johnson MJ, Rai R, Barathi S, Mendonca R, and Bustamante-Valles K, “Affordable stroke therapy in high-, low- and middle-income countries: From theradrive to rehab cares, a compact robot gym,” *J. Rehabil. Assistive Technol. Eng.*, vol. 4, 5 2017, Art. no. 2055668317708732.
- [26]. Beck A, Steer R, and Brown G, “Manual for the beck depression inventory-fast screen for medical patients,” *Psychol. Corp.*, San Antonio, TX, USA, Tech. Rep., 2000.
- [27]. D’Elia L, Satz P, Uchiyama CL, and White T, “Color trails test,” *Tech. Rep.*, 1996.
- [28]. Wechsler D, “Wechsler adult intelligence scale-revised (WAIS-R),” *Psychol. Corp.*, San Antonio, TX, USA, Tech. Rep., 1981.
- [29]. Nasreddine ZS et al., “The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment,” *J. Amer. Geriatrics Soc.*, vol. 53, no. 4, pp. 695–699, 4. 2005.
- [30]. Sacktor NC et al., “The international HIV dementia scale: A new rapid screening test for HIV dementia,” *Aids*, vol. 19, no. 13, pp. 1367–1374, 2005. [PubMed: 16103767]
- [31]. Maj M, “Evaluation of two new neuropsychological tests designed to minimize cultural bias in the assessment of HIV-1 seropositive persons: A WHO study,” *Arch. Clin. Neuropsychol.*, vol. 8, no. 2, pp. 123–135, 4. 1993. [PubMed: 14589670]
- [32]. Ettenhofer ML et al., “Aging, neurocognition, and medication adherence in HIV infection,” *Amer. J. Geriatric Psychiatry*, vol. 17, no. 4, pp. 281–290, 4. 2009.
- [33]. Fazeli PL, Casaletto KB, Paolillo E, Moore RC, Moore DJ, and HNRP Group, “Screening for neurocognitive impairment in HIV-positive adults aged 50 years and older: Montreal cognitive assessment relates to self-reported and clinician-rated everyday functioning,” *J. Clin. Experim. Neuropsychol.*, vol. 39, no. 9, pp. 842–853, 10. 2017.
- [34]. Lam B et al., “Criterion and convergent validity of the montreal cognitive assessment with screening and standardized neuropsychological testing,” *J. Amer. Geriatrics Soc.*, vol. 61, no. 12, pp. 2181–2185, 12. 2013.
- [35]. Mathiowetz V, Volland G, Kashman N, and Weber K, “Adult norms for the box and block test of manual dexterity,” *Amer. J. Occupational Therapy*, vol. 39, no. 6, pp. 386–391, 6. 1985.
- [36]. Ruff RM and Parker SB, “Gender- and age-specific changes in motor speed and eye-hand coordination in adults: Normative values for the finger tapping and grooved pegboard tests,” *Perceptual Motor Skills*, vol. 76, no. 3, pp. 1219–1230, 6. 1993. [PubMed: 8337069]
- [37]. Schrack JA et al., “Effect of HIV-infection and cumulative viral load on age-related decline in grip strength,” *AIDS*, vol. 30, no. 17, p. 2645, 2016. [PubMed: 27603294]
- [38]. Bui KD, Rai R, and Johnson MJ, “Using upper limb kinematics to assess cognitive deficits in people living with both HIV and stroke,” in *Proc. Int. Conf. Rehabil. Robot. (ICORR)*, 7. 2017, pp. 907–912.
- [39]. Balasubramanian S, Melendez-Calderon A, Roby-Brami A, and Burdet E, “On the analysis of movement smoothness,” *J. NeuroEng. Rehabil.*, vol. 12, no. 1, p. 112, 12. 2015. [PubMed: 26651329]
- [40]. Owen AM, Mcmillan KM, Laird AR, and Bullmore E, “N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies,” *Hum. Brain Mapping*, vol. 25, no. 1, pp. 46–59, 2005.
- [41]. Kessels RPC, van Zandvoort MJE, Postma A, Kappelle LJ, and de Haan EHF, “The Corsi block-tapping task: Standardization and normative data,” *Appl. Neuropsychol.*, vol. 7, no. 4, pp. 252–258, 12. 2000. [PubMed: 11296689]
- [42]. Brunetti R, Del Gatto C, and Delogu F, “ECorsi: Implementation and testing of the Corsi block-tapping task for digital tablets,” *Frontiers Psychol.*, vol. 5, p. 939, 9. 2014.
- [43]. Fellows RP, Byrd DA, and Morgello S, “Effects of information processing speed on learning, memory, and executive functioning in people living with HIV/AIDS,” *J. Clin. Experim. Neuropsychol.*, vol. 36, no. 8, pp. 806–817, 9. 2014.
- [44]. Elicer IM, Byrd D, Clark US, Morgello S, and Robinson-Papp J, “Motor function declines over time in human immunodeficiency virus and is associated with cerebrovascular disease, while HIV-associated neurocognitive disorder remains stable,” *J. Neuro Virol.*, vol. 24, no. 4, pp. 514–522, 8. 2018. [PubMed: 29696578]

- [45]. Cohen RA et al., “Neural response to working memory demand predicts neurocognitive deficits in HIV,” *J. NeuroVirol*, vol. 24, no. 3, pp. 291–304, 6. 2018. [PubMed: 29280107]
- [46]. Fogel GB et al., “Factors related to HIV-associated neurocognitive impairment differ with age,” *J. NeuroVirol*, vol. 21, no. 1, pp. 56–65, 2. 2015. [PubMed: 25404233]
- [47]. O’Brien KK et al., “Advancing research and practice in HIV and rehabilitation: A framework of research priorities in HIV, disability and rehabilitation,” *BMC Infectious Diseases*, vol. 14, no. 1, p. 724, 12. 2014. [PubMed: 25551619]
- [48]. Piggott DA, Erlandson KM, and Yarasheski KE, “Frailty in HIV: Epidemiology, biology, measurement, interventions, and research needs,” *Current HIV/AIDS Rep*, vol. 13, no. 6, pp. 340–348, 12. 2016.
- [49]. Johs NA et al., “Disability among middle-aged and older persons with HIV infection,” *Clin. Infectious Diseases*, vol. 65, no. 1, pp. 83–91, 2017.

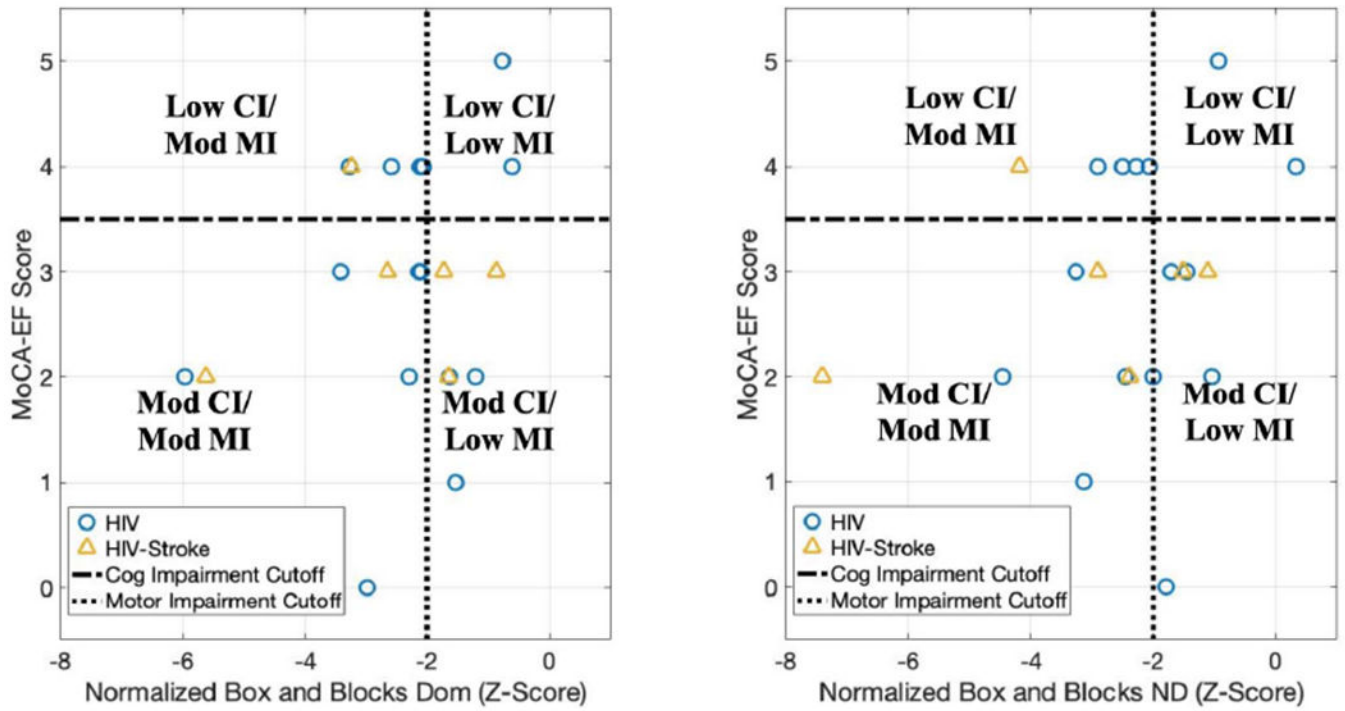


**Fig. 1.**  
The Haptic Theradrive, a one degree-of-freedom rehabilitation robot system used in this study. Image used with permission from [38].

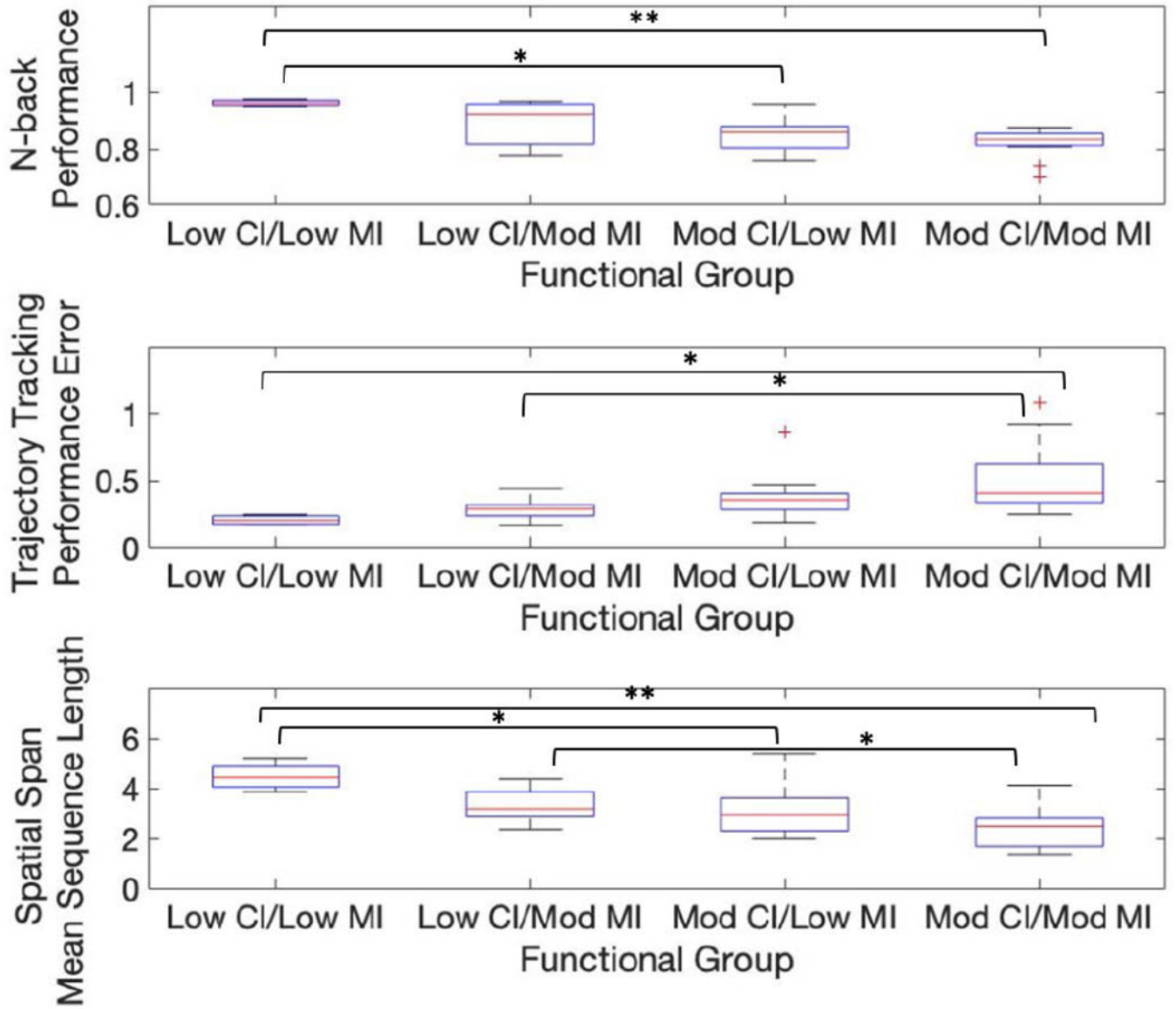


**Fig. 2.**

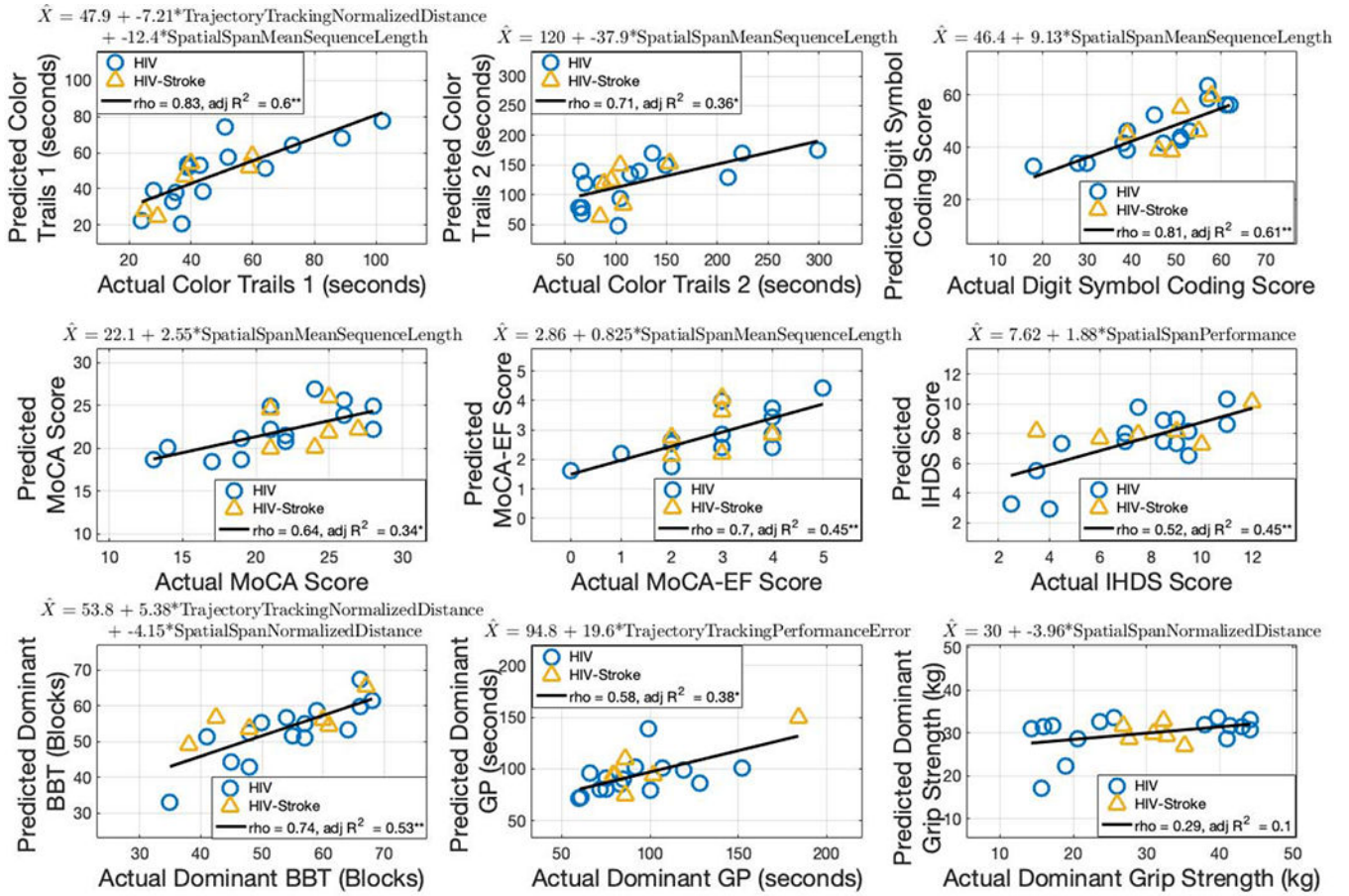
Left: The mean trajectory from the trajectory tracking task is shown for an example HIV (blue) and HIV-stroke (pink) subject. The expected trajectory is shown as a black dotted line. The shaded region represents the standard deviation across all the trials. Right: Histograms showing the distribution of sequence lengths on the spatial span task for the same HIV and HIV-stroke subject.



**Fig. 3.** Distribution of subjects by cognitive and motor function, using a score of 3.5 for the MoCA-EF cutoff and -2 as the BBT Z-score cutoff. The left figure is the distribution using the dominant limb BBT scores, while the right is from non-dominant limb BBT scores. (CI = cognitive impairment; MI = motor impairment; mod = moderate).

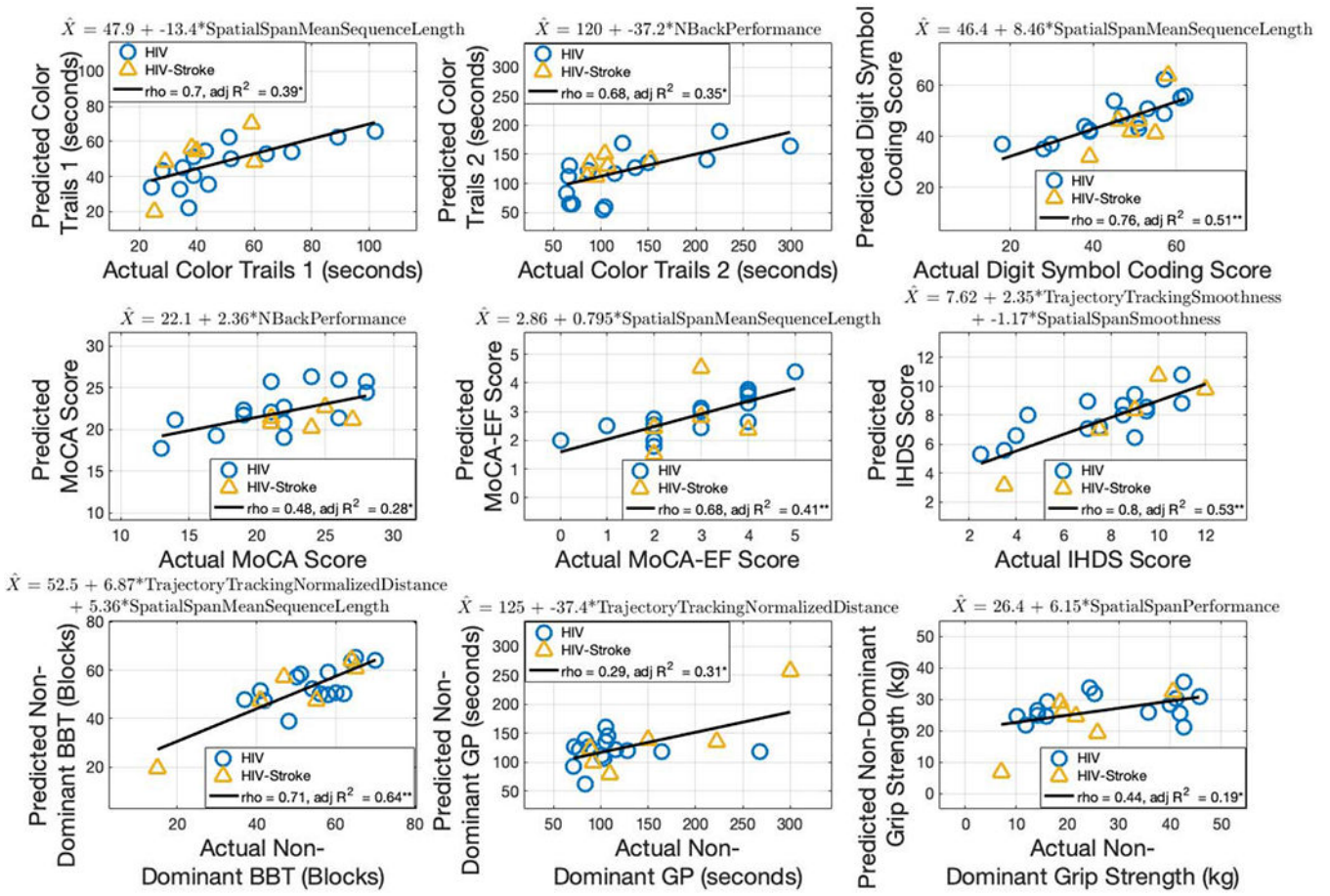


**Fig. 4.** Box plots for each of the functional subgroups on N-back performance (top), trajectory tracking performance error (middle), and spatial span mean sequence length (bottom). CI = cognitive impairment; MI = motor impairment (\* =  $p < 0.05$ , \*\* =  $p < 0.005$  after correcting for multiple comparisons).



**Fig. 5.** Multiple linear regression for clinical assessments using dominant limb robot-based metrics. The robot-based predictors for each model are included in the equation at the top of each subplot. Spearman's rho and adjusted R<sup>2</sup> are shown. (\* = p < 0.05, \*\* = p < 0.001).





**Fig. 6.** Multiple linear regression for clinical assessments using non-dominant limb robot-based metrics. The robot-based predictors for each model are included in the equation at the top of each subplot. Spearman's rho and adjusted R<sup>2</sup> are shown. (\* = p < 0.05, \*\* = p < 0.001).

**TABLE I:**

## Subject Demographics and Clinical Scores

Characteristics	HIV-only mean±standard deviation (n = 15)	HIV-stroke group mean ±standard deviation (n = 6)	Subject population mean ±standard deviation (n = 21)
Age (years old)	56.2 ± 5.4	54.2±8.1	55.5 ±6.3
Gender (Male/Female)	10M/5F	3M/3F	13M/8F
12 years edu (count)	10	5	15
Color Trails 1 (seconds)	50.27±21.74	41.83±13.48	47.86±20.10
Color Trails 2 (seconds)	125.53±67.30	105.50±22.60	119.81±58.85
Digit Symbol–Coding Score	45.07±12.48	49.67±6.16	46.38±11.24
MoCA (out of 30)	21.47±4.43	23.83 ±2.19	22.14±4.06
MoCA-EF (score out of 5)	2.87±1.31	2.83±0.69	2.86±1.17
IHDS (out of 12)	7.47±2.60	8.00±2.75	7.62±2.66
Dominant BBT (blocks)	54.20±9.73	52.75±11.57	53.79±10.01
Non-Dominant BBT (blocks)	54.40±9.54	47.83±18.62	52.52±12.63
Dominant GP (seconds)	91.60±26.36	102.92±40.55	94.83±30.41
Non-Dominant GP (seconds)	111.53±49.49	160.33±84.63	125.48±63.36
Dominant Grip Strength (kg)	29.56±12.15	30.93±3.22	29.95±10.31
Non-Dominant Grip Strength (kg)	28.08±13.60	22.13±10.97	26.38±12.93

**TABLE II:**

Group Robot Performance Results by Dominant (D) and Non-Dominant (ND) Limbs (Mean  $\pm$  Standard Deviation)

Robot Metrics	HIV-only	HIV-Stroke	Subject population
N-back performance	D: 0.86 $\pm$ 0.08 ND: 0.87 $\pm$ 0.07	0.85 $\pm$ 0.07 0.84 $\pm$ 0.03	0.86 $\pm$ 0.07 0.86 $\pm$ 0.06
Trajectory tracking performance	0.34 $\pm$ 0.15 0.34 $\pm$ 0.14	0.44 $\pm$ 0.23 0.54 $\pm$ 0.37	0.37 $\pm$ 0.18 0.39 $\pm$ 0.24
Trajectory tracking normalized distance traversed	1.01 $\pm$ 0.11 1.04 $\pm$ 0.10	1.06 $\pm$ 0.07 0.96 $\pm$ 0.29	1.02 $\pm$ 0.10 1.02 $\pm$ 0.17
Trajectory tracking smoothness	-9.19 $\pm$ 1.16 -9.56 $\pm$ 1.23	-10.26 $\pm$ 1.06 -10.41 $\pm$ 2.46	-9.50 $\pm$ 1.21 -9.80 $\pm$ 1.65
Spatial span mean sequence length	2.83 $\pm$ 0.97 3.22 $\pm$ 0.99	2.97 $\pm$ 0.88 3.02 $\pm$ 1.32	2.87 $\pm$ 0.93 3.16 $\pm$ 1.06
Spatial span performance	0.62 $\pm$ 0.13 0.69 $\pm$ 0.09	0.67 $\pm$ 0.06 0.59 $\pm$ 0.20	0.63 $\pm$ 0.12 0.66 $\pm$ 0.13
Spatial span normalized distance traversed	1.57 $\pm$ 0.56 1.61 $\pm$ 0.35	1.58 $\pm$ 0.26 1.50 $\pm$ 0.29	1.57 $\pm$ 0.48 1.59 $\pm$ 0.33
Spatial span smoothness	-2.18 $\pm$ 0.41 -2.35 $\pm$ 0.51	-2.36 $\pm$ 0.42 -2.59 $\pm$ 0.97	-2.23 $\pm$ 0.41 -2.42 $\pm$ 0.65