

Reliable and valid robot-assisted assessments of hand proprioceptive, motor and sensorimotor impairments after stroke

Monika Zbytniewska (✉ monika.zbytniewska@hest.ethz.ch)

ETH Zürich: Eidgenössische Technische Hochschule Zurich <https://orcid.org/0000-0002-2280-7636>

Christoph M. Kanzler

ETH Zürich: Eidgenössische Technische Hochschule Zurich

Lisa Jordan

ETH Zürich: Eidgenössische Technische Hochschule Zurich

Christian Salzmann

Kliniken Schmieder Allensbach: Kliniken Schmieder Stiftung und Co KG

Joachim Liepert

Kliniken Schmieder Allensbach: Kliniken Schmieder Stiftung und Co KG

Olivier Lambercy

ETH Zürich: Eidgenössische Technische Hochschule Zurich

Roger Gassert

ETH Zürich: Eidgenössische Technische Hochschule Zurich

Research

Keywords: robot-assisted assessments; neurorehabilitation; stroke; recovery; somatosensation; hand function

Posted Date: November 18th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-107703/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Journal of NeuroEngineering and Rehabilitation on July 16th, 2021. See the published version at <https://doi.org/10.1186/s12984-021-00904-5>.

RESEARCH

Reliable and valid robot-assisted assessments of hand proprioceptive, motor and sensorimotor impairments after stroke

Monika Zbytewska^{1*}, Christoph M. Kanzler¹, Lisa Jordan¹, Christian Salzmann², Joachim Liepert², Olivier Lamercy¹ and Roger Gassert¹

*Correspondence:

relab.publications@hest.ethz.ch

¹Rehabilitation Engineering Laboratory, Institute of Robotics and Intelligent Systems, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland
Full list of author information is available at the end of the article

Abstract

Background: Neurological injuries such as stroke often differentially impair hand motor and somatosensory function, as well as the interplay between the two, which leads to limitations in performing activities of daily living. However, it is challenging to identify which specific aspects of sensorimotor function are impaired based on conventional clinical assessments that are often insensitive and subjective. In this work we propose and validate a set of robot-assisted assessments aiming at disentangling hand proprioceptive from motor impairments, and capturing their interrelation (sensorimotor impairments).

Methods: For this purpose, a battery of five complementary assessment tasks was implemented on a one degree-of-freedom end-effector robotic platform acting on the index finger metacarpophalangeal joint. Specifically, proprioceptive impairments were assessed using a position matching paradigm. Fast target reaching, range of motion and maximum fingertip force tasks characterized motor function deficits. Finally, sensorimotor impairments were assessed using a dexterous trajectory following task. Clinical feasibility (duration), reliability (intra-class correlation coefficient ICC, smallest real difference SRD) and validity (Kruskal-Wallis test, Spearman correlations ρ) of these tasks were evaluated with 36 sub-acute stroke subjects and 31 age-matched neurologically intact controls.

Results: Eighty-three percent of stroke survivors with varied impairment severity (mild to severe) could complete all robotic tasks (duration: <15 min per tested hand). Further, the study demonstrated good to excellent reliability of the robotic tasks in the stroke population (ICC>0.7, SRD<30%), as well as discriminant validity, as indicated by significant differences (p -value<0.001) between stroke and control subjects on all metrics extracted from the robotic tasks. Concurrent validity was shown through moderate to strong correlations ($\rho = 0.4-0.8$) between the robotic outcome measures and clinical scales. Finally, robotic tasks targeting different deficits were not strongly correlated, thereby presenting complementary information about a patient's impairment profile.

Conclusions: The proposed robot-assisted assessments provide a clinically feasible, reliable, and valid approach to distinctly characterize impairments in hand proprioceptive and motor function, along with the interaction between the two. This opens new avenues to help unravel the contributions of unique aspects of sensorimotor function in post-stroke recovery, as well as to contribute to future developments towards personalized, assessment-driven therapies.

Keywords: robot-assisted assessments; neurorehabilitation; stroke; recovery; somatosensation; hand function

Background

At the level of the hand, somatosensory and motor function, as well as the interplay between the two, are essential for performing dexterous and skillful movements during activities of daily living (ADLs) [1–4]. For example when grasping a small object, proprioception is necessary to sense the current position of the limb [5, 6]. This sensory input is then integrated by the central nervous system to shape the motor output, a process called sensorimotor integration [7, 8]. Subsequently, the motor system is responsible for eliciting and executing the planned movement [9].

Neurological injuries such as stroke often disrupt specific aspects of this process, which consequently prevents affected individuals from performing ADLs [1, 10]. Often the exact impairments that cause activity limitations are unclear, although their detection would be a prerequisite to design appropriate rehabilitation strategies tailored to each patient's impairment profile [11]. Most commonly reported are motor impairments, with 80% of stroke survivors experiencing paresis [12–15]. However, some symptoms that seem to originate from a motor function impairment may be caused by disturbed proprioceptive feedback [7]. Somatosensory function is in fact frequently affected and has been shown to be associated with poor functional recovery and higher activity limitations, although the reporting prevalence varies between 23% and 67% [16–21].

The difficulty in accurately identifying each patients' impairment profile originates, among others, from the lack of sensitive assessment methods [22, 23]. Most widely used clinical assessments are observer-based and subjective, not optimal for providing reproducible stimuli, and prone to floor/ceiling effects [23, 24]. Further, a majority of clinical methods are focused on evaluating activity limitations (e.g. Action Research Arm Test, Box & Block Test [25, 26]) and only a few standardized tests target the underlying impairments (e.g. Fugl-Meyer [22]), making it difficult to understand the cause of decreased performance. Finally, multiple conventional

clinical measures are needed to holistically assess sensorimotor impairment profiles, which is time consuming and hence rarely performed at regular time intervals throughout rehabilitation [27].

Technology-driven solutions provide a promising complement to conventional clinical assessments [1, 11, 28]. Robot-assisted methods are objective (not relying on observer judgement), accurate (e.g. able to measure exact body position/force applied), as well as capable of delivering precise, reproducible stimuli (e.g. to assess sensory function or spasticity [29, 30]). Further, it becomes possible to evaluate different impairments with one single device through multiple robot-assisted assessment tasks, which results in time efficiency and a more comprehensive overview of impairments. Even though they are promising, the existing robotic approaches aiming at concurrent sensory and motor assessment remain in their infancy. The methods proposed so far focus on the proximal joint of the upper limb [31, 32], consist of tasks that target only a specific impairment modality (e.g. proprioception, without the possibility to concurrently assess motor function) [33–36], or fail to provide a detailed evaluation of clinimetric properties of their outcome measures (reliability, measurement error, validity) [11, 37]. Reporting of test-retest reliability and measurement error is essential to understand the sensitivity of an assessment metric to capture different impairments and detect changes over time [37], while the study of concurrent validity is important to relate a new technological approach to the commonly accepted assessment methods [38]. The current lack of standardized evaluations of reliability and validity in the target population makes new assessment technologies less likely to be clinically accepted and applied outside of research projects [39].

The objective of this work was to propose and evaluate clinimetric properties of a new set of assessments of hand proprioceptive, motor and sensorimotor impairments, implemented on a single, previously described robotic platform (ETH MIKE:

Motor Impairment and Kinesthetic Evaluation) [40, 41]. This one degree-of-freedom end-effector device can provide well-controlled movement stimuli to the index finger metacarpophalangeal (MCP) joint and sensitively measure kinematic and kinetic responses. Focusing on assessing motor and somatosensory impairments at the level of the hand is important, since the hand is often impaired after stroke, which leads to limitations in executing many ADLs, especially those requiring fine dexterity, where somatosensation plays a key role [42]. The index finger was selected due to its relevance in many ADLs (grasping, precision grip [43]) and also because the ability to actively move the MCP joint is often presented as an early predictor of functional recovery [44, 45]. Furthermore, from a practical perspective, focusing on a single joint also allows to simplify the technology to a one degree-of-freedom, increasing simplicity and clinical usability. In this paper we propose a battery of five behavioural tasks and their outcome measures, three of which address motor impairments, one targets proprioception and one measures combined sensorimotor deficits. We investigate the reliability and validity of these robot-assisted assessments in a group of 36 participants with stroke and in an age-matched group of 31 neurologically intact controls. We hypothesized that the newly proposed robot-assisted assessment metrics (i) are reliable due to the objective nature of the tasks, their repeatability and the standardized protocol; (ii) allow to distinguish stroke patients from control subjects and identify different impairment profiles; (iii) can separately quantify proprioceptive, motor and sensorimotor impairments and correlate with corresponding clinical scales.

This work aspires to contribute to the field of neurorehabilitation by providing novel objective assessments, which aim at disentangling different aspects of sensorimotor impairments in order to better understand the cause of observed activity limitations. In the long term the proposed robot-assisted assessments intend to help

in designing more effective therapies, as well as in tracking and predicting recovery of patients after neurological injuries.

Methods

Subjects

Thirty-six participants with stroke were recruited for this study among the patients receiving an inpatient neurological rehabilitation at the Kliniken Schmieder Al-lensbach, Germany. Inclusion criteria were: above 18 years old, diagnosis of stroke (ischemic or hemorrhagic), and the ability to passively move the subject's MCP joint by at least 20°. Exclusion criteria were: inability to understand instructions and pain when moving the MCP joint. Moreover, we designed the study to include a maximum of 40% of subjects with intact proprioception as measured by a conventional clinical scale. This design choice was made to allow for validating the newly proposed measure of proprioception. In addition, thirty-one age-matched neurologically intact control subjects were recruited. The inclusion criteria for this group were: right-handed and above 50 years old. The exclusion criteria was any history of neurological, orthopaedic or rheumatologic disease affecting wrist or hand function. In both groups, handedness was assessed using the Edinburgh Handedness Inventory, where stroke subjects were asked to evaluate their pre-stroke handedness retrospectively. All subjects gave written informed consent before participating in the experiment. The study was approved by the ETH Ethics Committee EK 2019-N-108 and the Ethics Commission of Baden-Württemberg F-2016-126 and retrospectively registered as a clinical trial ^[1].

^[1]https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00018681

Robot-assisted assessments

Apparatus

The ETH MIKE (Motor Impairment and Kinesthetic Evaluation) is a one degree of freedom end-effector robot, which can provide well-controlled stimuli to the index finger and sensitively measure subjects' kinematic and kinetic responses [40, 41]. The end-effector has its center of rotation aligned with the MCP joint of the index finger. Subjects are seated in front of the device, the hand is placed grasping an easily exchangeable, 3D printed handle, and the index finger is stretched and attached to the end-effector via Velcro straps (Figure 1a-1b). The device is suitable to test both hands, one hand at a time. A tablet computer with a touch screen is placed directly above the hand, displaying a Graphical User Interface (GUI) programmed in Unity (Unity Technologies, California, USA), that is used as a visual display during the assessment tasks. To minimize cognitive load, the GUI displays a simple gauge with colored indicators for all assessment tasks (Figure 1c).

The hardware of the robot consists of one actuator (a DC motor), as well as an incremental encoder, a tachometer and a force sensor. The device is controlled by a real-time embedded board (myRIO, National Instruments, Texas, USA) and programmed in LabVIEW (National Instruments, Texas, USA). The end-effector position, velocity and interaction force signals are recorded at a sampling frequency of 1 kHz. Velocity and force signals are smoothed in real-time using a 1st order Butterworth low-pass filter with 20 Hz cutoff frequency. Post-processing is implemented in MATLAB (Mathworks Inc., Massachusetts, USA).

Robotic assessment battery

The battery of robot-assisted assessments consists of tasks targeting proprioceptive and motor impairments, as well as the interplay between the two (sensorimotor impairments). Specifically, there is one task for proprioception assessment (gauge position matching), three tasks focused on motor impairments (range of motion,

maximum force generation and fast target reaching). Additionally, one task is designed to evaluate the ability to integrate proprioceptive information to execute a complex movement, i.e. sensorimotor impairments (trajectory following). All of these tasks are performed on the robotic platform ETH MIKE. A motivation from related literature, the task procedures, and sensor-based metrics extracted from each task are described below.

Gauge position matching task - assessment of proprioceptive impairments: the objective of this assessment is to evaluate the MCP joint proprioception, while minimizing possible confounds coming from motor impairments. The task is based on previous studies that optimized the gauge position matching task procedure [46, 47]. After the tested finger has been passively moved to a target angle by the robot, the user is prompted to indicate the perceived finger position on the tablet screen by moving a virtual gauge indicator to a position aligned with the tip of the tested index finger (Figure 2a). Their view of the hand is constrained by the location of the tablet, hence subjects can not compensate by visual feedback to complete the task. Every trial starts with the robot moving the finger from the resting position (30°) to one of 21 angles (integer values $[10 - 30^\circ]$ in flexion from the resting position) within 3 seconds. In one assessment, each angle is presented once, in a random order, to limit learning effects. There is no time constraint for the subjects to indicate the perceived position and no feedback is given about the subject's performance. To ensure that the task assesses one hand only and does not rely on subject's ability to indicate the perceived position on the screen with the other hand, the experimenter helps the subject to point to the perceived finger position on the screen. The experimenter first asks if the gauge indicator is below or above the reference position and then moves the gauge indicator slowly in that direction until the subject says "stop". Then the experimenter asks for confirmation and allows for final adjustments. For each trial, the absolute error is calculated by

taking the absolute value of the difference between the reported and the presented angle. The primary outcome measure is the mean value of this absolute error across all 21 trials, denoted *Position Matching Absolute Error*. The higher the absolute error, the worse the task performance.

Fast target reaching task - assessment of motor impairments (1): the objective of this task is to quantify subjects' ability to produce fast ballistic target reaching movements. Target reaching has been used before as an assessment method of motor function deficits [48–51]. However, in contrast to target reaching tasks typically implemented in literature [48–51], in the newly proposed task the velocity is of interest and the accuracy of the movement is not considered. By examining the initial fast gross movement when reaching to the target (i.e the first 2 s from movement onset), we focus on the feed-forward nature of the task, thereby minimizing the involvement of somatosensory feedback in generating such motion. Subjects are instructed to move as fast as possible from a starting position to a target (40° movement amplitude), each displayed on the tablet computer screen as a red and green gauge indicator respectively. Visual feedback of the actual finger position is not provided, to avoid fine position corrections at the end of the movement, thereby encouraging subjects to concentrate on the movement velocity. The movement is performed either in flexion or in extension direction, in a random order. First, the tested finger is passively moved to a starting position by the robot (20° is the starting position for flexion and 60° for extension trials). Then, after a 3-second countdown, subjects are instructed to move within four seconds to the target (displayed at 60° for flexion and at 20° for extension trials). This trial duration was chosen to standardize the protocol and ensure that subjects with a slower reaction have enough time to generate a movement. Subjects are instructed to remain at their position once they believe they have reached the target. One assessment consists of 20 trials (10 times each direction). The primary outcome measure is the

mean of the three maximum velocity values (in $^{\circ}/s$) over all 10 trials per movement direction (denoted *Maximum Velocity Flexion/Extension*). The higher the velocity, the better the task performance. Representative velocity profiles, demonstrating how the outcome measure is derived, are shown in Figure 3a.

Range of motion task - assessment of motor impairments (2): the purpose of this task is to measure the range of motion of the index finger in flexion and extension direction. The range of motion is regularly evaluated in clinical settings to describe hand impairments [52–54]. Stroke subjects often show limited range of motion and the ability to extend the finger early post-stroke has even been shown as a predictor of recovery [45]. In this task subjects are instructed to move the index finger (which is secured on the ETH MIKE finger interface) as far as possible first in flexion and then in extension direction. This is repeated three times. Subjects can see the visual feedback of their finger displayed on the tablet computer. Afterwards, the same task is repeated in a passive manner, meaning that the experimenter moves the subject’s finger in flexion and then in extension (until the subject says “stop” or the experimenter detects tension in the finger), while the subjects are instructed to relax their finger. Here, the tablet computer is removed so as to not obstruct the experimenter that induced the motion. For each repetition, the difference between the maximum position in flexion and the maximum position in extension (measured in degrees) is calculated (denoted as *Active/Passive Range of Motion – AROM/PROM*). The primary outcome measure is the mean value across three repetitions for both AROM and PROM. The higher the ROM, the better the task performance. Representative position profiles are shown in Figure SM1a.

Maximum fingertip force generation task - assessment of motor impairments (3): the objective of this task is to measure maximum fingertip force. Assessments of grip strength are often performed in clinical settings in patients after stroke [55, 56], as weakness is frequently present after stroke and is linked to the

damage to the corticospinal tract [12]. In this task procedure the end-effector is first blocked by a fixation mechanism, located at a 45° angle. The subjects are instructed to generate maximal force with their index finger for an indicated period of time (3 s), preceded by a 3 s preparation phase. Three repetitions are performed first in flexion and then in extension direction. The primary task metric is the mean of the maximum force over three trials for both the flexion and extension direction, measured in Newtons by the force sensor located at the end-effector (denoted *Maximum Force Flexion/Extension*). The higher the force, the better the task performance. Representative force profiles are shown in Figure SM2a.

Trajectory following task - assessment of sensorimotor impairments: the aim of this task is to assess finger dexterity, which requires integration of proprioceptive feedback to generate motor output (i.e. sensorimotor integration). Trajectory following has been used previously to evaluate fine motor control [57–59]. First, the index finger is passively moved to a starting position by the robot (45° flexion). After a three second countdown, a trajectory is displayed on the tablet screen in the form of a moving gauge indicator, which the subjects are instructed to follow as accurately as possible. The vision of the actual finger position is not displayed on the screen, to ensure that subjects rely on proprioception to guide the motion. Two trajectory scenarios are displayed (slow and fast) in order to diversify the task. Each trajectory consists of three superimposed sine waves, each of different frequency and the same amplitude (15°). The slow trajectory consists of the following sine wave frequencies: 0.03 Hz, 0.07 Hz and 0.13 Hz, while the fast trajectory is composed of 0.10 Hz, 0.20 Hz and 0.40 Hz. One trial lasts 30 seconds and in total there are six trials in one assessment (three times each trajectory, first 3 times slow, then 3 times fast). For each trial, the tracking error between the trajectory displayed on the screen and the performed motion is calculated (Root Mean Squared Error RMSE [57]). The primary outcome measure is the mean across the three trials for the slow

and the fast trajectory (denoted *Tracking Error RMSE Slow/Fast*). The higher the tracking error, the worse the task performance. Representative trajectories are shown in Figure 4a.

Clinical assessments

The following clinical assessments were performed as a part of the study protocol. The kinesthetic Up-Down Test (kUDT) as part of the Nottingham Sensory Assessment (NSA) was chosen as a measure of proprioception [60]. In order to keep the scoring system of the kUDT from the NSA consistent with the commonly used Erasmus modified Nottingham Sensory Assessment [61], scores 1 and 2 were grouped together as score 1 and the best score was assigned the value 2. To clinically evaluate motor function, the Fugl-Meyer Upper Limb Assessment (FMA) was used [22]. The Box & Block Test of Manual Dexterity (BBT) was selected as an assessment of combined sensorimotor function and activity limitations [26] and it was completed for both hands. To quantify cognitive function, the Montreal Cognitive Assessment (MoCA) was performed [62]. Finally, the Modified Ashworth Scale (MAS) was used as a measure of spasticity [63].

Experimental protocol

Two testing sessions on two separate days were conducted to evaluate test-retest reliability in stroke subjects. Clinical assessments were performed in a separate session. For the control subjects, the protocol consisted of only one experimental session with the robot.

Subjects were seated in front of the robotic device and the height of the chair and the armrests was adjusted to a comfortable seating position close to the robot (Figure 1a). A wrist splint was used to ensure that the MCP joint was tested in isolation without any compensatory movements from the wrist. The elbow of the subject was placed on the cushioned armrest and subjects were instructed to keep it close to their body and to avoid compensatory movements throughout the

trial. The hand was strapped to the handle after ensuring optimal alignment of the forearm and the wrist joint with the orientation of the handle (30° neutral handle orientation). The index finger was attached to the finger module. The robotic assessments were always started with the range of motion and maximum fingertip force generation tasks, as they were the least complex and helped subjects to get familiar with the device. The order of the other three tasks, as well as the starting hand were randomized. Afterwards, subjects performed the assessments in the same order with the other hand. There was a familiarization round before each task.

Data analysis

Descriptive statistics

The metrics of the robot-assisted assessments are reported as mean and standard deviation per hand per studied group. In order to analyze potential confounding effects on task outcome measures emerging from subject demographics and task protocol, linear mixed effect models (LME) were built based on the control subjects dataset, following the steps defined in Kanzler et al. [38]. The parameters suspected to have a confounding effect on robotic metrics and hence included in the model were age, gender, tested hand (left/right) and the task specific effect. The latter was considered since some of the robot-assisted tasks were performed multiple times under different conditions, for example in two directions (e.g. maximum force in flexion/extension) or at different speeds (e.g. slow/fast trajectory following). The statistical significance of these effects was tested using the t-statistic (significance level of 0.05). In addition, the ability of the models to represent the experimental data (model quality) was analyzed according to the criteria C1 and C2, which characterize the mean absolute error of the model and its variability [38, 64] (moderate quality: $C1 \leq 15\%$ and $C2 \leq 25\%$; good quality: $C1 \leq 10\%$ and $C2 \leq 20\%$).

Test-retest reliability

For a comprehensive evaluation of the reliability of the newly proposed robot-assisted assessments, different statistical measures were considered. First, the intraclass correlation coefficient $ICC(A,k)$ was used to calculate absolute agreement between test and retest based on a two-way analysis of variance, taking into account all individual trials on test and retest [38, 65, 66]. This statistical method characterizes how well it is possible to discriminate between subjects across testing days (taking into consideration inter-subject and intra-subject variability). Acceptable ICC values are above 0.7 [38, 67]. Secondly, smallest real difference (SRD) and SRD% (% with respect to the range across all trials of a task) were calculated. These measures describe how well it is possible to distinguish between measurement noise and an actual physiological change [68]. Previous work suggested a cut-off of 30% for the SRD% to identify metrics without strong measurement error [38]. Further, to identify potential learning effects, the presence of possible systematic shifts between test and retest was analyzed. This was expressed as a mean difference between test and retest normalized with respect to the range of observed values [38]. Previous work suggested to consider a range of systematic shifts of [-6.35 and 6.35] to identify metrics without strong learning effects [38]. Finally, as a general check of similarity between test and retest, Spearman rank-order correlations between test and retest were calculated. It is desired for the test and retest metrics for each task to be strongly correlated, since that shows that the task outcomes are comparable between test and retest.

Discriminant validity

In order to define if the robotic metrics are capable of capturing abnormal task performance and thus impairments, the task metrics were compared between the stroke and the control group. This comparison was performed using three statistical methods. Firstly, control subjects were compared to the affected and the less-affected side

of stroke subjects using the Kruskal-Wallis test (Bonferroni corrected). The same group comparison was also performed using the Area Under the Curve (AUC) of the Receiver Operating Characteristic [38]. This method defines true positive/true negative rates of classifying subjects into two groups (stroke/control). A metric can well discriminate between the two groups if AUC is above 0.7 [38]. Finally, z-score normalization was implemented to find the percentage of stroke subjects performing worse than the 95th percentile of control subjects, thereby allowing to identify individuals that are impaired according to a specific outcome measure. To avoid comparing, for example, maximum fingertip force generated by an older female to a control population that is on average younger and gender-mixed, potentially confounding effects were removed based on the LME analysis [38]. This was only implemented if the model was of moderate or good quality and therefore able to accurately describe the data and only for the z-score normalization analysis.

Concurrent validity

To determine if the newly proposed tasks are able to capture impairments they were designed to assess, each task outcome measure was correlated with the clinical score that was expected to best reflect the underlying physiological construct. Namely, Spearman correlation was used to find the relationship between each task metric and the three clinical assessments: BBT, FMA and kUDT. The correlation strength was defined as: $\rho < 0.1$ negligible, $0.1 < \rho < 0.39$ weak, $0.4 < \rho < 0.69$ moderate, $\rho > 0.7$ strong [69]. Further, the classified impairments based on the ETH MIKE metrics were compared to the impairments detected by the corresponding clinical scores. This was implemented on subject-per-subject basis, leading to an overall agreement between the subjects classified as impaired according to the robotic and the clinical measures. The clinical scores used for classifications were FMA and kUDT, as both of these measures have clearly defined impairment classification thresholds (impaired defined as FM < 60 [70] and kUDT < 2 [61]).

Independence of task outcome measures

Partial Spearman correlation was calculated between the outcome measures of tasks that aim to characterize different impairments (proprioceptive, motor, sensorimotor), in order to determine if the proposed battery of tasks presents complementary information. For this analysis, the task condition that had highest reliability and validity was chosen, as the analysis of task-specific effects was treated separately.

Results

Out of the recruited 36 subjects with stroke, 34 successfully completed the two sessions of robot-assisted assessments. Two stroke subjects dropped out of the study (ID 10 and 20 - Table SM5), because they both felt fatigued/unwell during the first robotic assessment session and decided not to participate in the second session. Additionally, three participants (ID 4, 23 and 31) had missing data in one of the robotic assessments. The reason was either data saving malfunction or because they had another urgent appointment in the clinic and the full robotic assessment protocol could not be completed. Finally, one subject (ID 9) had severe problems with task comprehension and was not able to correctly follow the robotic task instructions. Altogether 30 stroke subjects were included in the data analysis, aged 64.50 ± 14.02 years, 19 males, 4 left-handed before stroke, 12 with left hemispheric stroke, 21 with ischemic and 9 with hemorrhagic stroke. All except one (chronic, 98 weeks post-stroke) were sub-acute (8.17 ± 4.56 weeks post-stroke), as summarized in Table 1. The number of days between test and retest was 1.97 ± 0.18 . The group of stroke subjects that completed the protocol had a diverse range of impairments, from severe (FMA=4 and kUDT=0) to mild (FMA=65, kUDT=3). The average result of FMA was 34.43 ± 22.22 and of kUDT 1.20 ± 0.83 . The time to perform the robotic assessment protocol (excluding instructions and setup) was 13.74 ± 4.55 min on test and 12.34 ± 2.98 min on retest on the affected side (details in Tables 3 & SM4). It took a comparable amount of time to complete the tasks on the less

affected side (14.09 ± 3.32 min test and 13.03 ± 2.30 min retest). Thirty-one age-matched control subjects were recruited in this study. None were excluded from the statistical analysis. Their average age was 66.87 ± 7.92 years, 20 were male and all were right-handed (Table 1).

Descriptive statistics results and the outcomes of the LME are shown in Table SM2. Overall, model quality was moderate to good for all tasks except for the position matching task. The maximum fingertip force generation task was significantly affected by age ($t=-2.83$, $DF=366$, $p\text{-value}=0.0049$). The outcome measure of that task was also affected by gender ($t=-5.00$, $p\text{-value}<0.001$), which was also a confounding effect for the fast target reaching task ($t=-3.73$, $p\text{-value}<0.001$). The effect of the tested hand was significant for all three tasks within the motor function category. All metrics were significantly influenced by the task-specific effect.

Table 1: Participants information.

	Stroke subjects	Control subjects
N	30	31
Age (mean \pm SD)	64.50 ± 14.02	66.87 ± 7.92
Male Female	19 11	20 11
Handedness	4 Left	Right only
LHS RHS	12 18	
Ischemic Hemorrhagic	21 9	
Weeks since stroke	8.17 ± 4.56 (range [2.57, 20.86])	
FMA [0-66]	34.43 ± 22.22 ([4, 65])	
kUDT [0-2]	1.20 ± 0.83 ([0, 2])	
BBT affected side [# /min]	20.90 ± 20.16 ([0, 74])	
BBT less affected side [# /min]	60.30 ± 11.27 ([44, 80])	
MoCA [0-30]	22.00 ± 5.84 ([6, 30])	
MAS [0-6]	0.10 ± 0.55 ([0, 3])	

These results only consider subjects that were included in the data analysis (i.e. does not consider subjects that were excluded due to missing data, drop-outs etc.). Acronyms - LHS: Left Hemispheric Stroke; RHS: Right Hemispheric Stroke; FMA: Fugl-Meyer Assessment; kUDT: kinesthetic Up-Down Test; BBT: Box& Block Test; MoCA: Montral Cognitive Assessment; MAS: Modified Ashworth Scale.

Test-retest reliability

All task metrics had sufficient test-retest reliability according to the defined criteria for the affected side of stroke subjects (Table 2). Specifically, ICC was good to excellent (ranging from 0.86 for Tracking Error RMSE Fast to 0.98 for Maximum Velocity Extension), the measurement error was small (SRD% < 30% for all tasks) and the systematic shift was within the defined range (the smallest value of -3.63 for Position Matching Absolute Error was still above the threshold of -6.35). Test and retest were strongly correlated for all tasks on the affected side ($\rho = 0.87 \pm 0.08$). On the less-affected side, ICC and SRD% were within required thresholds for all task metrics except for Maximum Velocity Flexion (ICC = 0.59, SRD% = 40.40%), as summarized in Table SM3. Further, the systematic shift was equal to 7.27 for Maximum Force Extension and to -7.24 for Tracking Error RMSE Fast. For all other tasks it was within ± 6.35 . There was a moderate to strong significant correlation between test and retest for all tasks on the less affected side ($\rho = 0.62 \pm 0.12$). Graphical representations of the test-retest reliability of one example task per category (proprioception, motor and sensorimotor) for both affected and less affected sides are shown in Figures 2c-4c & SM1c-2c.

Table 2: Test-retest reliability results for all task metrics on the affected side.

Category	Metric	Mean±SD T1	Mean±SD T2	Mean±SD T2-T1	ICC(A,k)	SRD	SRD(%)	Shift	Correlation
Sensory	AE [deg]	14.63 ± 6.43	13.71 ± 6.24	2.68 ± 2.94	0.90 (0.88-0.91)	9.12	14.65	-3.63	0.74***
	Vel. Flex. [deg/s]	314.94 ± 180.30	315.16 ± 162.50	0.22 ± 107.56	0.89 (0.84-0.93)	159.54	17.83	-0.003	0.84***
Motor	Vel. Ext. [deg/s]	149.06 ± 146.27	149.80 ± 144.74	0.74 ± 43.80	0.98 (0.97-0.98)	60.68	11.62	-0.15	0.97***
	AROM [deg]	43.89 ± 36.20	42.11 ± 32.29	-1.78 ± 11.16	0.97 (0.96-0.98)	15.58	12.77	1.47	0.92***
	PROM [deg]	83.88 ± 10.48	85.41 ± 7.95	1.53 ± 5.89	0.89 (0.83-0.93)	8.83	16.57	-3.02	0.83***
	Force Flex. [N]	11.64 ± 11.55	10.72 ± 10.08	-0.91 ± 3.44	0.97 (0.96-0.98)	4.88	10.92	2.07	0.94***
	Force Ext. [N]	4.12 ± 4.52	3.55 ± 3.67	-0.57 ± 1.93	0.94 (0.91-0.96)	2.80	16.72	3.80	0.93***
Sensorimotor	RMSE Slow [deg]	21.09 ± 8.68	21.03 ± 9.40	-0.06 ± 6.12	0.87 (0.81-0.92)	9.19	25.72	-0.17	0.81***
	RMSE Fast [deg]	21.45 ± 7.80	22.06 ± 9.06	0.60 ± 5.87	0.86 (0.79-0.91)	8.82	24.75	1.95	0.82***

AE: Position Matching Absolute Error; Vel. Flex./Ext.: Maximum Velocity Flexion/ Extension; AROM/PROM: Active/ Passive Range of Motion; Force Flex./Ext.: Maximum Force Flexion/ Extension; RMSE Slow/ Fast: Tracking Error RMSE Slow/ Fast; T1: Test; T2: Retest; ICC: Intraclass Correlation Coefficient; SRD: Smallest Real Difference; SRD%: Smallest Real Difference as a percentage of the range of values across all trials; Shift: Systematic Shift; Correlation: Spearman Correlation between Test and Retest. The defined thresholds for reliability were: ICC>0.70, SRD%<30%, Shift<|6.35|. Statistical significance was defined as: p -value<0.05: *, p -value<0.01: **, p -value<0.001: ***.

Discriminant and concurrent validity

Summary results of discriminant and concurrent validity for each task metric are shown in Tables 3 & SM4. A graphical representation of group differences (with severity subgroups within the affected side of stroke subjects) are shown for one outcome measure per task in Figures 2d-4d & SM1d-2d. Details on the classification agreement are provided in the form of a matrix (Figures 2b-4b & SM1b-2b).

Position Matching Absolute Error allowed to discriminate between controls and the affected side of stroke subjects (p -value<0.001; AUC=0.95), as well as between controls and the less-affected side of stroke subjects (p -value<0.001; AUC=0.82). There was no significant difference between the affected and the less affected side of stroke subjects (p -value=0.324; AUC=0.61). 66.70% of subjects were classified as impaired on their affected and 56.70% on their less affected side. Position Matching Absolute Error correlated moderately with the clinical measure of proprioception (ρ =-0.48, p -value=0.007). Further, there was a weak significant correlation between the robotic task score and the BBT (ρ =-0.37, p -value=0.046) and no significant correlation with the FMA. The agreement between impairment classification by robotic and clinical measure was 70.00%.

There was a significant difference between control group and the affected side of stroke subjects for all task outcome measures within the motor impairments assessment category, although Passive Range of Motion AUC was just below the threshold (p -value=0.007; AUC=0.69). Further, on a group-level stroke subjects performed significantly worse also on their less affected side comparing to controls in the fast target reaching task (flexion: p -value<0.001; AUC=0.85, extension: p -value<0.001; AUC=0.85). Additionally, the following outcome measures allowed to discriminate between both sides of stroke subjects - Maximum Velocity Extension (p -value=0.036; AUC=0.79), Active Range of Motion (p -value<0.001; AUC=0.85), Maximum Force Flexion (p -value<0.001; AUC=0.82) and Extension

(p -value<0.001; AUC=0.78). Across all motor assessment tasks, 23.3-90.0% of subjects were classified as impaired on their affected side, and 3.30-60.00% on the less affected side. Strong significant correlations were found between all tasks within the motor category and the FMA ($\rho = 0.80 \pm 0.05$, p -value<0.001, details in Table 3), except for the Passive Range of Motion ($\rho=0.58$, p -value<0.001). Further, all task outcome measures were moderately to strongly correlated with the BBT (from $\rho=0.54$, p -value=0.0013 to $\rho=0.77$, p -value<0.001). None of the outcomes correlated significantly with the kUDT. For most tasks in the motor category there was a high level of agreement between subjects classified as impaired according to the FMA threshold and according to z-scores of the robotic metrics - Maximum Force Flexion, Active Range of Motion, as well as Maximum Velocity Flexion and Extension reached classification agreement of 80.00%, 83.00%, 76.70% and 80.00%, respectively. For Maximum Force Extension and Passive Range of Motion the agreement was lower - 50.00% and 47.00%.

It was possible to distinguish between the affected side of stroke subjects and control subjects according to the assessment of sensorimotor impairments (Tracking Error RMSE Slow: p -value<0.001, AUC=0.91 and Fast: p -value<0.001, AUC=0.92). Stroke subjects scored comparably to controls on their less affected side when following the slow trajectory (p -value=0.061; AUC=0.66), but significantly worse when following the fast trajectory (p -value=0.026, AUC=0.70). There was a significant difference between the affected and the less-affected side of stroke subjects (Slow: p -value<0.001; AUC=0.81 and Fast: p -value<0.001; AUC=0.80). 73.30% of stroke subjects were classified as impaired on their affected and 23.30% on the less affected side when following the slow trajectory. There was a moderate significant correlation between Tracking Error RMSE Slow and kUDT ($\rho=-0.44$, p -value=0.016), as well as a weak significant correlation between this metric and the BBT ($\rho=-0.39$, p -value=0.035), but it did not significantly correlate with the FMA. Tracking Error

RMSE Fast was moderately correlated with the BBT ($\rho=-0.45$, p -value=0.014), as well as with the FMA ($\rho=-0.43$, p -value=0.017) and weakly with kUDT ($\rho=-0.39$, p -value=0.0347). Impairment classification according to the task z-scores was compared to both the kUDT and FMA, since this task was designed to involve both proprioceptive and motor function components. For both slow and fast trajectories, agreement between robotic z-score and the kUDT classification was 70.00%. When comparing the robotic z-scores classification to the one obtained with the FMA, Tracking Error RMSE Slow and Fast resulted in an agreement of 76.70% and 63.30%, respectively.

Table 3: Validity and feasibility results for all task metrics on the affected side.

Category	Metric	AUC	% impaired	% class. agree.	Corr. FMA	Corr. BBT	Corr. kUDT	Time T1 [min]
Sensory	AE [deg]	0.95	66.70	70.00 (kUDT)	-0.32	-0.37*	-0.48*	3.52 ± 1.82
	Vel. Flex. [deg/s]	0.91	60.00	76.70 (FMA)	0.74***	0.62***	-0.13	3.25 ± 0.55
	Vel. Ext. [deg/s]	0.94	90.00	80.00 (FMA)	0.76***	0.66***	-0.07	
Motor	AROM [deg]	0.80	60.00	83.00 (FMA)	0.85***	0.77***	0.14	1.67 ± 0.46
	PROM [deg]	0.69	23.30	47.00 (FMA)	0.58***	0.56**	0.27	
	Force Flex. [N]	0.93	83.30	80.00 (FMA)	0.82***	0.69***	0.08	1.50 ± 0.89
	Force Ext. [N]	0.80	40.00	50.00 (FMA)	0.81***	0.73***	0.21	
Sensorimotor	RMSE Slow [deg]	0.91	73.30	70.00 (kUDT) 76.60 (FMA)	-0.33	-0.34*	-0.44*	3.80 ± 0.83
	RMSE Fast [deg]	0.92	46.70	70.00 (kUDT) 63.30 (FMA)	-0.43*	-0.45*	-0.39*	

AE: Position Matching Absolute Error; Vel. Flex./Ext.: Maximum Velocity Flexion/ Extension; AROM/PROM: Active/ Passive Range of Motion; Force Flex./Ext.: Maximum Force Flexion/ Extension; RMSE Slow/Fast: Tracking Error RMSE Slow/ Fast; FMA: Fugl-Meyer Assessment; kUDT: kinesthetic Up-Down Test; BBT: Box & Block Test, AUC: Area Under the Curve of the Receiver Operating Characteristic; % impaired: % of stroke subjects classified as impaired according to z-scores; % class. agree.: % classification agreement as compared to clinical scores; Corr.: Spearman correlation of robotic metric with a clinical score. The defined threshold for discriminant validity was: AUC>0.7. Statistical significance: p -value<0.05: *, p -value<0.01: **, p -value<0.001: ***.

Independence of task metrics

Partial Spearman correlations between the five different robotic tasks are shown in Table 4. There was a moderate significant correlation between tasks within the motor category (Maximum Force Flexion, Active Range of Motion, Maximum Velocity Extension), ρ reached 0.42, 0.45, 0.47 and p -values were equal to 0.030, 0.019 and 0.013. Further, there was a moderate significant correlation between Position Matching Absolute Error and Tracking Error RMSE Slow ($\rho=0.50$, p -value=0.007). No significant correlations were found between the proprioceptive function assessment task and the tasks from the motor category. Likewise, tasks from the motor category were not significantly correlated with the assessment of sensorimotor impairments.

Table 4: Partial Spearman correlation coefficient ρ between 5 different robotic task metrics on the affected side.

		Sensory	Motor			Sensorimotor
		AE	Vel. Ext.	AROM	Force Flex.	RMSE Slow
Sensory	AE	1.00***	0.32	-0.25	-0.11	0.50**
	Vel. Ext.	0.32	1.00***	0.47*	0.42*	-0.20
Motor	AROM	-0.25	0.47*	1.00***	0.45*	-0.02
	Force Flex.	-0.11	0.42*	0.45*	1.00***	0.09
Sensorimotor	RMSE Slow	0.50**	-0.20	-0.02	0.09	1.00***

AE: Position Matching Absolute Error; Vel. Ext.: Maximum Velocity Extension; AROM: Active Range of Motion; Force Flex.: Maximum Force Flexion; RMSE Slow: Tracking Error RMSE Slow. Statistical significance: p -value<0.05: *, p -value<0.01: **, p -value<0.001: ***.

Discussion

The aim of this study was to propose and validate a set of robot-assisted assessments of the hand, implemented on a 1-DOF robotic platform, in a group of 36 stroke and 31 control subjects. We demonstrated that the novel set of robotic task metrics has good to excellent clinimetric properties (reliability, discriminant and concurrent validity) and can independently inform of proprioceptive, motor and combined sensorimotor impairments. Taken together, the presented robot-assisted assessments

provide quantitative insights into different aspects of sensorimotor function deficits, which could contribute to the understanding of each patient's impairment profile and, in the future, allow to better track and predict recovery after neurological injuries, as well as personalize therapies.

Test-retest reliability

Overall, the newly proposed robot-assisted assessments were reliable, did not show strong measurement error and were not confounded by strong systematic shift/learning effects. Demonstrating the reliability of the proprioception assessment (gauge position matching task) is an important contribution, since it is a long-standing challenge to reliably evaluate proprioceptive hand impairments in neurological patients, using either conventional or robotic methods [23, 35]. In addition, the reliability results obtained for the motor and sensorimotor task are in line with the ones previously reported in literature (ICC 0.7-0.9 for metrics of good quality) [11, 71–73], although it is challenging to directly compare these, as some studies were performed with control subjects instead of the target population. Compared to conventional clinical assessments, the proposed robotic assessments generally result in higher measurement error (e.g. FMA 10% [74], our robotic assessments between 15% and 25%). This could partially be explained by the precise sensing capability of the robotic method, which is more prone to detecting behavioural variability than clinical scales [73]. Given that the observed SRD% were relatively low, we can further expect that the metrics are able to responsively capture longitudinal changes in impairments. The satisfying test-retest reliability results obtained in this study could be explained by the objective and sensitive nature of the measurement method. The robotic platform can deliver exact and repeatable stimuli and objectively measure the corresponding response, hence ensuring reproducibility of this method. Another factor contributing to the positive reliability results on the affected side is the large inter-subject variability of the stroke participants recruited in this

study (severe to mild motor and proprioceptive impairments according to clinical assessments) [66]. Lower inter-subject variability and higher intra-subject variability could be observed on the less-affected side of stroke subjects, leading to overall lower intraclass correlation coefficients (although still within moderate to good reliability). This is because more subjects may score within the same range of values, however those values are more prone to some subject-specific mood/fatigue/motivation confounds, since there is a wider spectrum of outcomes each subject can potentially achieve on their less affected side [66]. Taken together, these positive reliability results lay out the foundation for integrating the robotic assessments in longitudinal studies to sensitively monitor the recovery of post-stroke subjects.

Discriminant validity

On the group level, stroke subjects performed significantly worse on the affected side compared to age-matched control subjects in all robot-assisted assessments. This confirms that the robotic metrics capture abnormalities in task performance that can be related to specific post-stroke impairments. In addition, it is clinically relevant to provide information about existing impairments on a subject-by-subject basis. From the percentage of subjects classified as impaired, we see that not all stroke subjects scored worse than controls, depending on their impairment profile. In addition, from the graphical representation of different severity groups (Figures 2d-4d & SM1d-2d), a clear trend of decreasing task performance with increasing impairment severity can be observed. These results indicate the capability of the robotic assessments to discriminate between different impairment severities.

In more detail, according to Position Matching Absolute Error, 66.7% of stroke subjects were classified as impaired on their affected side according to the z-scores. This result is aligned with previous findings (robotic assessments 57-67% [21, 34], clinical - 21-54% [20]). However, it should be noted that, in this work, we set an inclusion criteria to assure that 60% of recruited subjects have some proprioceptive

deficits (as measured by the kUDT), hence the resulting prevalence might be higher than expected for an overall stroke subject population [14]. Position Matching Absolute Error was found to be significantly higher in stroke subjects than in control subjects also on the less affected side. Previous work has shown similar trends of proprioceptive deficits often occurring on both sides after stroke [34, 75]. Such recent findings are possible thanks to the novel sensitive methods of proprioception assessment, which can detect small impairments on the less affected side, previously undetectable by the conventional, observer-based ordinal scales [23].

Secondly, it is interesting to observe that in the motor function assessments targeting finger extension (e.g. Maximum Velocity Extension), the difference between severely affected and moderately-mildly affected subjects was particularly noticeable and statistically significant in the case of AROM (Figure SM1d)). Voluntary extension of the index finger is typically more impaired in the severely affected subjects [76], a measure which has been shown useful as a predictor of functional recovery and is hence recommended as a routine clinical assessment [45]. Therefore, the ability of our method to precisely capture limitations in finger extension early could aid clinical decision making. The percentage of subjects classified as impaired on their affected side varied between 23.3% and 90.0% depending on the motor task metric, which indicates that each task may be sensitive to a different aspect of a patient's motor impairments. Similar trends were observed on another robotic platform, where depending on the task metric, different percentages of subjects were classified as impaired (19-81% [49]), despite the fact that over 80% of stroke subjects typically suffer upper limb paresis [14, 15].

Finally, results of the sensorimotor function assessment are in line with previous findings indicating decreased upper limb/hand dexterity post-stroke [57, 77, 78]. However, such high-level group comparison analysis does not inform on the reasons for decreased performance in this task, i.e. whether subjects could not follow the

trajectory due to weakness, impaired sensory feedback or both. A more refined picture can be obtained by considering severity subgroups, here created according to the clinical measure of proprioception (Figure 4d). Some subjects within the group with no proprioceptive impairment as reported by the clinical test (kUDT = 2) showed high Tracking Error RMSE. This can be explained by the nature of this task, which not only measures proprioception, but the combination thereof with motor function. Therefore, these subjects most likely could not follow the trajectory due to weakness, although it is also possible that the process of integrating sensory input to generate motor output was affected [7]. This indicates that the results of the trajectory following task, being the most complex task out of the assessment battery, can be best interpreted when considered together with the two other categories of tasks, in order to understand the different components of the impaired performance.

Concurrent validity

Moderate to strong significant correlations were found between the tasks from each category (i.e. proprioceptive, motor, sensorimotor function assessments) and their corresponding clinical scores, which indicates the capability of robotic metrics to capture specific impairments. In more detail, a moderate significant relationship was found between Position Matching Absolute Error and the kUDT, confirming that this robotic task is able to assess proprioceptive deficits. These observed correlations are in the expected range [11] given the limitations of the reference clinical score (ordinal scale, 0-2 points only [23, 35]). It would be of interest to further investigate concurrent validity using more accurate measures of somatosensory impairment, e.g. somatosensory evoked potentials [79], which precisely and objectively measure the strength and latency of somatosensory responses. Moreover, a weak significant correlation was found with the BBT, which may indicate the important role of proprioceptive feedback in the execution of dexterous activities of daily life involving the hand [13]. The fact that this robotic metric was not correlated with the FMA

suggests that this task may be able to assess proprioceptive impairments independently of motor function. Further, Maximum Force, Active Range of Motion, as well as Maximum Velocity showed strong significant correlations with the FMA and no significant correlations with the kUDT. It is therefore likely that these robotic tasks reflect impairments in basic motor execution [12, 13]. In addition, these tasks might be sensitive to damage to the corticospinal tract [80, 81]. Given their moderate to strong correlation with the BBT, it can be claimed that each of the kinematic and kinetic subcomponents of movement generation described by these robotic tasks (i.e. ability to displace the finger, generate force and fast motion), are necessary to produce the functional behaviour of grasping and holding an object. Finally, the trajectory following task resulted in slightly different correlations for the slow and fast movement trials. Slow trials follow a similar pattern to the position matching task – there is a moderate significant correlation between Tracking Error RMSE and the kUDT and weak significant correlation with the BBT. This would indicate that successful completion of this task primarily requires intact proprioceptive function. The lack of strong correlation with the FMA may indicate that this task relies on dexterity, which has been shown to be dissociated from basic motor function, like strength, for mildly to moderately impaired stroke subjects [13, 81]. Fast trials show moderate correlation to the BBT and the FMA, yet only a weak correlation with the kUDT, but all significant. It suggests that this task might integrate both motor and sensory components of movement execution. Stronger relation to motor scores in the fast trials can be explained by higher engagement of the motor system to generate faster movements, while slow movements require more dexterity. Overall, the proposed battery of assessments provides functionally relevant information about patients' capabilities by dissecting different subcomponents of motor control (i.e. proprioception, strength, speed, sensorimotor integration), which are necessary

to execute activities of daily living (e.g. as in the BBT - grasping and lifting small objects).

Independence of robotic task metrics

The fact that correlations between robotic assessments were at most moderate indicates that each task, to some extent, presents independent information. Interestingly, Tracking Error RMSE was correlated with Position Matching Absolute Error, but to none of the metrics in the motor function assessment category. Possibly this task requires fine motor control which is not needed in the simple motor execution-based tasks. Here the choice of the task metric for the sensorimotor task is crucial, given that if we chose, e.g., the range of motion during trajectory following instead of tracking error, the task outcome would have likely been correlated with motor metrics. However, we purposely chose Tracking Error RMSE by hypothesizing that this error-based metric could best describe the utilization of proprioceptive feedback in motor execution [57]. Both trajectory following and gauge position matching tasks assess some aspects of proprioception; one involves active (kinesthesia) and the other passive trajectory/position matching (position sense), hence some relationship between the two was expected [82, 83]. Although given the complex trajectory, the active following scenario resulted in higher errors. The lack of correlation between the tasks from motor and proprioceptive function assessment categories indicated that these tasks, as desired, are capable of evaluating these types of impairments separately. This result is in line with previous findings, which highlighted the independence of motor and somatosensory deficits [32]. Overall, different tasks in the proposed battery of robotic assessments complement each other in providing a detailed characterization of each patient's impairment profile.

Clinical feasibility

The protocol was completed by the majority of the recruited stroke subjects (only 2/36 subjects dropped out of the study and in total 30/36, that is 83% of recruited

participants were included in data analysis) that had a broad range of different impairments (severe to mild impairments on the FMA and the kUDT). Hence, this assessment protocol is feasible to be performed by a majority of patients in the early stage of stroke rehabilitation, which could provide insights for designing appropriate therapy programs and predicting recovery. Further, the duration of the assessments was acceptable (recommended is less than 15 min per tested side [84]), although the reported time does not include set-up and rest. Taking together assessments of both sides, set-up, explanation and rest, the average session per patient was close to 1 hour, which is at the limit of feasibility for more severely affected patients. However, mainly the more affected side should be assessed at regular time intervals in the long term, making the overall assessment time shorter and hence clinically acceptable. Further, reducing the number of trials in some of the tasks could help decrease the overall assessment time, which is a possibility only if such reduction doesn't affect reliability of the task metrics. Analysis of the effect of trial number reduction on metric reliability will be done as future work in order to optimize the study protocol for prospective studies.

Limitations

While the results of this study underline the ability of our metrics to identify specific aspects of hand sensorimotor impairments, several limitations need to be considered. The first limitation relates to the constrained range of motion in flexion, which blocked the finger movement at 90°. As the hand was aligned at 30° in flexion, the setup allowed maximally 60° flexion, which is smaller than full range of motion achievable by the index finger [85]. This particularly affected the fast target reaching task, in which some subjects overestimated their position with respect to the target and instead of naturally decelerating, were stopped by the device's mechanical stop. That in turn affected reliability of that task, especially on the less affected side, where subjects were able to reach higher velocities. The solution is to expand the

range of motion of the device by approx. 15° for future studies. Another limitation concerns some of the task metrics. In the trajectory following task, subjects that could not move at all (N=4/30) scored better in Tracking Error RMSE than subjects that moved “randomly” (e.g. due to impaired proprioceptive feedback but good enough motor function to extend/flex the finger). This is not necessarily desired, because such a scoring system does not encourage subjects to do their best, since staying stationary may provide a better outcome. Similarly, in the gauge position matching task, subjects that always indicated the starting position as where they thought their finger was, because they did not perceive their finger position at all (N=2/30), scored better than those that tried to perform the task but were clearly off the target. These are shortcomings of how the metrics themselves are calculated. A possible solution could be to modify scores of subjects that did not move their finger or did not move the gauge indicator from the starting position, by e.g. assigning the worst possible score achieved by all other stroke subjects + 1 standard deviation, as has been done in other studies [32]. This method was not used in this study in order to allow analyzing clinimetric properties of the proposed metrics in their purest form, without the introduction of an arbitrary ceiling effect. Finally, this study only considered the index finger and it remains to be investigated how well these results generalize to the whole hand somatosensory and motor function. However, correlations with clinical assessments targeting the whole upper limb observed in this study suggest that the index finger function is indeed essential for performing daily life activities (i.e., grasping objects as in the BBT) and it may be related to impairments in other parts of the upper limb (i.e., as shown through the strong correlations of some robotic tasks with the FMA), which is in agreement with findings from previous studies involving individual finger movements [86].

Conclusions

In conclusion, this work successfully proposed and validated a set of robot-assisted assessments targeting proprioceptive, motor, and combined sensorimotor impairments in the hand. This contributes to addressing a long-standing gap in the neurorehabilitation domain, as such a comprehensive impairment profile of the hand could not be established in a reliable, valid, and clinically feasible manner before. Building a fine-grained picture of patients' deficits is important to sensitively track rehabilitation progress and effectively adapt therapies. Generally, this work addresses a strong need for more sensitive, accurate and objective assessments, which could positively impact therapy planning and outcomes.

Abbreviations

D: dominant, LA: less affected, A: affected, BBT: Box and Block Test, FMA: Fugl-Meyer Assessment, kUDT: kinesthetic Up-Down Test, ICC: intraclass correlation coefficient, SRD: smallest real difference.

Acknowledgements

The authors would like to thank Aida Sehle and Jana Stürner from Kliniken Schmieder Allensbach for their help in the development of the robotic assessments and in initial testing. We also thank Lucille Cazenave for the support in data analysis.

Authors' contributions

Study design: MZ, CK, JL, OL, RG. Data collection: MZ, LJ, CS. Data analysis: MZ, CK, LJ. Data interpretation: MZ, CK, JL, OL, RG. Manuscript writing: MZ, CK, OL, RG. All authors read and approved the final manuscript.

Funding

This work is supported by the Swiss National Science Foundation, project 320030L_170163 and by the ETH Zurich Foundation in collaboration with Hocoma AG.

Availability of data materials

The data presented in this manuscript are available upon reasonable request and under consideration of the ethical regulations.

Ethical approval and consent to participate

The study was approved by the ETH Ethics Committee EK 2019-N-108 and the Ethics Commission of Baden-Württemberg F-2016-126 and retrospectively registered as a clinical trial [nr DRKS00018681](#).

Consent for publication

Not applicable.

Competing interests

There are no conflicts of interest in the submission of this manuscript.

Author details

¹Rehabilitation Engineering Laboratory, Institute of Robotics and Intelligent Systems, Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland. ²Kliniken Schmieder Allensbach, Zum Tafelholz 8, 78476 Allensbach, Germany.

References

- Lambercy, O., Ranzani, R., Gassert, R.: Robot-assisted rehabilitation of hand function. *Rehabilitation Robotics*, 205–225 (2018). doi:[10.1016/B978-0-12-811995-2.00027-8](https://doi.org/10.1016/B978-0-12-811995-2.00027-8)
- Yousif, N., Cole, J., Rothwell, J., Diedrichsen, J.: Proprioception in motor learning: lessons from a deafferented subject. *Experimental Brain Research* **233**(8), 2449–2459 (2015). doi:[10.1007/s00221-015-4315-8](https://doi.org/10.1007/s00221-015-4315-8)
- Miall, R.C., Kitchen, N.M., Nam, S.-H., Lefumat, H., Renault, A.G., Ørstavik, K., Cole, J.D., Sarlegna, F.R.: Proprioceptive loss and the perception, control and learning of arm movements in humans: evidence from sensory neuronopathy. *Experimental Brain Research* **236**, 2137–2155 (2018). doi:[10.1007/s00221-018-5289-0](https://doi.org/10.1007/s00221-018-5289-0)
- Miall, R.C., Rosenthal, O., Ørstavik, K., Cole, J.D., Sarlegna, F.R.: Loss of haptic feedback impairs control of hand posture: a study in chronically deafferented individuals when grasping and lifting objects. *Experimental Brain Research* **237**(9), 2167–2184 (2019). doi:[10.1007/s00221-019-05583-2](https://doi.org/10.1007/s00221-019-05583-2)
- Dietz, V.: Proprioception and locomotor disorders. *Nature Reviews Neuroscience* **3**(10), 781–790 (2002). doi:[10.1038/nrn939](https://doi.org/10.1038/nrn939)
- Ghez, C., Gordon, J., Ghilardi, M.F., Christakos, C.N., Cooper, S.E.: Roles of Proprioceptive Input in the Programming of Arm Trajectories. *Cold Spring Harbor Symposia on Quantitative Biology* **55**(February 1990), 837–847 (1990). doi:[10.1101/SQB.1990.055.01.079](https://doi.org/10.1101/SQB.1990.055.01.079)
- Edwards, L.L., King, E.M., Buetefisch, C.M., Borich, M.R.: Putting the “sensory” into sensorimotor control: The role of sensorimotor integration in goal-directed hand movements after stroke. *Frontiers in Integrative Neuroscience* **13**, 16 (2019). doi:[10.3389/fnint.2019.00016](https://doi.org/10.3389/fnint.2019.00016)
- Wolpert, D.M., Goodbody, S.J., Husain, M.: Maintaining internal representations: the role of the human superior parietal lobe. *Nature Neuroscience* **1**(6), 529–533 (1998). doi:[10.1038/2245](https://doi.org/10.1038/2245)
- Schwartz, A.B.: Movement: How the Brain Communicates with the World. *Cell* **164**(6), 1122–1135 (2016). doi:[10.1016/j.cell.2016.02.038](https://doi.org/10.1016/j.cell.2016.02.038)
- Lawrence, E.S., Coshall, C., Dundas, R., Stewart, J., Rudd, A.G., Howard, R., Wolfe, C.D.A.: Estimates of the Prevalence of Acute Stroke Impairments and Disability in a Multiethnic Population. *Stroke* **32**(6), 1279–1284 (2001). doi:[10.1161/01.STR.32.6.1279](https://doi.org/10.1161/01.STR.32.6.1279)
- Schwarz, A., Kanzler, C.M., Lambercy, O., Luft, A.R., Veerbeek, J.M.: Systematic Review on Kinematic Assessments of Upper Limb Movements After Stroke. *Stroke* **50**(3), 718–727 (2019). doi:[10.1161/STROKEAHA.118.023531](https://doi.org/10.1161/STROKEAHA.118.023531)
- Raghavan, P.: The nature of hand motor impairment after stroke and its treatment. *Current Treatment Options in Cardiovascular Medicine* **9**(3), 221–228 (2007). doi:[10.1007/s11936-007-0016-3](https://doi.org/10.1007/s11936-007-0016-3)
- Xu, J., Haith, A.M., Krakauer, J.W.: Motor Control of the Hand Before and After Stroke. In: *Clinical Systems Neuroscience*, pp. 271–289. Springer, Tokyo (2015). doi:[10.1007/978-4-431-55037-2_14](https://doi.org/10.1007/978-4-431-55037-2_14)
- Rathore, S.S., Hinn, A.R., Cooper, L.S., Tyroler, H.A., Rosamond, W.D.: Characterization of incident stroke signs and symptoms findings from the atherosclerosis risk in communities study. *Stroke* **33**(11), 2718–2721 (2002). doi:[10.1161/01.STR.0000035286.87503.31](https://doi.org/10.1161/01.STR.0000035286.87503.31)
- Aqueveque, P., Ortega, P., Pino, E., Saavedra, F., Germany, E., Gómez, B.: After Stroke Movement Impairments: A Review of Current Technologies for Rehabilitation. *Physical Disabilities - Therapeutic Implications* (2017). doi:[10.5772/67577](https://doi.org/10.5772/67577)

16. Connell, L.A., Lincoln, N.B., Radford, K.A.: Somatosensory impairment after stroke: Frequency of different deficits and their recovery. *Clinical Rehabilitation* **22**(8), 758–767 (2008). doi:[10.1177/0269215508090674](https://doi.org/10.1177/0269215508090674)
17. Zandvliet, S.B., Kwakkel, G., Nijland, R.H.M., van Wegen, E.E.H., Meskers, C.G.M.: Is Recovery of Somatosensory Impairment Conditional for Upper-Limb Motor Recovery Early After Stroke? *Neurorehabilitation and Neural Repair* **34**(5), 403–416 (2020). doi:[10.1177/1545968320907075](https://doi.org/10.1177/1545968320907075)
18. Kusoffsky, A., Wadell, I., Nilsson, B.Y.: The relationship between sensory impairment and motor recovery in patients with hemiplegia. *Scandinavian journal of rehabilitation medicine* **14**(1), 27–32 (1982)
19. Nudo, R.J., Friel, K.M., Delia, S.W.: Role of sensory deficits in motor impairments after injury to primary motor cortex. *Neuropharmacology* **39**(5), 733–742 (2000). doi:[10.1016/S0028-3908\(99\)00254-3](https://doi.org/10.1016/S0028-3908(99)00254-3)
20. Meyer, S., De Bruyn, N., Lafosse, C., Van Dijk, M., Michielsen, M., Thijs, L., Truyens, V., Oostra, K., Krumlinde-Sundholm, L., Peeters, A., Thijs, V., Feys, H., Verheyden, G.: Somatosensory impairments in the upper limb poststroke: Distribution and association with motor function and visuospatial neglect. *Neurorehabilitation and Neural Repair* **30**(8), 731–742 (2016). doi:[10.1177/1545968315624779](https://doi.org/10.1177/1545968315624779)
21. Kenzie, J.M., Semrau, J.A., Hill, M.D., Scott, S.H., Dukelow, S.P.: A composite robotic-based measure of upper limb proprioception. *Journal of NeuroEngineering and Rehabilitation* **14**(1) (2017). doi:[10.1186/s12984-017-0329-8](https://doi.org/10.1186/s12984-017-0329-8)
22. Gladstone, D.J., Danells, C.J., Black, S.E.: The Fugl-Meyer Assessment of Motor Recovery after Stroke: A Critical Review of Its Measurement Properties. *Neurorehabilitation and Neural Repair* **16**(3), 232–240 (2002). doi:[10.1177/154596802401105171](https://doi.org/10.1177/154596802401105171)
23. Lincoln, N.B., Crow, J.L., Jackson, J.M., Waters, G.R., Adams, S.A., Hodgson, P.: The unreliability of sensory assessments. *Clinical Rehabilitation* **5**(4), 273–282 (1991). doi:[10.1177/026921559100500403](https://doi.org/10.1177/026921559100500403)
24. Valdes OTD, K., Naughton OTD, N., Algar OTD, L.: Sensorimotor interventions and assessments for the hand and wrist: A scoping review. *Journal of Hand Therapy* **27**, 272–286 (2014). doi:[10.1016/j.jht.2014.07.002](https://doi.org/10.1016/j.jht.2014.07.002)
25. Pike, S., Lannin, N.A., Wales, K., Cusick, A.: A systematic review of the psychometric properties of the Action Research Arm Test in neurorehabilitation. *Australian Occupational Therapy Journal* **65**(5), 449–471 (2018). doi:[10.1111/1440-1630.12527](https://doi.org/10.1111/1440-1630.12527)
26. Mathiowetz, V., Weber, K.: Adult Norms for the Box and Block Test of Manual Dexterity. *the American Journal of Occupational Therapy* **39**(6), 387–391 (1985)
27. Lang, C.E., Bland, M.D., Bailey, R.R., Schaefer, S.Y., Birkenmeier, R.L.: Assessment of upper extremity impairment, function, and activity after stroke: Foundations for clinical decision making. *Journal of Hand Therapy* **26**(2), 104–115 (2013). doi:[10.1016/j.jht.2012.06.005](https://doi.org/10.1016/j.jht.2012.06.005)
28. Scott, S.H., Dukelow, S.P.: Potential of robots as next-generation technology for clinical assessment of neurological disorders and upper-limb therapy. *The Journal of Rehabilitation Research and Development* **48**(4), 335 (2011). doi:[10.1682/JRRD.2010.04.0057](https://doi.org/10.1682/JRRD.2010.04.0057)
29. Lamercy, O., Robles, A.J., Kim, Y., Gassert, R.: Design of a robotic device for assessment and rehabilitation of hand sensory function. In: 2011 IEEE International Conference on Rehabilitation Robotics, pp. 1–6 (2011). doi:[10.1109/ICORR.2011.5975436](https://doi.org/10.1109/ICORR.2011.5975436)
30. Plantin, J., Pennati, G.V., Roca, P., Baron, J.-C., Laurencikas, E., Weber, K., Godbolt, A.K., Borg, J., Lindberg, P.G.: Quantitative Assessment of Hand Spasticity After Stroke: Imaging Correlates and Impact on Motor Recovery. *Frontiers in Neurology* **10**, 836 (2019). doi:[10.3389/fneur.2019.00836](https://doi.org/10.3389/fneur.2019.00836)
31. Dukelow, S.P., Herter, T.M., Moore, K.D., Demers, M.J., Glasgow, J.I., Bagg, S.D., Norman, K.E., Scott, S.H.: Quantitative assessment of limb position sense following stroke. *Neurorehabilitation and Neural Repair* **24**(2), 178–187 (2010). doi:[10.1177/1545968309345267](https://doi.org/10.1177/1545968309345267)

32. Semrau, J.A., Herter, T.M., Scott, S.H., Dukelow, S.P.: Examining differences in patterns of sensory and motor recovery after stroke with robotics. *Stroke* **46**(12), 3459–3469 (2015). doi:[10.1161/STROKEAHA.115.010750](https://doi.org/10.1161/STROKEAHA.115.010750)
33. Ingemanson, M.L., Rowe, J.B., Chan, V., Wolbrecht, E.T., Cramer, S.C., Reinkensmeyer, D.J.: Use of a robotic device to measure age-related decline in finger proprioception. *Experimental Brain Research* **234**(1), 83–93 (2016). doi:[10.1007/s00221-015-4440-4](https://doi.org/10.1007/s00221-015-4440-4)
34. Ingemanson, M.L., Rowe, J.R., Chan, V., Riley, J., Wolbrecht, E.T., Reinkensmeyer, D.J., Cramer, S.C.: Neural Correlates of Passive Position Finger Sense After Stroke. *Neurorehabilitation and Neural Repair* **33**(9), 740–750 (2019). doi:[10.1177/1545968319862556](https://doi.org/10.1177/1545968319862556)
35. Rinderknecht, M.D., Lambercy, O., Raible, V., Büsching, I., Sehle, A., Liepert, J., Gassert, R.: Reliability, validity, and clinical feasibility of a rapid and objective assessment of post-stroke deficits in hand proprioception. *Journal of NeuroEngineering and Rehabilitation* **15**, 47 (2018). doi:[10.1186/s12984-018-0387-6](https://doi.org/10.1186/s12984-018-0387-6)
36. Contu, S., Basteris, A., Plunkett, T.K., Kuah, C.W.K., Chua, K.S., Campolo, D., Masia, L.: Wrist Proprioception in Acute and Subacute Stroke: A Robotic Protocol for Highly Impaired Patients. In: 2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob), pp. 19–24 (2018). doi:[10.1109/BIOROB.2018.8488083](https://doi.org/10.1109/BIOROB.2018.8488083)
37. Lexell, J.E., Downham, D.Y.: How to assess the reliability of measurements in rehabilitation. *American Journal of Physical Medicine and Rehabilitation* **84**(9), 719–723 (2005). doi:[10.1097/01.phm.0000176452.17771.20](https://doi.org/10.1097/01.phm.0000176452.17771.20)
38. Kanzler, C.M., Rinderknecht, M.D., Schwarz, A., Lamers, I., Gagnon, C., Held, J.P.O., Feys, P., Luft, A.R., Gassert, R., Lambercy, O.: A data-driven framework for selecting and validating digital health metrics: use-case in neurological sensorimotor impairments. *npj Digital Medicine* **3**(1), 80 (2020). doi:[10.1038/s41746-020-0286-7](https://doi.org/10.1038/s41746-020-0286-7)
39. Shirota, C., Balasubramanian, S., Melendez-Calderon, A.: Technology-aided assessments of sensorimotor function: Current use, barriers and future directions in the view of different stakeholders. *Journal of NeuroEngineering and Rehabilitation* **16**(1), 53 (2019). doi:[10.1186/s12984-019-0519-7](https://doi.org/10.1186/s12984-019-0519-7)
40. Zbytnewska, M., Rinderknecht, M.D., Lambercy, O., Barnobi, M., Raats, J., Lamers, I., Feys, P., Liepert, J., Gassert, R.: Design and Characterization of a Robotic Device for the Assessment of Hand Proprioceptive, Motor, and Sensorimotor Impairments. In: 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR), pp. 441–446 (2019). doi:[10.1109/ICORR.2019.8779507](https://doi.org/10.1109/ICORR.2019.8779507)
41. Zbytnewska, M., Siegenthaler, M., Kanzler, C.M., Hermsdorfer, J., Gassert, R., Lambercy, O.: Design and Preliminary Evaluation of a Perturbation-based Robot-assisted Assessment of Hand Sensorimotor Impairments. In: 2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob), pp. 721–726 (2020). doi:[10.1109/BioRob49111.2020.9224412](https://doi.org/10.1109/BioRob49111.2020.9224412)
42. Ekstrand, E., Rylander, L., Lexell, J., Brogårdh, C.: Perceived ability to perform daily hand activities after stroke and associated factors: a cross-sectional study. *BMC neurology* **16**(1), 208 (2016). doi:[10.1186/s12883-016-0733-x](https://doi.org/10.1186/s12883-016-0733-x)
43. Dollar, A.M.: *Classifying Human Hand Use and the Activities of Daily Living*, pp. 201–216. Springer (2014). doi:[10.1007/978-3-319-03017-3_10](https://doi.org/10.1007/978-3-319-03017-3_10)
44. Smania, N., Paolucci, S., Tinazzi, M., Borghero, A., Manganotti, P., Fiaschi, A., Moretto, G., Bovi, P., Gambarin, M.: Active Finger Extension. *Stroke* **38**(3), 1088–1090 (2007). doi:[10.1161/01.STR.0000258077.88064.a3](https://doi.org/10.1161/01.STR.0000258077.88064.a3)
45. Winters, C., Kwakkel, G., Nijland, R., van Wegen, E.: When Does Return of Voluntary Finger Extension Occur Post-Stroke? A Prospective Cohort Study. *PLOS ONE* **11**(8), 0160528 (2016). doi:[10.1371/journal.pone.0160528](https://doi.org/10.1371/journal.pone.0160528)

46. Carey, L.M., Oke, L.E., Matyas, T.A.: Impaired limb position sense after stroke: A quantitative test for clinical use. *Archives of Physical Medicine and Rehabilitation* **77**(12), 1271–1278 (1996).
doi:[10.1016/S0003-9993\(96\)90192-6](https://doi.org/10.1016/S0003-9993(96)90192-6)
47. Rinderknecht, M.D., Popp, W.L., Lamercy, O., Gassert, R.: Reliable and Rapid Robotic Assessment of Wrist Proprioception Using a Gauge Position Matching Paradigm. *Frontiers in human neuroscience* **10**, 316 (2016).
doi:[10.3389/fnhum.2016.00316](https://doi.org/10.3389/fnhum.2016.00316)
48. Collins, K.C., Kennedy, N.C., Clark, A., Pomeroy, V.M.: Kinematic Components of the Reach-to-Target Movement After Stroke for Focused Rehabilitation Interventions: Systematic Review and Meta-Analysis. *Frontiers in neurology* **9**, 472 (2018). doi:[10.3389/fneur.2018.00472](https://doi.org/10.3389/fneur.2018.00472)
49. Coderre, A.M., Amr Abou Zeid, Dukelow, S.P., Demmer, M.J., Moore, K.D., Demers, M.J., Bretzke, H., Herter, T.M., Glasgow, J.I., Norman, K.E., Bagg, S.D., Scott, S.H.: Assessment of upper-limb sensorimotor function of subacute stroke patients using visually guided reaching. *Neurorehabilitation and Neural Repair* **24**(6), 528–541 (2010). doi:[10.1177/1545968309356091](https://doi.org/10.1177/1545968309356091)
50. Semrau, J.A., Herter, T.M., Kenzie, J.M., Findlater, S.E., Scott, S.H., Dukelow, S.P.: Robotic Characterization of Ipsilesional Motor Function in Subacute Stroke. *Neurorehabilitation and Neural Repair* (2017).
doi:[10.1177/1545968317704903](https://doi.org/10.1177/1545968317704903)
51. Koepfel, T., Pila, O.: Test-Retest Reliability of Kinematic Assessments for Upper Limb Robotic Rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **28**(9), 2035–2042 (2020).
doi:[10.1109/TNSRE.2020.3013705](https://doi.org/10.1109/TNSRE.2020.3013705)
52. Lang, C.E., Bland, M.D., Bailey, R.R., Schaefer, S.Y., Birkenmeier, R.L.: Assessment of upper extremity impairment, function, and activity after stroke: foundations for clinical decision making. *Journal of Hand Therapy* **26**(2), 104–115 (2013). doi:[10.1016/j.jht.2012.06.005](https://doi.org/10.1016/j.jht.2012.06.005)
53. Hayashi, H., Shimizu, H., Okumura, S., Miwa, K.: Necessary Metacarpophalangeal Joints Range of Motion to Maintain Hand Function. *Hong Kong Journal of Occupational Therapy* **24**, 51–55 (2014).
doi:[10.1016/j.hkjot.2014.10.001](https://doi.org/10.1016/j.hkjot.2014.10.001)
54. Beebe, J.A., Lang, C.E.: Active range of motion predicts upper extremity function 3 months after stroke. *Stroke* **40**(5), 1772–1779 (2009). doi:[10.1161/STROKEAHA.108.536763](https://doi.org/10.1161/STROKEAHA.108.536763)
55. Chen, H.-M., Chen, C.C., Hsueh, I.-P., Huang, S.-L., Hsieh, C.-L.: Test-Retest Reproducibility and Smallest Real Difference of 5 Hand Function Tests in Patients With Stroke. *Neurorehabilitation and Neural Repair* **23**, 435–440 (2009). doi:[10.1177/1545968308331146](https://doi.org/10.1177/1545968308331146)
56. Lindberg, P.G., Roche, N., Robertson, J., Roby-Brami, A., Bussel, B., Maier, M.A.: Affected and unaffected quantitative aspects of grip force control in hemiparetic patients after stroke. *Brain Research* **1452**, 96–107 (2012). doi:[10.1016/j.brainres.2012.03.007](https://doi.org/10.1016/j.brainres.2012.03.007)
57. Ao, D., Song, R., Tong, K.-y.: Sensorimotor Control of Tracking Movements at Various Speeds for Stroke Patients as Well as Age-Matched and Young Healthy Subjects. *PLOS ONE* **10**(6), 0128328 (2015).
doi:[10.1371/journal.pone.0128328](https://doi.org/10.1371/journal.pone.0128328)
58. Hu, X.L., Tong, K.Y., Song, R., Zheng, X.J., Lui, K.H., Leung, W.W.F., Ng, S., Au-Yeung, S.S.Y.: Quantitative evaluation of motor functional recovery process in chronic stroke patients during robot-assisted wrist training. *Journal of Electromyography and Kinesiology* **19**(4), 639–650 (2009). doi:[10.1016/J.JELEKIN.2008.04.002](https://doi.org/10.1016/J.JELEKIN.2008.04.002)
59. Colombo, R., Pisano, F., Micera, S., Mazzone, A., Delconte, C., Carrozza, M.C., Dario, P., Minuco, G.: Assessing Mechanisms of Recovery During Robot-Aided Neurorehabilitation of the Upper Limb. *Neurorehabilitation and Neural Repair* **22**(1), 50–63 (2008). doi:[10.1177/1545968307303401](https://doi.org/10.1177/1545968307303401)
60. Lincoln, N.B., Jackson, J., Adams, S.: Reliability and Revision of the Nottingham Sensory Assessment for

- Stroke Patients Reliability and Revision of the Nottingham Sensory Assessment for Stroke Patients. *Physiotherapy* **84**(8), 358–365 (1998). doi:[10.1016/S0031-9406\(05\)61454-X](https://doi.org/10.1016/S0031-9406(05)61454-X)
61. Stolk-Hornsveld, F., Crow, J.L., Hendriks, E.P., van der Baan, R., Harmeling-van der Wel, B.C.: The Erasmus MC modifications to the (revised) Nottingham Sensory Assessment: A reliable somatosensory assessment measure for patients with intracranial disorders. *Clinical Rehabilitation* **20**(2), 160–172 (2006). doi:[10.1191/0269215506cr932oa](https://doi.org/10.1191/0269215506cr932oa)
 62. Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H.: The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* **53**(4), 695–699 (2005). doi:[10.1111/j.1532-5415.2005.53221.x](https://doi.org/10.1111/j.1532-5415.2005.53221.x)
 63. Gregson, J.M., Leathley, M., Moore, A.P., Sharma, A.K., Smith, T.L., Watkins, C.L.: Reliability of the tone assessment scale and the modified Ashworth scale as clinical tools for assessing poststroke spasticity. *Archives of Physical Medicine and Rehabilitation* **80**(9), 1013–1016 (1999). doi:[10.1016/S0003-9993\(99\)90053-9](https://doi.org/10.1016/S0003-9993(99)90053-9)
 64. Roy, K., Das, R.N., Ambure, P., Aher, R.B.: Be aware of error measures. Further studies on validation of predictive QSAR models. *Chemometrics and Intelligent Laboratory Systems* **152**, 18–33 (2016). doi:[10.1016/j.chemolab.2016.01.008](https://doi.org/10.1016/j.chemolab.2016.01.008)
 65. Vaz, S., Falkmer, T., Passmore, A.E., Parsons, R., Andreou, P.: The Case for Using the Repeatability Coefficient When Calculating Test-Retest Reliability. *PLoS ONE* **8**(9), 1–7 (2013). doi:[10.1371/journal.pone.0073990](https://doi.org/10.1371/journal.pone.0073990)
 66. Koo, T.K., Li, M.Y.: A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine* **15**(2), 155–163 (2016). doi:[10.1016/j.jcm.2016.02.012](https://doi.org/10.1016/j.jcm.2016.02.012)
 67. Prinsen, C.A.C., Mokkink, L.B., Bouter, L.M., Alonso, J., Patrick, D.L., de Vet, H.C.W., Terwee, C.B.: COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research* **27**(5), 1147–1157 (2018). doi:[10.1007/s11136-018-1798-3](https://doi.org/10.1007/s11136-018-1798-3)
 68. Beckerman, H., Roebroeck, M.E., Lankhorst, G.J., Becher, J.G., Bezemer, P.D., Verbeek, A.L.M.: Smallest real difference, a link between reproducibility and responsiveness. *Quality of Life Research* **10**(7), 571–578 (2001). doi:[10.1023/A:1013138911638](https://doi.org/10.1023/A:1013138911638)
 69. Schober, P., Boer, C., Schwarte, L.A.: Correlation Coefficients. *Anesthesia & Analgesia* **126**(5), 1763–1768 (2018). doi:[10.1213/ANE.0000000000002864](https://doi.org/10.1213/ANE.0000000000002864)
 70. Meyer, S., De Bruyn, N., Krumlinde-Sundholm, L., Peeters, A., Feys, H., Thijs, V., Verheyden, G.: Associations between sensorimotor impairments in the upper limb at 1 week and 6 months after stroke. *Journal of Neurologic Physical Therapy* **40**(3), 186–195 (2016). doi:[10.1097/NPT.0000000000000138](https://doi.org/10.1097/NPT.0000000000000138)
 71. Simmatis, L.E.R., Early, S., Moore, K.D., Appaqaq, S., Scott, S.H.: Statistical measures of motor, sensory and cognitive performance across repeated robot-based testing. *Journal of NeuroEngineering and Rehabilitation* **17**(1), 86 (2020). doi:[10.1186/s12984-020-00713-2](https://doi.org/10.1186/s12984-020-00713-2)
 72. Mang, C.S., Whitten, T.A., Cosh, M.S., Scott, S.H., Wiley, J.P., Debert, C.T., Dukelow, S.P., Benson, B.W.: Test-retest reliability of the KINARM end-point robot for assessment of sensory, motor and neurocognitive function in young adult athletes. *PLOS ONE* **13**(4), 0196205 (2018). doi:[10.1371/journal.pone.0196205](https://doi.org/10.1371/journal.pone.0196205)
 73. Kanzler, C.M., Schwarz, A., Held, J.P.O., Luft, A.R., Gassert, R., Lamercy, O.: Technology-aided assessment of functionally relevant sensorimotor impairments in arm and hand of post-stroke individuals. *Journal of NeuroEngineering and Rehabilitation* **17**(1), 128 (2020). doi:[10.1186/s12984-020-00748-5](https://doi.org/10.1186/s12984-020-00748-5)
 74. Hsueh, I.-P., Hsu, M.-J., Sheu, C.-F., Lee, S., Hsieh, C.-L., Lin, J.-H.: Psychometric Comparisons of 2 Versions of the Fugl-Meyer Motor Scale and 2 Versions of the Stroke Rehabilitation Assessment of Movement. *Neurorehabilitation and Neural Repair* **22**(6), 737–744 (2008). doi:[10.1177/1545968308315999](https://doi.org/10.1177/1545968308315999)
 75. Ingemanson, M.L., Rowe, J.R., Chan, V., Wolbrecht, E.T., Reinkensmeyer, D.J., Cramer, S.C.: Somatosensory

- system integrity explains differences in treatment response after stroke. *Neurology* **92**(10), 1098–1108 (2019). doi:[10.1212/WNL.0000000000007041](https://doi.org/10.1212/WNL.0000000000007041)
76. Cruz, E.G., Waldinger, H.C., Kamper, D.G.: Kinetic and kinematic workspaces of the index finger following stroke. *Brain* **128**, 1112–1121 (2005). doi:[10.1093/brain/awh432](https://doi.org/10.1093/brain/awh432)
77. Térémetz, M., Colle, F., Hamdoun, S., Maier, M.A., Lindberg, P.G.: A novel method for the quantification of key components of manual dexterity after stroke. *Journal of NeuroEngineering and Rehabilitation* **12**(1) (2015). doi:[10.1186/s12984-015-0054-0](https://doi.org/10.1186/s12984-015-0054-0)
78. Kwakkel, G., Kollen, B.J., Van der Grond, J.V., Prevo, A.J.H.: Probability of regaining dexterity in the flaccid upper limb: Impact of severity of paresis and time since onset in acute stroke. *Stroke* **34**(9), 2181–2186 (2003). doi:[10.1161/01.STR.0000087172.16305.CD](https://doi.org/10.1161/01.STR.0000087172.16305.CD)
79. Zeman, B.D., Yiannikas, C.: Functional prognosis in stroke: Use of somatosensory evoked potentials. *Journal of Neurology Neurosurgery and Psychiatry* **52**(2), 242–247 (1989). doi:[10.1136/jnnp.52.2.242](https://doi.org/10.1136/jnnp.52.2.242)
80. Lindenberg, R., Renga, V., Zhu, L.L., Betzler, F., Alsop, D., Schlaug, G.: Structural integrity of corticospinal motor fibers predicts motor impairment in chronic stroke. *Neurology* **74**(4), 280–287 (2010). doi:[10.1212/WNL.0b013e3181ccc6d9](https://doi.org/10.1212/WNL.0b013e3181ccc6d9)
81. Thickbroom, G.W., Byrnes, M.L., Archer, S.A., Mastaglia, F.L.: Motor outcome after subcortical stroke: MEPs correlate with hand strength but not dexterity. *Clinical Neurophysiology* **113**(12), 2025–2029 (2002). doi:[10.1016/S1388-2457\(02\)00318-8](https://doi.org/10.1016/S1388-2457(02)00318-8)
82. Fuentes, C.T., Bastian, A.J.: Where is your arm? Variations in proprioception across space and tasks. *Journal of Neurophysiology* **103**(1), 164–171 (2010). doi:[10.1152/jn.00494.2009](https://doi.org/10.1152/jn.00494.2009)
83. Marini, F., Ferrantino, M., Zenzeri, J.: Proprioceptive identification of joint position versus kinaesthetic movement reproduction. *Human Movement Science* **62**, 1–13 (2018). doi:[10.1016/j.humov.2018.08.006](https://doi.org/10.1016/j.humov.2018.08.006)
84. Gresham, G.E., Duncan, P.W., Stason, W.B.: Poststroke rehabilitation: Assessment, referral, and patient management. *Topics in Stroke Rehabilitation* **3**(2), 1–26 (1996). doi:[10.1080/10749357.1996.11754112](https://doi.org/10.1080/10749357.1996.11754112)
85. Bain, G.I., Polites, N., Higgs, B.G., Heptinstall, R.J., McGrath, A.M.: The functional range of motion of the finger joints. *Journal of Hand Surgery (European Volume)* **40**(4), 406–411 (2015). doi:[10.1177/1753193414533754](https://doi.org/10.1177/1753193414533754)
86. Wolbrecht, E.T., Rowe, J.B., Chan, V., Ingemanson, M.L., Cramer, S.C., Reinkensmeyer, D.J.: Finger strength, individuation, and their interaction: Relationship to hand function and corticospinal tract injury after stroke. *Clinical Neurophysiology* **129**(4), 797–808 (2018). doi:[10.1016/j.clinph.2018.01.057](https://doi.org/10.1016/j.clinph.2018.01.057)

Figure 1: **Schematic of the robotic platform ETH MIKE.** (a) Subjects are seated in front of the ETH MIKE robot, with their elbow supported on an arm rest. A wrist splint is worn to avoid any compensatory movements at the wrist. The device is inclined by 20° to minimize parallax errors. b) The hand is wrapped around a handle, which is set up in a 30° deviation from the 0° position, ensuring a comfortable resting position for the wrist. The index finger is attached to an adjustable finger module by Velcro straps. The centre of rotation of the end-effector is aligned with the MCP joint. (b) A tablet computer is placed above the hand, removing visual cues from the tested hand and providing an interactive graphical user interface displaying a simple gauge with a red indicator.

Figure 2: **Gauge position matching task for the assessment of proprioceptive impairments.** a) The index finger is first passively moved to a target angle, and the subject then needs to indicate the perceived finger position on the tablet screen (no visual feedback is provided). The outcome measure is the error between indicated angle and target angle. b) There is a high agreement (70%) in impairment classification between the gauge position matching task metric and the clinical measure of proprioception (kUDT). c) The task metric has good ($ICC > 0.75$) reliability on the less affected and excellent reliability ($ICC > 0.90$) on the affected side. d) The box plot indicates a trend of increasing position matching absolute error with increasing proprioceptive impairment (according to kUDT). There is a significant difference between control ($N=62$) and all stroke impaired sub-groups (kUDT=2: $N=14$, kUDT=1: $N=8$, kUDT=0: $N=8$). Acronyms: LA: less affected, A: affected, kUDT: kinesthetic Up-Down Test, ICC: intraclass correlation coefficient, SRD: smallest real difference. Statistical significance: p -value <0.05 : *, p -value <0.01 : **, p -value <0.001 : ***.

Figure 3: Fast target reaching (extension) task for the assessment of motor impairments. a) Subject are instructed to move their index finger as fast as possible from a starting position (60°) to a target position (20°). The outcome measure is the maximum velocity reached during that movement. b) The task has good reliability ($ICC > 0.70$) on the less affected and excellent reliability ($ICC > 0.90$) on the more affected side. c) There is a high agreement (80%) in impairment classification between the task metric and the clinical measure of upper-limb motor function FMA. d) The box plot indicates a tendency of decreasing robotic task performance with increasing stroke severity (mild: $FMA \geq 54$, moderate: $54 > FMA \geq 35$, severe: $FMA < 35$). There is a significant difference between control ($N=62$) and all stroke sub-groups - severe ($N=14$), moderate ($N=8$) and mild ($N=8$). Acronyms: D: dominant, LA: less affected, A: affected, FMA: Fugl-Meyer Assessment, ICC: intraclass correlation coefficient, SRD: smallest real difference. Statistical significance: p -value <0.05 : *, p -value <0.01 : **, p -value <0.001 : ***.

Figure 4: Trajectory following (slow) task for the assessment of sensorimotor impairments. a) Subjects need to actively follow with their index finger a target trajectory displayed on the tablet screen. Control subjects can follow the target more accurately, while stroke subjects are further away from the target and their movement is more fragmented. b) There is a high agreement (70%) in impairment classification between this task metric and the clinical measure of proprioception (kUDT). c) The task has good reliability on both the less affected and on the more affected side ($ICC > 0.70$). d) The box plot indicates increasing tracking error RMSE with increasing proprioceptive impairment (according to kUDT). There is a significant difference between controls ($N=62$) and all stroke sub-groups on the affected side (kUDT=2: $N=14$, kUDT=1: $N=8$, kUDT=0: $N=8$), as well as between less affected side of stroke subjects ($N=30$) and the group with the most severely impaired proprioception ($N=8$). Acronyms: D: dominant, LA: less affected, A: affected, kUDT: kinesthetic Up-Down Test, ICC: intraclass correlation coefficient, SRD: smallest real difference. Statistical significance: p -value <0.05 : *, p -value <0.01 : **, p -value <0.001 : ***.

Figures

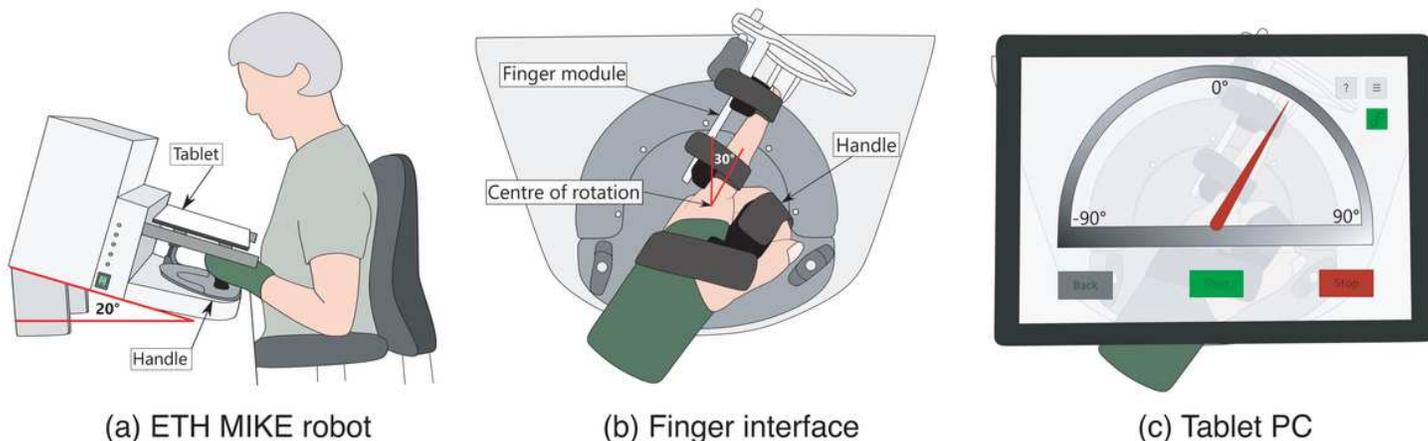
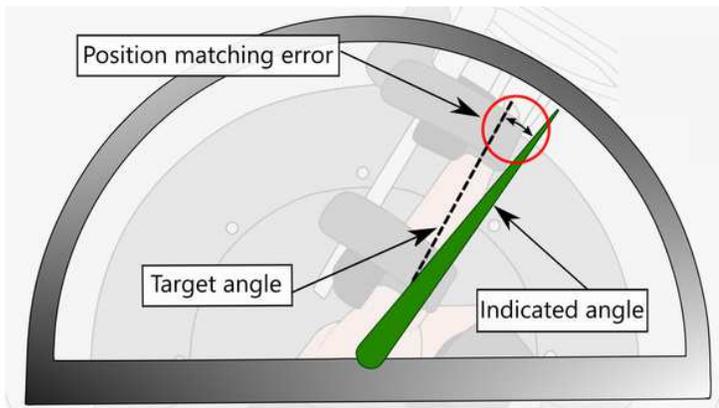


Figure 1

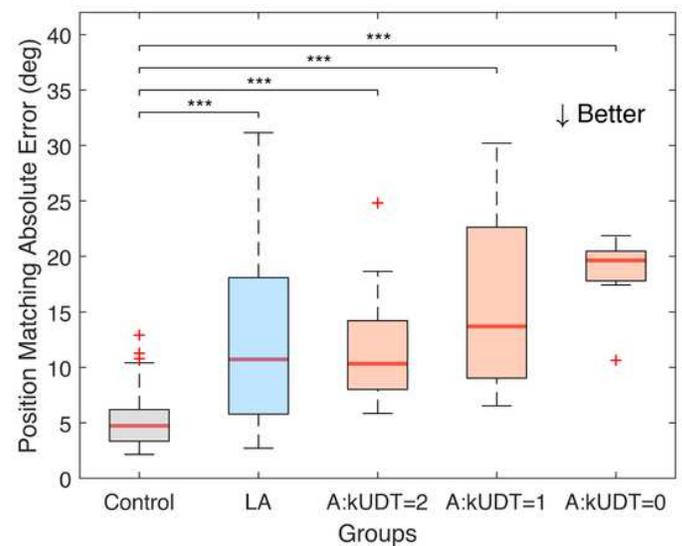
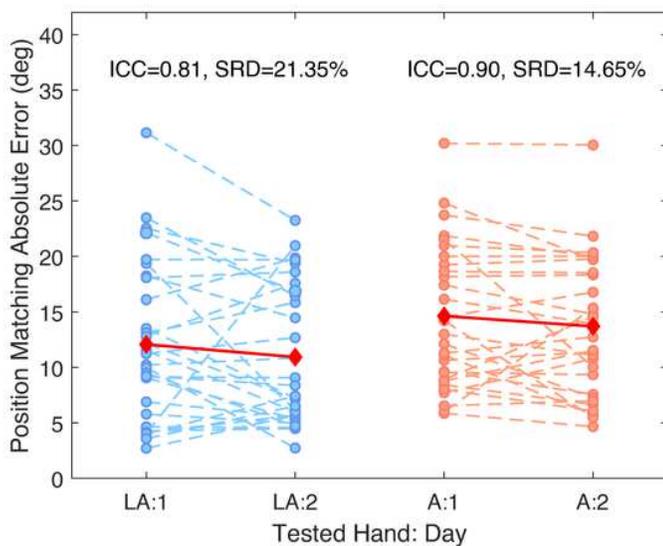
Schematic of the robotic platform ETH MIKE. (a) Subjects are seated in front of the ETH MIKE robot, with their elbow supported on an arm rest. A wrist splint is worn to avoid any compensatory movements at the wrist. The device is inclined by 20° to minimize parallax errors. (b) The hand is wrapped around a handle, which is set up in a 30° deviation from the 0° position, ensuring a comfortable resting position for the wrist. The index finger is attached to an adjustable finger module by Velcro straps. The centre of rotation of the end-effector is aligned with the MCP joint. (c) A tablet computer is placed above the hand, removing visual cues from the tested hand and providing an interactive graphical user interface displaying a simple gauge with a red indicator.



		Gauge Position Matching Task	
		Impaired	Non-impaired
Gauge Position Matching Task	Impaired	13	7
	Non-impaired	3	7
		Impaired	Non-impaired
		Kinesthetic Up-Down Test	

(a) Determination of the position matching error

(b) Impairment classification matrix

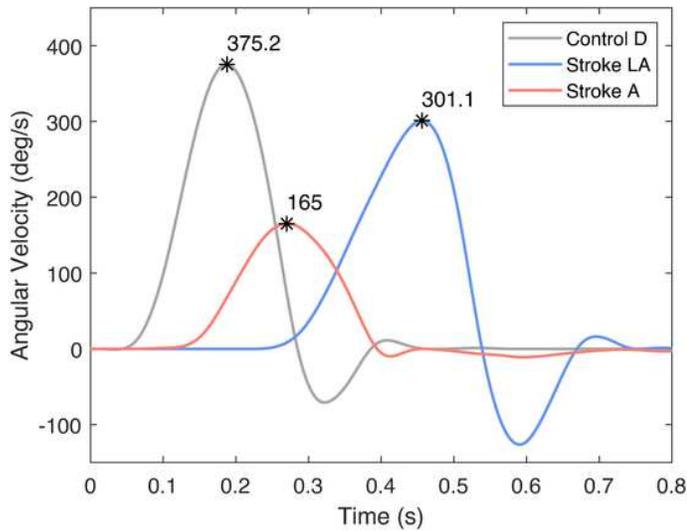


(c) Test-retest reliability

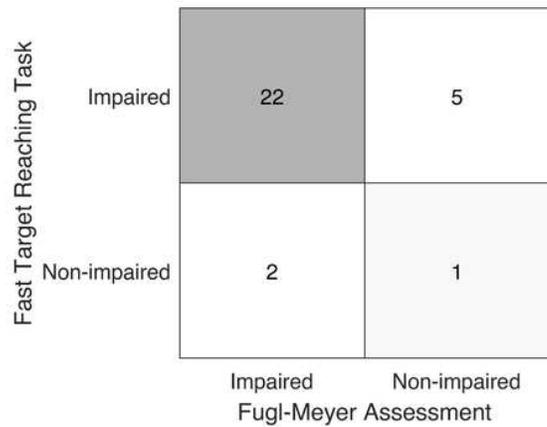
(d) Discriminant validity

Figure 2

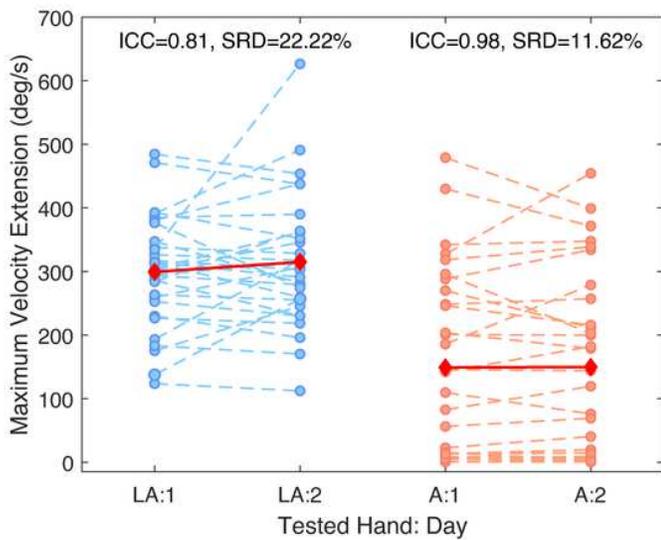
Gauge position matching task for the assessment of proprioceptive impairments. a) The index finger is first passively moved to a target angle, and the subject then needs to indicate the perceived finger position on the tablet screen (no visual feedback is provided). The outcome measure is the error between indicated angle and target angle. b) There is a high agreement (70%) in impairment classification between the gauge position matching task metric and the clinical measure of proprioception (kUDT). c) The task metric has good ($ICC > 0.75$) reliability on the less affected and excellent reliability ($ICC > 0.90$) on the affected side. d) The box plot indicates a trend of increasing position matching absolute error with increasing proprioceptive impairment (according to kUDT). There is a significant difference between control ($N=62$) and all stroke impaired sub-groups (kUDT=2: $N=14$, kUDT=1: $N=8$, kUDT=0: $N=8$). Acronyms: LA: less affected, A: affected, kUDT: kinesthetic Up-Down Test, ICC: intraclass correlation coefficient, SRD: smallest real difference. Statistical significance: p -value <0.05 : *, p -value <0.01 : **, p -value <0.001 : ***.



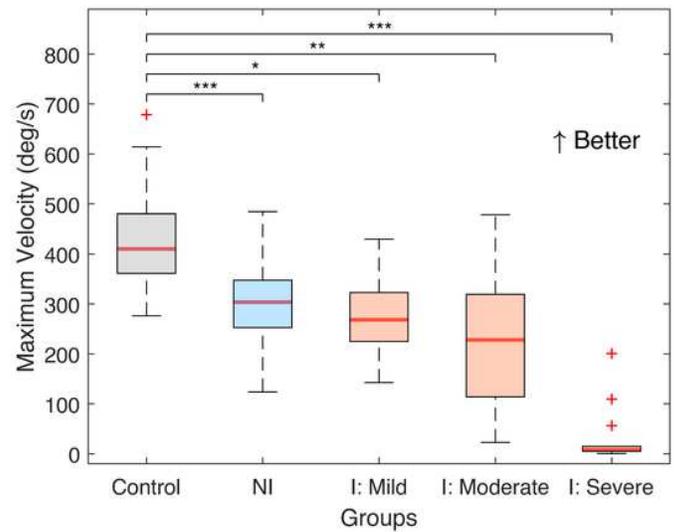
(a) Representative velocity profiles



(b) Impairment classification matrix



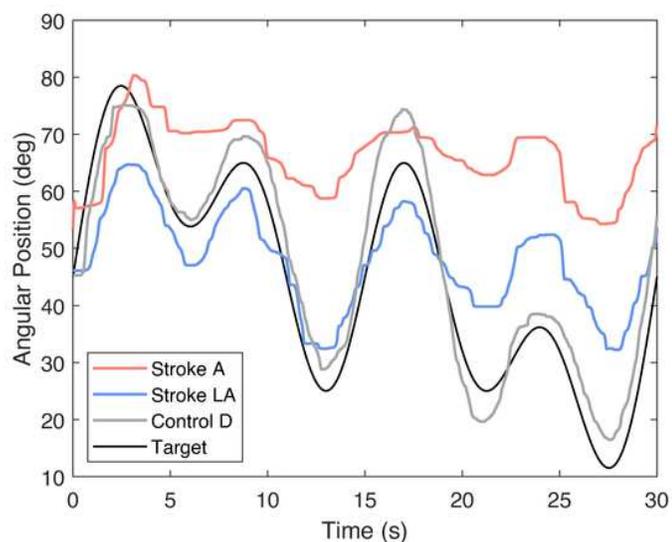
(c) Test-retest reliability



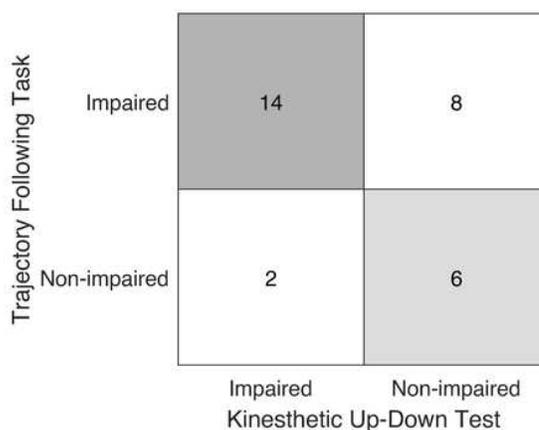
(d) Discriminant validity

Figure 3

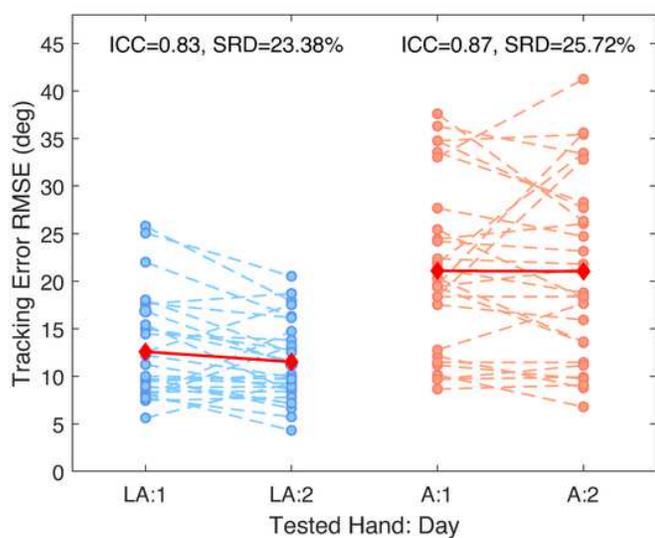
Fast target reaching (extension) task for the assessment of motor impairments. a) Subject are instructed to move their index finger as fast as possible from a starting position (60°) to a target position (20°). The outcome measure is the maximum velocity reached during that movement. b) The task has good reliability ($ICC > 0.70$) on the less affected and excellent reliability ($ICC > 0.90$) on the more affected side. c) There is a high agreement (80%) in impairment classification between the task metric and the clinical measure of upper-limb motor function FMA. d) The box plot indicates a tendency of decreasing robotic task performance with increasing stroke severity (mild: $FMA \geq 54$, moderate: $54 > FMA \geq 35$, severe: $FMA < 35$). There is a significant difference between control ($N=62$) and all stroke sub-groups - severe ($N=14$), moderate ($N=8$) and mild ($N=8$). Acronyms: D: dominant, LA: less affected, A: affected, FMA: Fugl-Meyer Assessment, ICC: intraclass correlation coefficient, SRD: smallest real difference. Statistical significance: $p\text{-value} < 0.05$: *, $p\text{-value} < 0.01$: **, $p\text{-value} < 0.001$: ***.



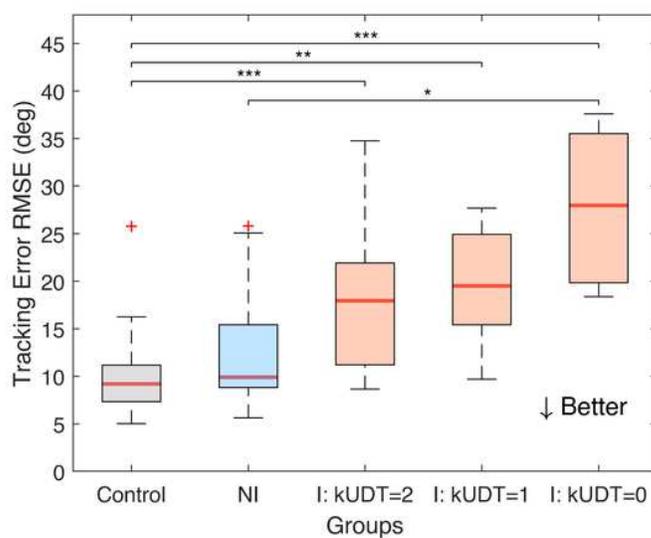
(a) Representative position profiles



(b) Impairment classification matrix



(c) Test-retest reliability



(d) Discriminant validity

Figure 4

Trajectory following (slow) task for the assessment of sensorimotor impairments. a) Subjects need to actively follow with their index finger a target trajectory displayed on the tablet screen. Control subjects can follow the target more accurately, while stroke subjects are further away from the target and their movement is more fragmented. b) There is a high agreement (70%) in impairment classification between this task metric and the clinical measure of proprioception (kUDT). c) The task has good reliability on both the less affected and on the more affected side ($ICC > 0.70$). d) The box plot indicates increasing tracking error RMSE with increasing proprioceptive impairment (according to kUDT). There is a significant difference between controls ($N=62$) and all stroke sub-groups on the affected side (kUDT=2: $N=14$, kUDT=1: $N=8$, kUDT=0: $N=8$), as well as between less affected side of stroke subjects ($N=30$) and the group with the most severely impaired proprioception ($N=8$). Acronyms: D: dominant, LA: less affected, A:

affected, kUDT: kinesthetic Up-Down Test, ICC: intraclass correlation coefficient, SRD: smallest real difference. Statistical significance: p-value<0.05: *, p-value<0.01: **, p-value<0.001: ***.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [20201111ZbytniewskaetalJNERSupplementaryMaterial.pdf](#)