Title
Development and validation of patient-friendly dexterity marker in multiple sclerosis

Author
Gopal, Arpita

Publication Date
2023

Peer reviewed|Thesis/dissertation

Development and validation of a patient-friendly dexterity marker in multiple sclerosis

by
Arpita Gopal

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Rehabilitation Science

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Diane D. Allen
B6D079BF87ED421...

Diane D. Allen
Chair

DocuSigned by:
Nora Fritz
DocuSigned by:345D...

Nora Fritz

DocuSigned by:
Jeffrey Gelfand
1079DFB1BC4D460...

Jeffrey Gelfand

Committee Members

Dedication and Acknowledgements

For the teachers, librarians, and professors who fostered my curiosity and love of learning.

To Logan, for making sure I don't work every weekend.

My parents, for loving their weird science-fair obsessed child.

And to my sister for always doing my homework.

I am grateful to Dr. Riley Bove for her support and mentorship during this dissertation. Many thanks to all the participants who participated in this study.

Epigraph

Take chances, make mistakes, and get messy!

-Ms. Frizzle

Contributions

Gopal A, Hsu WY, Allen DD, Bove R

Remote Assessments of Hand Function in Neurological Disorders: Systematic Review

JMIR Rehabil Assist Technol 2022;9(1):e33157

URL: https://rehab.jmir.org/2022/1/e33157

DOI: 10.2196/33157

Development and validation of a patient-friendly dexterity marker in multiple sclerosis

Arpita Gopal

Abstract

Hand function is critical to activities of daily living (ADL), and a change in function can significantly impact the ability to participate in self-care, occupational and recreational activities. In multiple sclerosis (MS), upper extremity dysfunction is highly prevalent. Current assessments of hand function are exclusively performed in the MS clinic, and do not adequately quantify the quality and variety of movement needed to identify specific dexterity impairments. Moreover, recent research has identified a relationship between diminished hand function and worsening disease progression, suggesting that assessments capable of capturing changes in hand function are needed. While there is an expanding focus on digital tools to meet this need, large-scale uptake has been stymied due to rapidly changing software and hardware, cost of devices such as wearable sensors, proprietary software, and low patient usability. In contrast, patient self-uploaded videos, or "selfies", could represent a more patient-friendly, usable, and economic method for collecting functional data regularly, without the technology burdens. The data generated in this body of work demonstrate that patient-generated videos are a simple, feasible, cost-effective, and scalable mode of data collection with high patient acceptability. From these videos, human pose estimation represents a viable method of quantifying dexterity in relationship to common clinical methods of assessment. Further, exploratory longitudinal data show that pose estimation is capable of identifying changes in dexterity and therefore has the potential to serve as a clinical marker for future trials. Overall, human pose estimation is capable of identifying changes in dexterity longitudinally and captures patient-reported change in daily tasks. Collectively, these findings advance our understanding of digital tools for frequent, granular, and user-friendly dexterity assessments in MS.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Multiple sclerosis (MS) is a progressive, neurodegenerative disease, and a majority of people with MS (pwMS) experience bilateral upper extremity dysfunction.[1,2] Typical age of diagnosis is between 20 and 50 years, when most individuals have demanding professional and personal obligations. Focal areas of demyelination in the brain and spinal cord, as well as subsequent axonal degeneration, result in impairments of upper extremity function, and specifically, of hand function in a majority of people with MS (PwMS).[1,2] Though dysfunction throughout the upper extremity is typically associated with later stages of the disease, loss of fine motor skills is observed in 46% of pwMS within the first 6 years of diagnosis.[3,4]

Hand function is critical to independence in activities of daily living (ADLs). As a result, it is increasingly being used as a marker of disease progression in clinical trials for MS disease modifying therapies (DMT).[5-7] Recent findings suggest that DMTs have an early, neuroprotective effect on hand function, as the slope of degeneration is more visibly slowed in shorter axon pathways due to their lower lesion burden.[6] Studies have noted a relationship between diminished hand function and worsening disease progression, suggesting that capturing changes in hand function could predict progression of overall disability.[2,8] Unfortunately, routine measurement of hand function in pwMS is stymied by a lack of reliable, valid, and remotely-accessible measures.

There are several in-clinic assessments for hand function for neurological populations including the 9 Hole Peg Test (9HPT), Action Research Arm Test (ARAT) and Box and Block Test (BBT)[9]; however, a thorough understanding of their use as measures of clinical disease progression as well as granular assessment of movement quality is lacking. For example, the 9HPT requires individuals to manipulate small pegs into holes on a board. While this task certainly requires finger dexterity and coordination, both essential for ADLs, the score recorded is limited to time to complete the task. Assessments of *quality and variety* of movement (e.g. handling objects of different sizes and weights,

each of which requires a different grip) are needed in order to identify specific impairments for intervention. Further, assessments that are correlated with patient-reported difficulties are needed.

Impairments of upper extremity function can be caused by lesions in pyramidal, somatosensory, and cerebellar structures. In order to categorize the extent of impact on function and disability, we will utilize the International Classification for Functioning, Disability, and Health (ICF)[10] framework. The ICF centers on three components: (1) body structure/function (e.g. grip and grasp impairments); (2) activity (e.g. reduced independence with ADLs like dressing, and instrumental ADLs (like cooking); and finally participation (e.g. limited ability to perform roles at work or for leisure activities).[11]

In other neurological conditions such as stroke and Parkinson's disease, hand function is commonly evaluated as a measure of disease status.[12,13] In MS as well, patients experience hand-related symptoms such as spasticity, weakness, and loss of sensation[14] suggesting that similar methods of assessment could be adopted for MS. Though most assessments of hand function are performed in-clinic, rehabilitative services are increasingly being delivered via telecommunication platforms in order to expand access and treatment dosing and reduce costs, thus requiring robust, in-home assessments of function and change.

Currently Available Remote Assessments of Dexterity

The shift toward telehealth services, especially in light of the COVID-19 pandemic, requires new methods of assessment adaptable to remote delivery. Remote monitoring can provide clinicians with valuable, real-time data to inform treatments.[15] However, there are no validated, patient-operated, remote assessments available to monitor all ICF domains of hand function in MS. Clinicians can assess gait[16], balance[17], and cognitive impairments[18] through remote platforms, but are similarly assessing hand function is not as readily available. Remote-based assessments, such as the Roche group's FLOODLIGHT tool, are being piloted[19] to evaluate hand function in MS, but do not encompass all domains of the ICF model, namely activity and participation limitations. Further, while smartphone-based

assessments are convenient, they require maintenance and updates to keep up with changing hardware which can lead to implementation challenges.

Preliminary studies have demonstrated good reliability and validity of remote assessments of dexterity against gold-standard, in-clinic assessments in MS.[19] Though there is evidence to show that hand function worsens with neurological disease progression,[8] most notably in progressive subtypes of MS, there is no *prospective analysis* of remote assessment of hand function to capture this. These assessments evaluate specific body structure/function impairments, and there is no comprehensive, simple, cost effective, patient-deployable tool available to quantitatively assess hand function and related disability.

Pose Estimation

In an effort to bridge this gap and develop a patient-friendly, cost efficient, functional, remote assessment of hand function, we employed pose estimation algorithms. Human pose estimation is a computer vision task that includes detecting, associating and tracking key points (e.g., body joints such as "right knee" and "left shoulder"). Specifically, computer vision is a field of artificial intelligence enabling computers to derive information from images and videos. By evaluating many images and videos of humans, computers can develop an understanding of human body language. The development of these networks allows researchers to quantify kinematics through videos of patients performing daily activities, especially those related to activity and participation domains of the ICF.

Implications

Important insights are to be gained from prospective, longitudinal studies given the slow rate of progression of hand function impairments. In addition to developing and validating more sophisticated assessment tools, such as human pose estimation, their feasibility and acceptability must also be determined, especially in a longer trial. Additionally, the frequency of assessment needs to be identified to capture changes in function without overburdening patients.

Focus of dissertation

My dissertation aims to demonstrate that remote, quantitative measures of hand function in pwMS can serve as a viable marker of disease progression. Accessible remote assessments will enable clinicians to examine hand function more regularly as part of a comprehensive disease management plan. In Chapter 2, I present data from a systematic review of currently available tools for remotely assessing hand function in people with neurological disorders. These data establish the paucity of assessments that are cost-effective, scalable, and evaluate key functional domains. In Chapter 3, I demonstrate the validity of human pose estimation as a method of quantifying dexterity in relationship to common clinical methods of assessment, such as the 9 hole peg test. These data demonstrate the simplicity of the methodology as well as the high-quality data generated via pose estimation algorithms. In Chapter 4, I present data demonstrating the feasibility, participant acceptability, and clinical implications of pose estimation as a method of dexterity assessment in pwMS. Altogether, these studies will demonstrate the importance of granular, feasible, functional and remotely accessible measures of hand function in pwMS.

Chapter 2: Remote Assessments of Hand Function in Neurological Disorders: A Systematic Review

Abstract

Objective. While assessment of hand function can monitor loss of fine motor skills in people with neurological disorders, concurrent mobility impairments may hinder regular access to tools in-clinic. Remote assessments could facilitate tracking of limitations, aiding in early diagnosis and intervention. We systematically reviewed existing evidence regarding remote assessment for hand function in populations with chronic neurological dysfunction. *Methods*. PubMed/MEDLINE, CINAHL, Web of Science, and Embase were searched for studies that reported remote assessment of hand function (i.e., outside of traditional in-person clinical settings) in adults with chronic, central nervous system disorders. We excluded studies that included participants with orthopedic upper limb dysfunction or utilized tools for intervention and treatment. We extracted data on hand function domains evaluated, validity and reliability, feasibility, and stage of development. *Results*. 74 studies met inclusion criteria: for Parkinson's disease (PD) (57 studies); stroke (9); multiple sclerosis (6); spinal cord injury (1); and amyotrophic lateral sclerosis (1). Three assessment modalities were identified: external device (e.g., wrist-worn accelerometer), smartphone/tablet, and telerehabilitation. Feasibility and overall participant acceptability were high. The most common hand function domains assessed included: 1) finger tapping speed (fine motor control and rigidity); 2) hand tremor (pharmacological and rehabilitation efficacy); 3) finger dexterity (manipulation of small objects required for daily tasks); and handwriting (coordination). While validity and reliability data were heterogenous across studies, statistically significant correlations with traditional in-clinic metrics were most commonly reported for telerehabilitation and smartphone/tablet applications. The most readily-implementable assessments were smartphone/tablet-based. *Conclusions*. Findings show that remotely assessing hand function is feasible in neurological disorders. While varied, the assessments allow clinicians to objectively record performance in multiple hand function domains, improving the reliability of traditional in-clinic assessments. Remote assessments, particularly via telerehabilitation and smartphone/tablet-based applications that align with in-clinic metrics, facilitate

clinic-to-home transitions, have few barriers to implementation, and prompt remote identification and treatment of hand function impairments.

Introduction

Normally functioning human hands allow for everyday participation in self-care, work, and leisure roles that involve precise grip and object manipulation.[20] Specifically, daily activities and fine motor tasks require finger dexterity, thumb-finger opposition, and hand opening-closing that adapt to task requirements, including those needed to navigate the 'digital world'.[21] Unfortunately, chronic disorders of the central nervous system (CNS) can impair hand function even during early stages of disease.[22] Damage to the CNS, including the spinal cord, can result in tremor, spasticity, sensory loss, weakness, and coordination loss in the upper limbs which can negatively impact the ability to adapt to task requirements, thus limiting independence with activities of daily living (ADLs) and quality of life (QoL).[22] For example, a majority of individuals with Parkinson's disease (PD) develop a hand tremor over the course of the disorder, leading to difficulty with precise finger and hand movements.[23] Also, ischemic strokes occur most commonly in the cortical regions supplied by the middle cerebral artery,[24] affecting areas of the motor and sensory cortices responsible for fine motor activity of the hands.[25] In these disorders and others, evaluating hand function at regular intervals can detect changes signaling neurological decline or monitor response to disease-modifying therapies, symptomatic therapies, and/or rehabilitation.

While assessments of hand function are routinely performed in the clinic, clinicians have an increasing interest in deploying tools to measure hand function remotely. In-home remote monitoring of function in general provides benefits to patients with increased convenience, reduced travel, and ability to capture data more frequently. Over the last decade, many studies have examined remote monitoring devices in healthy and diseased populations.[26,27] For example, in multiple sclerosis (MS), studies have shown that continuous remote monitoring of ambulatory step count can capture - and even predict changes in - MS-related disability, and can serve as a longitudinal outcome measure for targeted intervention.[15,28] To date, reviews have mainly focused on lower extremity function or overall physical

activity[29]; In fact, the methodological discrepancies in remote device use and reporting regarding hand function have yielded conflicting results in terms of validity, reliability, and ease of clinical use.

In this systematic review, we evaluate the existing evidence regarding remote assessment devices for hand function in populations with chronic, CNS disorders. We specifically examined evidence of validity, reliability, and feasibility for each domain of hand function and the stage of development of the assessments. We expect our findings to facilitate ready implementation of remote assessment of hand function in prevalent neurological disorders.

Methods

Eligibility Criteria

This review was structured using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[30] framework. Studies were included based on the following criteria: 1) participants had chronic, neurological pathologies of the CNS, 2) participants were aged 18 years or older, 3) studies were peer-reviewed and original, 4) designed to objectively assess hand function; 5) and the assessments were deployable remotely (i.e., outside of traditional in-person clinical settings). Studies were excluded if they were: 1) conducted in participants with orthopedic impairments of the wrist and/or hand, 2) conducted in non-human primates, 3) designed as an intervention to improve an aspect of hand function, or 4) not published in English.

Search Procedures

A literature search was performed using the following databases: PubMed, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Web of Science, and Embase. The search was conducted using both MeSH (Medical Subject Headings) terms and the following keywords independently and in combination: *remote, assessment, outcome, test, measurement, hand, upper extremity, arm, function*. Two researchers (AG and WYH) independently assessed articles for relevance

and adherence to eligibility criteria. Studies were recursively searched to identify cited and cited-by articles.

Data Extraction and Categorization

To evaluate the methodological quality of the included studies, we used the National Institutes of Health quality assessment for observational cohort and cross-sectional studies.[31] Each study was evaluated according to 8 criteria. Overall study quality was assessed as "good" (>5 criteria met), "fair" (4-5 criteria met), or "poor" (<5 criteria met).

Data were extracted (AG) and checked (WYH); discrepancies were resolved through discussion with the senior author (RB). Variables of interest included participant demographics, study design and duration, device type and modality, disease-specific severity levels, comparison assessments, and stage of development/implementation. Participant satisfaction with study protocol and assessment and time to complete the novel assessment were extracted when available. Statistical data extracted included concurrent validity (defined as the comparison between a new test and a well-established one[32]), and reliability (defined as a measure of stability or consistency[33]).

Selected studies evaluated many variables relating to hand function. To compare the most salient domains across studies, we classified assessments into the following hand function domains based on the Functional Repertoire of the Hand established by the American Journal of Occupational Therapy[34]: 1) finger tapping: speed and accuracy of finger taps onto a pre-specified target; 2) whole hand grasp: range of motion and coordination of full hand movement; 3) pincer grasp: range of motion and coordination of thumb to index finger movement; 4) hand tremor: quantification of tremor distal to the wrist at rest; 5) reaction time: time to respond to a pre-determined stimulus using only fingers; 6) pinch and grip strength: quantification of maximum pinching and gripping strength; 7) finger dexterity: in-hand manipulation of an object; 8) handwriting: clarity and accuracy in drawing or writing; 9) ADLs: tasks required for self-care independence;[35] and 10) Instrumental ADLs (IADLs): tasks required for household or community-level independence.[36]

Results

Search Strategy

A search of databases in June 2021 identified 1295 studies with an additional 33 studies through recursive searches. After title and abstract screen and removal of duplicates, 122 studies remained and the full texts were assessed for eligibility based on the inclusion and exclusion criteria. Fifty full-text studies were excluded for not meeting inclusion criteria. The final 74 studies were confirmed by a second reviewer (WYH) to have met all eligibility criteria. The PRISMA diagram of the search process is outlined in Figure 1. Individual studies are summarized in Table 1. Of the studies reviewed, a majority (N = 49) were rated "good" in terms of overall methodological quality; 14 were rated "fair; and 9 were rated "poor." Study quality is summarized in Table 2.

Modalities of hand function assessment

Three different modalities of assessment devices were used across the included studies, summarized in Table 1. The most frequently utilized assessment was an external device specific to hand assessment with the most common types being wrist-worn accelerometers[37-55] and specialized keyboards.[56-65] These designated external devices allowed for collection of information on reaction time, finger tapping speed, and finger dexterity. While many study authors noted that their external devices were able to capture granular, specific data, many devices were developed under proprietary agreements and are not commercially available currently. The second most common type of assessments were generic smartphone and/or tablet-based electronic devices adapted for hand assessment[19,66-76] or suites of assessments.[77-83] These assessments included an application designed to test finger tapping speed as well as accuracy of drawing and tracing various shapes. Such applications facilitated the gathering of data on specific hand function domains at a relatively low cost for people who already have these electronic devices. Finally, three studies utilized telerehabilitation platforms to validate remote administration of well-established in-clinic assessments.[84-86] For example, Amano et al.[86] validated administration of the

Fugl-Meyer Assessment (FMA) and Action Research Arm Test (ARAT) over telehealth platforms, allowing clinical researchers to gather standardized outcome data over secure telehealth tools.

A majority of the included studies (N = 51) performed same-day, cross-sectional validation experiments where participants completed novel and comparison assessments at the same time point. However, 21 studies[41,43,54,55,60,64,65,80,81,83,87-98] remotely monitored participants' hand function longitudinally. The duration of remote monitoring period was 3 days[55] to 3 years.[95] Participant retention and adherence were reported by four studies,[78,83,92,94] all of which had >90% participant retention.

Target population

The included studies targeted five populations of patients with neurological conditions. The majority of studies (N = 57) included individuals with PD.[55,63,82,84,85,87,90,99] Other populations evaluated were stroke (N = 9)[88,89,100] and MS (N = 6).[19,83,101] Neurological conditions designated spinal cord injury (SCI)[102] and amyotrophic lateral sclerosis (ALS)[65] were described in one study each.

Most included studies evaluated individuals with mild to moderate disease severity on average, as graded by established disease-specific metrics (e.g., Movement Disorder Society- Unified Parkinson's Disease Rating Scale (MDS-UPDRS); Expanded Disability Status Scale (EDSS) for people with MS).[19,55,63,82,84,86,87,90,100] Six studies specified in their inclusion criteria to limit recruitment to participants with mild to moderate disease severity.[55,59,71,76,80,88]

The sample sizes of studies varied between 1 (case study)[44] and 495 participants[83] in the experimental groups. A majority of studies (N = 41) included control groups of healthy individuals or those with non-neurological conditions to determine discriminant validity of the assessments (Table 1).

Validity and reliability

Fifty-four heterogeneous studies reported validity data in comparison to well-established in-clinic assessments (Table 3). Nine studies examining external devices reported high, statistically significant correlation with well-established assessments.[37,38,65,68,70,89,90,101,102] Six studies utilizing smartphone

assessment[46,67,70,83,97,103] and one study utilizing telerehabilitation[86] found moderate to high, statistically significant correlation with well-established assessments.

Fifteen heterogeneous studies reported reliability statistics. Two telerehabilitation assessments[85, 86] revealed high, statistically significant inter-rater reliability. One external device assessment[94] revealed high though statistically insignificant reliability.

Hand function domain, based on the Functional Repertoire of the Hand[34]

Finger tapping speed. The most common hand function domain assessed was finger tapping speed.[40,49,53,55-57,60-63,66,67,69-72,75-79,81-84,87,90,92,98,104-106] Finger tapping can provide clinicians with an understanding of fine motor control and stiffness, especially in individuals with spasticity. Of the included studies that examined finger tapping, Albani et al.[90] reported the highest correlation with MDS-UPDRS scores in participants with PD. In their study, the authors used an external device: a gesture-based tracking system involving a specialized depth-camera and gloves with colored markers to track and quantify fine hand movements. The MDS-UPDRS item on finger tapping relies on visual assessments of finger tapping (e.g., interruptions in tapping rhythm), and specialized equipment such as an external device aid in quantifying finger tapping capability.[90]

Hand tremor. The second most commonly assessed domain was hand tremor, a prevalent impairment in many neurological disorders. Quantifying tremor can help determine the efficacy of pharmacological and rehabilitative therapies. The included studies that examined this domain were conducted in participants with PD.[37-39,41,42,44-48,50-52,54,55,64,68,73,77,78,80,81,84,85,91,92,94,96,97,99,107] Hoffman et al.[85] found a 100% agreement of their visual examination of hand tremor at rest in their evaluation of telerehabilitation administration of the MDS-UPDRS assessment in comparison to in-clinic evaluation. Sigcha et al.[97] developed a novel smartphone application utilizing an internal gyroscope and accelerometer to measure resting hand tremor. This method had a strong correlation ($r = 0.97$) with in-clinic MDS-UPDRS resting hand tremor scores.

Finger dexterity. The third most commonly assessed domain was finger dexterity.[59,63,65,84,85,89,101,102,108] Finger dexterity assessment tasks included manipulation of small objects (e.g., 9 hole peg test (9HPT), coin rotation test) which are useful metrics of fine motor control required for ADLs such as buttoning clothing. Finger dexterity was examined in all five of the neurological conditions examined in this review. Of the included studies examining participants with PD, Cabrera-Martos et al.[84] found a mean (SD) of 0.3 (1.2) difference in scores between telerehabilitation and in-clinic administration of the coin rotation task[109] in the affected limb. Similarly, using telerehabilitation to examine the pinch domain for participants with stroke, Amano et al.[86] reported a Spearman's rho of 0.99 between telerehabilitation and in-clinic administered items. In participants with MS, Dubuisson et al.[101] validated an external device, a cardboard 9HPT with a correlation of 0.96 between this novel assessment tool and a standard, plastic 9HPT.

Handwriting. Six studies[19,66,74,85,95,98] examined handwriting accuracy, a specific and sensitive measure of fine motor coordination. The greatest accuracy in comparison to in-clinic assessments was reported by Hoffman et al.,[85] who found a high percentage of agreement (85%) between in-clinic measures and an external telemetry device of the MDS-UPDRS item? for handwriting.

Specific functions. Eight studies evaluated specific functional domains. Grip and pinch strength were examined in three studies,[85,89,102] using remote deployment of these standard in-clinic metrics. Prochazka et al.[102] evaluated the validity of a novel external device to collect force data from grip and pinch tasks, finding a coefficient of determination (R^2) of 0.88 between the remote device and in-clinic administered ARAT. Three studies[43,85,100] specifically examined ADLs and IADLs. Hoffman et al.[85] compared in-clinic and telerehabilitation-administered functional independence measures (FIM) and found 100% agreement in scores for eating, and 91.7% agreement for dressing. Bochniewicz et al.[100] developed a wrist worn accelerometer to capture and quantify disability in individuals post-stroke. The protocol simulated IADLs such as doing laundry and shopping in a grocery store, and the authors reported 88.4% accuracy compared to ARAT scores of upper extremity functional use.

Participant acceptability

In PD populations, seven studies reported participant acceptability and usability of assessments. Albani et al.[90] found that participants rated the hand gesture-based tracking system 5.9/7 on a post-study usability questionnaire, indicating ease of use, high interface quality, and usefulness. In three studies,[42,48,55] participants using wearable sensors to monitor hand tremor and finger tapping found the devices comfortable and easy to use. Both Goetz et al.[92] and Ferreira et al.[41] reported >80% participant satisfaction with external devices to examine hand tremor. Mitsi et al.[82] found that 76% of participants using a tablet-based assessment for finger tapping[82] and reaction time found it easy to use, with an additional 63% reporting willingness to use it long-term to monitor disease activity.

In stroke populations, Burdea et al.[88] asked both participants and caregivers for feedback on their videogame-like assessment and intervention using a 5-point study-specific Likert scale (higher scores indicating statement agreement). Participants reported that the device was moderately easy to use (mean score = 3.1/5.0), that they would encourage others to use it (mean score = 4.3/5.0), and that they liked the system overall (mean score = 4.2/5.0). However, participants did encounter some technical difficulties during use (mean score = 2.2/5.0). Caregivers also found the device setup appropriate for the home environment and easy to use (mean score = 3.5/5.0).

In people with MS, Dubuisson et al.[101] reported that 66.7 % of participants preferred the portable, in-home 9HPT in comparison to the standard in-clinic version.

Safety

Two studies reported safety data.[55,85] Hoffman et al.[85] reported that participants who received assessment via telerehabilitation were accompanied by a researcher to ensure safety. Boroojerdi et al.[55] employed a wearable patch and reported no adverse skin reactions at the application site or device malfunction. Adverse events were not reported in the included studies.

Stage of development and implementation

Because the assessments in this review were novel, availability for clinical implementation is varied. A majority of studies (N = 44) evaluated assessments requiring specialized equipment for implementation. These devices included specialized cameras, wearable devices, electromyography, and specialized keyboards. While not an application, the cardboard 9HPT developed by Dubuisson et al.[101] was designed specifically to be environmentally friendly, cost-effective, and used by patients at home. The remaining external devices evaluated in this review were designated as developmental, with a need for subsequent safety and prospective studies on usability prior to clinical use.

Two studies employing telerehabilitation methods required videoconferencing devices and stable internet connection for both provider and patient for implementation. However, though Hoffman et al.[85] similarly employed telerehabilitation methods, their protocol required participants to use clinical equipment during in-home assessments (e.g., hand dynamometer, 9HPT), potentially limiting widespread implementation

Twenty studies utilized a smartphone or tablet-based application to administer assessments. The FLOODLIGHT application studied by Creagh et al.[19] is currently available for download for iOS and Android devices. The remaining applications were study-specific developments, but given compatible devices and secure broadband internet connection availability, have limited barriers to implementation.

Discussion

The purpose of this review was to systematically gather available literature on remote assessments for monitoring hand function in people with central, chronic, neurological diseases. The search yielded 74 studies that met inclusion criteria, and 71 unique assessments were examined for validity, reliability, and clinical implementation. A wide variety of metrics were collected on a number of hand function domains including amplitude of finger tapping, finger dexterity, hand tremor, and ADL independence. Altogether, the studies provide a number of insights but to date no single tool, or

combination of tools, that validly and reliably captures hand function across these major neurological conditions.

Many of the studies were of good quality; several study characteristics were found to enhance their quality. Including healthy controls as a comparison, when available, helped to demonstrate the discriminant validity of the novel assessments examined. A majority of studies included participants with lower disability status, which likely allowed for more dynamic testing of hand function domains. Unfortunately, a majority of included studies reported statistically insignificant association with standard in-clinic metrics. Since prior literature suggests that traditional in-clinic assessments have limited granularity for upper limb function in neurological populations, differences between the novel assessments and these traditional in-clinic tests could indicate that the new tools capture additional aspects of function (e.g., quantifying pincer grasp) relative to the traditional in-clinic assessments, or vice versa. Additionally, few studies reported reliability, especially inter-rater reliability, suggesting the need for more research-- and that the included tools remain primarily in the development phase.

The most commonly assessed hand function domain was finger tapping speed, with moderate to high agreement across comparison assessments. The finger tapping test is a valid and reliable measure of bradykinesia in PD[110] and predictor of ADL independence in acute stroke.[111] It is relatively simple to quantify finger tapping in-clinic or via smartphone/tablet application by counting the number of finger taps within a specific time frame. While overall construct validity and participant satisfaction were high, further work in other hand function domains will help determine the most salient predictors of ADL independence and response to treatment and intervention.

This review highlights important aspects of feasibility of remote evaluations. Participant and caregiver satisfaction, when reported, were moderate to high for these technologically innovative assessments. This suggests that participants found the novel assessments easy to use and effective in evaluating their hand function despite being non-traditional. Further, 21 of the included studies demonstrated the feasibility of remotely monitoring hand function over multiple days. This is a key

finding, since long-term monitoring of hand function in a patient's natural environment has the potential to identify changes in real-time, allowing for timely intervention modifications.

Regarding patient safety, while the included assessments are non-invasive and pose a relatively low safety risk, ensuring secure transfer of data especially with internet-based communication (e.g., telerehabilitation, smartphone/tablet-based applications) between patient and clinician is critical to confidentiality and HIPAA compliance. Future studies should report data storage and encryption methodologies.

The assessments evaluated were in varied stages of development and implementation. The most readily implementable types of assessment were those utilizing telerehabilitation or smartphone/ tablet-based applications. According to 2019 data, 85% of Americans own a smartphone, and 93% use the internet regularly—of whom, 75% of whom use a home high-speed broadband network.[112] Given these statistics, telerehabilitation and application-based assessments, if interoperable across devices, might be relatively accessible to a majority of patients. Lower costs could make clinical implementation less of a challenge. Furthermore, with no specialized devices to purchase or distribute to patients, clinics could similarly benefit from these cost-effective measures.

One major limitation of the review is the heterogeneity of hand function domains evaluated, which when compounded with the methodological variability (in comparison assessments, inclusion criteria, and statistical approaches), made it difficult to compare the various tools. Future studies including more homogeneous patient populations and standardized reporting of correlation coefficients with comparison assessments will facilitate analysis across domains and assessment types. A second limitation was the paucity of studies conducting repeated trials of the assessments, limiting identification of any practice effects with use of a new device. In repeat trials of smartphone-based assessments, performance improved in the first 10 trials due to a practice effect, followed by a narrowing of variance as the practice effect waned and familiarity with the assessment increased.[113] Follow-up studies should include repeated trials, preferably over multiple days to capture these effects as well as fluctuations in disease progression. Third, the effect of confounding variables (e.g., disease-modifying therapies, age,

disease duration) were infrequently described in validity statistics; generalizability of this review should proceed with caution. Fourth, all tools included require active participant engagement as opposed to passive monitoring (e.g., collecting data on dexterity as a participant types to complete a survey). Passive monitoring may be able to capture similar metrics with reduced participant time burden. Finally, we may have missed relevant studies that were published in non-English languages.

This review suggests that remote assessments can be valid and reliable tools to measure hand function impairments in chronic neurological disease, and that doing so is clinically feasible and acceptable to them. In the last decade, personal smartphone and computer ownership has become commonplace—with it, patients and health care providers are able to communicate in real time, opening new avenues for care delivery and disease monitoring. We highlight the current potential to implement remote assessments via telerehabilitation and smartphone/tablet-based applications. As interventions for ambulation and lower extremity function become increasingly robust, these methods will allow clinicians to reliably assess multiple domains of hand function to monitor disease progression and response to intervention.

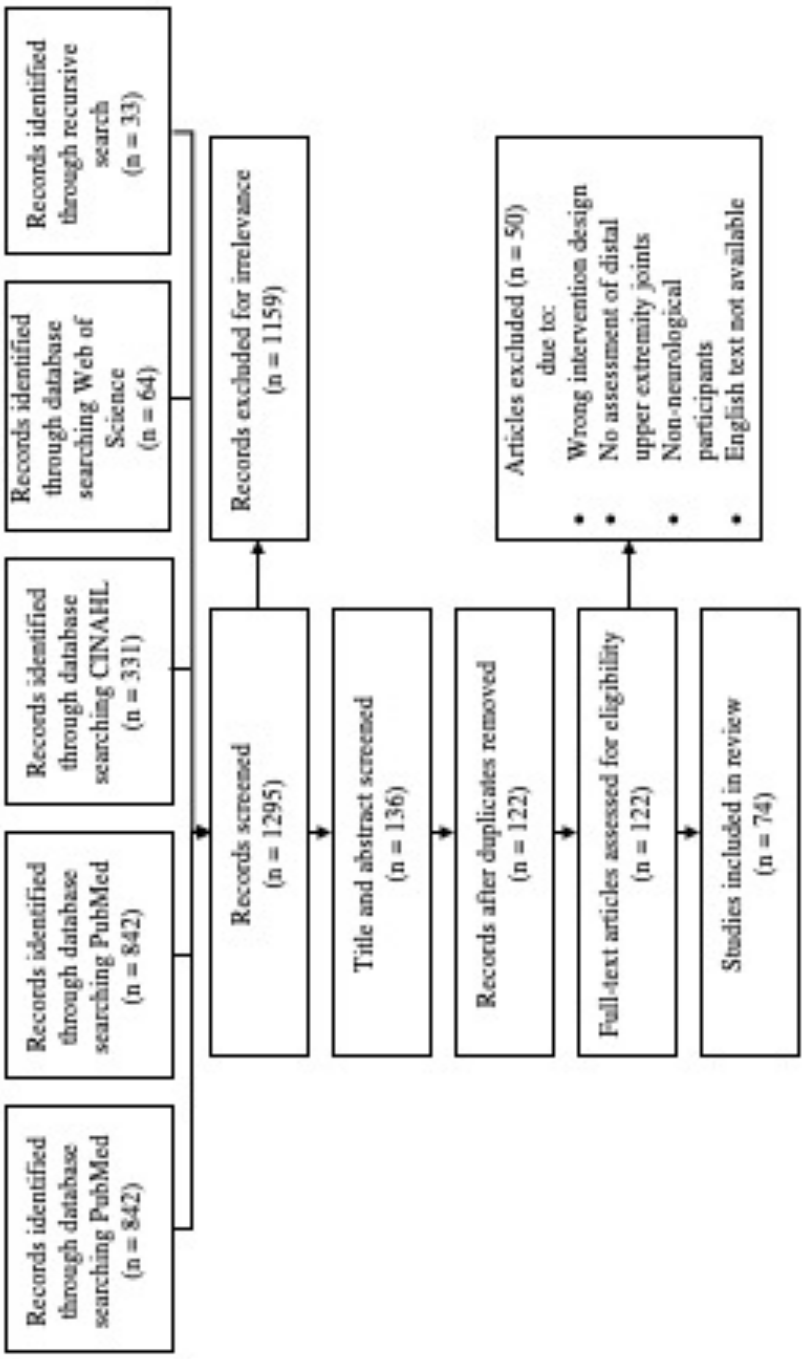


Figure 2. 1 PRISMA diagram outlining study selection

Table 2. 1. Summary of Studies

Author/Year	Study Design (duration)	Diagnosis/ phenotype	Modality of Assessment	Assessment Name	Hand function domains	N		Age [mean (SD)]		Sex	
						Control	Experimental	Control	Experimental	Control	Experimental
Adams (2018)	Longitudinal (8 months)	PD	Smartphone		Hand tremor	128	337				
Aghanavesi (2017)	Cross-Sectional	PD	External Device		Finger tapping Handwriting	22	19	64.2 (7.4)	71.4 (6.3)	14M, 5F	16M, 6F
Akram (2020)	Cross-Sectional	PD	Smartphone		Finger tapping	24	45	66.0 (11.8)	61.9 (7.25)	8M, 16F	26M, 19F
Albani (2019)	Longitudinal (12 weeks)	PD	External Device		Finger tapping Whole hand grasp	15	25	66.4	66.9	9M, 6F	15 M, 10 F
Amano (2018)	Cross-sectional	Stroke	Telerehabilitation		Finger dexterity Whole hand grasp Pincer grasp		30		65.5		18M, 12F
Arora (2015)	Longitudinal (34.4 days)	PD	Smartphone		Finger tapping	10	10	57.7 (14.3)	65.1 (9.8)	40M, 6F	7M, 3F
Arroyo-Gallego (2017)	Cross-Sectional	PD	Smartphone		Finger tapping	27	24	54.35 (13.95)	59.24 (11.43)	4M, 19F	10M, 11F
Bazgir (2018)	Cross-Sectional	PD	External Device		Hand tremor		36		54 (13)		20M, 16F
Bochniewicz (2017)	Cross-sectional	Stroke	External Device		IADLs	10	10	43 ± 15.9	56 ± 10.4	4M, 6F	8M, 2F
Boroojerdi (2019)	Longitudinal (3 days)	PD	External Device	NIMBLE patch	Finger tapping Hand tremor		25		64.2 (7.8)		15M, 10F

Table 2. 1. Summary of Studies

Author/Year	Study Design (duration)	Diagnosis/ phenotype	Modality of Assessment	Assessment Name	Hand function domains	N		Age [mean (SD)]		Sex	
						Control	Experimental	Control	Experimental	Control	Experimental
Burdea (2020)	Longitudinal (4 weeks)	Stroke	External Device	BrightBrainer system	Pincer grasp		7		64.1		4 M, 3F
Cabrera-Martos (2019)	Cross-sectional	PD	Telerehabilitation		Finger dexterity Finger tapping Hand tremor		21		70.9 ± 9.6		11M, 10F
Cai (2018)	Cross-Sectional	PD	External Device	A-WEAR bracelet	Hand tremor	14	34	61.6 (12.8)	64.0 (11.3)	10M, 4F	27M, 7F
Channa (2021)	Cross-Sectional	PD	External Device		Hand tremor	20	20	71.65 (6.87)	70.25 (6.31)	16M, 4F	5M, 15F
Cole (2014)	Cross-Sectional	PD	External Device	Visualbasic	Hand tremor	4	8	54 (16.6)	62.9 (5.3)	4M, 0F	7M, 1F
Creagh (2020)	Cross-sectional	MS	Smartphone	FLOODLIGHT	Handwriting	22	21	34 ± 9	40 ± 8	15M, 7F	49M, 22F
Cunningham (2011)	Longitudinal (4 days)	PD	External Device	electromagnetic tracking system	Hand tremor		10		68.2 (5.2)		9M, 1F
Dai (2021)	Cross-Sectional	PD	External Device/Smartphone	SENSE-PARK	Finger tapping Hand tremor	30	45	61.35 (6.56)	64.53 (9.31)	20M, 10F	35M, 10F
Dubuisson (2017)	Cross-sectional	MS	External Device		Finger dexterity	68	109	41 ± 16.12	49 ± 12.5	18M, 50F	38M, 71 F
Ferreira (2015)	Longitudinal (12 weeks)	PD	External Device	neuroQWERTY	Hand tremor	22		60.2 (9.8)		14M, 8F	

Table 2. 1. Summary of Studies

Author/Year	Study Design (duration)	Diagnosis/ phenotype	Modality of Assessment	Assessment Name	Hand function domains	N		Age [mean (SD)]		Sex	
						Control	Experimental	Control	Experimental	Control	Experimental
Giancardo (2016)	Cross-Sectional	PD	External Device	Kinesia	Finger tapping	43	42	60.1 (10.2)	59.0 (9.8)	17M, 26F	24M, 18F
Giuffrida (2009)	Cross-Sectional	PD	External Device		Hand tremor		60				
Goetz (2009)	Longitudinal (6 months)	PD	External Device		Finger tapping Hand tremor		52		63.8 (8.9)		32M, 20F
Halloran (2016)	Longitudinal (8 weeks)	Stroke	External Device	MOX5	ADLs		24		61.8 (14.3)		18M, 6F
Heijmans (2019)	Case study	PD	External Device		Hand tremor		1		65		1M
Hoffman (2008)	Cross-sectional	PD	Telerehabilitation		Finger dexterity Grip strength Hand tremor Handwriting Pinch strength	6	6		66.1 (8.5)		6M, 6F
Hssayeni (2019)	Cross-Sectional	PD	Smartphone		Hand tremor		24		58.9 (9.3)	14M, 10F	14M, 10F
Iakovakis (2018)	Cross-Sectional	PD	Smartphone	iPrognosis	Finger tapping	13	18	57 (3.9)	61 (8.4)	8M, 7F	14M, 4F
Iakovakis (2020)	Cross-Sectional	PD	External Device		Finger tapping	17	22	54.6 (9.4)	58.6 (8.4)	10M, 7F	16M, 6F

Table 2. 1. Summary of Studies

Author/Year	Jeon (2017)	Jha (2020)	Kim (2018)	Kleinholder mann (2021)	Kostikis (2015)	Lam (2020)	Lee, C (2016)	Lee, S (2018)	Lee, U (2016)	Lin (2019)
Study Design (duration)	Cross-Sectional	Crossover-Randomized	Cross-Sectional	Cross-Sectional	Cross-sectional	Cross-Sectional	Cross-Sectional	Cross-Sectional	Cross-Sectional	Cross-Sectional
Diagnosis/ phenotype	PD	PD	PD	PD	PD	MS	PD	Stroke	PD	Stroke
Modality of Assessment	Smartphone	External Device	External Device	External Device	Smartphone	Smartphone	External Device	Smartphone	External Device	Smartphone
Assessment Name	CloudUPDRS	SNUMAP	Myo armband	Neurokeys		smartphone tapper (SmT)	Kinect v2	HLTapper		mPower
Hand function domains	Hand tremor	Finger tapping Hand tremor	Hand tremor	Finger tapping	Hand tremor	Finger dexterity	Finger tapping	Whole hand grasp	Finger tapping	Finger dexterity Whole hand grasp
N Control					20	24	87		11	15
N Experimental	85	62	92	45	25	102	57	10	9	
N Control		68 (median)			67.2 (6.3)	45.2 (13.5)	53.4 (14.8)		83.8	
Age [mean (SD)] Experimental	65.9 (9.2)		67.1 (9.0)	59.2 (8.9)	70.9 (11.8)	46.4 (10.1)	65.4 (9.0)	58 (16.5)	85.8	59.6 (16.3)
Age [mean (SD)] Control		42M, 20F			10M, 10F	8M, 10F	34M, 53F		3M, 8F	
Sex Experimental	41M, 44F		45M, 47F	34M, 11F	PD: 11M, 12F PD de novo (before	21M, 64F	34M, 23F	6M, 4F	1M, 8F	9M, 6F

Table 2. 1. Summary of Studies

Author/Year	Orozco-Arroyave (2020)	Mera (2012)	Memedi (2015)	Matarazzo (2019)	Mahadevan (2020)	Lopez-Blanco (2019)	Londral (2016)	Lipsmeier (2018)
Study Design (duration)	Cross-Sectional	Cross-Sectional	Longitudinal (3 years)	Longitudinal (6 months)	Cross-Sectional	Longitudinal (12 months)	Cross-sectional	Longitudinal (6 months)
Diagnosis/ phenotype	PD	PD	PD	PD	PD	PD	ALS	PD
Modality of Assessment	Smartphone	External Device	External Device	External Device	External Device	External Device	External Device	External Device
Assessment Name	PD Dr	BRAIN		spiral drawing assessment		APDM, Biostamp		Smartwatch3
Hand function domains	Finger tapping Hand tremor	Finger tapping	Handwriting	Finger tapping	Hand tremor	Hand tremor	Finger dexterity	Finger tapping Hand tremor
N Control	60	10	10	30	50		26	35
N Experimental	23	10	65	29	31	22	19	43
Age [mean (SD)] Control	62.2 (10.2)	53.0 (17.3)	61 (7)	63.0 (56.48-69.44)	43.9 (10.02)		56.23 (7.83)	
Age [mean (SD)] Experimental	68.6 (11.3)	67.8 (8.8)	61.4 (7.4)	65 (11)	59.78 (54.19-68.60)	68.1 (8.13)	72 (7.6)	64 (median)	57.5 (8.45)
Sex Control	30M, 30F	32M, 61F	8M, 2F	5M, 5F	14M, 16F	23M, 27F		27M, 8F
Sex Experimental	11M, 12F	37M, 21F	10M, 9F	8M, 2F	43M, 22F	15M, 14F	20M, 11F	13M, 9F	3M, 16F	35M, 8F

Table 2. 1. Summary of Studies

Author/Year	Schallert (2020)	Sanchez-Perez (2018)	San-Segundo (2020)	Salarian (2007)	Rigas (2012)	Prochazka (2015)	Pratap (2020)	Powers (2021)	Papadopoulos (2021)	Pan (2015)
Study Design (duration)	Cross-Sectional	Cross-Sectional	Cross-Sectional	Cross-Sectional	Cross-Sectional	Cross-sectional	Longitudinal (12 weeks)	Longitudinal (6 months)	Cross-Sectional	Cross-Sectional
Diagnosis/ phenotype	MS, PD, stroke, cerebellar	PD	PD	PD	PD	SCI	MS	PD	PD	PD
Modality of Assessment	External Device	Tablet	External Device	External Device	External Device	External Device	Smartphone	External Device	Smartphone	Smartphone
Assessment Name	BRAIN			Kinesia		Rejoice Arm and Hand Function Test (RAHFT)	SymptoMS Screen			
Hand function domains	Finger tapping Handwriting	Hand tremor	Hand tremor	Hand tremor	Hand tremor	Finger dexterity Whole hand grasp Pincer grasp	Finger tapping	Hand tremor	Hand tremor	Hand tremor
N Control	25	57	12	10	5		134		14	
N Experimental	29			10	18	13	Self-referred:359 Confirmed: [illegible]	225	31	40
Age [mean (SD)] Control	58.3 (16.1)	66.4 (9.0)	62-85	63.6 (10.5)	63.9 (6.2)	24-56	36.9 (11.4)		55.4 (11.7)	
Age [mean (SD)] Experimental	46.6 (14.8)			61.5 (7.8)			Self-referred: 45.2 (11.6) Confirmed: 48.9 (11.2)	71.4 (8.9)	62.1 (7.3)	68.5 (9.5)
Sex Control	14M, 15F			5M, 5F			15M, 27F			
Sex Experimental	10M, 15F	35M, 22F		5M, 5F			Self-referred: 56M, 154F Confirmed: 14M, 78F	156M, 69F		35M, 5F

Table 2. 1. Summary of Studies

Author/Year	Yu (2016)	Wu (2020)	Wissel (2018)	Westin (2010)	Trager (2020)	Tavares (2005)	Stamatakis (2013)	Simonet (2021)	Sigcha (2021)	Shribman (2017)
Study Design (duration)	Longitudinal (12 weeks)	Cross-Sectional	Cross-Sectional	Longitudinal (1 week)	Cross-sectional	Pre/Post-Interventional	Cross-Sectional	Cross-Sectional	Longitudinal (8 weeks)	Cross-Sectional
Diagnosis/ phenotype	Stroke	PD	PD	PD	PD	PD	PD	PD	PD	MS
Modality of Assessment	External Device	External Device	External Device	Smartphone	External Device	External Device	External Device	External Device	Smartphone	External Device
Assessment Name	Quantitative Fugl-Meyer			iMotor					SMART	
Hand function domains	Finger dexterity Pinch strength	Hand tremor	Finger tapping	Handwriting	Finger dexterity Finger tapping	Finger tapping	Finger tapping	Finger tapping	Hand tremor	Finger tapping
N										
Control	24	17	11	60	11	62	36	26	18	39
Experimental										
Age [mean (SD)]										
Control	69.4 (12.8)	63.3 (6.6)	62.5 (10.5)	64.9 (7.3)	63.2 (6.6)	59.9 (8.8)	63.9 (9.1)	63.8 (7.2)	64.9 (7.6)	43.2
Experimental								59.6 (10.9)		
Sex										
Control	16 M, 8 F	11M, 6F	8M, 3F	39M, 21F	12 M, 4F		28M, 8F	11M, 19F	8M, 10F	10M, 29F
Experimental					5M, 6F			17M, 9F		

Table 2. 1. Summary of Studies

Author/Year	Study Design (duration)	Diagnosis/ phenotype	Modality of Assessment	Assessment Name	Hand function domains	N		Age [mean (SD)]		Sex	
						Control	Experimental	Control	Experimental	Control	Experimental
Zambrana (2019)	Cross-Sectional	Stroke	Smartphone	HopkinsPD	ADLs	15	6	31.2 (4.6)	55.3 (16.9)	7M, 8F	4M, 2F
Zhan (2016)	Longitudinal (6 months)	PD	External Device	Axivity AX3	Finger tapping Hand tremor	105	121	45.4 (15.5)	57.6 (9.1)	56M, 49F	71M, 50F
Zhang (2020)	Longitudinal (4 weeks)	PD	External Device		Hand tremor		12		65-85		8M, 4F

Legend **PD: Parkinson's disease; MS: multiple sclerosis; ALS: amyotrophic lateral sclerosis; SCI: spinal cord injury; IADLs: instrumental activities of daily living; ADLs: activities of daily living**

Table 2. 2. Quality Assessment of Studies

Study (Year)	Adams (2018)	Aghanavesi (2017)	Akram (2020)	Albani (2019)	Amano (2018)	Arora (2015)	Arroyo-Gallego (2017)	Bazgir (2018)	Bochniewicz (2017)
Was the research question/objective clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the study population clearly specified?	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was a sample size justification or power description?	No	No	No	No	Yes	Yes	No	No	No
Were the exposure measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the exposure assessed more than once over time?	Yes	Yes	Yes	Yes	No	Yes	Yes	No	No
Were the outcome measures clearly defined, valid,	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Were the assessors blinded to the exposure status of	No	No	No	Yes	Yes	No	No	No	No
Was loss to follow-up after baseline 20% or less?	No	NA	Yes	NR	NA	Yes	NA	NA	NA
Quality Rating (Good, Fair, or Poor)	Poor	Good	Good	Good	Good	Good	Good	Good	Good

Table 2. 2. Quality Assessment of Studies

Study (Year)	Boroojerdi (2019)	Burdea (2020)	Cabrera Martos (2019)	Cai (2018)	Channa (2021)	Cole (2014)	Creagh (2020)	Cunningham (2011)	Dai (2021)
Was the research question/objective clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the study population clearly specified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was a sample size justification or power description?	Yes	Yes	Yes	No	No	No	No	No	No
Were the exposure measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the exposure assessed more than once over time?	Yes	Yes	No	No	Yes	No	No	Yes	No
Were the outcome measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Were the assessors blinded to the exposure status of	No	Yes	Yes	No	No	No	No	No	No
Was loss to follow-up after baseline 20% or less?	Yes	Yes	NA	NA	No	NA	NA	No	NA
Quality Rating (Good, Fair, or Poor)	Good	Good	Good	Fair	Fair	Poor	Fair	Fair	Fair

Table 2. 2. Quality Assessment of Studies

Study (Year)	Dubuisson (2017)	Ferreira (2015)	Giancardo (2016)	Giuffrida (2009)	Goetz (2009)	Halloran (2016)	Heijmans (2019)	Hoffman et al (2008)	Hssayeni (2019)
Was the research question/objective clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the study population clearly specified?	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes
Was a sample size justification or power description?	No	No	No	No	No	No	No	No	No
Were the exposure measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Was the exposure assessed more than once over time?	No	Yes	No	No	Yes	Yes	Yes	No	No
Were the outcome measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Were the assessors blinded to the exposure status of	No	No	No	No	No	No	No	Yes	No
Was loss to follow-up after baseline 20% or less?	NA	No	NA	NA	No	NA	No	NA	NA
Quality Rating (Good, Fair, or Poor)	Fair	Good	Fair	Poor	Fair	Fair	Poor	Good	Good

Table 2. 2. Quality Assessment of Studies

Study (Year)	Lee, C (2016)	Lam (2020)	Kostikis (2015)	Kleinholdermann (2021)	Kim (2018)	Jha (2020)	Jeon (2017)	Iakovakis (2020)	Iakovakis (2018)
Was the research question/objective clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the study population clearly specified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Was a sample size justification or power description?	No	No	No	No	No	No	No	No	No
Were the exposure measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the exposure assessed more than once over time?	No	No	No	No	No	Yes	No	No	No
Were the outcome measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the assessors blinded to the exposure status of	No	No	No	Yes	No	Yes	No	No	No
Was loss to follow-up after baseline 20% or less?	NA	Yes	NA	NA	NA	NA	NA	NA	NA
Quality Rating (Good, Fair, or Poor)	Good	Good	Good	Good	Good	Good	Fair	Poor	Good

Table 2. 2. Quality Assessment of Studies

Study (Year)	Memedi (2015)	Matarazzo (2019)	Mahadeva n (2020)	Lopez-Blanco (2019)	Londral (2016)	Lipsmeier (2018)	Lin (2019)	Lee, U (2016)	Lee, S (2018)
Was the research question/objective clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the study population clearly specified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Was a sample size justification or power description?	No	Yes	No	No	No	No	No	No	No
Were the exposure measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the exposure assessed more than once over time?	Yes	Yes	No	Yes	No	No	No	No	No
Were the outcome measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the assessors blinded to the exposure status of	No	No	Yes	Yes	Yes	No	No	No	No
Was loss to follow-up after baseline 20% or less?	Yes	Yes	NA	Yes	NA	NA	NA	NA	NA
Quality Rating (Good, Fair, or Poor)	Good	Good	Good	Good	Good	Good	Good	Poor	Good

Table 2. 2. Quality Assessment of Studies

Study (Year)	Prochazka (2015)	Pratap (2020)	Powers (2021)	Papadopoulos (2021)	Pan (2015)	Orozco-Arroyave (2019)	Noyce (2014)	Mitsi et al (2017)	Mera (2012)
Was the research question/objective clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the study population clearly specified?	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No
Was a sample size justification or power description?	No	No	No	No	No	No	No	Yes	No
Were the exposure measures clearly defined, valid,	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Was the exposure assessed more than once over time?	No	Yes	Yes	No	Yes	No	No	No	Yes
Were the outcome measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the assessors blinded to the exposure status of	Yes	No	Yes	No	No	No	No	No	No
Was loss to follow-up after baseline 20% or less?	NA	Yes	Yes	NA	No	NA	NA	NA	Yes
Quality Rating (Good, Fair, or Poor)	Good	Good	Good	Poor	Poor	Good	Good	Good	Good

Table 2. 2. Quality Assessment of Studies

Study (Year)	Rigas (2012)	Salarian (2007)	San-Segundo (2020)	Sanchez-Perez (2018)	Schallert (2020)	Shribman (2017)	Sigcha (2021)	Simonet (2021)	Stamatakis (2013)
Was the research question/objective clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the study population clearly specified?	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Was a sample size justification or power description?	No	No	No	No	No	No	No	No	No
Were the exposure measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the exposure assessed more than once over time?	No	No	Yes	No	No	No	Yes	Yes	No
Were the outcome measures clearly defined, valid,	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the assessors blinded to the exposure status of	No	No	No	No	Yes	No	No	No	No
Was loss to follow-up after baseline 20% or less?	NA	NA	Yes	NA	Yes	NA	NA	NA	NA
Quality Rating (Good, Fair, or Poor)	Good	Good	Fair	Good	Good	Good	Good	Good	Good

Table 2. 2. Quality Assessment of Studies

Study (Year)	Zhang (2020)	Zhan (2016)	Zambrana (2019)	Yu (2016)	Wu (2020)	Wissel (2018)	Westin (2010)	Trager (2020)	Tavares (2005)
Was the research question/objective clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the study population clearly specified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was a sample size justification or power description?	No	No	No	No	No	Yes	No	No	No
Were the exposure measures clearly defined, valid,	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the exposure assessed more than once over time?	No	Yes	No	Yes	No	No	Yes	No	Yes
Were the outcome measures clearly defined, valid,	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Were the assessors blinded to the exposure status of	No	No	No	No	No	Yes	Yes	No	No
Was loss to follow-up after baseline 20% or less?	NA	No	NA	Yes	NA	NA	No	NA	Yes
Quality Rating (Good, Fair, or Poor)	Poor	Fair	Fair	Good	Fair	Good	Good	Good	Good

Table 2. 3. Validity and Reliability

Author (Year)	Comparison Assessment	Validity	Reliability
Adams (2018)		Hand tremor AUC = 0.76	
Aghanavesi (2017)	MDS-UPDRS	Finger tapping r = 0.23 Handwriting r = 0.46	**Inter-rater reliability** Finger tapping r = 0.61 Handwriting r =0.65
Akram (2020)	MDS-UPDRS	Finger tapping r = -0.49, p<0.001	
Albani (2019)	MDS-UPDRS	Finger tapping ICC= 0.73	
Amano (2018)	In-clinic assessment	Finger dexterity r = 0.99 Whole hand grasp r = 0.99 Pincer grasp r = 0.99	**Inter-rater reliability** Finger dexterity r = 0.99
Arora (2015)	MDS-UPDRS	Finger tapping mean error of 1.26 UPDRS points	
Arroyo-Gallego (2017)	MDS-UPDRS	Finger tapping AUC = 0.85, p<0.001	
Bazgir (2018)	MDS-UPDRS	Hand tremor 97% accuracy	
Bochniewicz (2017)	ARAT	IADLs r = -0.14, p = 0.697)	
Boroojerdi (2019)	MDS-UPDRS	Finger tapping r = 0.291 Hand tremor r= 0.746	
Burdea (2020)			
Cabrera-Martos (2019)	In-clinic assessment		**Inter-rater reliability:** Finger dexterity r = 0.89 Finger tapping r = 1.0 Hand tremor r = 0.99
Cai (2018)	MDS-UPDRS	Hand tremor $r^2 = 0.95$	
Channa (2021)	MDS-UPDRS	Hand tremor 91.7% accuracy	
Cole (2014)	MDS-UPDRS		
Creagh (2020)	9HPT	Handwriting Dominant hand: $r^2 = 0.39$ Non-dominant hand: $r^2 = 0.41$	
Cunningham (2011)			
Dai (2021)	MDS-UPDRS	Finger tapping r= -0.970, p<0.01 Hand tremor r= 0.93, p<0.001	**Inter-rater agreement (Kendall's W)** Finger tapping 0.86 Hand tremor 0.84

Table 2. 3. Validity and Reliability

Author (Year)	Comparison Assessment	Validity	Reliability
Dubuisson (2017)	9HPT	Finger dexterity r = 0.95, p < 0.001	
Ferreira (2015)	MDS-UPDRS		
Giancardo (2016)	MDS-UPDRS	Finger tapping AUC = 0.75	
Giuffrida (2009)	MDS-UPDRS	Hand tremor r = 0.89	
Goetz (2009)	MDS-UPDRS		
Halloran (2016)	CAHAI	ADLs r = 0.63 (p<0.001)	
Heijmans (2019)	ESM app (tremor questionnaire)	Hand tremor r = 0.43	
Hoffman (2008)	In-clinic assessment	Hand tremor 83.3% agreement Handwriting 41.6% agreement	**Inter-rater reliability** Finger dexterity r = 0.99
Hssayeni (2019)	MDS-UPDRS	Hand tremor r = 0.84	
Iakovakis (2018)	MDS-UPDRS	Finger tapping AUC = 0.92	
Iakovakis (2020)	MDS-UPDRS	Finger tapping r = 0.66	
Jeon (2017)	MDS-UPDRS	Hand tremor 85.5% agreement	
Jha (2020)	MDS-UPDRS	Hand tremor Kappa = 0.68 (p<0.00001 substantial) Finger tapping Kappa = 0.54 (p<0.00001 moderate)	**Inter-rater agreement** Hand tremor 96% Finger tapping 50%
Kim (2018)	MDS-UPDRS	Hand tremor 85% accuracy	**Inter-rater reliability** Hand tremor r = 0.78
Kleinholdermann (2021)	MDS-UPDRS	Finger tapping r = 0.445	
Kostikis (2015)	MDS-UPDRS	Hand tremor Right hand: r = 0.75 (p = 0.0) Left hand: r = 0.85 (p = 0.0)	
Lam (2020)	9HPT	Finger dexterity r = -0.553	**Test-retest reliability** Finger dexterity ICC 0.601
Lee, C (2016)	MDS-UPDRS	Finger tapping AUC = 0.92 (95% CI 0.88-0.96)	
Lee, S (2018)	FMA	Whole hand grasp 92% accuracy	
Lee, U (2016)	MDS-UPDRS		
Lin (2019)			

Table 2. 3. Validity and Reliability

Author (Year)	Comparison Assessment	Validity	Reliability
Lipsmeier (2018)	MDS-UPDRS	Finger tapping t = 2.18, p = 0.028 Hand tremor t = 2.17, p = 0.033	**Test-retest reliability** **Finger tapping** ICC = 0.64 **Hand tremor** ICC = 0.90
Londral (2016)			**Test-retest reliability** r = 0.96, p = 0.09
Lopez-Blanco (2019)	MDS-UPDRS	Hand tremor r = 0.81 (p<0.001)	**Inter-rater reliability** Hand tremor ICC= 0.89
Mahadevan (2020)	MDS-UPDRS	Hand tremor r = 0.67 (p<0.0001)	**Inter-rater reliability** Hand tremor ICC = 0.75
Matarazzo (2019)	UPDRS-III		
Memedi (2015)	Visual assessment	Handwriting 85% accuracy	**Test-retest reliability** Handwrting ICC = 0.69
Mera (2012)			
Mitsi (2017)	MDS-UPDRS		
Noyce (2014)	MDS-UPDRS	Finger tapping r = -0.53	
Orozco-Arroyave (2020)	UPDRS-III		
Pan (2015)	MDS-UPDRS	Hand tremor r = 0.81	
Papadopoulos (2021)	MDS-UPDRS		
Powers (2021)	MDS-UPDRS	Hand tremor r = 0.72	
Pratap (2020)	Longitudinal Neuro-QoL scores	Finger tapping Beta = 0.40, p < 0.001	
Prochazka (2015)	ARAT and FMA	Finger dexterity r^2 = 0.49 Whole hand grasp r^2 = 0.88 Pincer grasp r^2 = 0.88	**Test-retest reliability** 0.67% ± 3.6
Rigas (2012)	MDS-UPDRS	Hand tremor 87% accuracy	
Salarian (2007)	MDS-UPDRS	Hand tremor r = 0.87 (p<0.001)	
San-Segundo (2020)			
Sanchez-Perez (2018)	MDS-UPDRS		
Schallert (2020)			
Shribman (2017)	9HPT	Finger tapping r = 0.926	
Sigcha (2021)	MDS-UPDRS	Hand tremor r = 0.969	
Simonet (2021)	MDS-UPDRS	Finger tapping r = -0.49	

Table 2. 3. Validity and Reliability

Author (Year)	Comparison Assessment	Validity	Reliability
Stamatakis (2013)	MDS-UPDRS	Finger tapping Goodman-Kruskal Index= 0.961	
Tavares (2005)	MDS-UPDRS	Finger tapping r = 0.67 (p<0.001)	
Trager (2020)	MDS-UPDRS	Finger dexterity r = 0.14 (p=0.43) Finger tapping r = 0.58 (p <0.0001)	
Westin (2010)	MDS-UPDRS	Handwriting r = 0.41	**Test-retest reliability** Handwriting r = 0.71
Wissel (2018)	MDS-UPDRS	Finger tapping r = 0.55	**Test-retest reliability** Finger tapping r > 0.75
Wu (2020)	MDS-UPDRS	Hand tremor r = -0.798	
Yu (2016)	FMA	Finger dexterity r^2 = 0.70 Pinch strength r^2 = 0.72	
Zambrana (2019)			
Zhan (2016)	MDS-UPDRS	Finger tapping 71% +/- 0.4	
Zhang (2020)	MDS-UPDRS	Hand tremor 85.9% accuracy	

Legend AUC: area under the curve; MDS-UPDRS: Movement Disorder Society- Unified Parkinson's Disease Rating Scale; ICC: interclass coefficient; ARAT: Action Research Arm Test; IADLs: instrumental activities of daily living; 9HPT: 9 hole peg test; CAHAI: Chedoke Arm and Hand Inventory; ADLs: activities of daily living; FMA: Fugl-Meyer Assessment

Chapter 3: Assessing dexterity in people with MS using pose estimation from patient videos of activities of daily living

Abstract

Background. Current in-clinic assessments of hand function for multiple sclerosis (MS) are not well-equipped to monitor dexterity, especially as it pertains to the handling of various sized and shaped objects that is required to perform daily activities. As an alternative, dexterity may be better captured through videos and analyzed using human pose estimation. Human pose estimation is a set of algorithms trained on images of many people to detect body landmarks. *Objective*. The aim of this study was to validate pose estimation as a means of measuring dexterity from patient-uploaded videos in people with MS. *Methods*. 50 participants receiving care for their MS the University of California, San Francisco were enolled in the study. They were asked to complete 4 in clinic assessments: grip and pinch strength, 9 hole peg test (9HPT), Action Research Arm Test, and vibration sense. They were also asked to self-report their dexterity using the ABILHAND survey and to record videos of 3 self-care tasks at home: buttoning, brushing teeth, and eating. The videos were analyzed using the open access MediaPipe Hand pose estimation software, and position and velocity kinematic data were extracted. *Results*. Buttoning in the nondominant hand correlated strongly with the 9HPT (r=0.69, p=0.0) and moderately with vibration sense (r=0.46, p=0.02). The brushing and eating tasks in the nondominant hand were moderately correlated with the 9HPT (r=0.38, p=0.05 and r= 0.35, p=0.05, respectively). The ABILHAND, while not associated with the 9HPT (grip subscore: Spearman r= -0.05, p=0.69; pinch subscore: r= 0.02, p= 0.91), was moderately correlated with buttoning (r=-0.48, p=0.05) and eating (r=0.39, p=0.05) tasks. *Conclusions*. These findings show that assessing hand function using pose estimation has moderate to high validity with established in-clinic measures of dexterity in multiple sclerosis. Additionally, in contrast to the 9HPT, it is correlated with patient-reported dexterity.

Introduction

The hand is the most active part of the upper extremity, and hand function is critical to independence in daily activities.[114] Quality and independence of performance in daily tasks, ability to perform at work and remain employed, and engagement in recreational activities are determined to a large degree by hand function and manual dexterity.[11,115,116] A number of neurological conditions, such as multiple sclerosis (MS), can result in acute or gradual worsening in hand function, threatening affected patients' independence in daily life.

Assessments of quality and variety of movement (e.g., handling objects of different sizes and weights, each requiring different grips) that are affordable and time-efficient are needed to quantify impairments for individually tailored rehabilitation. However, current in-clinic assessments are limited in their ability to monitor such changes. In MS for example, where early changes in hand function are often subtle, the commonly used in-clinic 9 Hole Peg Test (9HPT) requires individuals to manipulate small pegs into holes on a board. While this task certainly requires finger dexterity and coordination, which are essential for activities of daily living (ADL), the 9HPT score is limited to a measure of time to complete the task. A few research-grade assessments, like the Action Research Arm Test (ARAT) and Graded Redefined Assessment of Strength, Sensation, and Prehension (GRASSP) examine ADL-like movements but require costly equipment and considerable time to complete.

As an alternative to the in-clinic evaluations of hand function, methods for remote, longitudinal, granular assessments are being developed – including using wearable devices, smartphone-based applications, and specialized keyboards.[117] A further type of assessment, video capture, offers a number of advantages but requires validation of algorithms to capture and quantify the movement that is recorded.

Human pose estimation – a set of algorithms trained on many images of different people, resulting in robust networks capable of detecting body landmarks and their movement in space —is one such solution. These algorithms are freely available and have the potential to expand researchers' and clinicians' abilities to analyze large datasets of upper extremity movement collected in any setting (including the

home or clinic) with minimal cost of time, money, or effort. Prior studies have utilized pose estimation in the lower extremity to quantify features of gait.[118-121] In the upper extremity, pose estimation has been used to compute finger kinematics,[122] characterize repetitive motions,[123] and perform sign language and gesture recognition.[124,125] To date, however, pose estimation has mostly been validated as a method for capturing precise data on movement in healthy human populations, and not to our knowledge in MS or other neurological populations. Further, to date, the videos analyzed were captured through a standard recording protocol-- whereas videos uploaded from a participant's personal device could provide further usability benefits.

The current study sought to validate pose estimation as a means of measuring hand function from patient-uploaded videos in individuals with MS. Leveraging recent progress in computer vision and video-based pose estimation, open access software (Open Source Computer Vision Library (OpenCV) and MediaPipe) was used, enabling automated analysis of human movement using only digital video input. The primary goal was to determine the association between dexterity as measured using pose estimation from patient-generated videos, and traditional in-clinic performance-based and patient-reported measures.

Methods

Recruitment

The current study examined the baseline paraclinical and clinical data collected from a larger longitudinal study of remote hand function assessment in people with MS. Participants were recruited via convenience sampling from the University of California, San Francisco (UCSF) Multiple Sclerosis Center. Participating neurologists referred interested patients who met initial eligibility criteria (age 18 years or above, a confirmed diagnosis of MS) to the study team. Participants were approached via email and then screened via phone call to determine if additional eligibility criteria (smartphone ownership) were met. Of note, all approached participants had the required technology to participate, i.e., no

participants were excluded due to lack of smartphone. Study activities were approved by the UCSF IRB (IRB# 20-557) and all participants provided written informed consent to enroll. Study participation involved a baseline in-clinic and 6-month in-clinic evaluation, as well as remotely collected participant videos and patient-reported outcomes, both through the REDCap electronic data capture platform.[126] The longitudinal data are described elsewhere (Chapter 4).

To validate pose estimation parameters and their relationship to in-clinic measures, the data analyzed here were those collected at the baseline in-clinic visit for all 50 participants enrolled in the study. *A priori* power analysis was conducted using G*Power[127] to test the correlation between in-clinic and remote metrics using a two-tailed test, a medium effect size (Cohen's d= 0.50), and α= 0.05. Results showed that N =38 would be required to achieve a power of 0.95. Based on this, a sample size of 50 individuals was selected to account for potential attrition in the longitudinal study.

Measures collected

Demographic and clinical data: Age, sex, race, ethnicity, MS diagnosis, diagnosis date, current disease modifying therapy (DMT), dominant hand (right or left, recorded via patient report), and type of smartphone (iOS or Android) were collected from participants.

Functional assessments: In-clinic assessments of hand function included grip and pinch dynamometer testing (Jamar Technologies), vibration testing (Vibratron II- Physitemp), 9HPT, and Action Research Arm Test (ARAT). Adequate grip and pinch strength are required to perform ADLs, and reductions in strength are linked to reduced independence and function in people with neurological diseases.[128] Quantitative measures of sensation, specifically those assessing vibration, show the greatest trend toward detecting subtle changes in functional performance early in the disease course.[129] The 9HPT has been considered the clinical standard for assessing dexterity in people with MS,[130,131] and the ARAT is validated in MS and assesses the ability to perform different ADL-like tasks requiring manipulation and transportation of objects using different grip, grasp and pinch functions.[130,132]

Patient-reported outcomes: Using REDCap,[126] a secure, web-based application for managing surveys, participants self-reported their overall MS disease status through a validated electronic patient-reported Expanded Disability Status Scale (e-prEDSS)[133] assessment, as well as their hand function via the ABILHAND questionnaire, which generates pinch and grip subscores.[134]

Video tasks

A set of self-recorded videos of performance of 3 basic ADLs that require hand function were directly uploaded by participants into the REDCap platform. The ADLs selected were dressing (buttoning a shirt), personal hygiene (brushing teeth), and eating (fork to mouth). Example videos were provided for each of these tasks within REDCap. Feeding and personal hygiene tasks were uploaded separately for each hand, for a total of 5 videos.

Video analysis

The goal of video analysis was to estimate 3D hand pose from 2D patient-uploaded videos. Analysis was completed using a machine learning solution, MediaPipe Hand, a part of the OpenCV library which is freely available. MediaPipe Hand detects 21 landmarks in the hand, including the wrist. Based on the variable grips required for each of the ADLs, the tip of the index finger and wrist landmarks were chosen for analysis (Figure 1).

For each landmark, position and velocity kinematic data was obtained by mapping the joints of interest into a xy-cartesian coordinate system (Figure 1). The following metrics were obtained: 1) The path length of the landmark, a measure of the joint movement in the cartesian plane; 2) the complexity of movement, a count of the number of local peaks in the position vs time curve; 3) the total distance the joint travels, the area under the velocity vs time curve (AUC); and 4) smoothness, average velocity divided by maximum velocity (Table 1).

A video was considered valid if ≥10 coordinates were generated during analysis. Sampling frequency was established based on the frame rate of the uploaded video.

As a quality control measure, to ensure that MediaPipe was generating accurate joint trajectories, a validation using Kinovea,[135] a video annotation software for motion analysis, was performed. Here, the landmark (tip of index finger, wrist) was manually selected in each frame to generate point trajectory. A random number generator was used to select 5 videos to validate with position and velocity data generated by MediaPipe. No statistically significant differences emerged between the manual analysis and MediaPipe ($p>0.05$).

AG and WOT developed the script to analyze videos using MediaPipe, which is available on GitHub. Once the script for video analysis was written, the program was capable of video analysis without additional expert training. In other words, the videos and video data were easily generated using the written code. Each video took less than a minute to analyze.

Statistical analyses

The outcomes for this analysis were video kinematics (Table 1) and the comparative clinical measures of hand function.

Spearman's correlation coefficients were calculated to determine associations between demographic data, video kinematics, patient reported measures, and clinical measures of hand function.

Video analyses were performed using Python 3.4; kinematic and statistical analyses, as well as data visualizations were performed in Matlab 2022.

Results

Demographics

Among the 50 participants consented and enrolled in the study, 62% were women, 73% were non-Hispanic White, mean age was 47.2 (SD 12.9), and median disability scale as measured by the ePR-EDSS was 3 (IQR 2, 5) (Table 2).

To characterize dexterity, this cohort had a mean dominant hand 9HPT score of 23.5 seconds (SD 9.7 seconds); 90% of participants had high dexterity (9HPT below 33.3 seconds); 10% of participants had low dexterity (9HPT above 33.3 seconds).[136] Four of the in-clinic dexterity measures (9HPT, grip and pinch strength from dynamometer, and vibration sense) showed significant correlations with various demographic and clinical features (Figure 2). Notably, a greater (worse) 9HPT score for the dominant hand was correlated with older age (r= -0.35, p=-0.01), and the score for the nondominant hand was correlated with disease duration (r=0.26, p=0.05). Grip and pinch strength were significantly negatively correlated with female sex in both hands (r=0.50, p=0.05). Decreased vibration sense was also associated with older age in the nondominant hand (r=0.30, p=0.02) and with longer disease duration in both the dominant (r=0.27, p=0.04) and nondominant hands (r=0.26, p=0.05).

For the patient reported outcome, ABILHAND, mean pinch score was 29.5 (SD 4.6) out of 30, and mean grip score was 35.5 (SD 4.4) out of 40, indicating high dexterity overall.[134]

Video quality

Of 250 total videos uploaded by the 50 participants (5 tasks each) at the baseline evaluation, 16 (6.4%) were not analyzed due to low video quality or participant error (e.g., video out of focus, hand not in camera frame), thus 93.6% of videos collected were included in analyses. Upon analysis with MediaPipe Hand, 36 (15.4%) videos did not generate sufficient data for kinematic analysis. This was likely due to the hand not being in the video for a sufficient amount of time for the kinematic metrics to be estimated. Exclusion of unusable videos left 198 (79%) videos to be analyzed, videos from 95% participants were usable. Mean video duration was 12.1 seconds (SD 5.6, median 7.2s, range 3-165 seconds).

Validation of video metrics against clinical measures

The video measures extracted using pose estimation showed significant correlations with standard in-clinic assessments of hand function. These are depicted in Figure 3.

Buttoning. In the nondominant hand, the buttoning task resulted in the strongest correlations with in-clinic measures. 9HPT was strongly correlated with wrist position path length (r=0.66, p=0.01), wrist position peaks (r= 0.65, p=0.01), index position path length (r=0.69, p=0.0), and index position peaks (r=0.68, p=0.01). Further, vibration sense was also correlated with wrist position path length (r= 0.39, p=0.04), index position path length (r=0.46, p=0.02), and wrist position peaks (r=0.35, p=0.02) while buttoning.

Brushing. In the nondominant hand, 9HPT was moderately correlated with index position peaks (r= 0.35, p=0.05).

Eating. In the dominant hand, 9HPT was moderately correlated with wrist position AUC (r=0.38, p=0.05).

The ARAT results were excluded from correlations due to the ceiling effect reached by the cohort. Low variance in scores (95% of participants achieved the maximum score) resulted in uninterpretable correlation values (r=0.03, p=0.99).

Video correlates of poor motor control

Motor control is the study of how the central nervous system interacts with the environment to produce purposeful, coordinated movements.[137] When probing to specifically evaluate video correlates of poor in-clinic motor control, namely high 9HPT time and low grip or pinch strength, the video metrics showed several significant associations. The buttoning task in the nondominant hand revealed statistically significant correlations between clinical measures of poor motor control (low pinch strength and high 9HPT time), and between diminished vibration sense (high threshold of vibration detection) and increased path of wrist movement (Figure 3F). For the eating and brushing tasks, the clinical measures of poor motor control (low grip strength and high 9HPT time) showed significant correlations with video metrics associated with the total distance travelled by the wrist and index finger (position area under the curve- index finger (r=0.40, p=0.02); velocity area under the curve- index finger (r= 0.37, p=0.03)).

Associations between patient-perceived dysfunction and objective performance in in-clinic and video measures

The ABILHAND questionnaire generates sub-scores for pinch and grip strength. Pinch tasks include threading a needle and fastening a zipper, while grip tasks include opening a jar and turning on a faucet. Overall, the video metrics showed stronger associations with ABILHAND scores than they did with clinical measures (Table 3). For example, self-reported difficulty on the ABILHAND was not significantly associated with the 9HPT (grip subscore: Spearman r= -0.05, p=0.69; pinch subscore: r= 0.02, p= 0.91), but it did correlate with video metrics. For the buttoning task, wrist position path length (r=-0.48, p=0.05), wrist velocity area under the curve (r=-0.47, p-0.05), and wrist velocity smoothness (r=-0.48, p=0.05) all showed moderate correlations with the self-reported pinch subscore. For the eating task, wrist position path length (r=0.39, p=0.05), index position path length (r=0.32, p=0.04), and index position peaks (r= -0.34, p=0.05) all showed moderate correlations with the self-reported pinch subscore.

Discussion

In this cohort of 50 adults with MS with overall low impairment in hand function, measures of dexterity as extracted from simple patient-generated videos showed significant associations with both performance-based and patient-reported dexterity measures obtained in the clinic.

To remotely assess dexterity in neurological populations, remote biosensor-based and/or smartphone applications have been developed and validated,[117] but to our knowledge, no methods utilizing patient-generated videos have been reported to date. Using patient-uploaded videos to garner information about dexterity represents a significant advance towards more patient-centered data collection. The current method for quantifying dexterity in near real time represents a proof-of-concept that dexterity can be captured using pose estimation, with precision likely to increase as pose estimation algorithms become more sophisticated. Over time, clinicians may be able to more precisely identify dexterity impairments to refine and standardize clinical protocols. Moreover, this approach is poised to be applied to other types of

movement dysfunction in neurological populations (such as Parkinsonian hand tremor, or stroke-related motor loss), and may be used to augment existing care in these domains to provide optimal care.

The association of the video-generated metrics and in-clinic functional measures differed for each ADL task. Strong correlations were found between the buttoning task and the in-clinic vibration assessment, which may be indicative of the need for unimpaired sensation and proprioception required to manipulate buttons. The relatively weaker associations between the brushing and eating tasks and clinical measures may be due to the larger, gross motor movements required to complete these tasks. The ARAT was chosen specifically as a benchmark for gross movement as it requires the participant to reach to grasp tasks, but unfortunately reached a ceiling effect in the current population. Limitations in gross motor coordination and accuracy may have been detected by the pose estimation algorithm in brushing and eating tasks, but the clinical measures included in analyses focused on fine motor abilities (e.g., 9HPT).

Patients with MS can experience subtle reductions in function (dexterity, cognitive or other) that are currently subclinical. Inclusion of the ABILHAND questionnaire allowed us to distinguish between functional capacity (what an individual is able to do or thinks they are able to do) and functional performance (what an individual actually does in their daily environment). Here, no association was noted between the self-reported ABILHAND scores and the clinical assessments, suggesting that these existing clinical measures (performed under supervision, in a non-naturalistic setting) may not be sensitive enough to capture subtle reduction in participants' self-reported functional capacities. This is analogous to typical cognitive batteries in MS that are often insensitive to subtle cognitive deficits noted first by patients.[138] Stronger associations between the ABILHAND and video tasks suggest that assessing function in an individual's daily environment, performing real-world tasks, may be more representative of their functional capacities.

Participants were provided identical instructions for video recording and upload: they were asked to place the smartphone on a stable surface an arm's length away. However, the recordings submitted were not always within these parameters, resulting in hands out of frame or videos that were too short to be analyzed. Other studies using video metrics have obtained videos of participants under supervision,

ensuring standardized videos across trials and between participants.[86,123] While some videos had to be excluded, a majority (79%) could be analyzed and yielded accurate pose estimation. Overall, these patient-generated videos collected without any guidance or supervision, were sources of high-quality data. Future studies might include a brief review of the 12-second videos for quality control, requesting a re-upload from participants after additional training or feedback.

Previous methods for objectively assessing dexterity outside of in-person clinical settings have involved direct human observation via secure video, or the use of a proprietary device (e.g., smartphone sleeve to assess hand tremor[66,99]). While each of these methods has its advantages, the data generated are fixed. In contrast, the current videos could be readily uploaded, with the raw video files from each participant stored securely within UCSF's REDCap platform, so they can be re-analyzed as technology advances and algorithms develop. The ability to continue to generate more granular data and maintain a repository of videos of patient function, represents a considerable advantage in comparison to existing methods. Further, the script is freely available to analyze videos, and can be used without extensive programming knowledge. This is an initial effort to provide access to the technology, as opposed to costly software, which will aid in overall scalability.

This study had several limitations. The inclusion of a largely high-functioning cohort demonstrated the validity of the approach even in patients with low disability, but it is not clear whether self-upload would have been as successful in a more impaired cohort. Second, inclusion of the monofilament test in addition to vibration sense could have enhanced clinical measures of fingertip sensation. Here, motor control, the coordination between sensory information and motor output, showed significant associations with video metrics; and including a more thorough and robust assessment of sensation (temperature, pain, sharp and dull sensations) could likely improve the associations between motor control and video metrics. Third, one advantage of pose estimation is the seemingly limitless quantification of data output – making more challenging the decision on which video metrics to be included in this initial validation study. In the future, additional metrics to consider could be those examining acceleration, change in velocity at various phases of the task, and tertiary time derivatives such as jerk. Fourth, while data on hand dominance were

collected, handedness did not always coincide with a participant's weaker or stronger side. MS-related motor and sensory symptoms often present asymmetrically, and it is possible that dexterity changes in the weaker side may be a more sensitive variable relative to hand dominance. Finally, more sophisticated data science approaches to analyzing the video measures and their relationship with clinical measures, such as principal component analysis or other decomposition methods, can be performed in larger datasets, where more sophisticated analytical models can be built.

We describe an accessible, low-cost, patient-focused, and clinically relevant approach to assessing dexterity in people with MS using pose estimation. This method's versatility and clinical validity represent a significant advantage over existing approaches to remote data capture. To support its use in research and clinical care, validation of longitudinal measures will be required.

A.

B.

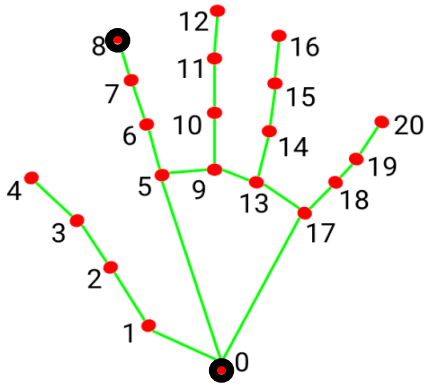
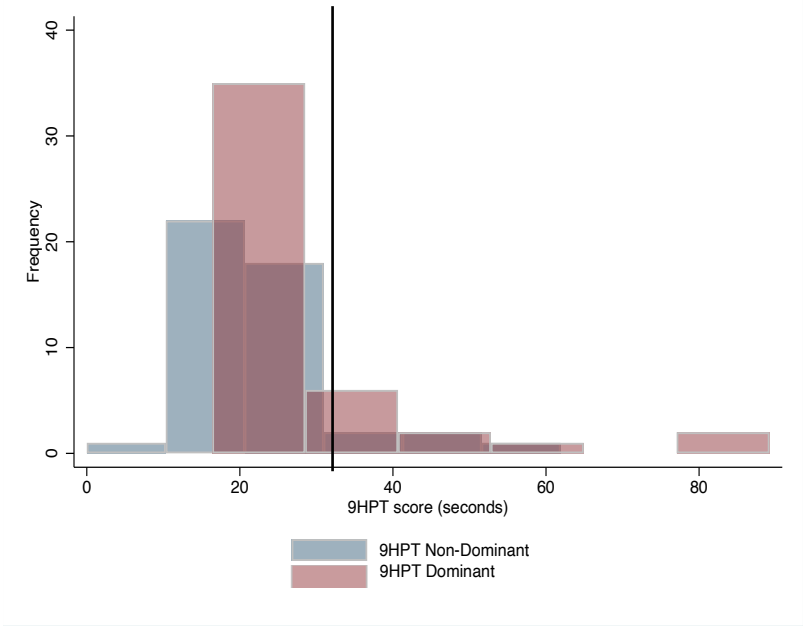


Figure 3. 1. Pose estimation landmarks for video analysis. *(A) Depiction of skeleton overlay in MediaPipe Hand. Index finger and wrist landmarks are highlighted in yellow. Coordinate system for video kinematics, with O at the origin. (B) All available landmarks in MediaPipe Hand. Point 8 represents the index finger, and Point 0 represents the wrist.*

A.



B.

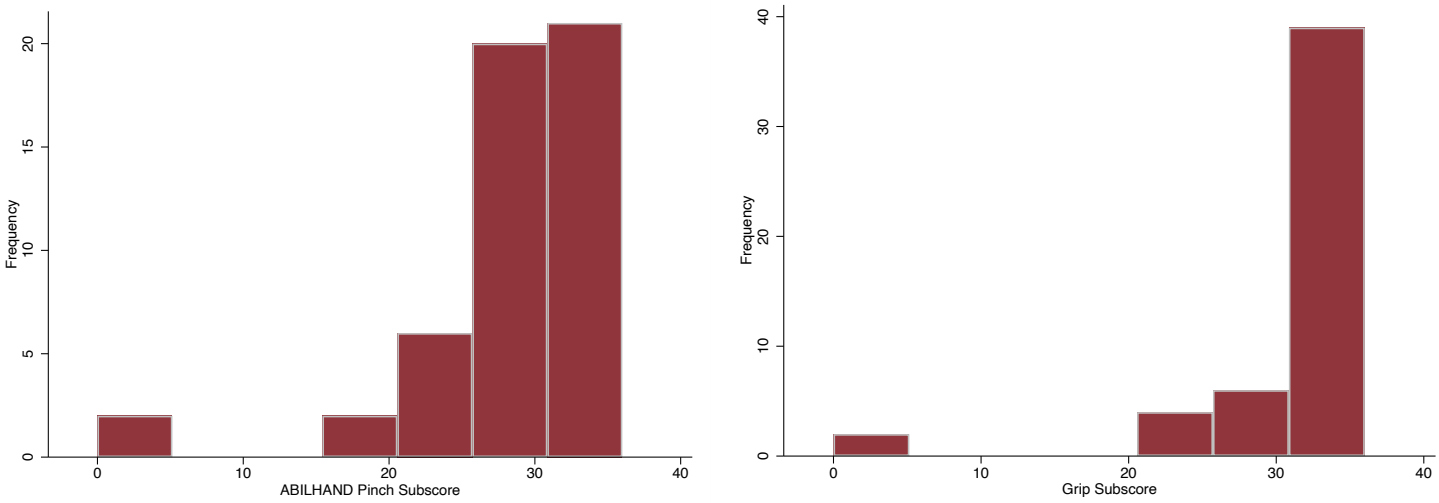
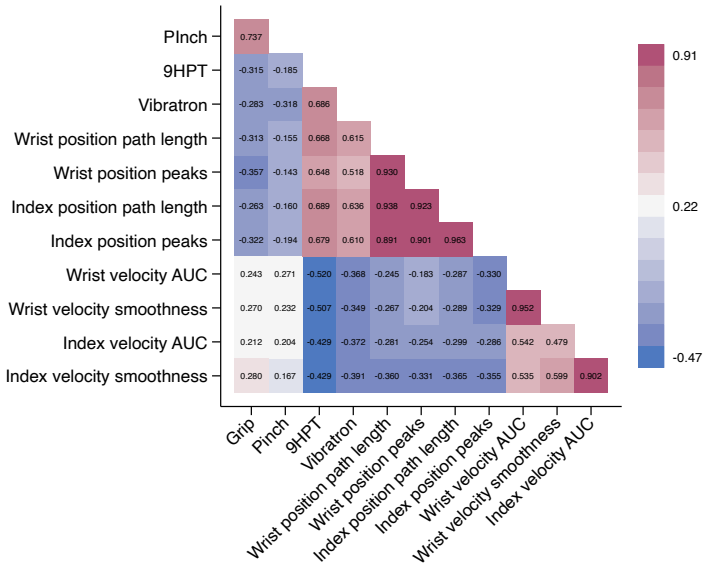
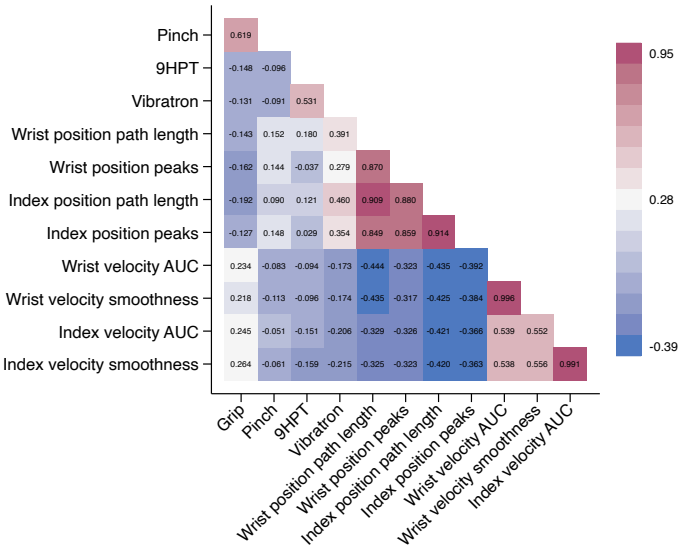


Figure 3. 2. Histograms of 9HPT and ABILHAND scores. *(A) 9HPT scores of participants. Blue indicates non-dominant hand, and pink represents dominant hand. Vertical line depicts 9HPT dexterity threshold, 33.3s. X axis is 9HPT score, with high scores indicating worse function. Y axis indicates the number of participants with a given scores. (B) Left: ABILHAND pinch subscores of participants. Higher scores indicating higher function. Right: ABILHAND grip subscores of participants. Higher scores indicating higher function.*

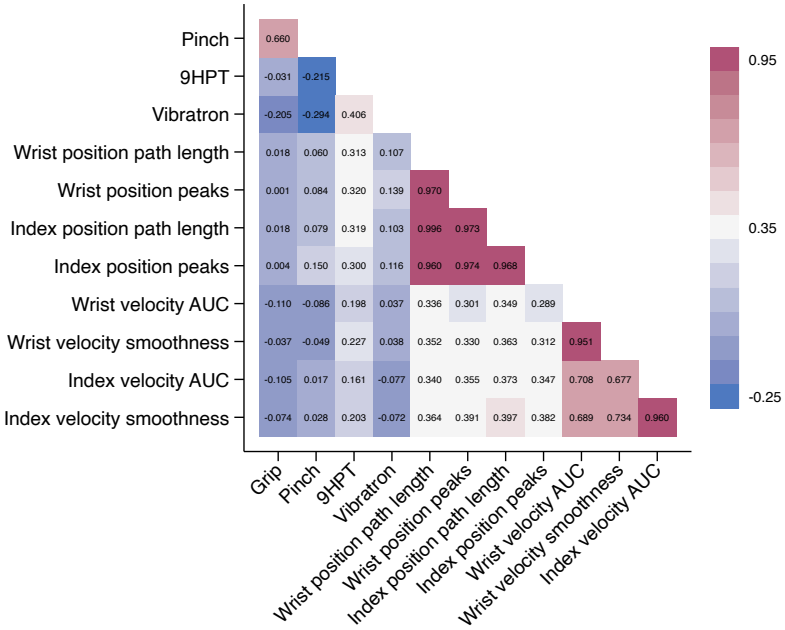
A.


PInch
9HPT
Vibratron
Wrist position path length
Wrist position peaks
Index position path length
Index position peaks
Wrist velocity AUC
Wrist velocity smoothness
Index velocity AUC
Index velocity smoothness
0.737
-0.315 -0.185
-0.283 -0.318 0.686
-0.313 -0.155 0.668 0.615
-0.357 -0.143 0.648 0.518 0.930
-0.263 -0.160 0.689 0.636 0.938 0.923
-0.322 -0.194 0.679 0.610 0.891 0.901 0.963
0.243 0.271 -0.520 -0.368 -0.245 -0.183 -0.287 -0.330
0.270 0.232 -0.507 -0.349 -0.267 -0.204 -0.289 -0.329 0.952
0.212 0.204 -0.429 -0.372 -0.281 -0.254 -0.299 -0.286 0.542 0.479
0.280 0.167 -0.429 -0.391 -0.360 -0.331 -0.365 -0.355 0.535 0.599 0.902
0.91
0.22
-0.47
Grip
Pinch
9HPT
Vibratron
Wrist position path length
Wrist position peaks
Index position path length
Index position peaks
Wrist velocity AUC
Wrist velocity smoothness
Index velocity AUC

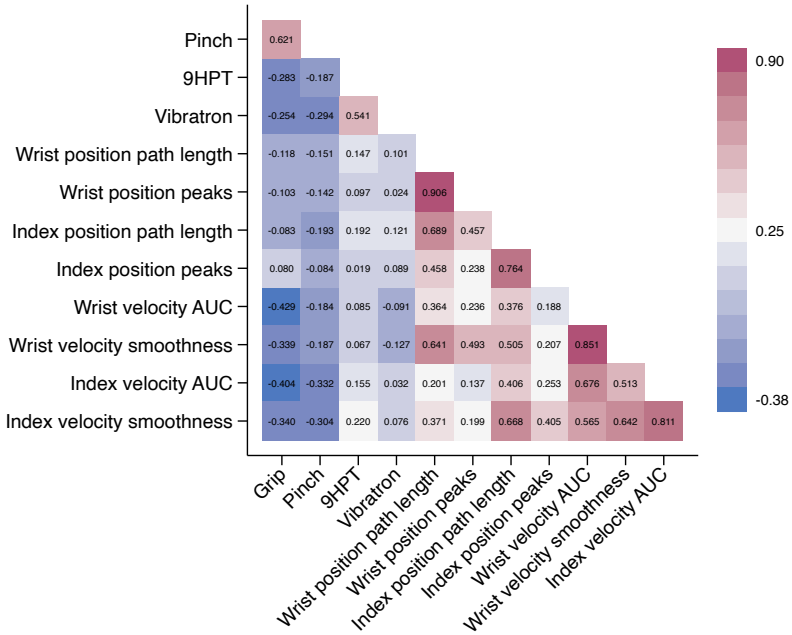
B.


Pinch
9HPT
Vibratron
Wrist position path length
Wrist position peaks
Index position path length
Index position peaks
Wrist velocity AUC
Wrist velocity smoothness
Index velocity AUC
Index velocity smoothness
0.619
-0.148 -0.096
-0.131 -0.091 0.531
-0.143 0.152 0.180 0.391
-0.162 0.144 -0.037 0.279 0.870
-0.192 0.090 0.121 0.460 0.909 0.880
-0.127 0.148 0.029 0.354 0.849 0.859 0.914
0.234 -0.083 -0.094 -0.173 -0.444 -0.323 -0.435 -0.392
0.218 -0.113 -0.096 -0.174 -0.435 -0.317 -0.425 -0.384 0.996
0.245 -0.051 -0.151 -0.206 -0.329 -0.326 -0.421 -0.366 0.539 0.552
0.264 -0.061 -0.159 -0.215 -0.325 -0.323 -0.420 -0.363 0.538 0.556 0.991
0.95
0.28
-0.39
Grip
Pinch
9HPT
Vibratron
Wrist position path length
Wrist position peaks
Index position path length
Index position peaks
Wrist velocity AUC
Wrist velocity smoothness
Index velocity AUC

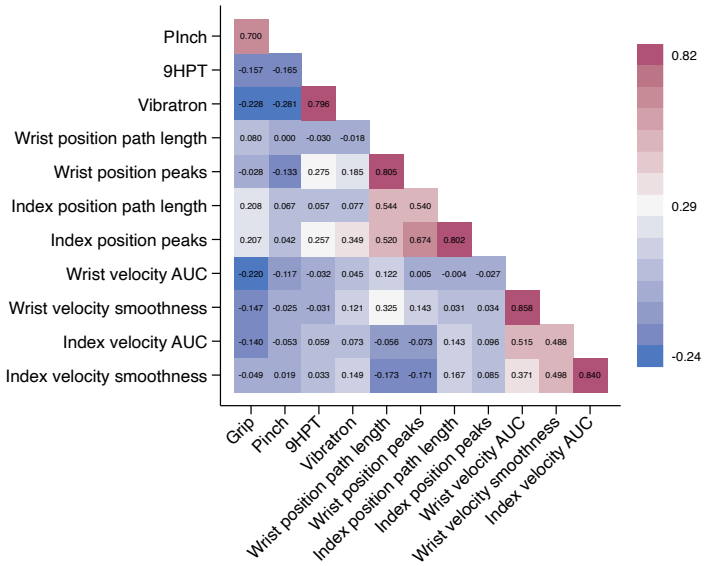
C.


Pinch
9HPT
Vibratron
Wrist position path length
Wrist position peaks
Index position path length
Index position peaks
Wrist velocity AUC
Wrist velocity smoothness
Index velocity AUC
Index velocity smoothness
0.660
-0.031 -0.215
-0.205 -0.294 0.406
0.018 0.060 0.313 0.107
0.001 0.084 0.320 0.139 0.970
0.018 0.079 0.319 0.103 0.996 0.973
0.004 0.150 0.300 0.116 0.960 0.974 0.968
-0.110 -0.086 0.198 0.037 0.336 0.301 0.349 0.289
-0.037 -0.049 0.227 0.038 0.352 0.330 0.363 0.312 0.951
-0.105 0.017 0.161 -0.077 0.340 0.355 0.373 0.347 0.708 0.677
-0.074 0.028 0.203 -0.072 0.364 0.391 0.397 0.382 0.689 0.734 0.960
Grip
Pinch
9HPT
Vibratron
Wrist position path length
Wrist position peaks
Index position path length
Index position peaks
Wrist velocity AUC
Wrist velocity smoothness
Index velocity AUC
0.95
0.35
-0.25

D.


Pinch
9HPT
Vibratron
Wrist position path length
Wrist position peaks
Index position path length
Index position peaks
Wrist velocity AUC
Wrist velocity smoothness
Index velocity AUC
Index velocity smoothness
0.621
-0.283 -0.187
-0.254 -0.294 0.541
-0.118 -0.151 0.147 0.101
-0.103 -0.142 0.097 0.024 0.906
-0.083 -0.193 0.192 0.121 0.689 0.457
0.080 -0.084 0.019 0.089 0.458 0.238 0.764
-0.429 -0.184 0.085 -0.091 0.364 0.236 0.376 0.188
-0.339 -0.187 0.067 -0.127 0.641 0.493 0.505 0.207 0.851
-0.404 -0.332 0.155 0.032 0.201 0.137 0.406 0.253 0.676 0.513
-0.340 -0.304 0.220 0.076 0.371 0.199 0.668 0.405 0.565 0.642 0.811
Grip
Pinch
9HPT
Vibratron
Wrist position path length
Wrist position peaks
Index position path length
Index position peaks
Wrist velocity AUC
Wrist velocity smoothness
Index velocity AUC
0.90
0.25
-0.38

E.



F.

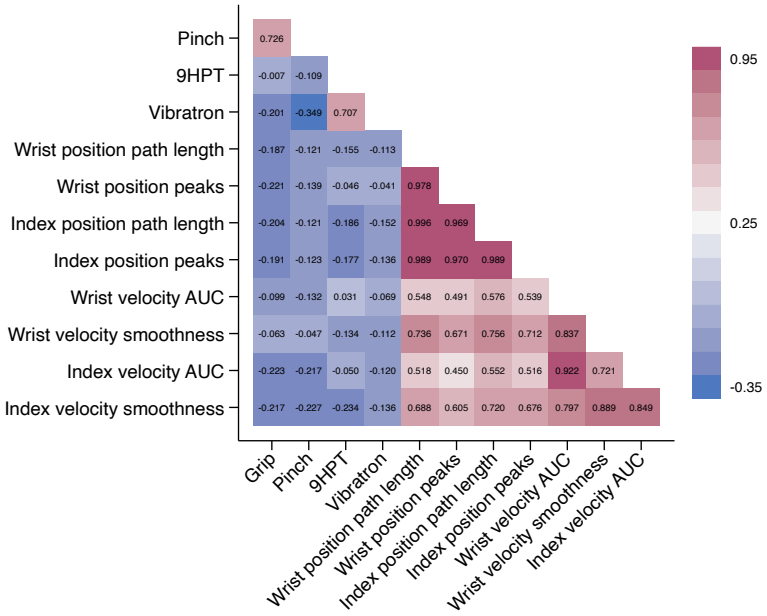


Figure 3. 3. Spearman's correlation coefficients heatmaps. *Darker pink colors indicate r values closer to 1, and darker blue colors indicate values closer to -1. (A) Buttoning dominant hand (B) buttoning nondominant hand (C) brushing dominant hand (D) brushing nondominant hand (E) eating dominant hand (F) eating nondominant hand.*

Table 3. 1.Video Kinematic Metrics Summary

Video Kinematic Metric	What is measured?	Formula/calculation
Position		
Path length	How much the joint is moving in 2D space	Sum of the magnitude of the vectors created in 2D space.
Local peaks	Movement complexity	MATLAB findpeaks function, which finds peaks from position magnitude curve with a minimum peak prominence of 0.05.
Velocity		
Area under the curve	Distance traveled by joint	MATLAB function trapz, which approximates integral of velocity magnitude using the trapezoidal method
Smoothness	Smoothness of movement	Mean velocity divided by max of non-normalized velocity

All video metrics except smoothness are normalized to a zero to one range

Table 3. 2. Demographics and baseline characteristics

Gender [N (%)]	
Men	18 (36%)
Women	31 (62%)
Non-binary	1 (2%)
Age [mean (SD)]	47.2 (12.9)
Race [N (%)]	
White, non Hispanic	44 (73%)
Black	1 (2%)
Asian	4 (8%)
Hispanic	1 (2%)
Baseline EDSS [median (IQR)]	3 (2, 5)
Disease duration (years) [mean (SD)]	14.3 (9.2)
MS Subtype [N (%)]	
Relapsing remitting	43 (86%)
Primary progressive	4 (8%)
Secondary progressive	3 (6%)
Dominant hand [N (%)]	
Right	45 (90%)
Left	5 (10%)
Smartphone type [N (%)]	
iOS	40 (80%)
Android	10 (20%)
Grip strength (kg/cm^2) [mean (SD)]	
Dominant	28.6 (9.4)
Non-dominant	27 (9.5)
Pinch strength (kg/cm^2) [mean (SD)]	
Dominant	5.8 (2.6)
Non-dominant	5.6 (2.5)
Nine-hole peg test (seconds) [mean (SD)]	
Dominant	23.5 (9.7)
Non-dominant	28.7 (14.9)
Vibration sense (Hertz) [mean (SD)]	
Dominant	2.4 (2.7)
Non-dominant	2.3 (2.6)
ARAT-Grip [mean (SD)]	
Dominant	11.9 (0.6)
Non-dominant	11.9 (0.8)
ARAT-Grasp [mean (SD)]	
Dominant	17.8 (0.7)
Non-dominant	17.9 (0.5)
ARAT-Pinch [mean (SD)]	
Dominant	17.3 (2.3)
Non-dominant	17.1 (2.7)
ARAT-Gross motor [mean (SD)]	
Dominant	9 (0)
Non-dominant	9 (0)
ABILHAND [mean (SD)]	
Pinch subscore	29.5 (4.6)
Grip subscore	35.5 (4.4)

Chapter 4: Self-care selfies: Capturing changes in dexterity in multiple sclerosis over time using patient-uploaded videos

Abstract

Background. Recent research has shown that upper extremity function may be a viable marker of disease-related changes in multiple sclerosis (MS). Given this, the ability to monitor changes in upper extremity dysfunction is critical to understanding the role of pharmaceutical therapies in slowing disease progression, and of rehabilitation therapies in improving function over time. Human pose estimation has been previously validated as a method of capturing dexterity in people with MS, with benefits including minimal cost, ease of use, and naturalistic data capture. *Objective*. The current study aimed to evaluate the feasibility of collection and validity of patient-uploaded videos of hand function over 6 months, analyzed via human pose estimation. *Methods*. 50 participants receiving care for their MS at the University of California, San Francisco were enrolled. They completed in-clinic visits at baseline and at 6 months, with 4 assessments performed: grip and pinch strength, 9 hole peg test (9HPT), Action Research Arm Test, and vibration sense. Participants were also asked to self-report their dexterity using the ABILHAND survey and to submit videos of 3 self-care tasks at regular intervals: buttoning, brushing teeth, and eating. Participant perceptions of feasibility and acceptability of the research modality were queried monthly using the Health Information Technology Usability Scale[139]. *Results*. Of 50 enrolled participants, 44 (88%) completed at least 3 video submissions, and 37 (74%) completed the 6 month study. 77% of participants strongly agreed (0% disagreed) that the assessments were easy to access, and 61% (0% disagreed) strongly agreed that the remote platform was easy to use. 86% of participants agreed (somewhat to strongly) that the study tasks were representative of their daily activities. Over the 6 month study period performance in all modalities worsened numerically. A reduction in pinch strength in the dominant hand was noted (mean change dominant hand: 0.8kg/cm^2, non-dominant hand: 0.6 kg/cm^2, p=0.05). This was corroborated by a change in the buttoning task path length and peak metrics in both

hands (dominant: p=0.01, non-dominant: p= 0.02). Overall, 5% of patients experienced 20% worsening on the 9HPT alone, 40% on the buttoning metrics alone, and 40% worsened on both 9HPT and buttoning metrics. *Conclusions*. Pose estimation using patient-uploaded videos represents a novel approach to quantify kinematics in key daily tasks. Its validity against in-clinic assessments, sensitivity to change, low cost and technological burden, and excellent acceptability by the intended users (patients) enhance its potential for dissemination for use in disease monitoring and treatment.

Introduction

Hand function is critical to quality of life and independence in performance of daily tasks. The ability of humans to perform at work, remain employed and participate in recreational activities are determined to a large degree by hand function and manual dexterity.[115,11,116] This ability is compromised by neurological conditions, whether suddenly after a stroke or more gradually, in the case of movement disorders such as Parkinson's Disease or demyelinating diseases such as multiple sclerosis (MS). Recently, disease modifying therapies in MS have been shown to reduce the progression of upper extremity dysfunction in clinical trials.[5] Therefore, the ability to monitor subtle and gradual changes in hand function represents an important tool for evaluating the effectiveness of pharmaceutical therapies in slowing disease progression - and rehabilitation therapies in improving function - over time.

Ideally, monitoring hand function should occur in the person's natural environment, which requires remote assessment. A number of modalities for remote assessment of dexterity in people with neurological dysfunction[117] have been developed and evaluated, including smartphone-based applications,[66,71,76] wrist-worn accelerometers,[40,43,48] and specialized keyboards.[56,60] Each of these, while informative, has its limitations. For example, body-worn sensors can be cumbersome, and purchasing specialized equipment can be costly and rapidly become obsolete. Smartphone-based applications require frequent software upgrades, must be actively maintained to be used long-term, and the data are often proprietary. Applications acquiring more passive data, such as keystrokes, require less user effort or

engagement outside of their usual activities, but are also limited as they do not capture functional limitations in daily tasks that are meaningful to patients' lives, thus limiting compliance and scalability.[67,140] Altogether, while these methods have demonstrated the feasibility of monitoring dexterity remotely, significant barriers to implementation and scale remain – particularly in contrast to other domains such as overall disability assessment,[141] gait,[142] fatigue[143], and cognition[144] where remote assessment of function in MS has made more gains. Further, no existing method of assessing dexterity remotely has been validated against activities of daily living. In addition, it is critical that research become more accessible in MS, and too often the research tools are designed without accounting for the preferences or convenience of their primary intended users, i.e., the patient and clinician. In fact, intended users' satisfaction and perceptions of usability are critical to determining the eventual widespread adoption of a digital tool.[145,146]

Patient-uploaded videos of themselves performing self-care tasks represent an entirely different approach to remote evaluation of hand function. These "self-care selfies" leverage accessibility (patients' personal preferred smartphone devices and software), usability (using a common modern communication medium, namely video "selfies"), and meaningfulness. Indeed, videos can capture function when engaging in self-care tasks of everyday life: grooming, dressing, and feeding; the ability to perform these tasks efficiently and *independently* makes them important targets for clinical observation and assessment. Sophisticated motion capture software has been validated in MS,[147,148] but measures derived from these simpler smartphone-captured videos using an open-source algorithm for pose estimation (OpenCV[149]), have been shown in a proof-of-concept cross-sectional analysis to show moderate to strong correlations with gold-standard in-clinic measures (the 9 hole peg test (9HPT)),[131] and patient-reported dexterity (ABILHAND).[134]

The goal of the current study was to evaluate the feasibility (adherence, technological validity, patient perspectives on usability (using the Health Information Technology Usability Scale[139]) and validity (changes in dexterity) of remote capture of patient-uploaded videos of hand function, paired with a patient-reported measure of hand function, over 6 months. The overarching goal is to provide a

granular, clinically meaningful metric of movement in the natural environment that can be used to measure progression and/or improvement in clinical trials and inform clinical care and neurological rehabilitation.

Methods

Recruitment

Participants were recruited via convenience sampling from the University of California, San Francisco (UCSF) Multiple Sclerosis Center. Participating neurologists referred interested patients who met initial eligibility criteria (age 18 years or above, a confirmed diagnosis of MS) to the study team. Participants were approached via email and then screened via phone call to determine if eligibility criteria were met. All study activities were approved by the UCSF IRB (IRB# 20-557) and all participants provided written informed consent to enroll. Baseline cross-sectional associations between in-clinic and video-extracted measures were previously reported (Chapter 3).

Study procedures

The total duration of the study was 6 months (Supplementary Figure 1). Participants attended in-person study visits after referral from their neurologist at baseline (0 months) and 6 months and completed assessments at home in between these visits.

Initial Visit

Demographic and clinical data: Age, sex, race, ethnicity, MS diagnosis date, diagnosis date, current disease modifying therapy (DMT), dominant hand (right or left, recorded via patient report), and type of smartphone (iOS or Android) were collected from participants. They additionally completed a smartphone literacy assessment which was marked as pass/fail for the following tasks: using a maps application to navigate to the Golden Gate Bridge, sending a new text message without using an existing

conversation window, and taking a picture. All participants were successfully able to complete these tasks ensuring adequate baseline understanding of smartphone usage.

Functional assessments: In-clinic assessments of hand function included grip and pinch dynamometer testing (Jamar Technologies), vibration testing (Vibratron II- Physitemp), 9HPT, and Action Research Arm Test (ARAT). Grip and pinch strength were tested because adequate strength is required to perform ADLs, and reductions in strength are linked to reduced independence and function in people with neurological diseases.[128] Vibration was tested because quantitative measures of sensation, specifically vibration, show the greatest trend toward detecting subtle changes in functional performance early in the disease course.[129] The 9HPT is considered the clinical standard for assessing dexterity in people with MS,[130,131] and the ARAT is validated in MS and assesses the ability to perform different ADL-like tasks requiring manipulation and transportation of objects using different grip, grasp and pinch functions.[130,132]

Patient-reported outcomes: Using REDCap,[126] a secure, web-based application for managing surveys, participants self-reported their overall MS disease status through a validated electronic patient-reported Expanded Disability Status Scale (e-prEDSS)[133] assessment. They also reported their hand function via the ABILHAND questionnaire, a 23-item questionnaire capturing self-perceived difficulty in completing bimanual activities; the results are summarized into pinch and grip subscores.[134]

Video tasks: A set of self-recorded videos of performance of 3 basic ADLs that require hand function were directly uploaded by participants into the REDCap platform. The ADLs selected were dressing (buttoning a shirt), personal hygiene (brushing teeth), and feeding (fork to mouth). Example videos were provided for each of these tasks within REDCap. Feeding and personal hygiene tasks were uploaded once for each hand, for a total of 5 videos per remote assessment timepoint. It was previously demonstrated (Chapter 3) that four metrics for the buttoning task (position of the wrist position path length, wrist position peaks, index finger path length, and index finger peaks) yielded moderate to high correlations ($r= \geq 0.65$, $p= 0.01$) with in-clinic assessments (9HPT and vibration).

Remote assessment training: Participants were trained to use the REDCap[126,150] to complete the remote patient reported outcomes (PRO), ABILHAND questionnaire,[134] and to upload their self-care videos. Training took approximately 5 minutes. All participants had the required technology (software and operating systems) to complete video uploads at enrollment, thus negating the need to supply new material or devices.

Remote assessment

Participants were sent an invitation to complete these remote assessments (PROs plus video upload) via email weekly for months 1-3 and monthly for months 4-6. Assessment frequency was intended to identify if practice effects occurred from repeated, early trials,[113] while reducing overall participant burden. Participants had 48 hours to complete each assessment, to ensure reasonably regular intervals between weekly assessments. Participants received one reminder email 24 hours after the initial email was sent.

Assessment of feasibility and acceptability

In addition to the remote assessments, feedback on the study activities was solicited monthly via REDCap survey to record participants' perceptions of the feasibility and acceptability of remote video capture as a research modality.

Feasibility questions were based on the Health Information Technology Usability Evaluation Scale (Health-ITUES)[139] questionnaire, which is designed to evaluate usability and ultimately inform the ease of adoption of new technologies. Participants were asked to rate their agreement on a 5-point Likert scale (strongly agree to strongly disagree). Statements covered the usability of the remote assessment ("the remote platform is easy to use"), accessibility ("the remote platform is easy to access"), and the relevance of video tasks to daily life ("the weekly hand function assessments are representative of the types of activities I engage in").

At the 6 month in-clinic visit, participants completed a semi-structured interview with study personnel to provide overall feedback on study procedures and design. A thematic analysis of interview

data was conducted, following the The COnsolidated criteria for Reporting Qualitative research (COREQ) Checklist[151] to minimize bias. Two investigators (AG and IW) coded the data. AG and IW independently developed code lists for salient themes after reviewing all data, then discussed and agreed upon a consolidated list. The code list was then reviewed with senior author (RB). A coding tree can be found in Figure 2.

Video analysis

Analysis was completed using computer vision and machine learning solutions, MediaPipe Hand and OpenCV. A video was considered valid if ≥10 coordinates were generated during initial analysis. Detailed methods on video acquisition and analysis were previously described (Chapter 3).

Statistical analyses

Participant feasibility was calculated by number of participants who completed the study and overall adherence to study protocol. Unpaired t-tests were used to compare participants who withdrew from the study and those who completed to determine if any demographics were predictors of adherence.

To quantify statistically significant changes in each clinical and video measure over the study period, paired t-tests were conducted between baseline and 6 months (primary time interval analyzed), as well as 3-month intervals (baseline and 3 months, 3 and 6 months). Spearman's correlation coefficients were calculated to determine correlations between clinical assessments and video metrics at both 0 and 6 months.

To evaluate predictors of study adherence, baseline demographic baseline data (e.g., age, sex, disease duration) were compared between participants who completed the study and those who withdrew. This was accomplished using an unpaired t-test.

Video analyses were performed using Python 3.4; kinematic and statistical analyses were performed in Matlab.

Results

Demographics

Among the 44 participants who provided remote data, 68% were women, 90% were non-Hispanic white, mean age was 47.2 (SD 12.9), and median disability scale as measured by the ePR-EDSS was 3 (IQR 2, 5) (Table 1).

Feasibility

Recruitment and retention: Of 365 patients who were contacted via email based on clinical referral, 87 agreed to a screening phone call. All 87 participants approached had the required technology (smartphone with built in camera). Of these 87 patients, 50 consented and enrolled in the study (Figure 1).

Engagement in remote activities: Following the enrollment/baseline visit, 44 participants provided at least 1 time point of video uploads. Of the 6 who did not, 2 experienced relapses and chose to withdraw from the study, and 4 were lost to follow up.

Adherence

Altogether, of the initial 50 participants enrolled, 37 (74%) completed the study, which was defined as: ≥3 weekly assessments uploaded and attending the 6 month in-person visit. Of note, no participants who experienced relapses provided remote or 6 month data. Of the 44 participants providing remote data, 2 participants completed >3 weekly assessments but chose to withdraw due to the time demands of the study; and 6 participants completed >3 weekly assessments but were lost to follow up (and did not complete the 6 month in person visit).

Predictors of Study Adherence: The demographic and clinical characteristics of the 37 completers are also outlined in Table 1. They did not differ significantly from the 13 who did not complete the study in terms of age, sex, race/ethnicity, disease duration or baseline 9HPT (p-value >0.05 for each variable compared).

Acceptability (participant feedback)

The study design was well-received by most participants both in their qualitative and quantitative responses (full responses in Supplementary Table 1).

Usability. A majority of participants strongly agreed (77%) that the assessments were easy to access, and a majority strongly agreed (61%) or agree (xx) that the remote platform was easy to use. No participants disagreed with these statements.

Study design. 80% of participants strongly agreed that the reminder emails from the study team were useful in their participation in the study. However, the weekly frequency of surveys in the first 3 months on-study was met with mixed reviews: 41% of participants neither agreed nor disagreed that these were too frequent, but 30% somewhat agreed that they were too frequent.

Relevance. A majority of participants strongly agreed (43%) or somewhat agreed (43%) that the tasks performed in the videos were representative of their daily activities.

Qualitative responses. The thematic analysis of the semi-structured interview data from the 37 participants completing the final study visit generated eight themes: time to complete, assessment frequency, number of questions, retention, accessibility, applicability of questions, self-reflection on symptoms and functional challenges.

Study burden. Survey: 12 participants (32%) mentioned the time to complete surveys. They reported that the survey was quick to complete, stating that it is, "very brief," "doesn't take too long." 5 participants (14%) felt that there were too many questions in the ABILHAND survey. Participants said, "it feels like too many questions," and "I wonder if the activity questions could be shortened to make [the assessment] easier to complete." Frequency: 7 participants (19%) noted the assessment frequency—of these 6 felt that the weekly assessments were too frequent especially since they felt that their PRO answers remained the same week to week. However, one participant felt that "the weekly cadence is not too much and is easy to complete."

Study platform. 12 participants discussed study retention and aspects of study design that aided in their participation. Of these, 6 noted that having paper instructions made it easier to complete the video uploads at home, and an additional 6 liked the reminder emails from the study team which encouraged them to submit their surveys on time. 23 participants discussed accessibility of the platform, the majority (82%) of whom felt the assessments were "easy to do", and that the "[remote platform] worked as expected without technical issues". The remaining participants wished the platform was more mobile friendly, and suggested radio buttons to answer the ABILHAND questions (as opposed to drop down menus).
Relevance. 9 participants discussed the overall applicability of questions, noting that the ABILHAND answers could be more nuanced. One stated, "there is a big jump between 'difficult' and 'easy' which makes the questions hard to answer." 6 participants highlighted that a benefit of study participation encouraged them to self-reflect on their symptoms and hand function. Two participants mentioned that it made them more mindful of the changes in their dexterity. One participant noted, "I get to practice using my weaker hand which normally gives me trouble." Finally, 3 participants highlighted functional challenges in completing required study activities. The buttoning task presented some issues. Three participants said that buttoning shirts was particularly challenging and one said, "I am unable to button shirts, it's more challenging than I remember."

Technological feasibility/quality of patient-uploaded videos

As noted above, 44/50 participants provided at least one timepoint of remotely acquired data. For each of the 15 timepoints, 5 videos were requested (brushing for each hand, eating for each hand, and buttoning), for a total of 75 videos for each participant. The mean and median number of timepoints uploaded for each of these 44 participants was 11 (out of total requested of 15), equivalent to 55 videos (out of total requested of 75).

Of 345 total videos uploaded by participants at baseline and 6 months, 12 videos (3.4%) were not analyzed due to low video quality or participant error (e.g., video out of focus, hand not in camera frame). Upon analysis with mediapipe hand, some videos did not generate sufficient data for kinematic analysis.

This was likely due to the hand not being in the video for a sufficient amount of time. For this reason, a further 54 (15.7%) videos were excluded. This left 279 (80.9%) videos to be analyzed.

Mean video duration was 12.1 seconds (SD 5.6).

Altogether, 44 participants uploaded at least 3 complete sets of videos that were subsequently analyzed. Of note, there were no associations between low (>33.5s)[131] 9HPT scores and difficulty using the survey platform or completing video uploads as measured by Pearson's correlation coefficient ($p>0.05$).

Worsening Dexterity and Changes in Clinical Assessments over 6 months (for 37 participants who provided baseline and 6 month data)

Change over study period was calculated between baseline and 3 months, 3 and 6 months, and baseline and 6 months. The two 3-month intervals revealed no statistically significant differences in collected metrics ($p<0.05$), nor were trends observed, therefore only the 6 month changes are presented here.

Clinical measures: Over the 6 month study period, a marginal reduction in pinch strength in the dominant hand was noted (mean change dominant hand: 0.8kg/cm^2, non-dominant hand: 0.6 kg/cm^2, $p=0.05$). For the 9HPT, mean change was 2.7 seconds for the dominant hand ($p=0.11$) and 2.4 seconds for the non-dominant hand ($p=0.52$); 7 participants (19%) had 9HPT scores that worsened ≥20% over the 6 month study period. The ABILHAND also yielded statistically significant differences ($p=0.05$) for the pinch subset but not for the grasp subset. The remaining clinical assessments (grip strength, vibration) were stable and did not change meaningfully over the study period (Table 2). Of the 13 participants whose pinch strength worsened in both hands over 6 months, 11 reported a worsening in the pinch subscore of the ABILHAND.

Changes in video measures over 6 months

In the current study, for this buttoning task, most metrics derived both for the dominant and non-dominant hand showed statistically significant worsening between the baseline and 6 month timepoints

(Table 3, Figure 3). This includes path position of the wrist (dominant: p=0.04, non-dominant: p=0.05), position peaks of the wrist (dominant: p=0.01, non-dominant: p= 0.02), velocity smoothness of the wrist (dominant: p=0.01), and position peaks of the index finger (dominant: p=0.02, non-dominant: p=0.04).

One participant was identified as an outlier who was potentially skewing the change in function identified in the cohort in terms of overall reduction in joint movement and movement complexity (pink line, Figure 3). However, even with this individual removed, changes over the 6 months in the four metrics mentioned above remained statistically significant ($p<0.05$ for each).

There were no significant changes in measures extracted from the brushing and feeding tasks (Table 4.3)

Change in key buttoning metrics over 6 months

Longitudinal patterns of change were explored in four video metrics for the buttoning task with statistically significant changes over the study period: wrist position path length, wrist position peaks, index finger position path length, and index finger peaks (Figure 3). This analysis revealed an overall reduction in joint movement and movement complexity for both the wrist and index finger landmarks. This is quantitatively demonstrated in Table 2.

The overlap between change demonstrated in 9HPT scores and video metrics was explored (Figure 4). First, numerical change was recorded, i.e., any change from baseline to 6 months. In participants whose measures changed over the study, a majority were captured by the video metrics (62-66%) or both the video metrics and 9HPT (27-34%). Change was only captured in very few participants solely by the 9HPT (6%). This trend was continued when the outcome was categorical, i.e., when a 20% change from baseline was evaluated. Only 6% of participants did not worsen by 20% in either video metrics or 9HPT scores, 42-46% of participants' change was captured by the video metrics, and 44-46% was captured by both the video metrics and 9HPT. Again, very few participants' change was captured by the 9HPT alone (2-6%).

Discussion

Monitoring hand function is important across many neurological diseases, including MS. The current study sought to overcome limitation of existing in-clinic and remote tools by evaluating a novel, patient-friendly form of data collection: patient-generated videos of themselves performing ecologically valid ADLs. This low-cost, low-technology solution was feasible, had high satisfaction and adherence and low barriers to completion, and was able to capture granular change in functional tasks – even in a population with low disability.

Some functional changes were observed in the cohort, noting a marginally significant worsening in pinch strength in the dominant hand corroborated by changes in the ABILHAND questionnaire self-reported worsening in pinch tasks. Overall 19% participants experienced 20% worsening in the 9HPT. In addition, over the 6 months, the video metrics revealed far more granular data. Of the included functional tasks, buttoning requires the most pinch strength, corroborating the results found in the clinical assessments. For buttoning, there were statistically significant differences for both hands in measures of total joint distance traveled (increasing over the study period), changes in movement pattern, and reduced velocity smoothness. The ability to generate more specific, detailed measures on the *aspects* of movement that changed (e.g., speed, distance traveled) can allow clinicians to better understand the extent of functional challenges a patient is experiencing.

This study involved both capacity measures (i.e. what someone is capable of doing, assessed through patient reported outcomes) such as the ABILHAND, and performance measures (i.e. the activity that someone actually does in an unstructured, free-living environment) in the video assessment.[152] A majority of rehabilitation effort focuses on *capacity* measures, though patients seek out rehabilitation services to improve *performance* in their daily lives.[152,153] Assessment through patient-generated videos allows for direct measurement of movement performance of tasks in a natural environment, which can provide relevant information specific to patient goals to enhance clinical decision-making. Both performance-based, functional and in-clinic assessments of hand function have characteristics of verisimilitude (i.e.,

reflect the demands of real-life hand function) and veridicality (i.e., the degree to which the tests are related to measures of everyday functioning), both of which could contribute to ecological validity.

Many groups have published on the benefits of telehealth and mobile health for people with MS, and the current study advances this conversation by evaluating a low-cost, low-technology solution. Intuitively, uploading "selfies" is near-ubiquitous in modern-day life, and is a technology accessible to many. While a digital divide exists, according to 2019 data, 85% of adult Americans own a smartphone, and 93% use the internet regularly—of whom, 75% of whom use a home high-speed broadband network.[112] In the current study, all 87 participants approached as part of study recruitment owned a smartphone compatible with study activities. Further, in contrast to other customized approaches to digital health data collection (e.g., wearable devices, computerized keyboards, smartphone and tablet-based applications, video conferencing with clinicians), the approach explored in this study is cost-effective, as consumer-grade video software is encoded in most smartphones, does not require the purchase of hardware, and no software maintenance beyond usual smartphone software upgrades. The assessment format resulted in high adherence, retention, participant satisfaction and acceptability, and high-quality video data, despite patient age. Overall, 80% of participants owned an iOS device, which is similar to the market share in the United States in 2019 (59.7% iOS, 40.1% Android). 44% of participants owned smartphones manufactured in 2018 or earlier, indicating that the latest technology is not a requirement for this method to be applicable. Further, the diversity of devices included overall speaks to the study's generalizability and ease of scale.

The feasibility and validity of Mediapipe Hands to analyze the uploaded videos suggests that computer vision, an aspect of artificial intelligence that allows computers to derive information from videos and images, may benefit any populations or conditions that require longitudinal monitoring and symptom observation. Regular data collection on functional status is difficult, especially when patients are required to travel to a clinic to complete an in-clinic assessment. Current telemedicine technology (e.g., webcam) mitigates the travel demand, but is not equipped to quantify movement dysfunction, although this is evolving rapidly, as evidenced with the 2023 Zoom feature that can capture movement

and identify a hand raised on video. The method proposed establishes the use of this open-source algorithm to generate high quality kinematic data for longitudinal monitoring. These data can be easily analyzed in a clinician's office with no special training. Additional benefits include the asynchronous aspect of "selfies", so the clinician does not need to be available to collect the data during the video appointment. The data can be uploaded whenever a patient has the time and analyzed at the researcher's convenience. Unlike other mobile health platforms, the raw video files are stored securely so that they can be re-analyzed as the pose estimation technology develops further. Finally, as participants' faces and recognizable features are included in the videos requested, facial blurring could be performed to preserve participant privacy and anonymity, with studies suggesting that this blurring may not impact the performance of models that estimate joint angles and kinematics.[154]

Hand function represents an ideal target for a number of reasons. First, due to the shorter length of upper extremity upper motor neuron axons relative to the lower extremity, measures of hand function may be more sensitive to neuroprotective treatments than measures of ambulation.[155] Second, functional measures such as grip strength are broadly relevant: a meta-analysis of 2 million healthy adults found that increased grip strength was associated with a reduced risk of mortality, regardless of age and sex.[156] Therefore, applications of the current platform can expand far beyond MS. Third, hand function is critical to independence in daily activities including grooming, dressing and feeding, which are crucial to maintaining a sense of self and purpose. This was corroborated in a survey of people with complete cervical spinal cord injuries, who resoundingly reported that they would rather have their arm function restored than lower extremity fuction.[157] Therefore, the results of this study are highly impactful and have important implications for quality of life.

Altogether, the advantages of capturing change through this modality – including accessibility, correlation with patient-reported difficulties and sensitivity to change, are particularly beneficial given the often insidious, "silent", disease progression in people with MS. Highly specific, easily attainable metrics such as those obtained from the current video analyses have the potential to allow clinicians to not only identify limitation in functional activities, but to detect early changes contributing to disease progression.

This study has several limitations. First, it was underpowered to conduct additional sub-analyses, such as principal component analyses. While an *a priori* power analysis identified led us to recruit 50 participants, missing data from video analysis prevented completion of sub-analyses. Future studies should investigate changes in video metrics by subgroup based on performance in clinical measures at baseline. Second, despite the written and digital instructions provided to participants, some had challenges with standardization of video recordings. Videos and hands out of frame, shaky images, and out of focus recordings limited some analyses, and reduced the overall number of usable videos to 80.9% of total uploaded. While this is a limitation of patient-generated videos, it remains advantageous in allowing participants to provide data in their homes as well as limiting in-person visits, and in the future, a brief review of uploaded videos would allow researchers to flag videos requiring re-filming. Reassuringly, most patients provided at least xxx sets of videos and their baseline dexterity did not influence their ability to upload videos. Third, while the EDSS was recorded for each participant, a complete neurological exam was not performed. Subsequent studies may expand the analyses to evaluate associations of video measures with other clinical, non-dexterity measures of upper extremity function such as the finger to nose test, dysdiadochokinesia, and finger tapping speed. Fourth, the overall health of this actively treated cohort of participants at baseline likely reduced the percentage of patients changing clinically over 6 months, and therefore influenced the conclusions of this research approach. The fact that many metrics from the buttoning task did change significantly is both concerning from a clinical standpoint and encouraging in the sense that the metric was sensitive to even sub-clinical change. Further, data collection at 3-month intervals may be sufficient rather than weekly or monthly. Finally, as noted in Chapter 3, while data on hand dominance were collected, dominance did not always coincide with a participant's weaker or stronger side. MS-related motor and sensory symptoms often present asymmetrically, and it is possible that dexterity changes in the weaker side may be a more sensitive variable relative to hand dominance.

Overall, "self-care selfies" represent a novel approach to quantify real-time kinematics in ADLs using pose estimation algorithms. Its validity against currently used in-clinic assessments, feasibility, and

participant acceptability enhance its potential to generate critical insights to promote effective and efficient early rehabilitation protocols.

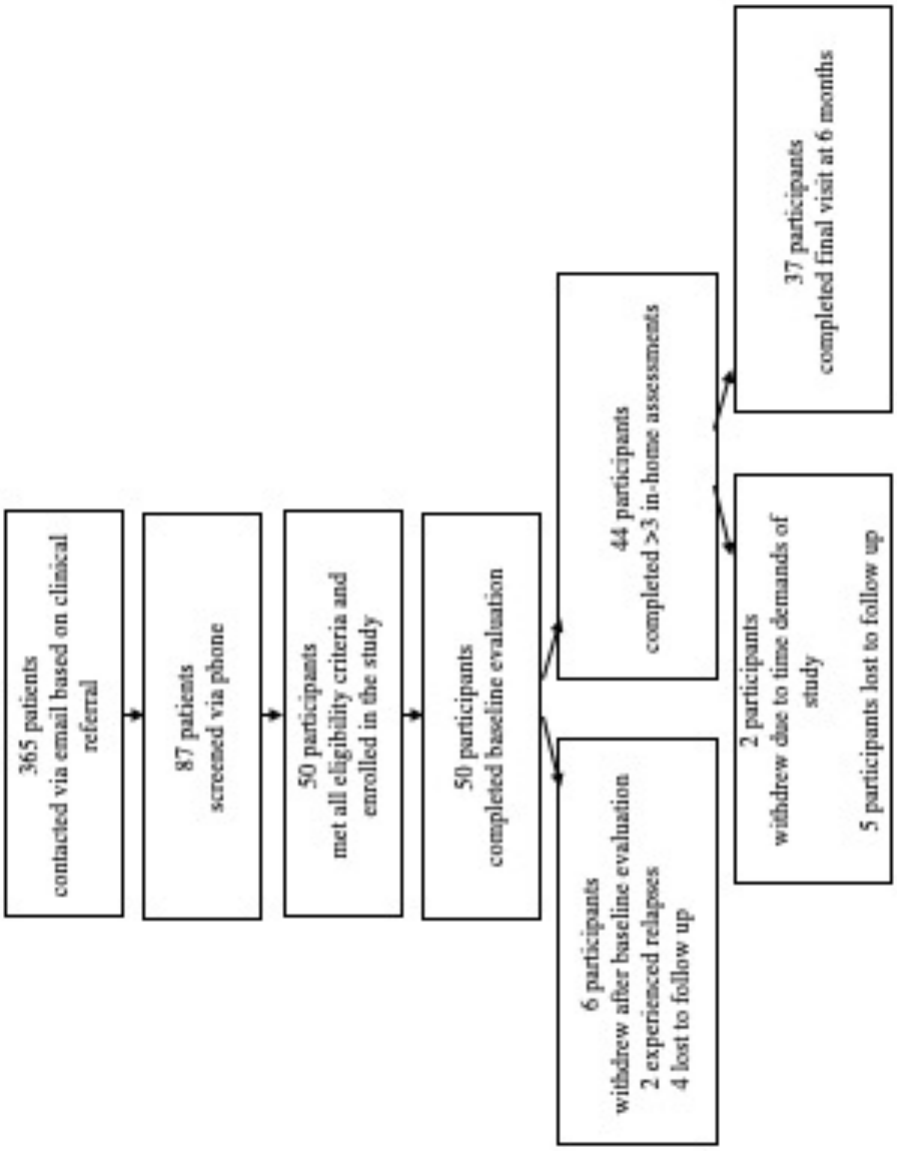


Figure 4. 1. Participant recruitment and retention. *Tracking patient contacts, enrollments, withdrawals, and completion.*

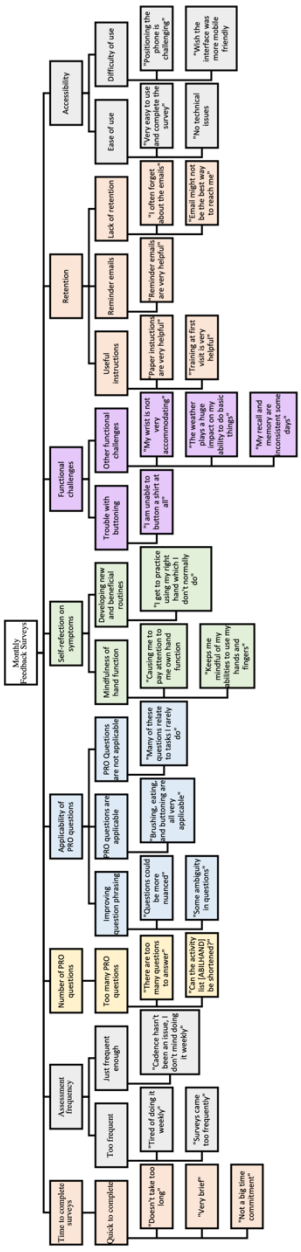


Figure 4. 2. Coding tree of qualitative analysis of participant interviews at 6 month visit. *Qualitative analysis followed the CORE-Q checklist. Themes are listed in primary branches, codes in secondary branches, and quotes in tertiary branches.*

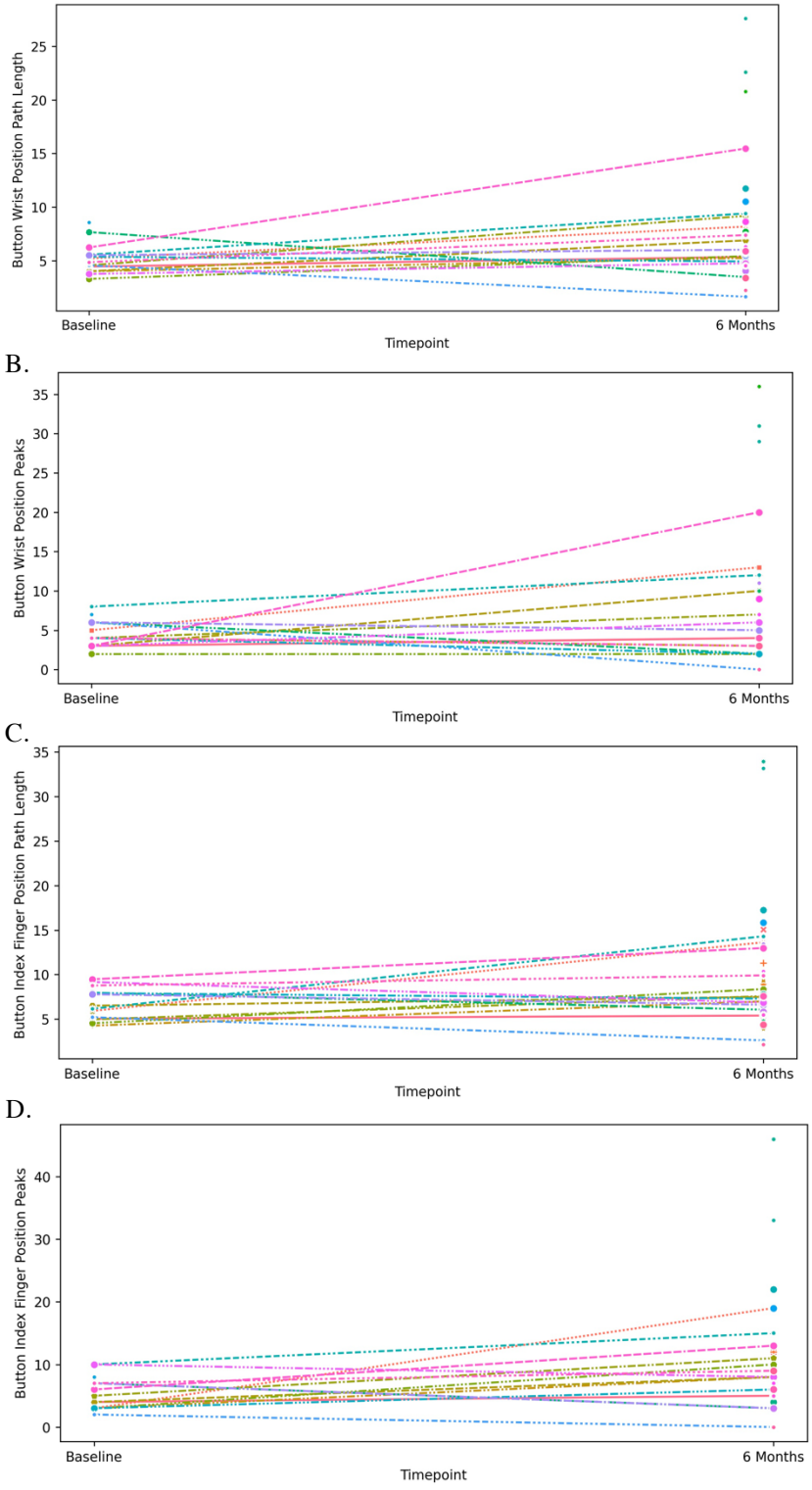


Figure 4. 3. Key buttoning metrics at baseline and 6 months in the nondominant hand. *(A) Wrist position path length, p=0.05, (B) wrist position peaks, p= 0.02, (C) index position path length, p= 0.10, (D) index position peaks, p= 0.04).*

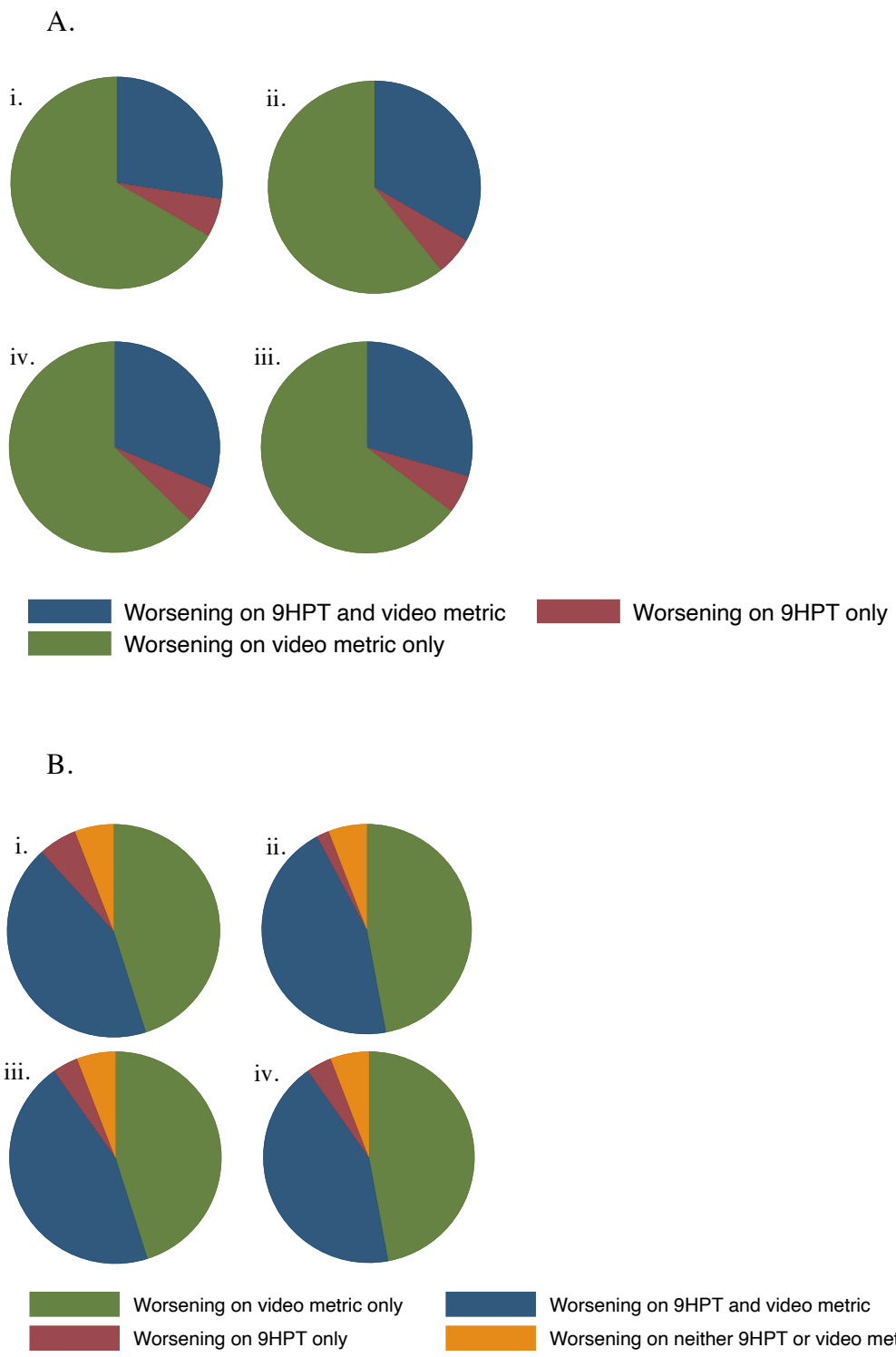
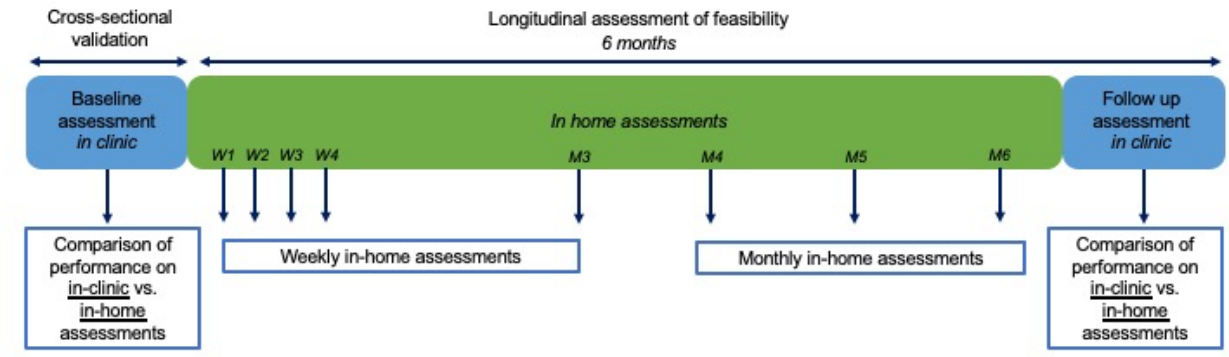


Figure 4. 4. Sensitivity of 9HPT and video metrics in change detection over study period*. i: wrist position path length, ii: wrist position peaks, iii: index position path length, iv: index position peaks. (A) Numerical change in video metrics and 9HPT scores (B) 20% change from baseline to 6 months in video metrics and 9HPT scores.*



Supp Figure 4. 1. Study timeline including in-clinic visits and frequency and administration of in-home assessments.

Table 4. 1. Baseline Demographics and Characteristics of Participants

	Remote completers (N=44)	Study completers (N=37)
Gender [N (%)]		
Men	14 (32%)	14 (37%)
Women	30 (68%)	23 (62%)
Non-binary	0	0
Age [mean (SD)]	47.2 (12.9)	48.1 (13.3)
Race [N (%)]		
White, non Hispanic	40 (90%)	33 (88%)
Black	0	0
Asian	3 (7%)	3 (7%)
Hispanic	0	0
Other	1 (2%)	1 (3%)
EDSS [median (IQR)]	3 (2, 5)	3 (2.5, 5)
Disease duration (years) [mean (SD)]	14.3 (9.2)	14.9 (8.4)
MS Subtype [N (%)]		
Relapsing remitting	36 (82%)	30 (81%)
Primary progressive	4 (9%)	3 (8%)
Secondary progressive	4 (9%)	4 (11%)
Dominant hand [N (%)]		
Right	38 (86%)	31 (84%)
Left	6 (14%)	6 (16%)
Smartphone type [N (%)]		
iOS	34 (77%)	29 (78%)
Android	10 (23%)	8 (22%)
Grip strength (kg/cm^2) [mean (SD)]		
Dominant	28.6 (9.4)	28.6 (9.5)
Non-dominant	27 (9.5)	26.8 (8.8)
Pinch strength (kg/cm^2) [mean (SD)]		
Dominant	5.8 (2.6)	5.4 (2.5)
Non-dominant	5.6 (2.5)	5.6 (2.5)
Nine-hole peg test (seconds) [mean (SD)]		
Dominant	23.5 (9.7)	24.4 (11.0)
Non-dominant	28.7 (14.9)	30.1 (16.9)
Vibration sense (Hertz) [mean (SD)]		
Dominant	2.4 (2.7)	2.7 (3.1)
Non-dominant	2.3 (2.6)	2.6 (2.9)
ARAT-Grip [mean (SD)]		
Dominant	11.9 (0.6)	11.9 (0.7)
Non-dominant	11.9 (0.8)	11.8 (1.0)
ARAT-Grasp [mean (SD)]		
Dominant	17.8 (0.7)	17.9 (0.5)
Non-dominant	17.9 (0.5)	17.9 (0.5)
ARAT-Pinch [mean (SD)]		
Dominant	17.3 (2.3)	17.6 (2.0)
Non-dominant	17.1 (2.7)	17.3 (2.7)
ARAT-Gross motor [mean (SD)]		
Dominant	9 (0)	9 (0)
Non-dominant	9 (0)	9 (0)
ABILHAND [mean (SD)]		
Pinch subscore	29.5 (4.6)	29.1 (4.4)
Grip subscore	35.5 (4.4)	35.3 (4.7)

No significant differences ($p>0.05$) were found between groups

Table 4. 2. Change in Hand Function Using Established Clinical and Patient-Reported Measures over 6 Months

Outcome	Baseline	6 months	Absolute Change from Baseline to 6 months (N)			p-value
			Worsened	Improved	No Change	
Participants (N)	44	37				
EDSS			10	16	11	
Median (IQR)	3.5 (2, 5)	2.5 (2, 5)				0.98
Patients with relapsing MS	43	27				
Grip strength (kg/cm^2) [mean (SD)]						
Dominant	28.6 (9.4)	28.2 (10.5)	12	11	14	0.84
Non-dominant	27 (9.5)	26.6 (10.1)	13	16	8	0.97
Pinch strength (kg/cm^2) [mean (SD)]						
Dominant	5.8 (2.6)	5.0 (2.7)	13	15	9	**0.05***
Non-dominant	5.6 (2.5)	5.0 (2.2)	13	14	10	0.16
9HPT (s)						
Dominant	23.5 (9.7)	20.8 (12.9)	15	9	13	0.11
Non-dominant	28.7 (14.9)	26.3 (18.8)	13	12	12	0.52
Vibration sense (Hertz) [mean (SD)]						
Dominant	2.4 (2.7)	3.0 (3.9)	12	11	14	0.35
Non-dominant	2.3 (2.6)	2.5 (3.5)	12	13	12	0.81
ABILHAND						
Pinch Subscore	29.5 (4.6)	29.1 (35.3)	11	9	17	**0.05***
Grip Subscore	35.5 (4.4)	35.3 (4.7)	7	6	24	0.17

Table 4. 3. Changes in video metrics derived from buttoning, eating, and brushing tasks over 6 months (N=37)

Outcome	Mean Difference	95% Confidence Interval	p-value
Buttoning task			
Wrist position AUC			
Dominant	-0.03	-0.12, 0.07	0.54
Non-Dominant	0.07	-0.04, 0.19	0.19
Wrist position path length			
Dominant	-1.50	-2.94, -0.07	**0.04***
Non-Dominant	-1.73	-3.61, 0.14	**0.05***
Wrist position peaks			
Dominant	-2.57	-5.91, 0.77	**0.01***
Non-Dominant	-2.00	-5.36, 1.35	**0.02***
Wrist velocity AUC			
Dominant	0.03	-0.01, 0.06	0.15
Non-Dominant	-0.01	-0.11, 0.09	0.83
Wrist velocity smoothness			
Dominant	0.03	-0.008, 0.07	**0.01***
Non-Dominant	-0.006	-0.11, 0.10	0.09
Index position AUC			
Dominant	-0.03	-0.13, 0.06	0.48
Non-Dominant	0.07	-0.04, 0.17	0.18
Index position path length			
Dominant	-1.28	-3.23, 0.67	0.18
Non-Dominant	-1.60	-3.57, 0.37	0.10
Index position peaks			
Dominant	-4.29	-7.87, -0.69	**0.02***
Non-Dominant	-3.14	-6.23, -0.06	**0.04***
Index velocity AUC			
Dominant	0.02	-0.05, 0.08	0.52
Non-Dominant	0.01	-0.04, 0.06	0.61
Index velocity smoothness			
Dominant	0.03	-0.04, 0.09	0.41
Non-Dominant	0.02	-0.04, 0.07	0.55
Brushing task			
Wrist position AUC			
Dominant	0.01	-0.08, 0.09	0.90
Non-Dominant	0.001	-0.11, 0.11	0.97
Wrist position path length			
Dominant	-6.57	-19.58, 6.45	0.29
Non-Dominant	0.84	-6.85, 8.53	0.81
Wrist position peaks			
Dominant	-2.69	-8.08, 2.70	0.29
Non-Dominant	-1.08	-5.37, 3.22	0.59

Table 4. 4. Changes in Video Metrics derived from buttoning, eating, and hair brushing tasks over 6 months (N=37)

Outcome	Mean Difference	95% Confidence Interval	p-value
Dominant	0.04	-0.10, 0.18	0.15
Non-Dominant	-0.08	-0.19, 0.03	**0.05***
Wrist velocity smoothness			
Dominant	0.02	-0.11, 0.16	0.71
Non-Dominant	-0.07	-0.17, 0.04	0.22
Index position AUC			
Dominant	-0.03	-0.13, 0.08	0.61
Non-Dominant	0.06	-0.04, 0.16	0.26
Index position path length			
Dominant	-4.91	-16.32, 6.5	0.36
Non-Dominant	1.88	-2.41, 6.16	0.35
Index position peaks			
Dominant	-2.85	-9.14, 3.45	0.34
Non-Dominant	1.84	-2.05, 5.73	0.32
Index velocity AUC			
Dominant	-0.03	-0.14, 0.08	0.50
Non-Dominant	-0.03	-0.12, 0.06	0.49
Index velocity smoothness			
Dominant	-0.03	-0.15, 0.09	0.61
Non-Dominant	-0.002	-0.08, 0.08	0.95
Eating task			
Wrist position AUC			
Dominant	-0.04	-0.17, 0.09	0.49
Non-Dominant	0.02	-0.17, 0.21	0.80
Wrist position path length			
Dominant	-1.18	-5.55, 3.19	0.56
Non-Dominant	0.21	-1.04, 1.46	0.71
Wrist position peaks			
Dominant	-1.45	-3.80, 0.89	0.19
Non-Dominant	0.89	-0.62, 2.39	0.21
Wrist velocity AUC			
Dominant	-0.03	-0.17, 0.11	0.65
Non-Dominant	-0.07	-0.21, 0.06	0.24
Wrist velocity smoothness			
Dominant	-0.09	-0.22, 0.05	0.19
Non-Dominant	-0.04	-0.14, 0.07	0.41
Index position AUC			
Dominant	0.007	-0.13, 0.14	0.91
Non-Dominant	-0.03	-0.19, 0.13	0.65
Index position path length			
Dominant	-0.34	-4.72, 4.05	0.87
Non-Dominant	-0.19	-0.98, 0.60	0.59

Table 4. 5. Changes in Video Metrics derived from buttoning, eating, and hair brushing tasks over 6 months (N=37)

Outcome	Mean Difference	95% Confidence Interval	p-value
Dominant	-0.45	-2.33, 1.43	0.60
Non-Dominant	-0.11	-1.09, 0.86	0.79
Index velocity AUC			
Dominant	0.02	-0.10, 0.14	0.71
Non-Dominant	-0.03	-0.17, 0.11	0.64
Index velocity smoothness			
Dominant	-0.01	-0.13, 0.19	0.83
Non-Dominant	-0.04	-0.12, 0.03	0.22

Supp Table 4. 1. Responses to Monthly Feedback Surveys

	N (%)				
	The assessments are easy to access.	The remote platform is easy to use.	The reminder emails from the study coordinator are helpful.	The weekly hand function assessments are too frequent.	The weekly hand function assessments are representative of the types of activities I engage in.
Strongly Agree	34 (77%)	27 (61%)	35 (80%)	5 (11%)	19 (43%)
Somewhat Agree	8 (18%)	16 (36%)	3 (7%)	13 (30%)	19 (43%)
Neither Agree nor Disagree	2 (5%)	1 (2%)	5 (11%)	18 (41%)	4 (9%)
Somewhat Disagree	0	0	1 (2%)	5 (11%)	2 (5%)
Strongly Disagree	0	0	0	3 (7%)	0

References

1. Luijten MAJ, Eekhout I, D'Hooghe M, Uitdehaag BMJ, Mokkink LB. Development of the Arm Function in Multiple Sclerosis Questionnaire-Short Form (AMSQ-SF): A static 10-item version. *Multiple Sclerosis Journal*. 2018;24(14):1892-1901.
2. Huertas-Hoyas E, Máximo-Bocanegra N, Diaz-Toro C, et al. A Descriptive Cross-Sectional Study of Manipulative Dexterity on Different Subtypes of Multiple Sclerosis. *Occupational Therapy International*. 2020:1-8.
3. Basak T, Unver V, Demirkaya S. Activities of daily living and self-care agency in patients with multiple sclerosis for the first 10 years. *Rehabil Nurs*. 2015;40(1):60-65.
4. Timmermans ST, de Groot V, Beckerman H. Ten-year disease progression in multiple sclerosis: walking declines more rapidly than arm and hand function. *Mult Scler Relat Disord*. 2020;45:102343.
5. Fox EJ, Markowitz C, Applebee A, et al. Ocrelizumab reduces progression of upper extremity impairment in patients with primary progressive multiple sclerosis: Findings from the phase III randomized ORATORIO trial. *Multiple Sclerosis Journal*. 2018;24(14):1862-1870.
6. Kapoor R, Ho P-R, Campbell N, et al. Effect of natalizumab on disease progression in secondary progressive multiple sclerosis (ASCEND): a phase 3, randomised, double-blind, placebo-controlled trial with an open-label extension. *The Lancet Neurology*. 2018;17(5):405-415.
7. Cambron M, Mostert J, D'Hooghe M, et al. Fluoxetine in progressive multiple sclerosis: The FLUOX-PMS trial. *Multiple Sclerosis Journal*. 2019;25(13):1728-1735.
8. Newsome SD, von Geldern G, Shou H, et al. Longitudinal assessment of hand function in individuals with multiple sclerosis. *Mult Scler Relat Disord*. 2019;32:107-113.
9. Lang CE, Bland MD, Bailey RR, Schaefer SY, Birkenmeier RL. Assessment of upper extremity impairment, function, and activity after stroke: foundations for clinical decision making. *J Hand Ther*. 2013;26(2):104-114;quiz 115.

10. *International classification of functioning, disability, and health : ICF* [computer program]. Version 1.0. Geneva : World Health Organization, [2001] ©2001; 2001.

11. Feys P, Romberg A, Ruutiainen J, Ketelaer P. Interference of Upper Limb Tremor on Daily Life Activities in People with Multiple Sclerosis. *Occupational Therapy In Health Care*. 2004;17(3-4):81-95.

12. Sabari JS, Lim AL, Velozo CA, Lehman L, Kieran O, Lai JS. Assessing arm and hand function after stroke: a validity test of the hierarchical scoring system used in the motor assessment scale for stroke. *Arch Phys Med Rehabil*. 2005;86(8):1609-1615.

13. Kandaswamy D, M M, Alexander M, Prabhu K, S MG, Krothapalli SB. Quantitative Assessment of Hand Dysfunction in Patients with Early Parkinson's Disease and Focal Hand Dystonia. *J Mov Disord*. 2018;11(1):35-44.

14. Reich DS, Lucchinetti CF, Calabresi PA. Multiple Sclerosis. *N Engl J Med*. 2018;378(2):169-180.

15. Block VJ, Lizée A, Crabtree-Hartman E, et al. Continuous daily assessment of multiple sclerosis disability using remote step count monitoring. *J Neurol*. 2017;264(2):316-326.

16. Venkataraman K, Amis K, Landerman LR, Caves K, Koh GC, Hoenig H. Teleassessment of Gait and Gait Aids: Validity and Interrater Reliability. *Physical Therapy*. 2020;100(4):708-717.

17. Hoffman NB, Prieto NM. Clinical Video Telehealth for Gait and Balance. *Fed Pract*. 2016;33(2):34-38.

18. Piau A, Wild K, Mattek N, Kaye J. Current State of Digital Biomarker Technologies for Real-Life, Home-Based Monitoring of Cognitive Function for Mild Cognitive Impairment to Mild Alzheimer Disease and Implications for Clinical Care: Systematic Review. *J Med Internet Res*. 2019;21(8):e12785.

19. Creagh AP, Simillion C, Scotland A, et al. Smartphone-based remote assessment of upper extremity function for multiple sclerosis using the Draw a Shape Test. *Physiol Meas*. 2020;41(5):054002.

20. InformedHealth.org. How do hands work? Institute for Quality and Efficiency in Health Care (IQWiG). Published 2010. Accessed June 17, 2021, 2021.

21. Chen CC, Kasven N, Karpatkin HI, Sylvester A. Hand strength and perceived manual ability among patients with multiple sclerosis. *Arch Phys Med Rehabil*. 2007;88(6):794-797.

22. Butler DP, Murray A, Horwitz M. Hand manifestations of neurological disease: some alternatives to consider. *Br J Gen Pract*. 2016;66(647):331-332.

23. Baumann CR. Epidemiology, diagnosis and differential diagnosis in Parkinson's disease tremor. *Parkinsonism Relat Disord*. 2012;18 Suppl 1:S90-92.

24. Ng YS, Stein J, Ning M, Black-Schaffer RM. Comparison of clinical characteristics and functional outcomes of ischemic stroke in different vascular territories. *Stroke*. 2007;38(8):2309-2314.

25. Jang SH, Chang MC. Motor outcomes of patients with a complete middle cerebral artery territory infarct. *Neural Regen Res*. 2013;8(20):1892-1897.

26. Soon S, Svavarsdottir H, Downey C, Jayne DG. Wearable devices for remote vital signs monitoring in the outpatient setting: an overview of the field. *BMJ Innovations*. 2020;6(2):55.

27. Ganeshan R, Enriquez AD, Freeman JV. Remote monitoring of implantable cardiac devices: current state and future directions. *Current Opinion in Cardiology*. 2018;33(1).

28. Block VJ, Bove R, Zhao C, et al. Association of Continuous Assessment of Step Count by Remote Monitoring With Disability Progression Among Adults With Multiple Sclerosis. *JAMA Netw Open*. 2019;2(3):e190570.

29. Block VA, Pitsch E, Tahir P, Cree BA, Allen DD, Gelfand JM. Remote Physical Activity Monitoring in Neurological Disease: A Systematic Review. *PLoS One*. 2016;11(4):e0154335.

30. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 2009;339:b2535.

31. National Heart L, and Blood Institute. Study Quality Assessment Tools. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools. Published 2014. Accessed May 12, 2021, 2021.

32. Glen S. Concurrent Validity Definition and Examples. https://www.statisticshowto.com/concurrent-validity/. Published 2015. Accessed April 6, 2021.

33. Glen S. Reliability and Validity in Research: Definitions, Examples. https://www.statisticshowto.com/reliability-validity-definitions-examples/. Published 2016. Accessed April 6, 2021.

34. Kimmerle M, Mainwaring L, Borenstein M. The functional repertoire of the hand and its application to assessment. *Am J Occup Ther*. 2003;57(5):489-498.

35. Edemekong PF, Bomgaars DL, Sukumaran S, Levy SB. Activities of Daily Living. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing StatPearls Publishing LLC.; 2021.

36. Guo HJ, Sapra A. Instrumental Activity of Daily Living. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing StatPearls Publishing LLC.; 2021.

37. Cai G, Lin Z, Dai H, et al. Quantitative assessment of parkinsonian tremor based on a linear acceleration extraction algorithm. *Biomedical Signal Processing and Control*. 2018;42:53-62.

38. Channa A, Ifrim R-C, Popescu D, Popescu N. A-WEAR Bracelet for Detection of Hand Tremor and Bradykinesia in Parkinson's Patients. *Sensors*. 2021;21(3).

39. Cole BT, Roy SH, De Luca CJ, Nawab SH. Dynamical learning and tracking of tremor and dyskinesia from wearable sensors. *IEEE Trans Neural Syst Rehabil Eng*. 2014;22(5):982-991.

40. Dai H, Cai G, Lin Z, Wang Z, Ye Q. Validation of Inertial Sensing-Based Wearable Device for Tremor and Bradykinesia Quantification. *IEEE Journal of Biomedical and Health Informatics*. 2021;25(4):997-1005.

41. Ferreira JJ, Godinho C, Santos AT, et al. Quantitative home-based assessment of Parkinson's symptoms: The SENSE-PARK feasibility and usability study. *BMC Neurology*. 2015;15(1):89.

42. Giuffrida JP, Riley DE, Maddux BN, Heldman DA. Clinically deployable Kinesia technology for automated tremor assessment. *Mov Disord*. 2009;24(5):723-730.

43. Halloran S, Tang L, Guan Y, Shi JQ, Eyre J. Remote monitoring of stroke patients' rehabilitation using wearable accelerometers. Proceedings of the 23rd International Symposium on Wearable Computers; 2019; London, United Kingdom.

44. Heijmans M, Habets J, Kuijf M, Kubben P, Herff C. Evaluation of Parkinson's Disease at Home: Predicting Tremor from Wearable Sensors. Paper presented at: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 23-27 July 2019, 2019.

45. Hssayeni MD, Jimenez-Shahed J, Burack MA, Ghoraani B. Wearable Sensors for Estimation of Parkinsonian Tremor Severity during Free Body Movements. *Sensors*. 2019;19(19).

46. Jeon H, Lee W, Park H, et al. Automatic Classification of Tremor Severity in Parkinson's Disease Using a Wearable Device. *Sensors*. 2017;17(9).

47. Kim HB, Lee WW, Kim A, et al. Wrist sensor-based tremor severity quantification in Parkinson's disease using convolutional neural network. *Computers in Biology and Medicine*. 2018;95:140-146.

48. Mahadevan N, Demanuele C, Zhang H, et al. Development of digital biomarkers for resting tremor and bradykinesia using a wrist-worn wearable device. *npj Digital Medicine*. 2020;3(1):5.

49. Mera TO, Heldman DA, Espay AJ, Payne M, Giuffrida JP. Feasibility of home-based automated Parkinson's disease motor assessment. *Journal of Neuroscience Methods*. 2012;203(1):152-156.

50. Rigas G, Tzallas AT, Tsipouras MG, et al. Assessment of tremor activity in the Parkinson's disease using a set of wearable sensors. *IEEE Trans Inf Technol Biomed*. 2012;16(3):478-487.

51. San-Segundo R, Zhang A, Cebulla A, et al. Parkinson's Disease Tremor Detection in the Wild Using Wearable Accelerometers. *Sensors*. 2020;20(20).

52. Sanchez-Perez LA, Sanchez-Fernandez LP, Shaout A, Martinez-Hernandez JM, Alvarez-Noriega MJ. Rest tremor quantification based on fuzzy inference systems and wearable sensors. *International Journal of Medical Informatics*. 2018;114:6-17.

53. Stamatakis J, Ambroise J, Crémers J, et al. Finger Tapping Clinimetric Score Prediction in Parkinson's Disease Using Low-Cost Accelerometers. *Computational Intelligence and Neuroscience*. 2013;2013:717853.

54. Zhang A, De la Torre F, Hodgins J. Comparing laboratory and in-the-wild data for continuous Parkinson's Disease tremor detection. *Annu Int Conf IEEE Eng Med Biol Soc*. 2020;2020:5436-5441.

55. Boroojerdi B, Ghaffari R, Mahadevan N, et al. Clinical feasibility of a wearable, conformable sensor patch to monitor motor symptoms in Parkinson's disease. *Parkinsonism Relat Disord*. 2019;61:70-76.

56. Akram N, Li H, Ben-Joseph A, et al. Developing and Validating a New Web-Based Tapping Test for Measuring Distal Bradykinesia in Parkinson's Disease. *medRxiv*. 2020:2020.2006.2030.20141572.

57. Giancardo L, Sánchez-Ferro A, Arroyo-Gallego T, et al. Computer keyboard interaction as an indicator of early Parkinson's disease. *Scientific Reports*. 2016;6(1):34468.

58. Papadopoulos A, Iakovakis D, Klingelhoefer L, et al. Unobtrusive detection of Parkinson's disease from multi-modal and in-the-wild sensor data using deep learning techniques. *Sci Rep*. 2020;10(1):21370.

59. Lam KH, Meijer KA, Loonstra FC, et al. Real-world keystroke dynamics are a potentially valid biomarker for clinical disability in multiple sclerosis. *Multiple Sclerosis Journal*. 2020:1352458520968797.

60. Matarazzo M, Arroyo-Gallego T, Montero P, et al. Remote Monitoring of Treatment Response in Parkinson's Disease: The Habit of Typing on a Computer. *Movement Disorders*. 2019;34(10):1488-1495.

61. Noyce AJ, Nagy A, Acharya S, et al. Bradykinesia-Akinesia Incoordination Test: Validating an Online Keyboard Test of Upper Limb Function. *PLOS ONE*. 2014;9(4):e96260.

62. Shribman S, Hasan H, Hadavi S, Giovannoni G, Noyce AJ. The BRAIN test: a keyboard-tapping test to assess disability and clinical features of multiple sclerosis. *Journal of neurology*. 2018;265(2):285-290.

63. Trager MH, Wilkins KB, Koop MM, Bronte-Stewart H. A validated measure of rigidity in Parkinson's disease using alternating finger tapping on an engineered keyboard. *Parkinsonism & Related Disorders*. 2020;81:161-164.

64. Adams WR. The detection of hand tremor through the characteristics of finger movement while typing. *bioRxiv*. 2018:385286.

65. Londral A, Pinto S, de Carvalho M. Markers for upper limb dysfunction in Amyotrophic Lateral Sclerosis using analysis of typing activity. *Clin Neurophysiol*. 2016;127(1):925-931.

66. Aghanavesi S, Nyholm D, Senek M, Bergquist F, Memedi M. A smartphone-based system to quantify dexterity in Parkinson's disease patients. *Informatics in Medicine Unlocked*. 2017;9:11-17.

67. Arroyo-Gallego T, Ledesma-Carbayo MJ, Á S-F, et al. Detection of Motor Impairment in Parkinson's Disease Via Mobile Touchscreen Typing. *IEEE Transactions on Biomedical Engineering*. 2017;64(9):1994-2002.

68. Bazgir O, Habibi SAH, Palma L, Pierleoni P, Nafees S. A Classification System for Assessment and Home Monitoring of Tremor in Patients with Parkinson's Disease. *J Med Signals Sens*. 2018;8(2):65-72.

69. Iakovakis D, Chaudhuri KR, Klingelhoefer L, et al. Screening of Parkinsonian subtle fine-motor impairment from touchscreen typing via deep learning. *Scientific Reports*. 2020;10(1):12623.

70. Iakovakis D, Hadjidimitriou S, Charisis V, Bostantzopoulou S, Katsarou Z, Hadjileontiadis LJ. Touchscreen typing-pattern analysis for detecting fine motor skills decline in early-stage Parkinson’s disease. *Scientific Reports*. 2018;8(1):7663.

71. Lee CY, Kang SJ, Hong SK, Ma HI, Lee U, Kim YJ. A Validation Study of a Smartphone-Based Finger Tapping Application for Quantitative Assessment of Bradykinesia in Parkinson's Disease. *PLoS One*. 2016;11(7):e0158852.

72. Lee U, Kang SJ, Choi JH, Kim YJ, Ma H-I. Mobile application of finger tapping task assessment for early diagnosis of Parkinson's disease. *Electronics Letters*. 2016;52(24):1976-1978.

73. Papadopoulos A, Kyritsis K, Klingelhoefer L, Bostanjopoulou S, Chaudhuri KR, Delopoulos A. Detecting Parkinsonian Tremor From IMU Data Collected in-the-Wild Using Deep Multiple-Instance Learning. *IEEE J Biomed Health Inform*. 2020;24(9):2559-2569.

74. Schallert W, Fluet M-C, Kesselring J, Kool J. Evaluation of upper limb function with digitizing tablet-based tests: reliability and discriminative validity in healthy persons and patients with neurological disorders. *Disability and Rehabilitation*. 2020:1-9.

75. Simonet C, Galmes MA, Lambert C, et al. Slow Motion Analysis of Repetitive Tapping (SMART) test: measuring bradykinesia in recently diagnosed Parkinson’s disease and idiopathic anosmia. *medRxiv*. 2021:2021.2003.2024.21254234.

76. Wissel BD, Mitsi G, Dwivedi AK, et al. Tablet-Based Application for Objective Measurement of Motor Fluctuations in Parkinson Disease. *Digital Biomarkers*. 2017;1(2):126-135.

77. Jha A, Menozzi E, Oyekan R, et al. The CloudUPDRS smartphone software in Parkinson’s study: cross-validation against blinded human raters. *npj Parkinson's Disease*. 2020;6(1):36.

78. Lipsmeier F, Taylor KI, Kilchenmann T, et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. *Movement Disorders*. 2018;33(8):1287-1297.

79. Orozco-Arroyave JR, Vásquez-Correa JC, Klumpp P, et al. Apkinson: the smartphone application for telemonitoring Parkinson's patients through speech, gait and hands movement. *Neurodegener Dis Manag*. 2020.

80. Pan D, Dhall R, Lieberman A, Petitti DB. A mobile cloud-based Parkinson's disease assessment system for home-based monitoring. *JMIR Mhealth Uhealth*. 2015;3(1):e29.

81. Zhan A, Little M, Harris DA, et al. High Frequency Remote Monitoring of Parkinson's Disease via Smartphone: Platform Overview and Medication Response Detection. *ArXiv*. 2016;abs/1601.00960.

82. Mitsi G, Mendoza EU, Wissel BD, et al. Biometric Digital Health Technology for Measuring Motor Function in Parkinson's Disease: Results from a Feasibility and Patient Satisfaction Study. *Front Neurol*. 2017;8:273.

83. Pratap A, Grant D, Vegesna A, et al. Evaluating the Utility of Smartphone-Based Sensor Assessments in Persons With Multiple Sclerosis in the Real-World Using an App (elevateMS): Observational, Prospective Pilot Digital Health Study. *JMIR Mhealth Uhealth*. 2020;8(10):e22108.

84. Cabrera-Martos I, Ortiz-Rubio A, Torres-Sánchez I, López-López L, Rodríguez-Torres J, Carmen Valenza M. Agreement Between Face-to-Face and Tele-assessment of Upper Limb Functioning in Patients with Parkinson Disease. *Pm r*. 2019;11(6):590-596.

85. Hoffmann T, Russell T, Thompson L, Vincent A, Nelson M. Using the Internet to assess activities of daily living and hand function in people with Parkinson's disease. *NeuroRehabilitation*. 2008;23(3):253-261.

86. Amano S, Umeji A, Uchita A, et al. Reliability of remote evaluation for the Fugl-Meyer assessment and the action research arm test in hemiparetic patients after stroke. *Top Stroke Rehabil*. 2018;25(6):432-437.

87. Arora S, Venkataraman V, Zhan A, et al. Detecting and monitoring the symptoms of Parkinson's disease using smartphones: A pilot study. *Parkinsonism Relat Disord*. 2015;21(6):650-653.

88. Burdea GC, Grampurohit N, Kim N, et al. Feasibility of integrative games and novel therapeutic game controller for telerehabilitation of individuals chronic post-stroke living in the community. *Top Stroke Rehabil*. 2020;27(5):321-336.

89. Yu L, Xiong D, Guo L, Wang J. A remote quantitative Fugl-Meyer assessment framework for stroke patients based on wearable sensor networks. *Comput Methods Programs Biomed*. 2016;128:100-110.

90. Albani G, Ferraris C, Nerino R, et al. An Integrated Multi-Sensor Approach for the Remote Monitoring of Parkinson's Disease. *Sensors (Basel)*. 2019;19(21).

91. Cunningham L, Mason S, Nugent C, Moore G, Finlay D, Craig D. Home-based monitoring and assessment of Parkinson's disease. *IEEE Trans Inf Technol Biomed*. 2011;15(1):47-53.

92. Goetz CG, Stebbins GT, Wolff D, et al. Testing objective measures of motor impairment in early Parkinson's disease: Feasibility study of an at-home testing device. *Movement disorders : official journal of the Movement Disorder Society*. 2009;24(4):551-556.

93. Lipsmeier F, Taylor KI, Kilchenmann T, et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. *Mov Disord*. 2018;33(8):1287-1297.

94. López-Blanco R, Velasco MA, Méndez-Guerrero A, et al. Smartwatch for the analysis of rest tremor in patients with Parkinson's disease. *Journal of the Neurological Sciences*. 2019;401:37-42.

95. Memedi M, Sadikov A, Groznik V, et al. Automatic Spiral Analysis for Objective Assessment of Motor Symptoms in Parkinson's Disease. *Sensors*. 2015;15(9).

96. Powers R, Etezadi-Amoli M, Arnold EM, et al. Smartwatch inertial sensors continuously monitor real-world motor fluctuations in Parkinson's disease. *Science Translational Medicine*. 2021;13(579):eabd7865.

97. Sigcha L, Pavón I, Costa N, et al. Automatic Resting Tremor Assessment in Parkinson's Disease Using Smartwatches and Multitask Convolutional Neural Networks. *Sensors*. 2021;21(1).

98. Westin J, Ghiamati S, Memedi M, et al. A new computer method for assessing drawing impairment in Parkinson's disease. *Journal of Neuroscience Methods*. 2010;190(1):143-148.

99. Kostikis N, Hristu-Varsakelis D, Arnaoutoglou M, Kotsavasiloglou C. A Smartphone-Based Tool for Assessing Parkinsonian Hand Tremor. *IEEE J Biomed Health Inform*. 2015;19(6):1835-1842.

100. Bochniewicz EM, Emmer G, McLeod A, Barth J, Dromerick AW, Lum P. Measuring Functional Arm Movement after Stroke Using a Single Wrist-Worn Sensor and Machine Learning. *J Stroke Cerebrovasc Dis*. 2017;26(12):2880-2887.

101. Dubuisson N, Bauer A, Buckley M, et al. Validation of an environmentally-friendly and affordable cardboard 9-hole peg test. *Mult Scler Relat Disord*. 2017;17:172-176.

102. Prochazka A, Kowalczewski J. A fully automated, quantitative test of upper limb function. *J Mot Behav*. 2015;47(1):19-28.

103. Lee S, Lee YS, Kim J. Automated Evaluation of Upper-Limb Motor Function Impairment Using Fugl-Meyer Assessment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2018;26(1):125-134.

104. Kleinholdermann U, Wullstein M, Pedrosa D. Prediction of motor Unified Parkinson's Disease Rating Scale scores in patients with Parkinson's disease using surface electromyography. *Clinical Neurophysiology*. 2021.

105. de Araújo ACA, Santos EGdR, de Sá KSG, et al. Hand Resting Tremor Assessment of Healthy and Patients With Parkinson's Disease: An Exploratory Machine Learning Study. *Frontiers in Bioengineering and Biotechnology*. 2020;8(778).

106. Taylor Tavares AL, Jefferis GS, Koop M, et al. Quantitative measurements of alternating finger tapping in Parkinson's disease correlate with UPDRS motor disability and reveal the improvement in fine motor control from medication and deep brain stimulation. *Mov Disord*. 2005;20(10):1286-1298.

107. Salarian A, Russmann H, Wider C, Burkhard PR, Vingerhoets FJ, Aminian K. Quantification of tremor and bradykinesia in Parkinson's disease using a novel ambulatory monitoring system. *IEEE Trans Biomed Eng*. 2007;54(2):313-322.

108. Lin SD, Butler JE, Boswell-Ruys CL, et al. The frequency of bowel and bladder problems in multiple sclerosis and its relation to fatigue: A single centre experience. *PLoS One*. 2019;14(9):e0222731.

109. Mendoza JE, Apostolos GT, Humphreys JD, Hanna-Pladdy B, O'Bryant SE. Coin rotation task (CRT): a new test of motor dexterity. *Arch Clin Neuropsychol*. 2009;24(3):287-292.

110. Růžička E, Krupička R, Zárubová K, Rusz J, Jech R, Szabó Z. Tests of manual dexterity and speed in Parkinson's disease: Not all measure the same. *Parkinsonism & Related Disorders*. 2016;28:118-123.

111. de Groot-Driessen D, van de Sande P, van Heugten C. Speed of finger tapping as a predictor of functional outcome after unilateral stroke. *Arch Phys Med Rehabil*. 2006;87(1):40-44.

112. Technology PRCIa. Mobile Fact Sheet. https://www.pewresearch.org/internet/fact-sheet/mobile/. Published 2021. Updated April 7, 2021. Accessed May 13, 2021.

113. Bove R, White CC, Giovannoni G, et al. Evaluating more naturalistic outcome measures: A 1-year smartphone study in multiple sclerosis. *Neurol Neuroimmunol Neuroinflamm*. 2015;2(6):e162.

114. Smith D. Grasping the Importance of Our Hands. In. *InMotion*. Vol 16: Amputee Coalition of America; 2006.

115. Carmeli E, Patish H, Coleman R. The Aging Hand. *The Journals of Gerontology: Series A*. 2003;58(2):M146-M152.

116. Cadden M, Arnett P. Factors Associated with Employment Status in Individuals with Multiple Sclerosis. *International Journal of MS Care*. 2015;17(6):284-291.

117. Gopal A, Hsu W-Y, Allen DD, Bove R. Remote Assessments of Hand Function in Neurological Disorders: Systematic Review. *JMIR Rehabil Assist Technol*. 2022;9(1):e33157.

118. Kidziński Ł, Yang B, Hicks JL, Rajagopal A, Delp SL, Schwartz MH. Deep neural networks enable quantitative movement analysis using single-camera videos. *Nat Commun*. 2020;11(1):4054.

119. Sato K, Nagashima Y, Mano T, Iwata A, Toda T. Quantifying normal and parkinsonian gait features from home movies: Practical application of a deep learning-based 2D pose estimator. *PLoS One*. 2019;14(11):e0223549.

120. Ong A, Harris IS, Hamill J. The efficacy of a video-based marker-less tracking system for gait analysis. *Comput Methods Biomech Biomed Engin*. 2017;20(10):1089-1095.

121. Sandau M, Koblauch H, Moeslund TB, Aanæs H, Alkjær T, Simonsen EB. Markerless motion capture can provide reliable 3D gait kinematics in the sagittal and frontal plane. *Med Eng Phys*. 2014;36(9):1168-1175.

122. Gionfrida L, Rusli WMR, Bharath AA, Kedgley AE. Validation of two-dimensional video-based inference of finger kinematics with pose estimation. *PLoS One*. 2022;17(11):e0276799.

123. Cornman HL, Stenum J, Roemmich RT. Video-based quantification of human movement frequency using pose estimation: A pilot study. *PLoS One*. 2021;16(12):e0261450.

124. Cheok MJ, Omar Z, Jaward MH. A review of hand gesture and sign language recognition techniques. *International Journal of Machine Learning and Cybernetics*. 2019;10(1):131-153.

125. Rautaray SS, Agrawal A. Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review*. 2015;43(1):1-54.

126. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208.

127. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175-191.

128. Bae JH, Kang SH, Seo KM, Kim DK, Shin HI, Shin HE. Relationship Between Grip and Pinch Strength and Activities of Daily Living in Stroke Patients. *Ann Rehabil Med*. 2015;39(5):752-762.

129. Zackowski KM, Wang JI, McGready J, Calabresi PA, Newsome SD. Quantitative sensory and motor measures detect change over time and correlate with walking speed in individuals with multiple sclerosis. *Multiple Sclerosis and Related Disorders*. 2015;4(1):67-74.

130. Lamers I, Feys P. Assessing upper limb function in multiple sclerosis. *Mult Scler*. 2014;20(7):775-784.

131. Feys P, Lamers I, Francis G, et al. The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis. *Mult Scler*. 2017;23(5):711-720.

132. Platz T, Pinkowski C, van Wijck F, Kim IH, di Bella P, Johnson G. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multicentre study. *Clin Rehabil*. 2005;19(4):404-411.

133. Romeo AR, Rowles WM, Schleimer ES, et al. An electronic, unsupervised patient-reported Expanded Disability Status Scale for multiple sclerosis. *Mult Scler*. 2021;27(9):1432-1441.

134. Barrett LE, Cano SJ, Zajicek JP, Hobart JC. Can the ABILHAND handle manual ability in MS? *Mult Scler*. 2013;19(6):806-815.

135. Puig-Diví A, Escalona-Marfil C, Padullés-Riu JM, Busquets A, Padullés-Chando X, Marcos-Ruiz D. Validity and reliability of the Kinovea program in obtaining angles and distances using coordinates in 4 perspectives. *PLOS ONE*. 2019;14(6):e0216448.

136. Lamers I, Cattaneo D, Chen CC, Bertoni R, Van Wijmeersch B, Feys P. Associations of upper limb disability measures on different levels of the International Classification of Functioning, Disability and Health in people with multiple sclerosis. *Phys Ther*. 2015;95(1):65-75.

137. Latash ML, Levin MF, Scholz JP, Schöner G. Motor control theories and their applications. *Medicina (Kaunas)*. 2010;46(6):382-392.

138. Sumowski JF, Benedict R, Enzinger C, et al. Cognition in multiple sclerosis: State of the field and priorities for the future. *Neurology*. 2018;90(6):278-288.

139. Schnall R, Cho H, Liu J. Health Information Technology Usability Evaluation Scale (Health-ITUES) for Usability Assessment of Mobile Health Technology: Validation Study. *JMIR Mhealth Uhealth*. 2018;6(1):e4.

140. Térémetz M, Colle F, Hamdoun S, Maier MA, Lindberg PG. A novel method for the quantification of key components of manual dexterity after stroke. *J Neuroeng Rehabil*. 2015;12:64.

141. Romeo AR, Rowles WM, Schleimer ES, et al. An electronic, unsupervised patient-reported Expanded Disability Status Scale for multiple sclerosis. *Mult Scler*. 2020:1352458520968814.

142. Supratak A, Datta G, Gafson AR, Nicholas R, Guo Y, Matthews PM. Remote Monitoring in the Home Validates Clinical Gait Measures for Multiple Sclerosis. *Frontiers in Neurology*. 2018;9.

143. Block VJ, Bove R, Nourbakhsh B. The Role of Remote Monitoring in Evaluating Fatigue in Multiple Sclerosis: A Review. *Frontiers in Neurology*. 2022;13.

144. Barcellos LF, Horton M, Shao X, et al. A validation study for remote testing of cognitive function in multiple sclerosis. *Mult Scler*. 2021;27(5):795-798.

145. Ross J, Stevenson F, Dack C, et al. Developing an implementation strategy for a digital health intervention: an example in routine healthcare. *BMC Health Services Research*. 2018;18(1):794.

146. Wienert J, Zeeb H. Implementing Health Apps for Digital Public Health – An Implementation Science Approach Adopting the Consolidated Framework for Implementation Research. *Frontiers in Public Health*. 2021;9.

147. Ellermeyer T, Otte K, Heinrich F, et al. Ranking of Dystonia Severity by Pairwise Video Comparison. *Mov Disord Clin Pract*. 2016;3(6):587-595.

148. Smith VM, Varsanik JS, Walker RA, et al. Movement measurements at home for multiple sclerosis: walking speed measured by a novel ambient measurement system. *Mult Scler J Exp Transl Clin*. 2018;4(1):2055217317753465.

149. Kadamba V. Face and Hand Landmarks Detection using Python – Mediapipe, OpenCV. https://www.geeksforgeeks.org/face-and-hand-landmarks-detection-using-python-mediapipe-opencv/. Published 2023. Updated January 4, 2023. Accessed January 4, 2023.

150. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-381.

151. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*. 2007;19(6):349-357.

152. Lang CE, Holleran CL, Strube MJ, et al. Improvement in the Capacity for Activity Versus Improvement in Performance of Activity in Daily Life During Outpatient Rehabilitation. *Journal of Neurologic Physical Therapy*. 2023;47(1).

153. Moore JL, Potter K, Blankshain K, Kaplan SL, O'Dwyer LC, Sullivan JE. A Core Set of Outcome Measures for Adults With Neurologic Conditions Undergoing Rehabilitation: A CLINICAL PRACTICE GUIDELINE. *J Neurol Phys Ther*. 2018;42(3):174-220.

154. Jiang J, Skalli W, Siadat A, Gajny L. Effect of Face Blurring on Human Pose Estimation: Ensuring Subject Privacy for Medical and Occupational Health Applications. *Sensors*. 2022;22(23):9376.

155. Javed K, Reddy V, Lui F. Neuroanatomy, Lateral Corticospinal Tract. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing
Copyright © 2022, StatPearls Publishing LLC.; 2022.

156. García-Hermoso A, Cavero-Redondo I, Ramírez-Vélez R, et al. Muscular Strength as a Predictor of All-Cause Mortality in an Apparently Healthy Population: A Systematic Review and Meta-Analysis of Data From Approximately 2 Million Men and Women. *Archives of Physical Medicine and Rehabilitation*. 2018;99(10):2100-2113.e2105.

157. Bhandari T. Restoring arm, hand function after spinal cord injury focus of clinical trial. *Washington University School of Medicine St Louis*. 2019. Published December 12, 2019. Accessed March 29, 2022.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Arpita Gopal
64FAAD225CB54F4...
Author Signature

5/11/2023
Date