A Rapid Tapping Task on Commodity Smartphones to Assess Motor Fatigability

Liliana Barrios[1], **Pietro Oldrati**[1], **David Lindlbauer**[1], **Marc Hilty**[2], **Helen Hayward-Koennecke**[2],
Christian Holz[1], **Andreas Lutterotti**[2]
[1] Department of Computer Science, ETH Zurich, Switzerland
[2] University of Zurich & University Hospital Zurich, Switzerland

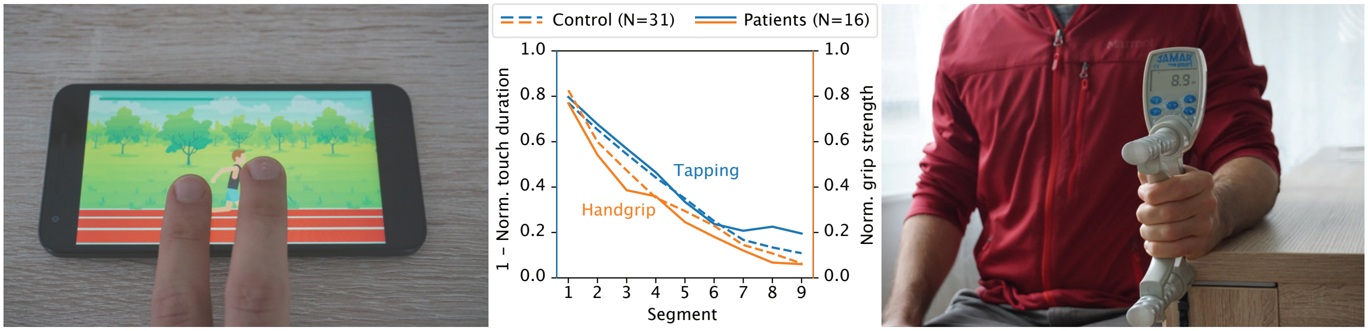

Figure 1. Multiple sclerosis patients typically monitor the progression of their condition using specialized hardware to measure motor fatigue such as handgrip dynamometers (*right*). To allow such assessment to become more widely available and be performed more frequently, we propose a smartphone-based method to measure fatigability and assess motor fatigue (*left*). We show that our rapid alternating tapping task, over the first 30 seconds of a trial (*center*), is strongly correlated with a standard hand dynamometer for patients ($\rho = 0.78$) and the control group ($\rho = 0.84$).

ABSTRACT

Fatigue is a common debilitating symptom of many autoimmune diseases, including multiple sclerosis. It negatively impacts patients' every-day life and productivity. Despite its prevalence, fatigue is still poorly understood. Its subjective nature makes quantification challenging and it is mainly assessed by questionnaires, which capture the magnitude of fatigue insufficiently. Motor *fatigability*, the objective decline of performance during a motor task, is an underrated aspect in this regard. Currently, motor fatigability is assessed using a handgrip dynamometer. This approach has been proven valid and accurate but requires special equipment and trained personnel. We propose a technique to objectively quantify motor fatigability using a commodity smartphone. The method comprises a simple exertion task requiring rapid alternating tapping. Our study with 20 multiple sclerosis patients and 35 healthy participants showed a correlation of $\rho = 0.8$ with the baseline handgrip method. This smartphone-based approach is a first step towards ubiquitous, more frequent, and remote monitoring of fatigability and disease progression.

Author Keywords

fatigability; smartphones; mobile health

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04 ...$15.00.
http://dx.doi.org/10.1145/3313831.3376588

CCS Concepts

•**Human-centered computing** → **Mobile devices;** *User studies;* Touch screens;

INTRODUCTION

Mobile health technologies, including wearable sensors and smartphones, provide great potential to improve the understanding of physiologic parameters in health and disease. Particularly in chronic diseases, like multiple sclerosis (MS), the widespread availability of smartphones, and the ability to collect continuous data are a unique opportunity to improve patient care. MS is an autoimmune disease characterized by recurrent areas of inflammation in the central nervous system [31], comprising the brain, spinal cord, and optic nerves. With more than 2 million patients worldwide [43], MS is one of the leading causes of neurological disability in young adults.

Fatigue is a common symptom in MS and 75–95% of patients have reported fatigue at some point [1, 17, 22, 27]. Fatigue has been defined as "a subjective lack of physical and/or mental energy that is perceived by individuals or caregivers to interfere with the usual and desired activities" [15]. MS fatigue has a significant impact on the quality of life patients, and it affects work performance as well as personal interactions [17]. To this day, fatigue in MS patients is not well understood despite being highly prevalent and being a debilitating symptom [37]. Development of therapies is also hampered by the lack of objective parameters to assess fatigue.

Recently, researchers have started to investigate the concept of *fatigability* and its relation to fatigue [21, 51]. Motor fatigability is defined as "an objective decline in strength as routine use of muscle groups proceeds" [9]. Bruce et al. [9] argued that fatigability can redefine our understanding of fatigue, because many symptoms of fatigue may be a consequence of demonstrable fatigability, but this has rarely been assessed.

Current clinically-used methods only evaluate fatigue and fatigability retrospectively using questionnaires like the Fatigue Severity Scale (FSS). Although objective measurements to quantify fatigability have been proposed, none is sufficiently researched to be established in clinical routine [41]. Tests using isokinetic dynamometers that measure peak isometric torque [18] on the knee or hand, for example, or measures of electrically-induced torque have been proposed [42]. Those devices, however, are expensive and bulky, and typically require supervision by professionals to properly perform tests [9]. Finding ubiquitous and inexpensive ways to assess fatigability would enable optimized treatment options that currently lack objective outcome parameters to prove their efficacy. Furthermore, regular assessment of fatigability in clinical routine would not only allow evaluating disease progression, but also add to the so-far limited options for quality of life measures. Therefore, finding ubiquitous and inexpensive ways to measure fatigability would be beneficial for a better understanding of fatigue and to guide therapeutic interventions.

We propose a commodity approach to monitor fatigability as a proxy metric to continuously assess disease progression. Our approach utilizes the prevalence of smartphones in conjunction with a simple tapping task, designed as an exertion technique to assess the user's motor fatigability. Finger tapping is commonly used to assess motor impairment [34, 47, 35]. In our work, we re-purpose this task to quantify motor fatigability. The rapid tapping task requires barely any instructions (other than "please tap as fast as possible"), and can be performed on any commercially available smartphone. In an experiment with 20 multiples sclerosis patients and 35 healthy participants as control group, we compare our approach with a standard fatigability assessment done with a handgrip dynamometer. Participants performed 500 alternating taps, which on average took patients roughly 2 minutes and healthy participants 75 seconds to complete. We show that participants' performance decreases during the tapping task, and correlates ($\rho = 0.8$) with the decrease in grip strength measured with the handgrip dynamometer for patients and control. We further show that this correlation is also present in the first 30 seconds of performing the tapping task with $\rho = 0.78$ for patients and $\rho = 0.84$ for controls. This suggests that performing the simple tapping task for 30 seconds is sufficient to measure motor fatigability.

Contributions

- We present an exertion technique on commodity devices that involves simple alternating rapid tapping. The task is fast, easy to implement, and can potentially be performed everywhere and anytime.
- A metric to assess such tapping tasks to represent motor fatigability: the increase of time that a user keeps a finger on the screen. We detail the specifics of the metric and our processing pipeline to extract a metric for motor fatigability from the data collected through the mobile app.
- An evaluation of our approach on 20 multiple sclerosis patients and 35 healthy participants as control group. We show that our proposed task strongly correlates with data collected from a standard handgrip dynamometer, and that performing the task for 30 seconds is sufficient.

RELATED WORK

Fatigue and fatigability

Fatigue is a common symptom of many neurological and autoimmune diseases [23]. Typically, it is measured using questionnaires [3, 24], such as the Fatigue Severity Scale [23], Fatigue Impact Scale [11], and Modified Fatigue Impact Scale [48]. Assessing fatigue and especially the shortcomings of such means has been discussed in several studies [38, 21, 6]. Schwid et al. [38] state that current fatigue-assessment methods rely on self-reporting questionnaires that are subjective, can be confounded by other symptoms, and require assessments in retrospect, which can be difficult. Kluger et al. [21] pointed out the lack of agreed-upon definition of the term fatigue in many of those studies. They state that progress in understanding fatigue is hampered by the lack of a unified taxonomy and assessment methods. We follow Kluger et al.'s definition of fatigue as the subjective sensations of weariness and increased sense of effort. Kim et al. [20] proposed a real-time digital fatigue score (RDFS) to overcome the retrospective assessment introduced by questionnaires by actively querying patients four times a day through notifications.

Researchers have suggested that fatigue is associated with *fatigability* [9, 44, 29, 49]. Wolkorte et al. [49] have highlighted the importance of including fatigability in the models to explain perceived fatigue in patients with MS. Kluger et al. define fatigability as "the magnitude or rate of change in a performance criterion relative to a reference value over a given time of task performance" [21]. While there is no established methodology to measure fatigability, it has been assessed through walking (e. g., a 6-minute walking test [14]), handgrip strength [40], and a knee dynamometer [45]. Most studies have applied maximal voluntary contractions (MVCs) within a given time limit to assess motor fatigability [41, 40, 39, 8, 44, 7], requiring patients to exert pressure on a handgrip over time. All of these assessment tasks require a special-purpose measurement apparatus and personnel for conducting observations [9, 41]. We believe finding ubiquitous and inexpensive ways to measure fatigability could be beneficial to understand fatigue in the broader population that may not have access to such assessments.

Finger tapping and motor impairment

Finger tapping is a commonly used clinical test to evaluate disease progression in Parkinson' Disease (PD) [34, 47, 35]. Prince et al. [34] quantify PD related disability with an alternating finger tap on a smartphone screen for 20 seconds (counting total number of taps). Tavares et al. [47] use a repetitive alternating finger-tapping (RAFT) task over 30 seconds on a physical keyboard to quantify motor impairment. Similarly, Lou et al. [28] use alternately pressing two spaced out piano keys to measure fatigue in PD patients. This Fitts' law-style task on a real piano does not provide the benefits of our more ubiquitous approaches. The finger tapping tasks used in PD patients assesses dysfunction of the *extrapyramidal motor* *system*, which leads to impairment in maintaining alternating movements. The tapping test thus assesses PD related impairment, but is not suitable to assess motor fatigability given the PD-specific confounding. Differences on a computerized single finger tapping in relation to gender, hand dominance and age are examined by Hubel et al. [16]. Their results suggest that the task can be used a diagnostic tool and that changes in tapping rate overtime can be due to fatigue or other factors. Notermans et al. [32] use a finger tapping test to measure ataxia. Perhaps closest to our work is the work of Boukhvalova et al. [5], who hypothesized that tapping with the index finger on a smartphone may allow to measure motor fatigability. However, they only used their approach to differentiate MS patients from healthy participants, and did not compare their task to the commonly used baseline of motor fatigability (hand dynamometer). Similarly, Tanigawa et al. [46] observed motor fatigue during fast tapping with the index finger on a custom button.

Smartphone-based health monitoring

Smartphones have been popular for monitoring chronic conditions due to their ubiquity. Much of the previous work has focused on self-reporting apps to track the development and manage these conditions (e. g., El-Gayar et al. [10] and Preuveneers et al. [33] for Diabetes, Lakshminarayana et al. [26] for Parkinson's disease). MS patients have described their interest in the use of mobile apps for tracking their condition. Ayobi et al. [2] described that when individuals faced the unpredictable and degenerative nature of MS, they regained a sense of control by intertwining self-care practices with different self-tracking technologies. Giunti et al. [13] presented a systematic review of MS health applications and could only find a small number of MS-specific applications compared to other equally prevalent diseases. Moreover, large pharmaceuticals companies have as well showed their interest on the use of mobile device for tracking MS. Genentech, Inc. a member of Roche group has launch the application Floodlight, with the aim of monitoring MS symptoms and health over time using a smartphone [12]. Similarly, Biogen launched their mobile application called Aby [4] that offers a variety of tools and resources to support patients living with MS, such as informational videos and self-reporting diaries. Psychomotor vigilance testing (PVT) has been suggested as potential standardized assessment tool for an important aspects of MS-related fatigue et al. [36]. Kay et al. [19] introduced and validated PVT-Touch, a smartphone based version of PVT. PVT is an alertness test and thus not within the scope of motor fatigability. Our short motor task is independent of visual stimuli, hence not influenced by reaction time.

ASSESSING FATIGABILITY THROUGH RAPID TAPPING

We aim to specify a task that can accurately quantify fatigability and meet a set of requirements specified by healthcare professionals (i.e., neurologists and neuropsychologists). The task should 1) be exhausting in terms of motor fatigue; however, it should not strain users' muscles for a prolonged time (i. e., enable quick recovery). The task should be 2) easy to learn and simple enough that it can be performed without an experimenter present, and 3) not require any specialized equipment to enable future in-the wild studies. Lastly, it should 4) not be prone to the speed-accuracy trade-off, as described by Zhai et al. [50]. For classical Fitts' law tasks such as pointing, users typically perform a task with high accuracy but slow (i. e., low exertion), or fast but with low accuracy. For many tasks, this means participants need to be well instructed to disregard errors and solely focus on speed. Even with clear instructions, however, participants might not completely disregard errors, thus might not perform the task as fast as possible. Since motor fatigability is measured when participants perform an exhausting task, other tasks such as pointing that are subject to this trade-off would not be well suited. We thus resorted to *rapid alternating tapping* as the task.

The simple user interface of the task is shown in Figure 1 (*left*). Users perform alternating taps to complete the task. The avatar moves forward as users perform the task with speed depending on the tapping speed. A progress bar on the top indicates completion. We do not display any indication of accuracy to avoid the speed-accuracy trade-off. During our preliminary tests, we found that a goal of 500 alternating taps suffices to measure motor fatigability, as described in the Method section. We implemented the task on a commodity smartphone (Nexus 5X), however porting it to other devices and operating systems would be trivial. To measure motor fatigability, we only require that the API has a measure for touch duration or time between taps.

METHOD

To analyze the validity of our exertion task as an indicator of motor fatigability, we compared our proposed tapping task with a standard handgrip dynamometer task, performed by a control group and patients with MS. Each participant performed both tasks with their dominant and non-dominant hand. The study was approved by the ethics review board of the local University. MS patients performed the experiment at a local hospital under supervision of healthcare professionals.

Participants

We recruited 35 participants as control group (14 female, 21 male), ages 20–55 ($\mu = 31$, $\sigma = 7.7$), all staff or students from a local university, and 20 MS patients (11 female, 9 male), aged 20–62 ($\mu = 43.1$, $\sigma = 12.9$). The inclusion criteria for the control group included: no known or suspicions of illness, autoimmune disease, fatigue or depression, based on self-reports. MS patients were included if they had a confirmed diagnosis. The Expanded Disability Status Scale (EDSS) [25] scores ranged from 0 to 8 ($\mu = 3$, $\sigma = 2.5$).

Apparatus

The control group performed the study in a calm experimental room with a chair with arm rest and a desk, shown in Figure 2. Patients performed the study in a room in an examination room at the local hospital, also with a chair with arm rest and a desk. Maximal voluntary handgrip contraction (MCV) was recorded using a digital Jamar handgrip dynamometer, configured using the Jamar iOS tablet application. We used a Nexus 5X to run our fatigability application.

Figure 2. Apparatus of our experiment for the handgrip task (*left*) and tapping task (*right*).

Design

We used a within-subject design with *Task* and *Hand* as independent variable with two levels each: *Handgrip* and *Tapping*; and *dominant* and *non-dominant* hand respectively. Order of *Task* and *Hand* was alternated, starting order was counterbalanced. Between each task, there was a resting period of three minutes to allow participant's muscles to recover.

Tasks

Handgrip

We used the handgrip dynamometer in the standard procedure to assess motor fatigability (cf. [41]). Participants were asked to sit upright, with both feet touching the ground and their forearms resting at a 90° angle on the armrests of the chair or the desk (see Figure 2). The dynamometer was held with the thumbs facing upwards in line with the forearm and the grip size was adjusted to comfort. After a short period of familiarization, participants were asked to perform the MVC task for 30 seconds.

Tapping

Participants performed a rapid alternating tap on the smartphone screen while the hand was resting on a desk. The smartphone was placed on landscape mode. The exertion movement was performance with the index and middle finger. *Participants were asked to perform the tapping task as fast as possible without stopping until the app indicated completion.* An initial assessment with six healthy participants performing 1500 alternating taps (naive about the end of study) showed a clear decrease in performance after 500 taps. Hence, for the final experiment, participants were asked to perform 500 valid alternating taps, i. e., 250 taps per finger. A tap was considered valid and counted towards the goal of 500 taps if exactly one finger was on the display.

Hypotheses

We performed the experiment with respect to the following hypotheses. First, we expected a decrease in grip strength when using the handgrip dynamometer, as reported by previous studies (cf. [41]). Secondly, we expected a decrease in performance as time progresses during the tapping task if the task is performed at maximal effort (speed). That is, users will take longer to alternate the fingers correctly on the screen, and the touch duration of each tap will increase with time. We analyze our data with respect to these hypothesis and also analyze the connection between the handgrip and the tapping task, to quantify and attribute rapid tapping to motor fatigability.

Procedure

Participants were briefly introduced to the setup and the experiment, and completed a demographic questionnaire. They then completed a short training session for the tapping task where they performed 40 alternating taps. Subsequently, they received instructions on how to use the handgrip dynamometer, including a demonstration by the experimenter. During the handgrip task, the experimenter instructed participants when to start and stop the MVC. After the introduction, participants completed all tasks with their dominant and non-dominant hand with counterbalanced order. Between tasks, they were asked to rest the arm and hand for three minutes.

Data collection

During each 30-second trial of the handgrip dynamometer, we collected 10 samples, which is the maximum sampling rate of the device we used. For the tapping task, we collected touch data of the smartphone using the Android API. We stored all timestamped touch down coordinates and up events, from which we compute touch duration (i. e., how long did the finger touch the screen). Each sample in our dataset contains the finger position on the screen, touch duration, area size and pressure. *We define task performance for the tapping task as the average time participant's finger stayed on the display (i. e., average touch duration).* This means that for fast tapping (high performance), touch duration will be low, whereas for slow taps (low performance), touch duration increases.

RESULTS

In summary, our results show that the performance in the rapid tapping task correlates strongly with the fatigability measurements of the handgrip dynamometer. Tapping performance decreased significantly over the course of a full trial (500 taps). For both groups, performing the tapping task for 30 seconds is sufficient to reliably measure motor fatigability.

Data processing

We use touch duration as the primary performance metric to assess the tapping task. To account for outliers and noise in the tapping data, we performed the following data processing steps for each trial separately. We removed samples with a touch duration of more than three standard deviations away from the mean (1.2% for patients, 0.9% for healthy control). These outliers occurred when participants did not alternately lift fingers but instead left one in contact with the screen and tapped only with the other. Outliers are evident, and thus we classified and removed them. Tapping duration was low-pass filtered with a moving average of 20% of the trial data and normalized per participant and trial. We normalized tapping trials and handgrip trials separately, resulting in two motor fatigability slopes. The fatigability slope of the tapping task is positive (as duration increases), whereas that of the gripper is negative (as force decreases). To make trends comparable, our results are computed as *1 - normalized touch duration*.

To compare participants' performance in the handgrip task (10 samples per trial) and the tapping task (500 taps), we split the measurements of touch duration over time in 10 segments. Each segment contains samples from 10% of the total duration of the task. The final value for each segment is defined as the

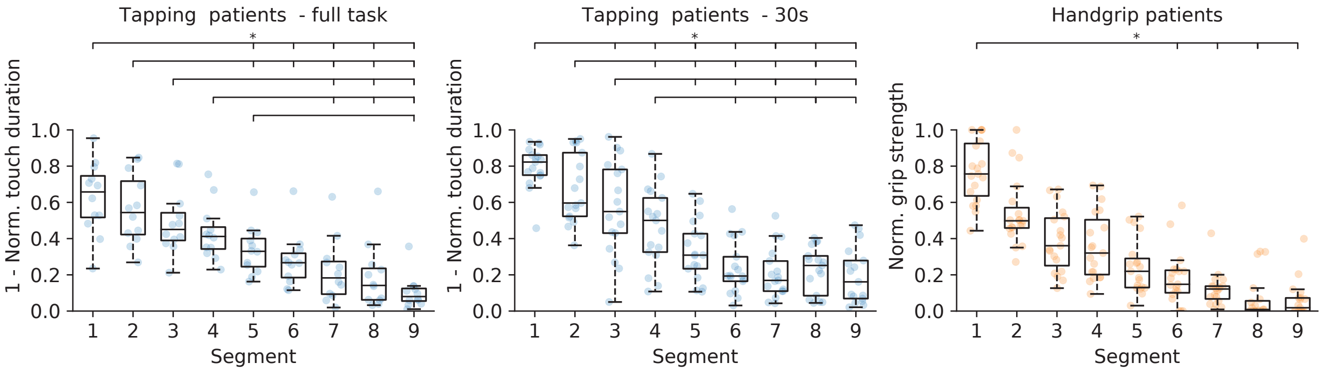

Figure 3. Patient group data: full tapping task (*left*), the first 30 seconds (*center*), and the handgrip task (*right*).

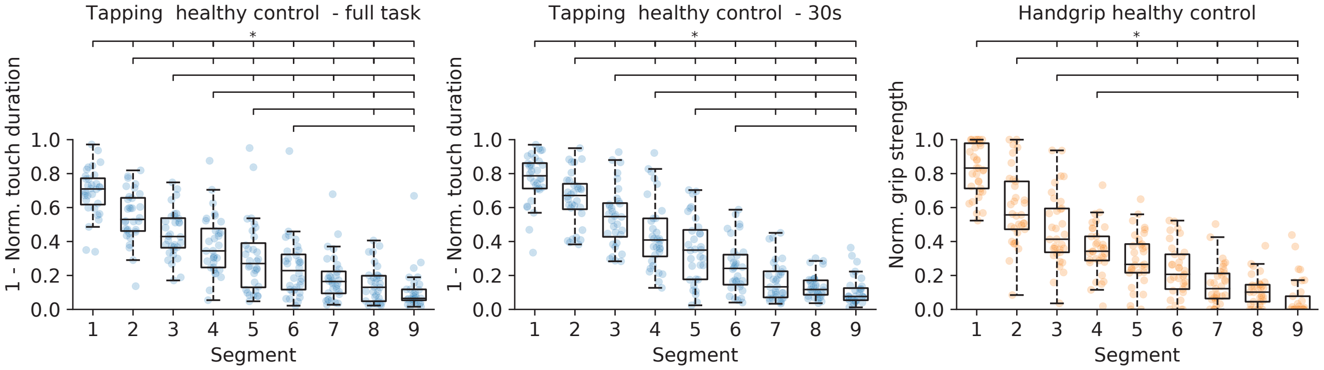

Figure 4. Control group data: full tapping task (*left*), the first 30 seconds (*center*), and the handgrip task (*right*).

mean value of the data in that segment. To account for inertia when participants start both the handgrip and the tapping task, we discard the first segment, and perform our analysis on the remaining 9 segments.

Dominant vs. non-dominant hand

We analyzed the data from both tasks for participants' dominant and non-dominant hand. For the tapping task using the non-dominant hand, the data showed large variability. In contrast to the dominant hand, the decrease in performance, while present, was less pronounced for the non-dominant hand. While data cleaning and statistical analysis as described in section *Performance results* yielded a main effect for segment ($F_{8,117} = 10.592, p < .001$), Bonferroni adjusted Tukey's post-hoc tests showed less statistically significant differences between segments as for the dominant hand. From observation, we believe this is due to challenges in coordinating the two fingers when performing the task. Since participants struggled to perform the task reliably, their speed decreased less. We therefore believe that the tapping task should be performed with the dominant hand. *We thus performed all following analyses on the data collected from dominant hand trials.*

Valid trials

For a small number of trials, both groups (control and patients) did not follow instructions when performing the handgrip and tapping task. The MVC task requires participants to evoke maximal potential from the beginning of the task. If participants successfully activate their muscles maximally, no further increase in force is evoked during the task [44]. Similarly to Steens et al. [44], we conservatively removed trials where participants failed at achieving maximal performance. For the 30-second handgrip task, we discarded the trials where consecutive measurements increased more than 50% of the maximum strength. This happened in 2 of 35 trials for the control group, and in 1 of 20 trials for patients.

To validate the tapping task, we fit a linear regression to the segment values and define trials as valid if 1) the slope of the regression is positive (i. e., touch duration increases), verifying an overall decrease in performance; and 2) consecutive segments do not have a duration decrease of more than 50% (i. e., participants' performance increases). Trials that fail these requirements suggest that participants did not perform the task as fast as possible, meaning they did not evoke maximal performance. The number of discarded trials for this condition depends on the analyzed time frame (Figure 7, *bottom*).

Analyzing the fully completed task, 63% and 90% of trials were valid for patients and the control group, respectively. Restricting the window to the first 30 seconds of the task, however, resulted in 84% valid trials for patients and 93% valid trials for the control group. For shorter durations (e. g., 10 seconds), not enough data is available to accurately quantify fatigability. For durations of 60 seconds or longer, participants seem to pace themselves, recover, and then speed up again. We thus believe that the first 30 seconds of the tapping task represents a suitable excerpt to assess motor fatigability.

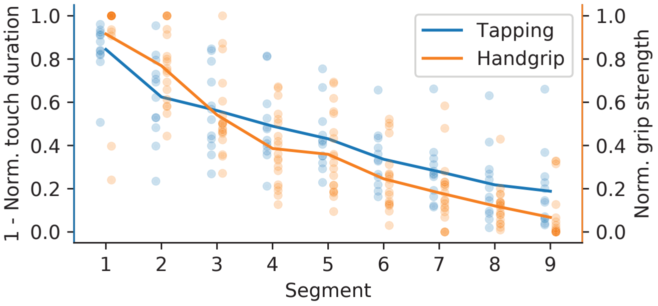

Figure 5. Complete task recordings of the patients group for handgrip and tapping. The solid line indicates each segment's mean value.

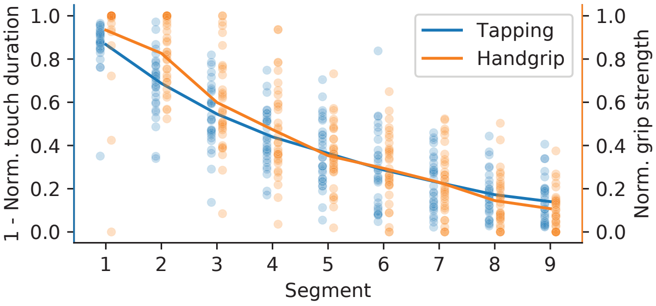

Figure 6. Complete task recordings of the control group for handgrip and tapping. The solid line indicates each segment's mean value.

Performance results

We performed individual ANOVAs on the handgrip and tapping data with *segment* as independent variable (9 levels) for both the control group and patients. Statistically significant differences between segments demonstrate an actual, non-random decline in performance during a task. For the tapping task, we found a main effect of *segment* on average duration for the control group $F_{8,279} = 50.918,\ p < .001$ and for patients $F_{8,99} = 14.211,\ p < .001$. To analyze the temporal progression of participants' performance, we performed a series of Bonferroni-corrected Tukey's post-hoc tests. Results are illustrated in Figure 3 for patients and Figure 4 for the control group. For the control group, segments are mostly significantly different from segments after the subsequent one. Only after Segment 7, average performance flattens and subsequent segments are no longer significantly different. Results for patients show a similar pattern, however most segments are not significantly different from their direct successor, but 2 or 3 segments thereafter. Performance flattens after Segment 6. This decline in performance indicates that both tasks can successfully invoke motor fatigue for both the control group and patients.

Handgrip vs. tapping

To verify the relation between the handgrip and tapping task, we computed their Spearman's rank correlation coefficient ρ. Figure 5 illustrates the patient data and Figure 6 shows the measurements of the control group. For both, the average correlation coefficient is $\rho = 0.8$.

Trial duration

Each trial took on average of 75 seconds for the control group (σ = 23.6 s) and 126.1 seconds for patients (σ = 81.4 s). To determine the optimal number of taps per trial leading to comparable results, we performed the same analysis as before on the first 10, 30, 60 and 90 seconds of the recordings data as shown in Figure 7.

During the first 30 seconds, participants of the control group performed on average 249.1 taps (σ = 56.6 taps), while patients performed on average 170.6 taps (σ = 60.3 taps, leaving a large number of data points for analysis. We performed similar processing on the data (discarding outliers and invalid trials), but only used the first 30 seconds, and split them into 10 segments.

We performed the same analysis as with the full task duration. That is, two ANOVAs with *segment* as independent variable and touch duration as dependent variable, one for the control group and one for the patients. We again found a main effect of *segment* on the data of the control group $F_{8,279} = 79.606,\ p < .001$ and the patients $F_{8,144} = 28.39,\ p < .001$. A series of Tukey's post-hoc tests revealed a similar pattern between segments as with the full data (Figure 3 and Figure 4). The statistically significant differences between the segments of the trials confirm that the non-random decline in performance is also present during a short task of 30s. Analyzing the correlation between the handgrip and this shortened tapping task revealed a correlation between the two tasks of $\rho = 0.78$ ($p < .001$) for patients and $\rho = 0.84$ ($p < .001$) for the control group. The similar correlation score indicates that a rapid tapping task of 30 seconds suffices to measure motor fatigue.

Finally, we analyze the agreement of our tapping task and handgrip dynamometer by comparing the rate of fatigue development captured by each method. Similarly to Lou et al. [28], we use the slopes of the regression line of touch duration and handgrip strengths to assess fatigue rate. Figure 8 shows the

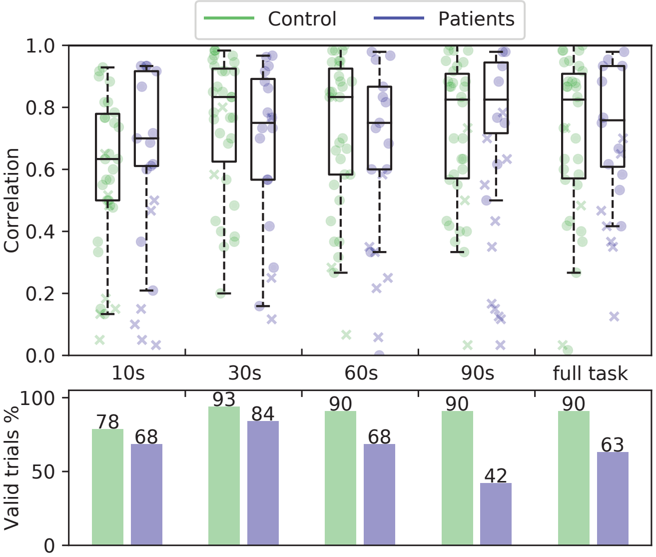

Figure 7. (Top) Spearman's correlation between *1 - normalized touch duration* and handgrip (*top*) by task duration. Crosses represent invalid trials. (Bottom) The bar chart shows the percentage of valid trials.

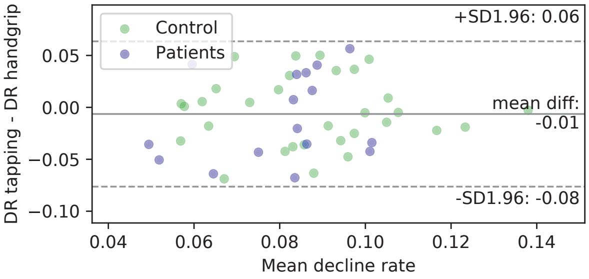

Figure 8. Bland-Altman plot for mean decline rate (DR) of normalized touch duration (30 sec) and normalized handgrip strength shows a mean bias of -0.01 with LoA [0.06,-0.08].

Bland-Altman plot comparing the decline rate of the tapping task and the handgrip dynamometer. The plot shows no particular pattern on the data. The mean difference of almost zero (0.01) and all data within two within two standard deviations from the mean with limits of agreement between (LoA) [-0.08, 0.06] confirms the good agreement between both approaches. Normality of the differences was verified using the Shapiro-Wilk test ($p = .08$)

DISCUSSION

The analysis of our evaluation showed that our simple rapid tapping task can be used to quantify motor fatigability. The task is easy to implement, runs on unmodified commodity smartphones, and, more importantly, the task is easy to perform for users. We thus believe that our method will allow moving beyond specialized hardware (e. g., handgrip dynamometers) and subjective feedback to make assessing motor fatigability ubiquitous and more accessible for all patients.

Through our experiment, we examined the appropriate task length to quantify motor fatigability using a tapping task. Our initial target was 500 taps and resulted in varying completion tasks for participants. On average, patients took 168% longer than healthy controls to complete the task, which was not unexpected. The slowest patient completed the taps in 7.68 minutes, while the slowest healthy participant took 2.29 minutes. The variance in completion times shows the importance of limiting trial durations, because performing the tapping task for up to 7 minutes does not only cause physical strain, but also makes motivated compliance challenging. The results of our evaluation shows that analyzing the first 30 seconds of our rapid tapping task is a suitable assessment of motor fatigability, which may be important to enable more frequent and ideally continual monitoring in a straightforward manner.

Time to complete the tapping task for patients was partly governed by the severity of their condition as measured by the EDSS scores. Patients with higher EDSS scores tended to take longer. We did not, however, observe differences between patients in terms of measured motor fatigability with either task. This, however, needs to be investigated further, since we did not have enough patients for each score to perform reliable statistical analysis on this data. Preliminary results with three groups of MS disability based on the EDSS show these Spearman correlations: EDSS = 0, N = 4: .75, EDSS in [1,3], N = 4: .85, EDSS in [4,8], N = 8: .77; all $p < .001$

Other metrics for tapping performance

We initially investigated the suitability of alternative metrics to evaluate tapping performance, such as number of taps per segment, tap pressure and area size. The average duration between taps was noisier due to the occurrence of simultaneous or quasi-simultaneous taps (which we counted as invalid). We, therefore, decided to use touch duration as primary performance metric. The number of taps shows slightly lower correlations than touch duration and a comparable number of invalid trials, hence it would also be suitable metric. Analyzing pressure and area size, we found low correlation with the handgrip measurements, possibly because participants' finger placement on the touchscreen is too person-specific.

LIMITATIONS AND FUTURE WORK

Even though participants performed a training session before measurements were taken, some still did not start their trials with maximum speed. Instead of exhibiting fatigue, some participants showed an increase in performance, which resulted in a limited number of invalid trials (16% for patients, 7% for the control group). This indicates that while the task is generally well suited to measure fatigability, further interventions are needed to ensure that participants follow instructions closely. We see potential in offering incentives to complete the tapping task with maximum effort, such as by further gamification or the use of scoring systems. We believe that such measures would decrease the number of outliers, and potentially eliminate the need for dedicated outlier removal. The percentage of trials that are invalid and how this might vary under different environments and without supervision needs further investigation. We plan to explore other methods and analysis strategies to ensure higher rates of valid trials. Additionally, intrinsic motivation is needed to perform the tapping task on a regular basis, and we cannot derive an estimate of potential learning effects so far. Extending our research to longitudinal in-the-wild evaluations with within-subject comparisons will allow us to assess the use of fatigability to judge disease progression in MS populations. Moreover, comparative tests are needed to discriminate between fatigability and disability. We plan on using the 9-Hole Peg Test [30] to assess fine motor skills of patients and use the fatigue scale for motor and cognitive functions (FSMC) to categorize mild and severe fatigue in patients. Cognitive fatigability measure (e. g., the N-Back test) can help discriminate cognitive and motor fatigability.

CONCLUSION

We introduced a novel approach to assess motor fatigability on a commodity smartphone using a simple rapid tapping tasks. Our experiment with 20 patients with multiple sclerosis and 35 healthy participants showed a significant correlation between the tapping tasks and grip strength measurements from a special-purpose handgrip dynamometer. We believe that our work is a first step towards measuring motor fatigability without having to rely on specialized equipment, which can be expensive and require professional supervision. We also think that our method may help quantify fatigue and complement the current use of subjective feedback through questionnaires, enabling patients to frequently and ubiquitously monitor their condition, and react to changes accordingly.

REFERENCES

[1] Lerdal Anners, Celius G. Elisabeth, Krupp B. Lauren, and Dahl A. Alv. 2007. A prospective study of patterns of fatigue in multiple sclerosis. *European Journal of Neurology* 14 (2007), 1338–1343. DOI: http://dx.doi.org/10.1111/j.1468-1331.2007.01974.x

[2] Amid Ayobi, Paul Marshall, Anna L. Cox, and Yunan Chen. 2017. Quantifying the Body and Caring for the Mind: Self-Tracking in Multiple Sclerosis. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 6889–6901. DOI: http://dx.doi.org/10.1145/3025453.3025869

[3] Rohit Bakshi. 2003. Fatigue associated with multiple sclerosis: Diagnosis, impact and management. *Multiple Sclerosis* 9, 3 (2003), 219–227. DOI: http://dx.doi.org/10.1191/1352458503ms904oa

[4] Biogen. 2019. Above MS - Meet Aby. (2019). https://www.abovems.com/en_us/home/ms-support-events/aby-app.html

[5] Alexandra K. Boukhvalova, Emily Kowalczyk, Thomas Harris, Peter Kosa, Alison Wichman, Mary A. Sandford, Atif Memon, and Bibiana Bielekova. 2018. Identifying and Quantifying Neurological Disability via Smartphone. *Frontiers in Neurology* 9 (2018), 740. DOI: http://dx.doi.org/10.3389/fneur.2018.00740

[6] Tiffany J. Braley and Ronald D. Chervin. 2010. Fatigue in Multiple Sclerosis: Mechanisms, Evaluation, and Treatment. *Sleep* 33, 8 (2010), 1061–1067. DOI: http://dx.doi.org/10.1093/sleep/33.8.1061

[7] David H. Clarke. 1986. Sex Differences in Strength and Fatigability. *Research Quarterly for Exercise and Sport* 57, 2 (1986), 144–149. DOI: http://dx.doi.org/10.1080/02701367.1986.10762190

[8] Ruth Djaldetti, Ilan Ziv, Anat Achiron, and Eldad Melamed. 1996. Fatigue in multiple sclerosis compared with chronic fatigue syndrome. *Neurology* 46, 3 (1996), 632–635. DOI:http://dx.doi.org/10.1212/WNL.46.3.632

[9] Bruce H. Dobkin. 2008. Fatigue Versus Activity-Dependent Fatigability in Patients With Central or Peripheral Motor Impairments. *Neurorehabilitation and Neural Repair* 22, 2 (2008), 105–110. DOI: http://dx.doi.org/10.1177/1545968308315046

[10] Omar El-Gayar, Prem Timsina, Nevine Nawar, and Wael Eid. 2013. Mobile Applications for Diabetes Self-Management: Status and Potential. *Journal of Diabetes Science and Technology* 7, 1 (2013), 247–262. DOI:http://dx.doi.org/10.1177/193229681300700130

[11] John D. Fisk, Paul G. Ritvo, Lynn Ross, David A. Haase, Thomas J. Marrie, and Walter F. Schlech. 1994. Measuring the Functional Impact of Fatigue: Initial Validation of the Fatigue Impact Scale. *Clinical Infectious Diseases* 18, Supplement 1 (1994), S79–S83. DOI: http://dx.doi.org/10.1093/clinids/18.supplement_1.s79

[12] Inc. Genentech. 2019. Floodlight Open. (2019). https://floodlightopen.com/en-US/

[13] Guido Giunti, Estefania Guisado-Fernandez, and Brian Caulfield. 2017. Connected Health in Multiple Sclerosis: A Mobile Applications Review. In *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*. 660–665. DOI: http://dx.doi.org/10.1109/CBMS.2017.27

[14] Myla D. Goldman, Ruth Ann Marrie, , and Jeffrey A. Cohen. 2008. Evaluation of the six-minute walk in multiple sclerosis subjects and healthy controls. *Multiple Sclerosis Journal* 14, 3 (2008), 383–390. DOI: http://dx.doi.org/10.1177/1352458507082607

[15] Multiple Sclerosis Clinical Practice Guidelines. 1998. *Fatigue and multiple sclerosis: evidence-based management strategies for fatigue in multiple sclerosis*. Paralyzed Veterans of America.

[16] Kerry A. Hubel, Bruce Reed, E. William Yund, Timothy J. Herron, and David L. Woods. 2013. Computerized Measures of Finger Tapping: Effects of Hand Dominance, Age, and Sex. *Perceptual and Motor Skills* 116, 3 (2013), 929–952. DOI: http://dx.doi.org/10.2466/25.29.PMS.116.3.929-952

[17] Fisk D. John, Pontefract Alexandra, Ritvo G. Paul, Archibald J. Catherine, and Murray J. Thomas. 1994. The impact of fatigue on patients with multiple sclerosis. *The Canadian Journal of Neurological Science* 21 (1994), 9–14. DOI: http://dx.doi.org/10.1017/S0317167100048691

[18] Alon Kalron, Anat Achiron, and Zeevi Dvir. 2011. Muscular and Gait Abnormalities in Persons With Early Onset Multiple Sclerosis. *J Neurol Phys Ther.* 35, 4 (2011), 164–169. DOI: http://dx.doi.org/10.1097/NPT.0b013e31823801f4

[19] Matthew Kay, Kyle Rector, Sunny Consolvo, Ben Greenstein, Jacob Wobbrock, Nathaniel Watson, and Julie Kientz. 2013. PVT-Touch: Adapting a Reaction Time Test for Touchscreen Devices. IEEE. DOI:http://dx.doi.org/10.4108/icst.pervasivehealth.2013.252078

[20] Edward Kim, Jesus Lovera, Laura Schaben, Dennis Bourdette, and Ruth Whitham. 2010. Novel method for measurement of fatigue in multiple sclerosis: Real-Time Digital Fatigue Score. *J Rehabil Res Dev* 47, 5 (2010), 477–84. DOI: http://dx.doi.org/10.1682/JRRD.2009.09.0151

[21] Benzi M. Kluger, Lauren B. Krupp, and Roger M. Enoka. 2013. Fatigue and fatigability in neurologic illnesses. *Neurology* 80, 4 (2013), 409–416. DOI: http://dx.doi.org/10.1212/WNL.0b013e31827f07be

[22] Gisela Kobelt, Alan Thompson, Jenny Berg, Mia Gannedahl, Jennifer Eriksson, the MSCOI Study Group, and the European Multiple Sclerosis Platform. 2017. New insights into the burden and costs of multiple sclerosis in Europe. *Multiple Sclerosis Journal* 23, 8 (2017), 1123–1136. DOI: `http://dx.doi.org/10.1177/1352458517694432`

[23] Lauren Krupp, Nicholas Larocca, Joanne Muir-Nash, and Alfred D. Steinberg. 1989. The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. 46 (1989), 1121–3. DOI:`http://dx.doi.org/10.1001/archneur.1989.00520460115022`

[24] Lauren B. Krupp. 2003. Fatigue in Multiple Sclerosis. *CNS Drugs* 17, 4 (2003), 225–234. DOI: `http://dx.doi.org/10.2165/00023210-200317040-00002`

[25] John F. Kurtzke. 1983. Rating neurologic impairment in multiple sclerosis. *Neurology* 33, 11 (1983), 1444–1444. DOI:`http://dx.doi.org/10.1212/WNL.33.11.1444`

[26] Rashmi Lakshminarayana, Duolao Wang, David Burn, K. Ray Chaudhuri, Clare Galtrey, Natalie Valle Guzman, Bruce Hellman, Ben James, Suvankar Pal, Jon Stamford, Malcolm Steiger, R. W. Stott, James Teo, Roger A. Barker, Emma Wang, Bastiaan R. Bloem, Martijn van der Eijk, Lynn Rochester, and Adrian Williams. 2017. Using a smartphone-based self-management platform to support medication adherence and clinical consultation in Parkinson's disease. *npj Parkinson's Disease* 3, 1 (2017), 2. DOI: `http://dx.doi.org/10.1038/s41531-016-0003-z`

[27] Krupp B. Lauren and Pollina A. Dean. 1996. Mechanisms and management of fatigue in progressive neurological disorders. *Current Opinion on Neurology* 9 (1996), 456–460. DOI: `http://dx.doi.org/10.1097/00019052-199612000-00011`

[28] Jau-Shin Lou, Greg Kearns, Theodore Benice, Barry Oken, Gary Sexton, and John Nutt. 2003. Levodopa improves physical fatigue in Parkinson's disease: A double-blind, placebo-controlled, crossover study. *Movement Disorders* 18, 10 (2003), 1108–1114. DOI: `http://dx.doi.org/10.1002/mds.10505`

[29] Bryan D. Loy, Ruby L. Taylor, Brett W. Fling, and Fay B. Horak. 2017. Relationship between perceived fatigue and performance fatigability in people with multiple sclerosis: A systematic review and meta-analysis. *Journal of Psychosomatic Research* 100 (2017), 1 – 7. DOI: `http://dx.doi.org/10.1016/j.jpsychores.2017.06.017`

[30] Virgil Mathiowetz, Karen Weber, Nancy Kashman, and Gloria Volland. 1985. Adult Norms for the Nine Hole Peg Test of Finger Dexterity. *The Occupational Therapy Journal of Research* 5, 1 (1985), 24–38. DOI: `http://dx.doi.org/10.1177/153944928500500102`

[31] Sospedra Mireia and Martin Roland. 2005. Immunology of multiple sclerosis. *Annual Review of Inmmunology* 65, 23 (2005), 683–747. DOI:`http://dx.doi.org/10.1146/annurev.immunol.23.021704.115707`

[32] N C Notermans, G W van Dijk, Y van der Graaf, J van Gijn, and J H Wokke. 1994. Measuring ataxia: quantification based on the standard neurological examination. *Journal of Neurology, Neurosurgery & Psychiatry* 57, 1 (1994), 22–26. DOI: `http://dx.doi.org/10.1136/jnnp.57.1.22`

[33] Davy Preuveneers and Yolande Berbers. 2008. Mobile Phones Assisting with Health Self-care: A Diabetes Case Study. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '08)*. ACM, New York, NY, USA, 177–186. DOI: `http://dx.doi.org/10.1145/1409240.1409260`

[34] John Prince, Siddharth Arora, and Maarten de Vos. 2018. Big data in Parkinson's disease: using smartphones to remotely detect longitudinal disease phenotypes. *Physiological Measurement* 39, 4 (2018), 044005. DOI: `http://dx.doi.org/10.1088/1361-6579/aab512`

[35] Blake P. Printy, Lindsey M. Renken, John P. Herrmann, Isac Lee, Bryant Johnson, Emily Knight, Georgeta Varga, and D. Diane Whitmer. 2014. Smartphone application for classification of motor impairment severity in Parkinson's disease. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. DOI: `http://dx.doi.org/10.1109/EMBC.2014.6944176`

[36] D. Rotstein, P. O'Connor, L. Lee, and B. J. Murray. 2012. Multiple Sclerosis Fatigue is Associated with Reduced Psychomotor Vigilance. 39, 2 (2012), 180–184. DOI:`http://dx.doi.org/10.1017/s0317167100013196`

[37] Ayache S. Samar and Chalah A. Moussa. 2017. Fatigue in multiple sclerosis - Insights into evaluation and management. *Clinical Neurophysiology* 47 (2017), 139–171. DOI: `http://dx.doi.org/10.1016/j.neucli.2017.02.004`

[38] Steven R. Schwid, Melissa Covington, Benjamin M. Segal, and Andrew D. Goodman. 2002. Fatigue in multiple sclerosis: Current understanding and future directions. *Journal of Rehabilitation Research and Development* 39, 2 (2002), 211–224.

[39] Steven R. Schwid, Charles A. Thornton, Shree Pandya, Karishma L. Manzur, Mohammed Sanjak, Marie D. Petrie, Michael P. McDermott, and Andrew D. Goodman. 1999. Quantitative assessment of motor fatigue and strength in MS. *Neurology* 53, 4 (1999), 743–743. DOI:`http://dx.doi.org/10.1212/WNL.53.4.743`

[40] Deborah Severijns, Ilse Lamers, Lore Kerkhofs, and Peter Feys. 2015. Hand grip fatigability in persons with multiple sclerosis according to hand dominance and disease progression. *J Rehabil Med.* (2015), 154–160. DOI:`http://dx.doi.org/10.2340/16501977-1897`

[41] Deborah Severijns, Inge Zijdewind, Ulrik Dalgas, Ilse Lamers, Caroline Lismont, and Peter Feys. 2017. The Assessment of Motor Fatigability in Persons With Multiple Sclerosis: A Systematic Review. *Neurorehabilitation and Neural Repair* 31, 5 (2017), 413–431. DOI: `http://dx.doi.org/10.1177/1545968317690831`

[42] Albertas Skurvydas, Marius Brazaitis, Julija Andrejeva, Dalia Mickeviciene, and Vytautas Streckis. 2011. The effect of multiple sclerosis and gender on central and peripheral fatigue during 2-min MVC. *Clinical Neurophysiology* 122, 4 (2011), 767 – 776. DOI: `http://dx.doi.org/10.1016/j.clinph.2010.10.005`

[43] National Multiple Sclerosis Society. 2018. Who Gets MS? (Epidemiology). Available at `https://www.nationalmssociety.org/What-is-MS/Who-Gets-MS` (2018/02/22). (2018).

[44] Anneke Steens, Astrid de Vries, Jolien Hemmen, Thea Heersema, Marco Heerings, Natasha Maurits, and Inge Zijdewind. 2012. Fatigue Perceived by Multiple Sclerosis Patients Is Associated With Muscle Fatigue. *Neurorehabilitation and Neural Repair* 26, 1 (2012), 48–57. DOI: `http://dx.doi.org/10.1177/1545968311416991`

[45] Jukka Surakka, Anders Romberg, Juhani Ruutiainen, Arja Virtanen, Sirkka Aunola, and Kari Mäentaka. 2004. Assessment of muscle strength and motor fatigue with a knee dynamometer in subjects with multiple sclerosis: a new fatigue index. *Clinical Rehabilitation* 18, 6 (2004), 652–659. DOI: `http://dx.doi.org/10.1191/0269215504cr781oa`

[46] Makoto Tanigawa, Jason Stein, John Park, Peter Kosa, Irene Cortese, and Bibiana Bielekova. 2017. Finger and foot tapping as alternative outcomes of upper and lower extremity function in multiple sclerosis. *Multiple sclerosis journal - experimental, translational and clinical* 3, 1 (2017). DOI: `http://dx.doi.org/10.1177/2055217316688930`

[47] Ana Lisa Taylor Tavares, Gregory S.X.E Jefferis, Mandy Koop, Bruce C. Hill, Trevor Hastie, Gary Heit, and Helen M. Bronte-Stewart. 2005. 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. *Movement Disorders* 20, 10 (2005), 1286–1298. DOI: `http://dx.doi.org/10.1002/mds.20556`

[48] Nieves Téllez, Jordi Río, Mar Tintoré, Carlos Nos, Inés Galán, and Xavier Montalban. 2005. Does the Modified Fatigue Impact Scale offer a more comprehensive assessment of fatigue in MS? *Multiple Sclerosis Journal* 11, 2 (2005), 198–202. DOI: `http://dx.doi.org/10.1191/1352458505ms1148oa`

[49] Ria Wolkorte, Dorothea J. Heersema, and Inge Zijdewind. 2015. Muscle Fatigability During a Sustained Index Finger Abduction and Depression Scores Are Associated With Perceived Fatigue in Patients With Relapsing-Remitting Multiple Sclerosis. *Neurorehabilitation and Neural Repair* 29, 8 (2015), 796–802. DOI: `http://dx.doi.org/10.1177/1545968314567151`

[50] Shumin Zhai, Jing Kong, and Xiangshi Ren. 2004. Speed-accuracy Tradeoff in Fitts' Law Tasks: On the Equivalency of Actual and Nominal Pointing Precision. *Int. J. Hum.-Comput. Stud.* 61, 6 (2004), 823–856. DOI: `http://dx.doi.org/10.1016/j.ijhcs.2004.09.007`

[51] Inge Zijdewind, Roeland Prak, and Ria Wolkorte. 2016. Fatigue and Fatigability in Persons With Multiple Sclerosis. *Exercise and Sport Sciences Reviews* 80 (2016), 123–128. DOI: `http://dx.doi.org/10.1249/JES.0000000000000088`