


# Potentials and barriers of using digital tools for collecting daily measurements in multiple sclerosis research

Digital Health  
Volume 7: 1–12  
© The Author(s) 2021  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/20552076211055552  
journals.sagepub.com/home/dhj  


Westergaard Katrine<sup>1</sup> , Signe Baattrup Ritzel<sup>1</sup>, Krogh Caroline<sup>1</sup>,  
Lynning Marie<sup>1</sup>, Bergien Sofie Olsgaard<sup>1</sup> and Skovgaard Lasse<sup>1</sup>

## Abstract

**Objective:** Digital tools offer new ways of collecting outcome data in intervention research. Little is known about the potentials and barriers of using such tools for outcome measurement in multiple sclerosis trials. This study aimed to examine reporting adherence and barriers experienced by people with multiple sclerosis in an intervention study using three different digital tools for outcome measurement.

**Methods:** This was a mixed-methods study conducted in the context of a randomized controlled trial. Data collected during the randomized controlled trial were analysed to assess reporting adherence. Twenty-three semi-structured, in-depth interviews were conducted to investigate randomized controlled trial participants' experiences.

**Results:** Reporting adherence was high for all three measurement tools, but lower in the control group. Four main barriers were defined: (1) the self-monitoring aspect and repeated tests imbedded in the digital tools affected participants' behavior during the randomized controlled trial. (2) Self-monitoring caused some participants to worry more about their health. (3) Passively collected data did not always correspond with participants' own experiences, which caused them to question the validity of the collected data. (4) Daily reporting using different digital tools placed a significant burden on participants.

**Conclusion:** The study indicates a high reporting adherence using digital tools among people with multiple sclerosis. However, future studies should carefully consider the overall burden imposed on participants when taking this approach. Measures should be taken to avoid the potential unintended effects of the self-monitoring and gamification aspects of using digital tools. These measures could include passive monitoring, reducing the frequency of reporting and blinding participants to their own data.

## Keywords

Multiple sclerosis, data collection, daily measurements, digital tools, mixed methods research

Submission date: 20 April 2021; Acceptance date: 7 October 2021

## Introduction

Multiple sclerosis (MS) is an incurable chronic disease. In many cases, medical treatment can halt its progress and accompanying symptoms can be treated medically. However, disease-modifying treatments are only partially effective and may cause several side effects.<sup>1</sup> Many people with MS (PwMS) receive treatments and engage in activities to maintain functions and alleviate symptoms

such as fatigue, poor sleep quality, walking disabilities, spasticity and cognitive challenges.<sup>2</sup>

<sup>1</sup>The Danish Multiple Sclerosis Society, Valby, Denmark

### Corresponding author:

Katrine Westergaard, The Danish MS Society, Poul Bundgaardsvvej 1, DK-2500 Valby, Denmark.

E-mail: kwe@scleroseforeningen.dk



A wide range of interventions is comprised in studies related to MS, however, methodological challenges have been identified which can lead to issues concerning the validity and precision of symptom data collected within these studies. The lack of continuously collected data from patients may entail an unnuanced picture of disease and symptom fluctuation, distorting the research results when applying pre–post design.<sup>3–5</sup> Additionally, data validity could be compromised due to recall bias when performing studies on a patient group with severe cognitive challenges.<sup>6</sup>

Ecological momentary assessment (EMA) has gained currency over the past few years, not least within clinical psychology.<sup>7–9</sup> EMA aims to accommodate the above-mentioned challenges by collecting data on subjects' behaviour and experiences in real-time and in their natural environments.<sup>10,11</sup> Similarly, intervention studies relating to MS must aim to accommodate methodology to secure research that captures outcomes important to patients from a daily perspective while simultaneously securing high validity.<sup>12–14</sup>

New digital tools for data collection have been developed over the past few years, offering solutions to the challenges mentioned. Wearables have enabled passive data collection relating to, for example, physical activity and sleep, and have enabled daily reporting of levels of functions and symptoms without imposing too large a burden on participants.<sup>3,15–17</sup> Moreover, the gamification element often embedded in digital tools may positively impact motivation and the response and reporting rates.<sup>16</sup>

Some research has been performed concerning the usability of specific tools, as well as the validity of data collected via such digital tools from PwMS.<sup>3,4,13–18</sup> A minor part of the research conducted within the use of digital tools among PwMS has successfully focused on detecting possible associations between various MS symptoms and/or between MS symptoms and activities.<sup>12,14</sup> Greenhalgh et al.<sup>18</sup> underline the potential of applying digital tools in clinical MS trials but emphasize at the same time the need for further research on study feasibility as well as the utility of data for clinicians/researchers. However, the minimal focus has in research been applied to the issue of balancing, on the one hand, the researchers' and clinicians' wish for high data validity/high reporting adherence and, on the other hand, the patients' need for manageable demands when participating in an intervention study. Hence, the present study aims at investigating the overall potentials and barriers when applying digital tools to capture repeated measurements daily among PwMS.

Based on a separative mixed methods design, including analyses of adherence and interview data, our research question is as follows: which overall potentials and barriers may be entailed by using daily data, collected via digital tools, as outcome measures in an intervention study on PwMS?

## Methods

### Study design

This study emerged from a randomized controlled trial (RCT), where different types of digital tools were used to collect outcome data. During the RCT, daily outcome measures were collected using three digital tools. Following the RCT, this present post hoc study was conducted to investigate participants' digital adherence and their experiences with daily use of digital tools. The post hoc study unfolded based on comments from the participants, which inspired a further investigation of the methodological implications of using digital tools for data collection.

### RCT study

A total of 28 PwMS participated in an 8-week RCT study. The study explored the effects of a self-help mind-body therapy called tension and trauma releasing exercises (TRE), primarily on fatigue and secondarily on other MS-related symptoms in PwMS. The recruitment process and specific inclusion and exclusion criteria for the study are described by Westergaard et al.<sup>19</sup> The participants were randomized into two groups. The intervention group followed a TRE intervention facilitated by a certified TRE therapist Supplementary to weekly sessions with the TRE therapist, the intervention group participants were asked to perform TRE daily at home. Other medications and therapies were maintained during the study. Participants in the control group did not receive any intervention and were asked to continue their usual care.

Participants were introduced to three different digital tools for outcome data registration: passive, test-based and self-reported. This data collection method aimed to collect data on a daily basis to obtain an overall insight into the potential fluctuating effects of TRE. Participants were asked to integrate the digital tools into their daily lives in a way that felt most convenient to them.

**MS patient-reported outcome tool.** Participants were asked to rate nine symptoms levels daily during the study period, using a recently developed patient-reported outcome (PRO) tool.<sup>20</sup> The MS PRO tool is a questionnaire made accessible via a smartphone app. The included symptom items are fatigue, gait disturbances, spasticity, muscle weakness, bladder and bowel dysfunction, pain, walking difficulty and sensory disturbances. Daily levels regarding all items are registered on an 11-point numerical rating scale scoring from 0 to 10, where 0 represents the specific symptom being non-existent and 10 represents the specific symptom appearing in its worst possible form. All registrations were blinded; hence, participants could not track previous registrations during the study.

**Floodlight Open.** As explorative outcome measures, we used the app Floodlight Open (FLO) as a test-based measure of two daily hand and motor function tests (the pinching test and the draw a shape test) and three daily gait and posture tests (the static balance test, the five U-turn test and the two-minute walk test). Additionally, they were asked to answer one mood question to assess their perceived overall state and perform a weekly cognitive test (the information processing speed test). FLO has been developed to continuously monitor MS progression and is accessible via smartphones.<sup>21</sup> Further details on the individual FLO measurements are described by Midaglia et al.<sup>22</sup> Participants had access to previous test scores from FLO so they could monitor potential progress. FLO is still undergoing testing and validation, so the measurements were considered explorative.<sup>22</sup>

**Fitbit.** Finally, all participants were provided with a Fitbit Charge 3 smart watch to passively monitor sleep duration and quality. Studies have shown promising results regarding the accuracy of sleep-wake states and sleep stage compositions measured via previous editions of Fitbit.<sup>23</sup> However, further validation is needed to fully assess its reliability and potential to replace gold standard tests such as polysomnography.<sup>23</sup> For this reason, sleep quality measures were considered explorative measurements in the RCT. Fitbits also automatically record physical activity measured as the number of daily steps and active minutes; however, only sleep data from the Fitbit were used for outcome measurement in the RCT study.<sup>19</sup> Physical activity was not defined as an outcome in the RCT study as it has seldom been emphasized by patients as an experienced effect. Furthermore, measurement of physical activity has shown to possibly entail a substantial element of gamification among PwMS,<sup>24,25</sup> which would have added an undesired level of complexity to the study. Fitbit allowed participants to view their data retrospectively and thus track progress in sleep and activity levels.

### Data collection

The present study includes quantitative data collected through digital tools during the RCT study as well as qualitative data derived from interviews. The quantitative data were used to analyse participants' reporting adherence. The qualitative methodology was applied to illuminate perceived barriers and potentials for participants when using different digital tools daily, thereby elaborating the quantitative results regarding reporting adherence.

**Quantitative data.** We collected quantitative data from the MS PRO tool, FLO and Fitbit for all 28 participants. These data included the time and date for each registration within all three digital tools.

**Qualitative data.** All participants in the RCT study were asked to participate in an in-depth interview within a month after completing the intervention. A total of 23 interviews with 12 participants from the intervention group and 11 participants from the control group were completed. These 23 informants constituted 82% of the participants. The remaining five participants from the RCT study were unable or unwilling to take part in the interviews. No incentives were provided to the participants. An interview guide was developed and small adjustments were made during the interview process. These adjustments primarily involved ensuring that both barriers and potentials linked to the use of wearables were articulated in the interviews. The final interview guide can be provided by the authors on request. The interviews were semi-structured, allowing for unexpected themes to occur. Due to practical reasons, the interviews were completed by four different researchers. All researchers who interviewed the participants had a background in the field of health care research and were involved in the development and/or implementation of the RCT study. All interviews were conducted in person. The duration of the interviews was 20 to 45 min.

All interviews were audio-recorded with the participant's consent. The study followed the European Union General Data Protection Regulations. Written consent from the participants was collected prior to the interview.

### Data analysis

**Overall approach.** Following the epistemological landscape suggested by Creswell and Clark,<sup>26</sup> the study was carried out as a mixed-methods study within a pragmatic worldview. The pragmatic worldview is exploratory and it is first and foremost characterized by focusing on research questions asked rather than on methods. Hence, the research process is oriented towards a dynamic application of methods of data collection and data analysis, rather than based on specific hypotheses or guided by specific pre-selected theories.<sup>26</sup> In the present mixed-methods study, quantitative data originating from an RCT study, were combined with qualitative data collected post hoc with the aim of generating a deeper and more nuanced understanding of the issue in question.

**Descriptive statistics.** The quantitative data were cleaned for double registrations within the same date per participant by assessing the exact time of the registration. Double registrations occurred due to registrations past midnight, where the digital tools recorded the registration for the following day rather than the day the participant was referring to. Therefore, the registration date was changed to the previous day if no registration was recorded for the day in question.

The assessment of reporting adherence was inspired by previous studies using daily measurements with similar digital tools; reporting adherence for each participant was

**Table 1.** Characteristics of participants.

	Participants from RCT study			Interviewed participants		
	Intervention group (n = 14)	Control group (n = 14)	All (n = 28)	Intervention group (n = 12)	Control group (n = 11)	All (n = 23)
<i>Age</i>						
Years, mean (SD)	50.1 (6.4)	49.4 (8.6)	49.8 (7.5)	49.3 (5.9)	47.2 (7.1)	48.3 (6.4)
Range	39-62	34-64	34-64	39-59	34-55	34-59
<i>Sex</i>						
Male n (%)	5 (35.7)	6 (42.9)	11 (39.3)	4 (33.3)	3 (27.3)	7 (30.4)
Female n (%)	9 (64.3)	8 (57.1)	17 (60.7)	8 (66.7)	8 (72.7)	16 (69.6)
<i>Disease duration</i>						
Years, mean (SD)	9.3 (7.4)	4.4 (3.9)	6.9 (6.3)	8.5 (7.6)	4.1 (3.8)	6.4 (6.4)
Range	1-21	1-12	1-21	1-21	1-12	1-21
<i>MS type</i>						
Relapsing-remitting MS n (%)	14 (100)	8 (57.1)	22 (78.6)	12 (100)	7 (63.6)	19 (82.6)
Progressive MS n (%)	0 (0.0)	6 (42.9)	8 (21.4)	0 (0.0)	4 (36.4)	4 (17.4)

RCT: randomized controlled trial; MS: multiple sclerosis; SD: standard deviation.

defined as the proportion of study weeks including at least 3 days of interaction with the digital tools.<sup>22,27</sup> Descriptive analyses of calculated adherence were reported for the digital tools separately and combined. The latter represented the proportion of study weeks with at least 3 days of interaction with all three digital tools. Due to a missing FLO identification number from one of the participants, only 27 participants were included when assessing adherence within FLO as well as the overall adherence.

All statistical analyses were conducted using Stata/IC 16 and MS Excel 2010.

**Qualitative analysis.** Interviews were transcribed verbatim, analysed and thematized using open coding. The coding was inductive, following the overall exploratory approach and performed after the data collection and separately by authors KW and SBR. Themes generated from the coding were compared, discussed and further analysed for possible new themes until concurrence between the authors was achieved. The thematic analysis was inspired by Nowell et al.<sup>28</sup> was conducted in striving towards their criteria of trustworthy research. The characteristics of all participants who were interviewed are presented in Table 1.

All coding was completed using the NVivo 12 software package.

**Merging of quantitative and qualitative data.** A separate approach was used for merging the quantitative and qualitative data. When applying a separate approach to mixed methods, quantitative and qualitative data are collected and analysed separately and integrated into the discussion of the results.<sup>29</sup>

## Results

The characteristics of the participants from the RCT study as well as interviewed participants are displayed in Table 1.

### Reporting adherence

Regarding the overall adherence calculated as the proportion of study weeks with at least 3 days of interaction with all three digital tools, a clear tendency was detected towards markedly lower overall adherence in the control group. Hence, the average overall adherence was 81% in the intervention group and 52% in the control group.

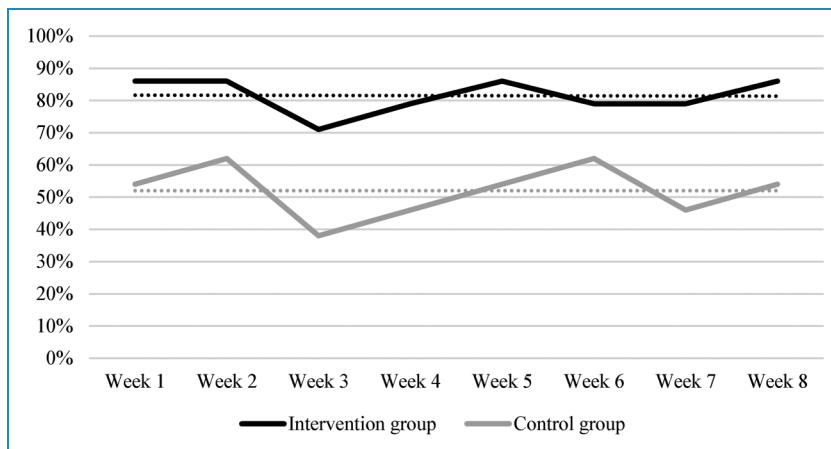


Figure 1. Overall reporting adherence over time.

Table 2. Overall reporting adherence over time.

	Week 1 (%)	Week 2 (%)	Week 3 (%)	Week 4 (%)	Week 5 (%)	Week 6 (%)	Week 7 (%)	Week 8 (%)	Average (%)
Intervention group	86	86	71	79	86	79	79	86	81
Control group	54	62	38	46	54	62	46	54	52

The overall reporting adherence over time, illustrated in Table 2 and Figure 1, demonstrated that the reporting adherence was stable.

The adherence analyses regarding the three digital tools separately indicated that the reporting adherence, calculated as study weeks with at least 3 days of interaction, was generally high. This is presented in Table 3 as average. However, the results indicated that adherence was lower in the control group and for the FLO tool in both groups.

The development in reporting adherence over time, related to the three registration tools individually (Table 4), indicated that adherence was stable but consistently lower in the control group. This tendency applied especially to the FLO tool, where a fall in adherence over time was detected. Adherence related to the MS PRO tool was high and stable.

The results of the quantitative data analyses suggested that reporting adherence was high and stable over time in the intervention group. Adherence was lower in the control group; especially, the FLO tool showed a lower and less stable adherence. The quantitative data indicated that the overall registration burden constituted a barrier, particularly to the control group.

In the following paragraphs, the interview data are presented, allowing the elaboration of tendencies regarding the adherence data.

Table 3. Reporting adherence separately for the three tools.

	Intervention group (%)	Control group (%)
MS PRO tool	95	89
Fitbit	93	81
FLO	83	64

MS PRO: multiple sclerosis patient-reported outcome; FLO: Floodlight Open.

### Interview findings

Four themes were generated from the qualitative analysis.

**Theme 1: self-monitoring affects behavior.** About half of the participants viewed the self-monitoring embedded in the two tools, Fitbit and FLO, as an opportunity to obtain further knowledge about their health and well-being. However, especially participants from the intervention group expressed how these new insights motivated health behavior and elicited a competitive mindset.

*Motivation derived from new realizations about symptoms, lifestyle and well-being.* Participants found that self-monitoring gave them new realizations about patterns between symptom burden, lifestyle and well-being.

**Table 4.** Reporting adherence over time.

	Week 1 (%)	Week 2 (%)	Week 3 (%)	Week 4 (%)	Week 5 (%)	Week 6 (%)	Week 7 (%)	Week 8 (%)	Average (%)
<i>MS PRO tool</i>									
Intervention group	93	100	93	93	93	93	93	100	95
Control group	92	85	92	85	100	85	85	92	89
<i>Fitbit</i>									
Intervention group	93	93	93	86	93	93	100	93	93
Control group	77	92	69	77	85	85	77	85	81
<i>FLO</i>									
Intervention group	86	86	71	86	86	86	79	86	83
Control group	77	69	62	62	62	62	62	62	64

MS PRO: multiple sclerosis patient-reported outcome; FLO: Floodlight Open.

Several participants especially showed interest in sleep patterns obtained from Fitbit; however, insights into other associations between behavior and well-being were also valued by the participants.

I think it was scary. I could easily see a common thread; when I got bad sleep, I also felt bad. I could also see a connection with physical activity. The more physically active I was, the better I slept. And the day and the mood were better (female, 34 years, control group).

Participants described knowledge about these associations as an opportunity to be more conscious about their sleep patterns and act to change these. Thus, self-monitoring became a motivator for some participants, who changed their health behavior based on Fitbit data.

I have been better to go to bed when I have time to go to bed early [...] I have been more conscious about not getting enough sleep, and I also activated an alarm to notify me to go to bed. So, I do not forget. In this way, I have also used it [Fitbit] for something else than wearing it (male, 44 years, intervention group).

Nevertheless, some participants indicated, when asked about how time-consuming the registration was, that the motivational impact the Fitbit had at the beginning of the

study disappeared quickly, as they experienced a discrepancy between their effort and the Fitbit results.

I think it was very motivating, but unfortunately it passed quite quickly when I found out that it [Fitbit] did not work. At first, I thought it was the coolest thing that I could see how much I had moved ... Even out just walking. Until I discovered that Fitbit was not so stable (female, 46 years, intervention group).

*Access to test results elicited a competitive attitude.* The interviews illustrated that participants perceived the digital clinical tests embedded in FLO as small challenges rather than clinical tests to provide evidence for disease status. This left participants in competition with themselves, aiming to exceed the score from the previous day. Participants described how they used the visible test scores from FLO to follow their progress or lack of it. Most of the participants stated that they saw an improvement in their test results over time.

It has been ... in fact, it has been a bit of competition with myself, just to see if I could get better over time. So, it was not just to perform these tasks, but constantly trying to get better at them [...] I have a feeling that I actually got better with time at those exercises. And, particularly, I can see my motor skills have been improved, especially during the exercise where you had to squeeze the tomatoes. In the

beginning, I got about 20 within the time period. It got up to almost 50 at the end (male, 51 years, intervention group).

However, according to some of the participants, these improvements appeared to be limited to the specific tests in FLO and were not rediscovered in their physical abilities in daily life. The participants attributed the effects to the routine and continuous practice of completing the test daily and not to actual improvements.

It has not changed anything in my way of walking or standing. There are, of course, some of those exercises we have done with Floodlight, where you walk and turn, where I may have a feeling that I may have gotten better at it, but in reality, I believe that it is because you do a routine and you just get better at that routine, and therefore I really think that is the effect (male, 53 years, intervention group).

*Theme 2: self-monitoring can cause worry.* Although several participants were enthusiastic about the health information acquired through Fitbit and FLO under some circumstances, these insights led to increased worries by a few participants in both groups.

Some participants even mentioned that they tried not to focus too much on the data because the new insights made them worry too much about their health.

I hope it is more of a coping strategy than a repression. I don't have the strengths to think about MS every day. I also think it [self-monitoring with Floodlight] is a way to give it [MS] more space than I'm interested in (female, 45 years, intervention group).

Especially where FLO and Fitbit made them aware of unhealthy lifestyle patterns, participants explained that they found the newly acquired understanding of their health situation to cause a disturbing stress response.

However, it is, or for me, it has been stressful after getting the Fitbit watch, when knowing I don't sleep enough. Because if I didn't have the Fitbit, I knew I didn't sleep so much. But I didn't think it was a problem [...] But it's not Fitbit's fault or that you have to keep an eye on it; it is more the understanding of one's own situation or what you should call it. Then, I'm being stressed because I'm not sleeping enough (female, 50 years, control group).

*Theme 3: discrepancy between data and experiences.* As mentioned in Theme 1, several participants described the motivational impact of using Fitbit and FLO. Regarding this, about half of the participants were curious about whether the data from Fitbit and FLO matched their experiences. Slightly more participants from the control group experienced a discrepancy between their own experiences

and the Fitbit data, which led them to question their own experiences and the data revealed by the digital tools.

The only thing I became mad about was the sleep one [Fitbit]. I didn't think it worked. When I slept like a baby, it said I was wide awake. And when I sat down reading, then it said I was asleep. So, I didn't think it worked. And I also believe it to be very optimistic with counting steps (female, 46 years, intervention group).

Moreover, the participants explained that several times, they had experienced that the data collected by Fitbit was affected by various factors, for example, which hand the smart watch was attached to.

It was actually a bit difficult to compare Fitbit and Floodlight with how I felt. Sometimes, I think it was contradictory because it did not quite recognize the numbers Fitbit revealed (female, 50 years, control group).

Other participants described how they wanted to confirm experienced symptoms on FLO. This was unsuccessful; however, rather than ascribing the discrepancy to FLO, the participants questioned their perception of symptoms.

Because I think I can't use my right hand at all today, so it was very interesting to be able to measure it on the app [Floodlight] by drawing the figure or squeezing the tomatoes. I think it has not been completely clear, but it may be because of my own perception of the symptoms is different than they are ... It may also be that it measures some things that are not so relevant, but it can also be that you have some mental perceptions ... No, I do not know, it goes wildly bad today, but in reality, it is actually not that big (female, 39 years, intervention group).

*Theme 4: the burden of daily measurements affects participants' views on self-registration.* Daily measurements were by some experienced as stressful and a daily inconvenience. Some types of measurements were regarded as more burdensome than others.

*Daily measurements contributed to daily inconveniences.* Many participants found it moderately stressful to complete the intervention and measurements daily; thus, they became a daily inconvenience. As illustrated below, the daily measurements became time-consuming and difficult to incorporate into daily life.

However, it has actually been a stress factor instead. [...] It quickly became an hour, and it was not easy to incorporate into my daily life other than when the kids were asleep or something. So, it has been a lot of time to spend on it compared to how my everyday life is like. (male, 44 years, intervention group)

The participants further mentioned that the feeling of stress related to the daily measurements was not always due to the time it took to register but because they felt obliged to use all three digital tools. To a few participants, the burden of the daily measurements became so strong that the end of the study period was perceived as a liberation.

I want to say, when we stopped [using the digital tools], it was a relief not to do it every evening. [...] It does not take long, but it was still a burden to actually do it. (female, 50 years, control group)

Moreover, participants viewed FLO as especially burdensome because it required more energy to complete daily.

I have registered in yours every day [MS PRO tool], but there have been days where I have not had the energy to complete Floodlight. And I was totally hooked when I began the study. (male, 54 years, control group)

*Some types of measurements were more burdensome than others.* Participants emphasized that some tests, especially those from FLO requiring them to get up or leave the house, were challenging to incorporate into daily life. Particularly, the two-minute walk was described by participants as difficult to incorporate into everyday life and the test was perceived as more inconvenient than many of the other tests.

Sometimes I missed it, especially the two-minute walk test Because you need to be outside ... and if I already got home from work and up on the fourth floor. Then it became ... well, not today. (male, 44 years, intervention group)

Moreover, the FLO tests were viewed as too monotonous, contributing to a lack of motivation to continuously complete the tests daily.

It [Floodlight] was too repetitive when doing it every day. It was fine when it was just once a week, and I think more assignments could be added, so you only completed them once a week. It got annoying to run around the same two chairs every day. It was so backbreaking that I didn't want to do it. Well, I did it anyway, but I was tired of it. (female, 55 years, control group)

Participants explicitly outlined FLO as the most burdensome of the digital tools used, whereas no participants explicitly ascribed any burden to Fitbit or the MS PRO tool.

## Discussion

The purpose of this study was to examine potentials and barriers to using daily data, collected via digital tools, as outcome measures in an intervention study on PwMS.

The study used a separative mixed methods approach and data derived from an RCT using passive, test-based and self-reported measurements for outcome data registration.

The quantitative data suggested a high and stable reporting adherence in the intervention group regarding all three measurement types. Conversely, adherence in the control group was lower and less stable, especially regarding the overall registration burden. The qualitative analysis illuminated barriers and potentials of using digital data collection tools as viewed by the study participants. The findings indicated that self-monitoring via digital tools provided participants with new insights into their health and well-being. In some cases, self-monitoring elicited a competitive mindset in participants, primarily in the intervention group, and motivating behavior change unrelated to the RCT intervention. For a few participants, self-monitoring also led to concerns about health and well-being. Many participants found it burdensome to incorporate all the reporting activities into their daily life, especially the functional tests from FLO and not least due to the requirement of specific physical performances linked to the registrations. Some participants, primarily in the control group, found the registration stressful. Finally, some participants – from both groups – experienced discrepancies between their own experiences and registered data from some of the digital tools. However, such discrepancies did not seem to affect markedly their adherence to wearing the devices and/or accomplishing the registrations.

### Reporting adherence and burden in relation to daily measurements

It is known that missing data can lead to potential bias. In both blinded and unblinded RCT studies, missing data occurs if subjects drop out of the study or if subjects from one group are more likely to be compliant.<sup>30,31</sup> The present study suggests that digital tools may contribute to a high and stable adherence; however, the quantitative and qualitative data also indicate that adherence was lower in the control group, especially if the overall registration burden was high.

Studies using daily measurements emphasized strengths as well as limitations related to reporting adherence and reporting burden. Speier et al.<sup>32</sup> found high adherence and low attrition in their study of cardiac patients when using activity trackers. Based on their review on physical activity measurements related to cancer and chronic respiratory disease, Maddocks et al.<sup>33</sup> found adherence challenging and reported mixed effects from physical activity interventions based on, for example, pedometers. Tsianakas et al.<sup>34</sup> reported barriers related to adherence and registration burden in community-based walking intervention studies on metastatic cancer patients using questionnaires and pedometers. For both the questionnaire and



the pedometer, the authors experienced low adherence in the control and the treatment groups. Particularly, the study participants described the daily questionnaire as burdensome and time-consuming.<sup>34</sup> The present study supports the complexity found in the literature when evaluating the use of digital tools for data collection, indicating high adherence and low attrition as well as a high registration burden, particularly linked to comprehensive tests.

The present study did not find that participants experienced a burden regarding using Fitbit for daily measurements, suggesting an interesting potential in passive data collection. However, studies have emphasized that although digital tools as pedometers and Fitbits can measure daily activity passively without input from the participant, they can still entail adherence barriers. Results from a cohort study of older people suggested that studies using wearables to monitor daily activity levels should be designed as short-term studies to maintain high adherence and obtain reliable data.<sup>35</sup> Maddocks et al.<sup>33</sup> argue that measurements from wearables such as Fitbit may underestimate physical activity compared to validated accelerometers and may overestimate sleep duration compared to sleep monitors. This perspective is supported by our findings, as some participants experienced a discrepancy between Fitbit data and their perception of sleep and physical activity. Maddocks et al.<sup>33</sup> suggest that a combined approach using both objective and subjective measurements is required to make data more reliable. A non-peer-reviewed report on experiences of PwMS with wearables in Denmark emphasizes that lack of confidence in the validity of wearables may lead participants to discontinue their use.<sup>24</sup> However, our overall mixed methods findings indicated that the discrepancy between the participants' perception and the data showed on the Fitbit did not markedly affect the registration adherence among the participants in this study. This may partly be due to a feeling of obligation towards the research project, or a questioning of the accuracy of own observations but it may also represent a pragmatic approach towards data accuracy in self-tracking among the participants. Observational studies covering a more extensive period of time may elucidate this issue further.

### Implications of self-monitoring

Using digital solutions with an element of gamification is a focus within MS rehabilitation, as these make physical training less tedious and increase motivation to perform exercises.<sup>36</sup> However, when used for data collection in intervention studies, gamification in digital tools may constitute a challenge if the aim is to objectively measure changes in functioning over time. Regarding the FLO tool, participants expressed a wish to compete with themselves to increase their test scores. In a controlled trial, a

training effect pertaining to the outcome measurement instrument should be equal in both groups and thus should not lead to overinterpretation (a false-positive result) of the intervention effects. However, it may cause outcome measures to be less sensitive to changes caused by the intervention and thus could lead to 'false negative' results.<sup>37</sup>

The Fitbit allowed participants to monitor their sleep data continuously. Analyses of the interviews indicated that these self-monitoring motivated participants to change their behavior by, for example, going to bed earlier to achieve better sleep quality. Studies have shown that self-monitoring with wearable devices may benefit lifestyle change and disease management.<sup>25,38</sup> However, in this intervention study, behavior change resulting from self-monitoring was not an intended effect. If the participants change health behavior due to self-monitoring, the changes observed in the data could reflect these behavioral changes rather than the studied intervention. This may entail implications for the interpretation of study results similar to those described regarding gamification. In the light of this finding, interpretations of the results from the RCT<sup>19</sup> based on Fitbit data (sleep duration and sleep quality) have been interpreted with caution as Fitbits were not blinded to the participants – this implication was not fully anticipated. Blinding is one solution to the challenges described above; by applying blinding to the data collection tools the aspect of self-monitoring could be limited among the participants. One possible disadvantage entailed by this solution is that blinding participants may have the adverse effect of decreasing adherence, as the interviews indicated that the gamification and self-monitoring elements increased participants' motivation to use the tools. However, the quantitative results show that among the three digital data collection tools, the MS PRO tool, which did not have a self-monitoring component, showed the best reporting adherence. This may indicate that continuous access to registered data is not necessarily essential for reporting adherence. However, it should be considered that participants in the present study had access to registered data from the Fitbit and the FLO, which might have accommodated the need for access to data.

*New ethical considerations.* Using daily measurements in research may entail ethical considerations. The present study indicates that some participants experienced continuous access to data regarding their health as challenging. Similar barriers were described by Andersen et al.<sup>38</sup> in a study on cardiac patients' experiences with wearable data. In this study, access to data on physical activity-induced concerns and anxiety in some participants, indicating complexity in such data experiences of people with chronic illness. A non-peer-reviewed report on experiences of PwMS with wearables also articulates the potential barriers embedded in access to daily measurements of health status

for participants who do not have the qualifications to understand and act on the data. The study emphasizes the need for support for these patients.<sup>24</sup> Andersen et al.<sup>38</sup> further suggest that patients may struggle to understand and act appropriately on wearable data, questioning the ethical justifiability of enrolling patients in studies using outcome measures that provide health status information without offering them the opportunity to discuss these insights with a health professional. In a paper on the application of wearable technologies in MS, Bricchetto et al.<sup>39</sup> stress the need to appropriately educate patients engaged in research trials involving using wearable technologies on the physiopathological mechanisms of MS to help them interpret changes and patterns in the data. Bricchetto et al.<sup>39</sup> suggest that lack of appropriate education might cause PwMS to pursue unsafe coping strategies. Some participants in the present study referred to stressful situations when having to act on data from the daily measurements without experiencing any effects of their effort. Furthermore, for a smaller number of participants in this study, the availability of daily measurements constituted an unwelcome increased attention to MS in daily life.

Moreover, it may be important to consider the burden caused by daily measurements for research participants. Using daily measurements may impose new expectations on research participants, as they must report data frequently. The qualitative results in this study indicate that some participants experienced stress relating to the overall burden of daily measurements. Hence, some of the participants referred to the measurements as time-consuming, especially when they were also asked to learn and integrate a new self-care intervention in their daily lives; however, this experience was not shared by all. The results of a recent preliminary study investigating the perceived burden of PRO measures for people with chronic pain indicate that less frequent measurements are less burdensome: weekly measurements less burdensome than daily and daily less burdensome than twice daily measurements.<sup>40</sup> Therefore, a better understanding of the acceptable burden threshold for PwMS is needed to apply digital tools to data collection in research. This threshold may depend on patients' disease status, health literacy and the intervention in question.

**Strengths and limitations.** The main strength of this study is that the qualitative interviews were designed after the RCT study and thereby inspired by participants' remarks during the study. Another strength is that both qualitative and quantitative data have been used in the study to explore the research question. Integrating both kinds of data makes it possible to generate knowledge that quantitative or qualitative data could not have produced independently,<sup>41</sup> for example, linked to the finding of experienced discrepancies between experiences and registered data not leading to low adherence. A relatively large number of

interviews were conducted (82% of the population from the RCT study), why the findings may be regarded as representing a wide range of patient perspectives. The credibility of the qualitative analysis was enhanced by the involvement of multiple researchers, who reached a consensus regarding the themes identified.<sup>42</sup> However, a limitation of the study may be entailed by the interviews being conducted by several different interviewers, potentially compromising the uniformity of the interviews. Although the interviews were based on a semi-structured interview guide, the variation in interviewers may have affected the credibility of the qualitative findings.<sup>42</sup> The transferability of the findings was improved by displaying the characteristics of the participants and explaining the context of the study setting. Agreement between the authors among the generated coding demonstrated the dependability of the study.

Three different types of digital tools were included for daily measurements: passive, test-based and self-reported, which constituted a strength of the study, as it allowed comparisons regarding adherence and registration burden. However, many tools exist with different modes of administration and degrees of participation; thus, the findings in this study are not applicable to all types of daily measurement tools.

The analyses regarding reporting adherence were based on calculations of weeks with at least 3 days of interaction, which may result in a broad variation of interaction.

## Conclusion and future directions

Digital tools can generate useful daily data regarding symptoms and functions for PwMS; a high and stable reporting adherence in the intervention group regarding all three types of measurements was found, whereas adherence in the control group was lower and less stable, especially regarding the overall registration burden. However, several barriers exist regarding integration into daily life as well as motivation factors. The latter aspect may ultimately affect the interpretation of the outcome data.

Future use of digital tools to collect outcome data in intervention studies related to MS may benefit from acknowledging the potential in acquiring daily data. Such data may enrich the comprehensiveness of the outcomes achieved, not least strengthening the integration of participants' experienced outcomes. However, future use of digital tools in intervention studies could also benefit from carefully adjusting the overall registration burden to the specific group of participants, considering the participants' resources regarding energy and cognitive capacity and the risk of evoking concerns regarding health status and functionality. This latter aspect is especially relevant when conducting studies on people with a progressive illness. The possibility of reducing the reporting burden by limiting registrations to every second or third day rather than every day, or by applying several limited

registration periods rather than an entire study period, should be considered. Future use of digital tools in intervention studies should involve consideration of possible elements of gamification. While such elements may strengthen the adherence, they may simultaneously compromise data quality due to motivation factors that are questionable within a research context.

**Contributorship:** KW and SBR conducted the data analyses wrote the first draft of the methods and findings of the manuscript. CSK was involved in patient recruitment and data collection. ML was involved in patient recruitment, reviewed study findings and wrote part of the discussion. SOB reviewed the study findings and wrote part of the discussion. LS reviewed the study findings, wrote the introduction and part of the discussion. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

**Declaration of conflicting interests:** The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

**Ethical approval:** The RCT study as well as the following interviews could not formally be reviewed by the administration of the Danish National Committee on Health Research Ethics since the intervention was *mild and non-invasive*. Formal written consent was collected by the researchers both prior to the RCT study and prior to the conducted interviews.

**Funding:** The authors received no financial support for the research, authorship and/or publication of this article.

**Guarantor:** LS

**Informed consent:** Not applicable, because this article does not contain any studies with human or animal subjects.

**Trial registration:** Not applicable, because this article does not contain any clinical trials.

**ORCID iD:** Westergaard Katrine  <https://orcid.org/0000-0001-9951-4463>

## References

- Dobson R and Giovannoni G. Multiple sclerosis – a review. *Eur J Neurol* 2019; 26: 27–40.
- Barin L, Salmen A, Disanto G, et al. The disease burden of multiple sclerosis from the individual and population perspective: Which symptoms matter most? *Mult Scler Relat Disord* 2018;25:112–121.
- Powell DJH, Liossi C, Schlotz W, et al. Tracking daily fatigue fluctuations in multiple sclerosis: Ecological momentary assessment provides unique insights. *J Behav Med* 2017; 40: 772–783.
- Romberg A, Ruutiainen J, Puukka P, et al. Fatigue in multiple sclerosis patients during inpatient rehabilitation. *Disabil Rehabil* 2008; 30: 1480–1485.
- Kratz AL, Murphy SL and Braley TJ. Ecological momentary assessment of pain, fatigue, depressive, and cognitive symptoms reveals significant daily variability in multiple sclerosis. *Arch Phys Med Rehabil* 2017; 98: 2142–2150.
- Grzegorski T and Losy J. Cognitive impairment in multiple sclerosis – a review of current knowledge and recent research. *Rev Neurosci* 2017; 28: 845–860.
- Schueller SM, Aguilera A and Mohr DC. Ecological momentary interventions for depression and anxiety. *Depress Anxiety* 2017; 34: 540–545.
- Davidson CL, Anestis MD and Gutierrez PM. Ecological momentary assessment is a neglected methodology in suicidology. *Arch Suicide Res* 2017; 21: 1–11.
- Bell IH, Lim MH, Rossell SL, et al. Ecological momentary assessment and intervention in the treatment of psychotic disorders: A systematic review. *Psychiatr Serv* 2017; 68: 1172–1181.
- McKeon A, McCue M, Skidmore E, et al. Ecological momentary assessment for rehabilitation of chronic illness and disability. *Disabil Rehabil* 2018; 40: 974–987.
- Dunton GF. Ecological momentary assessment in physical activity research. *Exerc Sport Sci Rev* 2017; 45: 48–54.
- Kratz AL, Fritz NE, Braley TJ, et al. Daily temporal associations between physical activity and symptoms in multiple sclerosis. *Ann Behav Med* 2019; 53: 98–108.
- Kim E, Lovera J, Schaben L, et al. Novel method for measurement of fatigue in multiple sclerosis: Real-time digital fatigue score. *J Rehabil Res Dev* 2010; 47: 477–484.
- Kratz AL, Murphy SL and Pain BT. Fatigue, and cognitive symptoms are temporally associated within but not across days in multiple sclerosis. *Arch Phys Med Rehabil* 2017; 98: 2151–2159.
- Babbage DR, van Kessel K, Drown J, et al. MS energize: Field trial of an app for self-management of fatigue for people with multiple sclerosis. *Internet Interv* 2019; 18: 100291.
- Giunti G, Mylonopoulou V and Rivera Romero O. More stamina, a gamified mHealth solution for persons with multiple sclerosis: Research through design. *JMIR Mhealth Uhealth* 2018; 6: e51.
- Motl RW, Sandroff BM and Sosnoff JJ. Commercially available accelerometry as an ecologically valid measure of ambulation in individuals with multiple sclerosis. *Expert Rev Neurother* 2012; 12: 1079–1088.
- Greenhalgh J, Ford H, Long AF, et al. The MS symptom and impact diary (MSSID): Psychometric evaluation of a new instrument to measure the day to day impact of multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2004; 75: 577–582.
- Westergaard K, Lynning M, Hanehøj K, Skovgaard L. Tension and Trauma Releasing Exercises for People with Multiple Sclerosis: A Randomized Controlled Trial [In review]. *Advances in mind-body medicine*.
- Knudsen A Karnoe. An investigation of diet and lifestyle factors in multiple sclerosis and their relations to symptom severity. PhD thesis, University of Copenhagen, 2019.
- Floodlight Open. Understanding MS together [Internet]. [henvist 11. juni 2020]. Tilgængelig hos, <https://floodlightopen.com/en-US> (2020, accessed 20 July 2020).

22. Midaglia L, Mulero P, Montalban X, et al. Adherence and satisfaction of smartphone- and smartwatch-based remote active testing and passive monitoring in people with multiple sclerosis: Nonrandomized interventional feasibility study. *J Med Internet Res* 30 2019; 21: e14863.
23. Haghayegh S, Khoshnevis S, Smolensky MH, et al. Accuracy of wristband Fitbit models in assessing sleep: Systematic review and meta-analysis. *J Med Internet Res* 28 2019; 21: e16273.
24. Bergien, Sofie Olsgaard. MS Life Logging. A research project on the possibilities of using wearables among people with Multiple Sclerosis [Internet]. [14. december 2020]. <https://www.scleroseforeningen.dk/sites/default/files/2020-10/MS%20Life%20Logging%20-%20English.pdf>.
25. Bergien SO, Fuglsang CH, Kayser L, Lynning M, Skovgaard L. MS Life Logging: How wearables can empower and benefit people with multiple sclerosis in thier everyday life. *Mult Scler* 2019; 25: 1052.
26. Creswell JW and Clark VLP. *Designing and conducting mixed methods research*. London: Sage Publications, 2011.
27. Montalban X, Mulero P, Midaglia L, Graves J, Hauser S, Julian L, et al. FLOODLIGHT: Remote self-monitoring is accepted by patients and provides meaningful, continuous sensor-based outcomes consistent with and augmenting conventional in-clinic measures. *Neurology* 2018;90: P4.382.
28. Nowell L S, Norris JM, White DE, et al. Thematic analysis: Striving to meet the trustworthiness criteria. *Int J Qual Methods* 2017; 16: 1–13.
29. Moseholm E and Fetters MD. Conceptual models to guide integration during analysis in convergent mixed methods studies. *Methodol Innov* 2017; 10: 1–11.
30. Lewis SC. How to spot bias and other potential problems in randomised controlled trials. *J Neurol Neurosurg Psychiatry* 2004; 75: 181–187.
31. Schulz KF and Grimes DA. Blinding in randomised trials: Hiding who got what. *Lancet* 2002; 359: 696–700.
32. Speier W, Dzibur E, Zide M, et al. Evaluating utility and compliance in a patient-based eHealth study using continuous-time heart rate and activity trackers. *J Am Med Inf Assoc* 2018; 25: 1386–1391.
33. Maddocks M and Granger CL. Measurement of physical activity in clinical practice and research: Advances in cancer and chronic respiratory disease. *Curr Opin Support Palliat Care* 2018; 12: 219–226.
34. Tsianakas V, Harris J, Ream E, et al. Canwalk: A feasibility study with embedded randomised controlled trial pilot of a walking intervention for people with recurrent or metastatic cancer. *BMJ Open* 2017; 7: e013719.
35. Kocherginsky M, Huisingh-Scheetz M, Dale W, et al. Measuring physical activity with Hip accelerometry among U.S. older adults: How many days are enough? Heymann D, redaktør. *PLoS ONE* 2017; 12: e0170082.
36. Lavorgna L, Russo A, De Stefano M, et al. Health-related coping and social interaction in people with multiple sclerosis supported by a social network: Pilot study with a new methodological approach. *Interact J Med Res* 2017; 6: e10.
37. Koivisto J and Hamari J. The rise of motivational information systems: A review of gamification research. *Int J Inf Manage* 2019; 45: 191–210.
38. Andersen TO, Langstrup H and Lomborg S. Experiences with wearable activity data during self-care by chronic heart patients: Qualitative study. *J Med Internet Res* 2020; 22: e15873.
39. Bricchetto G, Pedullà L, Podda J, et al. Beyond center-based testing: Understanding and improving functioning with wearable technology in MS. *Mult Scler* 2019; 25: 1402–1411.
40. Bodart S, Byrom B, Crescioni M, et al. Perceived burden of completion of patient-reported outcome measures in clinical trials: Results of a preliminary study. *Ther Innov Regul Sci* 2019; 53: 318–323.
41. Fetters MD and Freshwater D. The 1+ 1 = 3 integration challenge. *J Mix Methods Res* 2015; 9: 115–117.
42. Greenhalgh T and Taylor R. How to read a paper: Papers that go beyond numbers (qualitative research). *Br Med J* 1997; 315:740–743.