Review

Mobile health interventions in multiple sclerosis: A systematic review

Christoph Heesen iD, **Thomas Berger, Karin Riemann-Lorenz, Nicole Krause, Tim Friede** iD, **Jana Pöttgen, Björn Meyer and Dagmar Lühmann**

Abstract

Background: Persons with multiple sclerosis (pwMS) might be particularly well suited to benefit from digital health applications because they are, on average, younger and less severely disabled than patients with many other chronic diseases. Many digital health applications for pwMS have been developed.

Objectives: Analysis of the evidence of digital health applications to improve health outcomes from a patient perspective.

Methods: A systematic review was performed on all randomized controlled trials (RCTs) that have studied mobile health interventions for pwMS, that is, which can be applied with a smartphone, tablet, or laptop to improve patient-reported outcomes.

Results: Of the 1127 articles identified in the literature search, 13 RCTs fit the inclusion criteria. Two trials studied messaging systems, two depression interventions, one addressed MS fatigue, five cognition, and three mobility issues, of which two focused on spasticity management. One trial aimed to enhance physical activity. Most were pilot studies that cannot yield definitive conclusions regarding efficacy. One depression intervention and one fatigue intervention showed significant results across several outcomes.

Conclusion: Several mobile self-guided digital health applications for pwMS have been tested in RCTs, and two interventions targeting depression and fatigue have demonstrated significant effects. Challenges remain regarding implementation into routine care.

Correspondence to:
C Heesen
Clinical and Rehabilitative
MS Research, Institute
of Neuroimmunology
and Multiple Sclerosis,
University Medical Center,
Martinistr. 52, Hamburg
20246, Germany.
Heesen@uke.de

**Christoph Heesen
Karin Riemann-Lorenz
Nicole Krause
Jana Pöttgen**
Clinical and Rehabilitative
MS Research, Institute of
Neuroimmunology and
Multiple Sclerosis, University
Medical Center, Hamburg,
Germany

Thomas Berger
Department of Clinical
Psychology, University of
Bern, Bern, Switzerland

Tim Friede
Department of Medical
Statistics, University Medical
Center Göttingen, Göttingen,
Germany

Björn Meyer
GAIA AG, Hamburg,
Germany

Dagmar Lühmann
Department of General
Practice and Primary Care,
University Medical Center
Hamburg-Eppendorf,
Hamburg, Germany

Introduction

Digitalization is among the greatest contemporary societal challenges and has a profound impact on medicine and healthcare delivery. The term eHealth is often invoked in this context, referring to the use of information and communication technologies to support health and healthcare services. The term mHealth (mobile health) refers to a specific delivery mode, such as using short messaging services (SMS) and smartphone applications (apps) to convey health information or interventions to patients. mHealth interventions, delivered via the Internet, might be particularly relevant for persons with multiple sclerosis (pwMS), for several reasons.[1–3] Young age, complex presentation, and evolution of MS, and numerous decisions to be made needing decision support tools are major arguments to develop and deliver mHealth tools.

Despite the opportunities of easy access and scalability, mHealth applications also have limitations. For example, older age and low education can function as barriers to the uptake of mHealth applications. A wealth of digital information of uncertain quality can also be experienced as overwhelming, trigger fears,[4] or lead to unrealistic, inflated hopes, incorrect self-diagnoses, and inappropriate therapies. In addition, the attitudes of physicians and other care providers play a key role in the use of mHealth technologies, such that patients may be more likely to engage with mHealth tools if these are endorsed by their clinicians.[5,6] The recently established Digital Delivery of Care Act (DVG) in Germany[7] (https://www.bgbl.de) provides criteria to evaluate digital health applications in terms of their efficacy, as shown in methodologically sound trials, and their data security. Once approved and listed in the digital health application

directory, these mHealth tools are eligible for reimbursement by statutory health insurance companies in Germany, which opens a pathway for the implementation of digital health applications in the German healthcare system.

Although several reviews have been performed on digital health interventions in MS,[3,8] no systematic review has focused specifically on self-guided mHealth interventions in MS. Following the idea that digital interventions that do not require clinician support could be disseminated widely and cost-effectively, this review includes randomized controlled trials (RCTs) of self-guided digital interventions that target any stage of MS and that do not require personal support beyond that which may be offered in the context of usual treatment. Only studies with patient-reported outcomes were included to estimate the extent to which patients experience benefit from the interventions.

Methods

Inclusion criteria

We included all RCTs as pilot studies or fully powered trials in which pwMS were the target population or were investigated as a defined subgroup. The intervention needed to be a program (application, app) to support the self-management of the target group, which can be used on common mobile (location-independent) devices (smartphone, tablet, and laptop). "Self-management" includes all interventions implemented by the patient himself or herself to support coping with the disease and its consequences. Endpoints of interest are patient-relevant outcomes, such as morbidity parameters, quality of life, or adverse effects. Studies only examining performance-based tests, such as cognitive tests, mobility tests, or biomarkers, such as magnetic resonance imaging (MRI) or blood markers, were not considered. At least one patient-based outcome as a secondary or tertiary outcome had to have been reported to qualify for inclusion in this systematic review. Programs that require additional equipment beyond a smartphone or tablet for their implementation were excluded (example: a wheelchair that can be controlled via a smartphone or virtual reality training is not subject of the evaluation). Corresponding to the guidelines for use of the Digital Delivery of Care Act,[9] only interventions that do not require additional support from healthcare providers or caregivers were examined. The mHealth applications were used in addition to standard care. Comparison interventions were (1) usual care (treatment as usual, TAU) or waitlist or (2)

use of active controls, that is, electronic or paper-based applications that do not contain the component presumed to be effective and are ideally indistinguishable from the intervention under investigation.

Database search

A systematic literature search for studies (Figure 1), initial search on May 5, 2020 and updated on May 11, 2022, was performed in the MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials databases (see search strategy in Supplemental Appendix 1). In parallel, a search for relevant systematic reviews was conducted in MEDLINE, Embase, Cochrane Database of Systematic Reviews, and HTA Database. The selection of relevant studies was performed independently by two persons (C.H. and D.L.). Discrepancies were resolved by discussion between the two. Data extraction was performed into standardized tables.

Risk of bias assessment

We applied Risk of bias (ROB) 2 from the Cochrane Collaboration to assess the qualitative certainty of results.[10] Cross-endpoint and endpoint-specific criteria of bias potential were assessed. The endpoint-specific assessment focused on patient-reported outcomes as a group. The bias potential was classified as "low," "some concerns," or "high" in each case.

Estimates of efficacy

If not given in the publications, clinical relevance of reported statistically significant between-group differences was determined by calculating standardized mean differences (SDM, Cohen's *d*) to estimate clinical relevance.[11]

Funding source

This work is an update of a health technology assessment under the auspices of the Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWIG) in 2020 and 2021, published in 2022 on the website of the IQWIG.[12]

Results

In total, 13 RCTs (see Table 1) could be identified. Two studies addressed messaging systems, two depression interventions, one study was performed to ameliorate MS fatigue, five studies worked on interventions to improve cognition, and three studies addressed mobility issues, of which were two
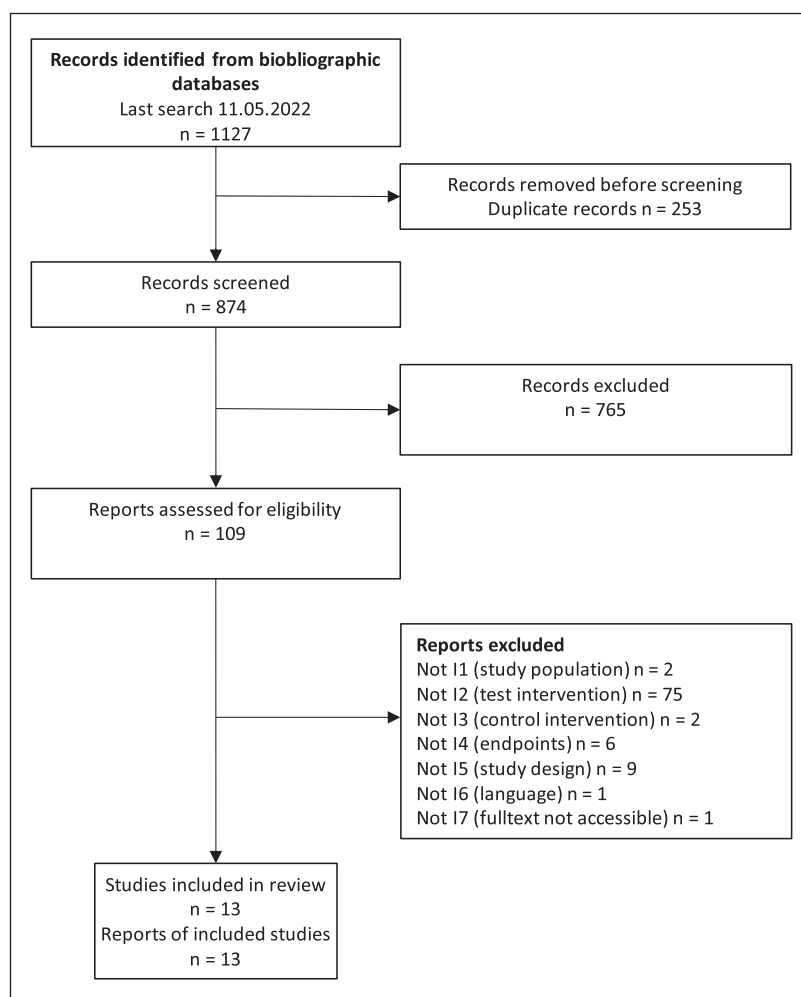
Figure 1. Flowchart of included studies.

spasticity management interventions and one aimed to enhance physical activity.

ROB ratings

Most studies were rated as having two or more areas of concern (Table 2). Outcome assessment (D4) was a key domain because five studies only used Patient-Reported Outcome Measures (PROMS). Furthermore, only four studies[13,19,20,22] provided active control interventions which were indistinguishable from the experimental intervention for pwMS. Therefore, participant blinding and blinded outcome assessment were not ensured in most studies (D2). Another area of concern is the lack of published protocols or analysis plans, so that selective reporting cannot be ruled out (D5). In addition, the randomization procedure was often insufficiently described (D1). Some studies also had high dropout rates (D3). In addition, treatment fidelity and adherence reflected in frequency

and usage time were seldomly reported in detail (for overview of adherence definitions applied and results, see Supplemental Appendix 2).

Messaging systems

Two studies investigated the effectiveness of messaging systems for self-organization.[13,14] Miller et al.[13] compared an established messaging concept system with an enhanced version with self-management tools in one of the largest studies ($n = 206$) running over 12 months. Goodwin et al.[14] performed a small ($n = 38$) and short-term pilot crossover study with a similar intervention approach comparing the messaging system NeuroPage® with an enhanced version with individualized goals. Primary endpoints in the study by Miller were validated PROMS, such as the perceived sickness impact, functioning in three key disability domains (multiple sclerosis functional composite, MSFC), and self-efficacy. Goodwin et al. chose an

Table 1. mHealth intervention randomised-controlled trials in MS—overview of study design and findings.

Author country	Number design	Patients' age (M, SD), disease duration (DD)	Intervention (IG)	Control (CG)	Intervention duration in months (FU)	Findings* (CI)
Miller et al.[13] USA	206 RCT	IG: 48.1 (9.1) CG: 48.1 (9.7) DD: no data	Internet-based messaging + self-monitoring and management tools	Internet-based messaging	12 (12)	Sickness-Impact-Profile, Multiple-Sclerosis-Functional-Composite, Multiple-Sclerosis-Self-Efficacy-Scale (control subscale), Seniors'- General-Satisfaction-and-Physician-Quality-of-Care-General Satisfaction-with-Medical-Care ns, European-Quality-of-life 5 Dimensions (EQ5D) sig. favoring control
Goodwin et al.[14] UK	38 Pilot RCT	IG: 48.8 (12.9) CG: 46.7 (9.7) DD: 10.6/9.7	Reminder text messages, individually tailored (NeuroPage)	Non-memory text messages	2 (2)	Everyday-Memory-Questionnairens, General-Health-Questionnaire 30 sig. Cohen's $d = 0.84$ (CI: −0.03 to 1.10*), EQ5D ps, Adaptation-to-Memory-Difficulties-Outcome-questionnairens, diary sig.
Fischer et al.[15] Germany	90 RCT	IG: 45.4 (12.6) CG: 42.2 (10.6) DD: no data	Fully automated, internet-based, cognitive behavioral therapy (CBT) program (deprexis); 10 modules up to 60 minutes	Waitlist	2 (6)	Beck-Depression-Inventory (BDI) mean diff −4.02 $p = 0.01$ Cohen's $d = 0.53$ (CI: 0.11 to 0.95), WHO-Quality-of-live BREF psych. sig., Hamburg-Quality-of-life-Scale in MS, Fatigue-Scale-Motor-and-Cognition, Suicidal-Behavior-Questionnaire-Revised ns
Cooper et al.[16] UK	24 Pilot RCT	IG: 48.0 (7.79 CG: 42.0 (7.0) DD: no data	Computerized CBT (Beating the Blues); eight modules á 50 minutes	Treatment as usual (TAU)	2 (3)	BDI mean diff 4.16, Multiple-Sclerosis-Impact-Scale 29 diff. phys. 5.18, diff. psychol. 6.06
Pöttgen et al.[17] Germany	275 RCT	IG: 40.8 (11.1) CG: 41.9 (9.4) DD: 8.9/9.2	Fully automated, internet-based, CBT program (elevida); eight modules up to 30–45 minutes	Waitlist	2 (3 and 6)	Chalder mean diff −2.74, $p = 0.0007$ Cohen's $d = 0.53$ (CI: 0.22 to 0.84), FSMC sig., Hospital-Anxiety-Depression-Scale (HADS)A (Anxiety) sig., HADS-D (Depression) ns, HAQUAMS ps, Frenchay-Activity-Index sig.
Di Giglio et al.[18] Italy	35 Pilot RCT	IG. 44.6 (7.6) CG: 42.9 (9.4) DD: 13.3/11.4	Video game-based cognitive rehabilitation program (Dr Kawashima's brain training); 30 minutes on 5 days/week	Waitlist	2 (2)	Stroop mean diff. 4.16 $p = 0.034$ Cohen's $d = 0.51$ (CI: −0.16 to 1.19), Paced-Auditory-Serial-Addition-Test (PASAT) ns, Symbol-Digit-Modalities-Test (SDMT) mean diff. 8.85 $p = 0.049$, Modified-Fatigue-Impact-Scale (MFIS) ns, Multiple-Sclerosis-Quality-of-life 54 mental subscale med diff. 7.47 $p = 0.002$
Charvet et al.[19] US	20 Pilot RCT	IG: 38 (10.6) CG: 42 (12.5) DD: no data	Adaptive cognitive remediation program; 30 minutes on 5 days/week	Active control (computer games); 30 minutes on 5 days/week	3 (3)	Compliance, general cognitive composite z change 0.5 $p = 0.02$ Cohen's $d = 1.21$ (CI: 0.65 to 1.59)

(Continued)

Table 1. (Continued)

Author country	Number design	Patients' age (M, SD), disease duration (DD)	Intervention (IG)	Control (CG)	Intervention duration in months (FU)	Findings* (CI)
Charvet et al.[20] US	135 RCT	IG: 48 (13) CG: 52 (11) DD: 11.9/13.5	Adaptive cognitive remediation program; 60 min/day on 5 days/week	Active control (computer games); 60 min/day on 5 days/week	3 (3)	General cognitive composite z change 0.16 p=0.028, Cohen's d=0.36 (CI: 0.04 to 0.73)
Campbell et al.[21] UK	38 Pilot RCT	IG: 46.2 (6.6) CG: 48.5 (9.6) DD: 10.5/12.7	Home-based, computer-assisted cognitive rehabilitation (RehaCom); 45 minutes on 3 days/week	DVDs with natural history films	1.5 (1.5)	Symbol-Digit-Modalities-Test mean diff 4.56 p=0.005 Cohen's d=1.067 (CI: 0.39 to 1.75), California-Verbal-Learning-Test (CVLT), Brief-Visuspatial-Memory-Test (BVMT), EQ5D, Functional-Assessment-Measure, HADS, Fatigue-Severity-Scale, Multiple-Sclerosis-Neuropsychological-Questionnaire
Bove et al.[22] UK	44 Pilot RCT	IG: 52.9 (14.0) CG: 49.2 (10.9) DD: 11.2/16.1	Game (AKL-T03, Akili Interactive) with sensory and motor tasks adapting to individual; 25 minutes on 5 days/week	Game (AKL-T09) with a word generation task; 25 minutes on 5 days/week	1.5 (1.5)	Mean diff: SDMT, PASAT, BVMT, CVLT, Center-for Epideimologica-Studies-Depression-Scale, State-Trait-Anxiety-Index, MFIS ns
Ehling et al.[23] Austria	20 Pilot RCT	IG: 46.6 CG: 50.5 DD: 12.6/16.3	Individualized training program (MS-spasticity app), including reminder; 2 × 15 minutes on 6 days/week	Paper-based instructions	3 (3)	All ns: Neurological-Rating-Scale (NRS), Ashworth, Spasticity Visaul-Analogue-Scale (VAS), pain VAS, HADSD/A, Würzburger-Erschöpfungs-Inventar-Multiple-Sklerose (WEIMUS), Motricity Index, Quality-of-liveVAS, SF36, Timed-25-Footwalk-Test
Ehling et al.[24] Austria	94 RCT	IG: 50.8 CG: 46.4 DD: 13.3/12.5	Individualized training program (MS-spasticity app), including reminder; presumable as Ehling et al.[23]	Paper-based instructions	3 (3)	NRS median diff: −1.0 p=0.009, Ashworth, motricity index, performance tests, WEIMUS, HADSA/D ns
Nasseri et al.[25] Germany	38 Pilot RCT	IG: 49.6 (8.5) CG: 52.5 (7.3) DD: 13.1/20.1	Smartphone app with health information, accelerometry	Information sheet	3 (3)	All ns: Performance tests, Expanded-Disability-Status-Scale, Multiple-Sclerois-Walking-Scale 12, Goodin-Leisure-Time-Exercise-Questionnaire, HAQUAMS

*key significant outcome with data and effect size estimate.

Table 2. ROB ratings of mHealth intervention RCTs focusing PROMs.

Author year country	Randomization process	Deviation from intended intervention	Missing outcome data	Measurement of outcomes (PROMs)	Selection of the reported results	Overall
Messaging						
Miller et al.[13]	+	+	+	!	!	!
Goodwin et al.[14]	+	+	!	!	!	!
Depression						
Fischer et al.[15] Germany	+	+	+	!	+	!
Cooper et al.[16] UK	!	!	+	!	!	!
Fatigue						
Pöttgen et al.[17] Germany	+	+	+	!	+	!
Cognition						
Di Giglio et al.[18] Italy	+	+	+	!	!	!
Charvet et al.[19] US	!	+	+	!	!	!
Charvet et al.[20] US	+	+	+	!	+	!
Campbell et al.[21] UK	+	!	+	+	!	!
Bove et al.[22] UK	-	!	!	+	+	!
Mobility						
Ehling et al.[23] Austria	-	!	+	+	!	!
Ehling et al.[24] Austria	!	+	+	!	!	!
Nasseri et al.[25] Germany	+	!	+	!	+	!

+ Low risk.   ! Some concerns.   - High risk.

PROMs = Patient-reported outcome measures. This ROB assessment focuses study evidence based on PROMs even if they had not been primary outcomes. Therefore, overall ROB is also driven by the quality of PROM assessment.

everyday memory questionnaire as primary outcome with other PROMS as secondary endpoints. Both studies did not show an effect on their primary endpoints. However, Goodwin reported a significant effect on the General Health Questionnaire with its clinical relevance remaining uncertain (Cohen's $d = 0.84$, CI: 0.03 to 1.10). In addition, some evidence can be derived for a relevant greater benefit of the NeuroPage intervention compared to the active control intervention with regard to the endpoint "forgotten target behaviors."

Depression management

A study with 90 participants[15] by Fischer et al. and a pilot study by Cooper et al.[16] with 24 participants investigated the effectiveness of online programs based on cognitive behavioral therapy (CBT; Fischer: deprexis®, 9 weeks; Cooper: Beating the Blues®, 8 weeks) for the treatment of depressive symptoms in comparison to a waitlist control group. Neither program was specifically designed for patients with MS. While any self-reported depressive symptoms were sufficient for inclusion in the Fischer study, the

Cooper study required a Beck Depression Inventory (BDI) score of at least 14 points (at least mild depressive symptoms) for inclusion, patients with BDI scores above 28 were excluded. Results were collected at 9 weeks (post-treatment, primary endpoint) and 6 months after the end of the intervention (follow-up assessment) in the study by Fischer and at 8 weeks (post-treatment) and 21 weeks (follow-up; 13 weeks after the end of the intervention) in the study by Cooper. In the later, the primary goal was to analyze recruitment, and the clinical endpoint was self-reported depression measured by the BDI. Clinical endpoints in this study were measured as a basis for a valid sample size calculation for a larger trial. The study had a dropout rate of 25% (3 out of 12 participants at 8 weeks) in the intervention group and none in the control group. The study by Fischer measured the severity of depression symptoms by BDI-2 as the primary endpoint to evaluate intervention effectiveness. Here, the dropout rate was 22% at 9 weeks in the intervention group and 20% in the control group.

The pilot study by Cooper et al.[16] was able to show the feasibility of the intervention. Regarding the endpoint depression, the study demonstrated numerically favorable results at the end of intervention and at follow-up for the intervention group. Given the small sample size of $n = 12$ per group, no conclusion on efficacy is possible. Fischer et al.[15] reported a statistically significant effect in favor of the intervention group compared to the waitlist control group in the German study.[15] The effect size (Cohen's $d = 0.53$, CI: 0.11 to 0.95) indicates at least a small and, based on the effect size, medium and clinically relevant effect in favor of deprexis. Regarding the other endpoints considered in the study by Fischer et al., statistically significant differences between the study groups could be shown in two secondary outcomes (psychological well-being and motor fatigue). In terms of potential adverse effects, no evidence of new occurrences of suicidal ideation was observed in either group in the German study. In the British study, no results were reported for the protocol-reported endpoints of anxiety and generic quality of life.

Fatigue management

In an RCT with 275 participants, a 12-week digital health application based on CBT techniques (elevida) was tested against a waitlist control group to reduce fatigue.[17] The key inclusion criterion was a total fatigue score of more than 43 points on the Fatigue Scale for Motor and Cognitive Functioning (FSMC).[26] Dropout rates after 12 and 24 weeks were 22% and 30% in the intervention group versus 11% and 15% in the control group, respectively. Significant

intervention effects on the primary outcome (Chalder Fatigue Scale)[27] and for the endpoint fatigue, measured with the FSMC scale, were observed at post-intervention (12 weeks) and follow-up (24 weeks). The point estimate for the effect size (Cohen's d) after 12 weeks (primary endpoint) was 0.53 (CI: 0.22 to 0.84). The confidence interval indicates at least a small clinically relevant effect; in fact, the point estimate suggests that the effect is medium in magnitude. Statistically significant effects in favor of the intervention group were also found with regard to the endpoints reduction of anxiety, increase in activity level, and in three out of six subscales of the applied quality of life scale.

Cognition management

A total of five studies investigated the effects of interventions to support cognitive functioning. All studies, except the Charvet et al.'s[19] pilot study, recruited MS patients with documented cognitive deficits in at least one dimension. All studies excluded patients with pre-existing psychiatric and dementia conditions. An Italian study with 35 participants[18] compared the effects of an 8-week computer-based training program (Dr Kawashima's Brain Training®) against a waitlist control group. They addressed three cognitive tests, fatigue, and quality of life, as endpoints, without a defined primary endpoint. De Giglio et al.[18] found significant but possibly not relevant differences in favor of the intervention group in one (Stroop test) of the three functional tests for the endpoint cognitive function. Also significant but possibly not certainly, relevant results were found for a subscale of MS-specific quality of life. A pilot study with 20 participants and a confirmatory study with 135 participants, both published by Charvet et al.,[19,20] investigated the effects of adaptive cognitive reeducation programs over 12 weeks. In the confirmatory study, the training time per day was doubled to 60 minutes each on 5 days/week. In both studies, the control groups trained for the same amount of time with non-adaptive computer games as an active control group. In the study by Charvet et al.,[19] feasibility was addressed as the primary outcome, measured as compliance defined as 50% or more use of the intervention per patient. In the full study,[20] a composite cognitive outcome parameter of six cognitive tests was used as the primary outcome. This battery was also applied in the pilot but has not been validated prior to the studies. In addition, both studies used simple, non-validated three- and two-item self-assessment and peer-assessment scales. Charvet et al.[19,20] were able to demonstrate the feasibility of the intervention in their pilot study (primary endpoint), which was confirmed in the confirmatory study. Both studies showed statistically significant

changes in the self-developed composite cognitive outcome measure. However, while in the pilot study, the composite cognitive outcome indicates a clinically meaningful change the magnitude of the effect could not be confirmed in the larger confirmatory study.

In another study by Campbell et al.[21] with 38 participants, the intervention group received a 6-week computer-based rehabilitation program (RehaCom®), while the control group watched natural history DVDs for the same amount of time. In this study, outcomes were measured after a follow-up at 6 weeks after completion of the intervention and at 12 weeks. Authors defined BICAMS (Brief International Cognitive Assessment in Multiple Sclerosis, composite score of cognitive function tests), and two quality-of-life measures as primary outcomes in the protocol. In the publication, only the BICAMS is mentioned as a primary outcome. In addition, they assessed perceived cognitive function and self-efficacy. Only in one of the three cognitive tests of the BICAMS, authors report a significant result in favor of the intervention group after 6 weeks. Data for calculating an effect size are not available. The two other cognitive measurement parameters, and all results after 12 weeks, showed no significant differences between the study groups. No significant effects could be obtained for perceived cognitive functioning self-efficacy and patient empowerment.

Bove et al.[22] investigated two different computer games one (intervention group) with combined sensory and motor tasks adapting to the individual performance level for 6 weeks. Primary outcome was change in information processing speed as measured by the symbol digit modalities test but also other BICAMS measures were assessed and patient-perceived deficits, depression, stress, and fatigue. The study did not find a differential effect of the adaptive intervention while both of the gaming groups improved. Effects persisted to some extent to Month 3.

In summary, no conclusive evidence is derived for a benefit of cognitive training programs compared to active control interventions or care as usual (CAU) with regard to cognitive functions. However, four of the five studies were just pilot studies.

Mobility
Three studies investigated the effects of mobility performance support programs. A pilot study by Ehling et al.[23] with 20 participants evaluated the effects of an app-controlled therapy program as a follow-up to inpatient rehabilitation treatment in MS patients with proven moderate spasticity of the lower extremities.

In 2022, Ehling et al.[23,24] published a full RCT with the same approach but only selected patients had gained benefits during the rehabilitation stay. The intervention included individually composed exercises for mobility, strength, and coordination of the lower extremities and aimed at reducing spasticity. Outcome measurement was obtained 3 months after the start of the intervention in both studies. In the 2017 study, Ehling et al.[23] did not define a primary outcome in their pilot study. In the full RCT, the Neurological Rating Scale for spasticity[21] was used as a primary outcome measure. A couple of other mobility measures and measures of fatigue and depression were applied. In the pilot study, Campbell et al.[21] could not demonstrate an effect of the Neurological-Rating-Scale (NRS)[25] or any other measure. The full RCT, however, showed a significant effect on the NRS, while all other measures did not show significant effects.

In the pilot study by Nasseri et al.,[25] the effectiveness of a smartphone app including a self-programmed accelerometry (including feedback) and evidence-based patient information on physical activity was tested against a leaflet with general information on physical activity and health. Patients with primary or secondary progressive MS and an EDSS score < 6.5 were included. The outcome was measured after 3 months. An increase in step count of at least 20% was chosen as the primary outcome. In addition, mobility tests and measures of daily physical activity and activity in general were included. Nasseri et al.[25] could not demonstrate a 20% increase in step count or any change in other measures.

In summary, data indicate a possible effect of a spasticity app in ameliorating spasticity. No effect of an educative digital intervention combined with a smartphone accelerometry to increase step counts in progressive MS could be shown.

Discussion
Based on the high potential of digital health in managing MS, we performed a systematic review on self-guided digital interventions, that is, applications without additional human support studied in RCTs.

Reviews on digital health in MS have already been performed. Nasseri et al.[25] and Lavorgna et al.[8] summarized mHealth in MS, Fricke and Apolinario-Hagen[28] MS apps, and Yeroushalmi et al.[3] telemedicine in MS. These papers address digital apps in MS in a very general way and especially also in the context of rehabilitative interventions with additional

interactive parts. The results point to their potential but conclude that most mHealth interventions are not yet implementable due to a lack of data on the benefit-risk profile. In telemedical and especially telerehabilitative studies, digital aids are combined with interactive components, which was not the subject of the current systematic review.

Based on our aim to focus on self-guided digital interventions, we identified 13 RCTs, 8 of which were feasibility or pilot RCTs. Statistically significant and clinically relevant benefits were observed in the area of neuropsychiatric functioning (i.e. depression and fatigue). The study by Fischer et al.[15] on depression examined the deprexis program and reached its primary endpoint with a likely clinically relevant effect size. A recent large multinational study of an MS-adapted version of deprexis, called amiria, funded by the National MS Society of the United States has shown efficacy among $n = 279$ pwMS with mild to moderate depression, with effect sizes on depression reduction of around $d = 0.95$.[29] In addition, a recent meta-analysis of 12 RCTs reported robust evidence for the efficacy of deprexis in mild to severe depression, with somatic comorbidity present in 2 of the 12 trials.[30] Another depression intervention, the Beating the Blues program, was examined in a small feasibility study by Cooper et al.,[16] also suggesting a potential benefit.

Also, the fatigue trial by Pöttgen et al.[17] with the digital application elevida reached its primary endpoint, with a moderate clinical-relevant effect size. Like deprexis, the elevida intervention has been approved by the German Federal Institute for Drugs and Medical Devices and is, therefore, the first German "DIGA" for MS that is permanently included in the directory of reimbursable digital health applications. Given the largely disappointing efforts to treat MS fatigue with medication, CBT-based interventions and exercise currently seem to be the most effective strategies. Therefore, a further development for elevida with more focus on increasing physical activity and possibly including sensoring technology might be an approach to enhance efficacy. However, while sensoring and multimodal monitoring is strongly advocated in MS,[31] we are not aware that this approach yet has shown to lead to better health outcomes in MS.

Approaches to enhance cognitive functioning show feasibility and some studies show improved cognitive performance tests but the only larger RCT[20] did not reach is primary endpoint. As the active control group received an intervention similar to the key intervention, the barrier to show a differential effect was high in this study. This may apply also for other studies with active control groups as, for example, the studies on messaging systems. Especially the large RCT by Miller et al.[13] might have failed because of the similarity of the interventions. In addition, studies to improve cognitive functioning are hampered by the open question of clinical relevance of changes in cognitive tests, which is not clear for all applied tests. Inclusion of established perceived deficits questionnaire, although showing other limitations, should be more encouraged in these interventions.

Interestingly, little effort has been made to enhance general physical activity through self-guided mHealth interventions in MS. The only study in this area[25] could not show an increase in step counts. However, as the intervention in this study only included evidence-based patient information and did not provide a behavior change program, it might have been too ambitious to change walking habits in progressive pwMS just through providing a smartphone accelerometry.[32,33] Much more comprehensive behavior change interventions should be studied here.

Most studies from this review had relevant methodological problems as in very few protocol data or detailed study registration data were available, making conclusions on selective reporting impossible. Randomization was often poorly described without documented concealed allocation. Although this was not the case for all trials examined here, a general recommendation for future research in this area is to improve the methodological rigor, in line with current standards, as delineated in the CONSORT E-Health Statement, among others.[34]

Reported adherence rates in the 13 trials examined here differed. However, as some studies were advertised, recruited, and organized purely online without any personal contact, the question should be discussed what could be considered a minimum dosage for different interventions and target conditions. Methodological concepts are needed to account for the possible selection bias of higher dropouts, especially in studies conducted unguided and completely online. However, comparative data of patient characteristics in online studies compared with the general population have shown substantial representativeness.[35]

This leads to one of the most critical questions in studying behavioral interventions: blinding. Patients can only be blinded when an active control intervention is very similar to the intervention to be studied. In this case, also the control intervention can have a treatment effect or even a negative impact leading to an under- or

overestimation of the benefits of the intervention. Especially in smaller studies, it will be therefore very difficult to show differential effects, which may have happened to some of the RCTs summarized in this review. An NIH expert panel in 2019 presented a concept to handle the methodological problem of controls in behavioral interventions.[36] The recommendation here is to start with a treatment as usual or waitlist control group. Having gathered some effectiveness data, further studies with active control conditions can attempt to disentangle the mechanisms and mediators by which treatment effects unfold over time.

This review is limited as it addresses only self-guided interventions. Guided interventions show a broad range from single email contacts to regular telephone–video–live contacts. However, scalability is substantially much more challenging. Furthermore, ROB assessment focused on self-reported outcomes, which were not always used as primary study endpoints. Therefore, ROB information on the study in the domain PROMS and the overall rating are to some extent selective. However, it remains an open question whether the studies with significant outcomes in performance-based measures but missing or non-significant PROMs have a relevant impact on the activity and participation of pwMS.

Searching for registered but not yet published studies, we found additional RCTs indicating that the currently limited evidence might improve in the near future.

Conclusion

Although the evidence on the effectiveness of self-guided digital interventions in MS is limited, studies especially in the area of neuropsychiatric symptoms (depression and fatigue) have demonstrated the utility of such interventions.

Author Contributions

C.H. provided the concept, was involved in the screening process and data interpretation, and wrote the initial paper draft. D.L. provided analysis methodology and supervised and performed data extraction and interpretation. She commented and revised the paper. K.R.-L. was involved in data extraction and interpretation. J.P. and T.F. were involved in effect size calculations and interpretation. T.B., K.R.-L., N.K., T.F., J.P., and B.M. provided substantial comments to the paper draft.

ORCID iDs

Christoph Heesen https://orcid.org/0000-0001-8131-9467

Tim Friede https://orcid.org/0000-0001-5347-7441

Supplemental Material

Supplemental material for this article is available online.

References

1. Haase R, Schultheiss T, Kempcke R, et al. Modern communication technology skills of patients with multiple sclerosis. *Mult Scler* 2013; 19(9): 1240–1241.

2. Marziniak M, Brichetto G, Feys P, et al. The use of digital and remote communication technologies as a tool for multiple sclerosis management: Narrative review. *JMIR Rehabil Assist Technol* 2018; 5(1): e5.

3. Yeroushalmi S, Maloni H, Costello K, et al. Telemedicine and multiple sclerosis: A comprehensive literature review. *J Telemed Telecare* 2020; 26(7–8): 400–413.

4. Marrie RA, Salter AR, Tyry T, et al. Preferred sources of health information in persons with multiple sclerosis: Degree of trust and information sought. *J Med Internet Res* 2013; 15(4): e67.

5. Apolinario-Hagen J, Menzel M, Hennemann S, et al. Acceptance of mobile health apps for disease management among people with multiple sclerosis: Web-based survey study. *JMIR Form Res* 2018; 2(2): e11977.

6. Seboka BT, Yilma TM and Birhanu AY. Factors influencing healthcare providers' attitude and willingness to use information technology in diabetes management. *BMC Med Inform Decis Mak* 2021; 21(1): 24.

7. Bundesministerium für Gesundheit. Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation (Digitale-Versorgung-Gesetz—DVG). *Bundesgesetzblatt Teil I* 2019; 49: 18.

8. Lavorgna L, Brigo F, Moccia M, et al. E-Health and multiple sclerosis: An update. *Mult Scler* 2018; 24(13): 1657–1664.

9. DiGA-Leitfaden, https://www.bfarm.de/SharedDocs/Downloads/DE/Medizinprodukte/diga_leitfaden.html

10. Higgins JPT, Savovic J, Page MJ, et al. Chapter 8: Assessing risk of bias in a randomised trial. In: Higgins JPT, Thomas J, Chandler J, et al. (eds) Cochrane handbook for systematic reviews of interventions: Version 6.3, 2022, www.training.cochrane.org/handbook

11. Kieser M and Hauschke M. Assessment of clinical relevance by considering point estimates and associated confidence intervals. *Pharm* 2005; 4: 101–107.

12. Heesen C, Riemann-Lorenz K, Lühmann D, et al. Multiple Sklerose: Führt die Nutzung von mhealth-Lösungen (z. B. Apps) im Selbstmanagement der Betroffenen zu besseren Ergebnissen? https://www.iqwig.de/download/ht19-03_apps-zum-selbstmanagement-bei-multipler-sklerose_hta-bericht_v1-0.pdf

13. Miller DM, Moore SM, Fox RJ, et al. Web-based self-management for patients with multiple sclerosis: A practical, randomized trial. *Telemed J eHealth* 2011; 17(1): 5–13.

14. Goodwin RA, Lincoln NB, das Nair R, et al. Evaluation of NeuroPage as a memory aid for people with multiple sclerosis: A randomised controlled trial. *Neuropsychol Rehabil* 2020; 30(1): 15–31.

15. Fischer A, Schroder J, Vettorazzi E, et al. An online programme to reduce depression in patients with multiple sclerosis: A randomised controlled trial. *Lancet Psychiatry* 2015; 2(3): 217–223.

16. Cooper CL, Hind D, Parry GD, et al. Computerised cognitive behavioural therapy for the treatment of depression in people with multiple sclerosis: External pilot trial. *Trials* 2011; 12: 259.

17. Pöttgen J, Moss-Morris R, Wendebourg JM, et al. Randomised controlled trial of a self-guided online fatigue intervention in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2018; 89(9): 970–976.

18. De Giglio L, De Luca F, Prosperini L, et al. A low-cost cognitive rehabilitation with a commercial video game improves sustained attention and executive functions in multiple sclerosis: A pilot study. *Neurorehabil Neur Repair* 2015; 29(5): 453–461.

19. Charvet LE, Shaw MT, Haider L, et al. Remotely-delivered cognitive remediation in multiple sclerosis (MS): Protocol and results from a pilot study. *Mult Scler J Exp Transl Clin* 2015; 1: 5609629.

20. Charvet LE, Yang J, Shaw MT, et al. Cognitive function in multiple sclerosis improves with telerehabilitation: Results from a randomized controlled trial. *PLoS ONE* 2017; 12(5): e0177177.

21. Campbell J, Langdon D, Cercignani M, et al. A randomised controlled trial of efficacy of cognitive rehabilitation in multiple sclerosis: A cognitive, behavioural, and MRI study. *Neur Plast* 2016; 2016: 4292585.

22. Bove R, Rowles W, Zhao C, et al. A novel in-home digital treatment to improve processing speed in people with multiple sclerosis: A pilot study. *Mult Scler* 2021; 27(5): 778–789.

23. Ehling R, Edlinger M, Hermann K, et al. Successful long-term management of spasticity in patients with multiple sclerosis using a software application (APP): A pilot study. *Mult Scler Relat Disord* 2017; 17: 15–21.

24. Ehling R, Seebacher B, Harsányi A, et al. Successful long-term management of spasticity in people with multiple sclerosis using a software application: Results from a randomized, controlled, multicenter study. *Eur J Neurol* 2022; 29(6): 1697–1707.

25. Nasseri NN, Ghezelbash E, Zhai Y, et al. Feasibility of a smartphone app to enhance physical activity in progressive MS: A pilot randomized controlled pilot trial over three months. *PeerJ* 2020; 8: e9303.

26. Penner IK, Raselli C, Stöcklin M, et al. The fatigue scale for motor and cognitive functions (FSMC): Validation of a new instrument to assess multiple sclerosis-related fatigue. *Mult Scler* 2009; 15(12): 1509–1517.

27. Cella M and Chalder T. Measuring fatigue in clinical and community settings. *J Psychosom Res* 2010; 69: 17–22.

28. Fricke U and Apolinario—Hagen J. E—Health—Dienste und M—Health Apps zur Verbesserung

des Selbstmanagements bei Multipler Sklerose—Ein Review zur aktuellen Forschungsliteratur zur Online-Selbsthilfe für MS-Betroffene. *eBeratungsNet* 2016; 12(2): 59.

29. Gold SM, Friede T, Meyer B, et al. Online intervention to reduce depressive symptoms in multiple sclerosis: An international multicenter randomised controlled phase III trial. *Lancet Dig Health* 2023 (in press).

30. Twomey C, O'Reilly G and Meyer B. Effectiveness of an individually-tailored computerised CBT programme (Deprexis) for depression: A meta-analysis. *Psychiatry Res* 2017; 256: 371–377.

31. Farrar JT, Troxel AB, Stott C, et al. Validity, reliability, and clinical importance of change in a 0-10 numeric rating scale measure of spasticity: A post hoc analysis of a randomized, double-blind, placebo-controlled trial. *Clin Ther* 2008; 30: 974–985.

32. Dillenseger A, Weidemann ML, Trentzsch K, et al. Digital biomarkers in multiple sclerosis. *Brain Sci* 2021; 11(11): 1519.

33. Haase R, Vogt I, Scholz M, et al. Profiles of eHealth adoption in persons with multiple sclerosis and their caregivers. Brain Sci 2021; 11(8): 1087.

34. Eysenbach G. CONSORT-EHEALTH. Improving and standardizing evaluation reports of web-based and mobile health interventions. *J Med Internet Res* 2011; 13(4): e126.

35. Spath C, Hapke U, Maske U, et al. Characteristics of participants in a randomized trial of an Internet intervention for depression (EVIDENT) in comparison to a national sample (DEGS1). *Internet Interv* 2017; 9: 46–50.

36. Freedland KE, King AC, Ambrosius WT, et al. The selection of comparators for randomized controlled trials of health-related behavioral interventions: Recommendations of an NIH expert panel. *J Clin Epidemiol* 2019; 110: 74–81. *(Continued)*