Title

Pragmatic phase II clinical trial to improve depression care in a real-world diverse MS cohort from an academic MS centre in Northern California: MS CATCH study protocol.

Journal

BMJ Open, 14(2)

Authors

Henderson, Kyra
Reihm, Jennifer
Koshal, Kanishka
et al.

Publication Date

2024-02-24

DOI

10.1136/bmjopen-2023-077432

Peer reviewed

BMJ Open

Pragmatic phase II clinical trial to improve depression care in a real-world diverse MS cohort from an academic MS centre in Northern California: MS CATCH study protocol

Kyra Henderson ,[1] Jennifer Reihm,[1] Kanishka Koshal,[1] Jaeleene Wijangco,[1] Nicolette Miller,[1] Narender Sara,[1] Marianne Doyle,[1] Alicia Mallory,[1] Judith Sheridan,[2] Chu-Yueh Guo,[1] Lauren Oommen,[1] Anthony Feinstein,[3] Christina Mangurian,[4] Ann Lazar,[5] Riley Bove[1]

To cite: Henderson K, Reihm J, Koshal K, *et al*. Pragmatic phase II clinical trial to improve depression care in a real-world diverse MS cohort from an academic MS centre in Northern California: MS CATCH study protocol. *BMJ Open* 2024;**14**:e077432. doi:10.1136/bmjopen-2023-077432

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2023-077432).

Received 04 July 2023
Accepted 25 January 2024

For numbered affiliations see end of article.

Correspondence to
Dr Riley Bove;
Riley.bove@ucsf.edu

ABSTRACT

Introduction Depression occurs in over 50% of individuals living with multiple sclerosis (MS) and can be treated using many modalities. Yet, it remains: under-reported by patients, under-ascertained by clinicians and under-treated. To enhance these three behaviours likely to promote evidence-based depression care, we engaged multiple stakeholders to iteratively design a first-in-kind digital health tool. The tool, MS CATCH (Care technology to Ascertain, Treat, and engage the Community to Heal depression in patients with MS), closes the communication loop between patients and clinicians. Between clinical visits, the tool queries patients monthly about mood symptoms, supports patient self-management and alerts clinicians to worsening mood via their electronic health record in-basket. Clinicians can also access an MS CATCH dashboard displaying patients' mood scores over the course of their disease, and providing comprehensive management tools (contributing factors, antidepressant pathway, resources in patient's neighbourhood). The goal of the current trial is to evaluate the clinical effect and usability of MS CATCH in a real-world clinical setting.

Methods and analysis MS CATCH is a single-site, phase II randomised, delayed start, trial enrolling 125 adults with MS and mild to moderately severe depression. Arm 1 will receive MS CATCH for 12 months, and arm 2 will receive usual care for 6 months, then MS CATCH for 6 months. Clinicians will be randomised to avoid practice effects. The effectiveness analysis is superiority intent-to-treat comparing MS CATCH to usual care over 6 months (primary outcome: evidence of screening and treatment; secondary outcome: Hospital Anxiety Depression Scale-Depression scores). The usability of the intervention will also be evaluated (primary outcome: adoption; secondary outcomes: adherence, engagement, satisfaction).

Ethics and dissemination University of California, San Francisco Institutional Review Board (22-36620). The findings of the study are planned to be shared through conferences and publishments in a peer-reviewed journal. The deidentified dataset will be shared with qualified collaborators on request, provision of CITI and other certifications, and data sharing agreement. We will share the results, once the data are complete and analysed, with the scientific community and patient/clinician participants through abstracts, presentations and manuscripts.

Trial registration number NCT05865405.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ A novel tool designed with extensive stakeholder engagement to deliver data from patients between visits to the point of care and to alert clinicians of severe depression in real-time.
⇒ Numerous resources are made available to both patients and clinicians for comprehensive, behaviourally informed depression care in multiple sclerosis (MS).
⇒ A pragmatic trial design enrolling diverse people with MS with mild-moderately severe depression.
⇒ One delayed-start arm allowing 6 months of observing and defining 'usual care'.
⇒ Neurologists randomised to delayed start could experience contamination if some patients begin months 6–12 before others.

INTRODUCTION

Despite the high burden of illness, there is a large depression care gap for patients with multiple sclerosis (MS) who experience depression. Depression is prevalent among people with MS (PwMS): PwMS are 2–3 times more likely to be diagnosed with a depressive disorder than the general population, with depression affecting about 50% PwMS over the course of their lifetime.[1] Clinical guidelines for MS care include depression management, for which non-sedating antidepressant drugs, suicidal intent screening, and psychotherapy are all usually recommended.[2 3] Yet despite the high prevalence

of evidence-based screening and treatment modalities—depression in MS remains under-reported, under-evaluated and under-treated.[4]

Barriers to reporting, evaluating and treating depression are many. To treat depression in the MS neurology clinic, it must be reported by the patient and screened for by the clinician, then it must be comprehensively evaluated to establish patient safety and implement interventions likely to be effective. In clinical settings, stigma may further reduce a patient's comfort with sharing their psychiatric symptoms.[1 4] Further, the standardised validated tools for depression screening are seldom used to identify depression in MS, reducing the effectiveness of interventions.[5] Treating depression in PwMS often requires a complex, individualised plan rather than a one size fits all approach. When other contributing MS-specific conditions are not fully considered, such as fatigue, cognitive impairment and urinary retention—these may interfere with the efficacy of treatments tested in non-MS patients and could worsen symptom burden.[5] In addition, antidepressants are prescribed often at inadequate dose, duration and ineffective mechanism of action.[6] After the visit, patients may face barriers to following through on treatment plans that they develop with their doctor, such as difficulties with access, insurance or finding specialists close to home, as well as the many competing demands on their time. Finally, although better integration is associated with improved psychiatric care, integration between psychiatric and non-psychiatric care in the electronic health record (EHR) is often low.[7]

MS CATCH (Care technology to Ascertain, Treat and engage the Community to Heal depression in patients with MS) is a behaviourally informed, digital health, closed-loop intervention that seeks to overcome some of the aforementioned challenges (figure 1). In simple terms, the tool includes a patient-facing mood survey and self-management tool, that loops with a comprehensive depression management dashboard that launches from the electronic medical record. This intervention avoids a 'one-size-fits-all' treatment model common in trials targeting MS symptoms, and focuses not on a specific new device (eg, app or neurostimulation) or medication (eg, antidepressant), but rather on improving information flows between patients, clinicians and depression-relevant resources. MS CATCH was designed using several established principles. The first was an extensive process of human-centred design in all phases of development to refine the overarching infrastructure. To promote sustainability and generalisability, the tool was developed using industry standards, such as REDCap for electronic data capture,[8] and Substitutable Medical Applications and Reusable Technologies on Fast Healthcare Interoperability Resources application programming interfaces.[9] Each individual care component and visualisation was then developed and refined using extensive stakeholder engagement. Throughout, decisions were made based on how they might promote changes in behaviours likely to improve depression reporting, screening, comprehensive treatment and follow-through. For this, the COM-B principles of behavioural change were followed. This extensive, iterative process is described elsewhere.[10–12] In its final version, the tool was rated as easily implementable, useful and valuable to screen for and catch depression and mood related symptoms in MS prior to the worsening health of a patient.

Trial objectives

The hypothesis tested is whether MS CATCH reduces depressive symptoms in PwMS by closing the communication loop between patients and clinicians. The goal of this single-site pragmatic clinical trial is to evaluate the effectiveness of MS CATCH to inform effect size, trial duration and expected 'adherence' for a possible larger, multicentre trial. Trial outcomes include: behaviours intended to reduce depression, patient-reported depression scores, clinical diagnoses of depression and adoption of/adherence to the tool.

METHODS AND ANALYSIS

Overall study design

MS CATCH is a phase II 12-month pragmatic clinical treatment trial enrolling diverse PwMS with mild to moderately severe self-reported mood using simple stepped wedge design (figure 2).

Patient and public involvement

Patients were involved in the design of the study as stakeholders.

Setting

The primary clinical setting is the University of California, San Francisco (UCSF) MS and Neuroinflammation Center in Northern California, USA, which provides specialised care to over 5000 adults with MS annually, and has an extensive track record of pivotal trials for MS,[13 14] remote monitoring[15 16] and treatment.[17] The start date is 1 August 2023. The end date is expected to be 15 May 2025.

Participants

MS clinicians and MS patients will be enrolled and sign written electronic informed consent forms. Clinicians will then be randomised 1:1 to arm 1 or arm 2 using an automated permuted block randomisation scheme to start using MS CATCH.

Intervention

MS CATCH consists of a simple patient-facing mood reporting and self-management tool (patient-facing electronic data capture of Patient Health Questionnaire-9 (PHQ-9)) that can be accessed via a smartphone, tablet or web browser as well as a comprehensive dashboard that launches from the medical record. The process for design of these intervention components is more fully described elsewhere.[18]

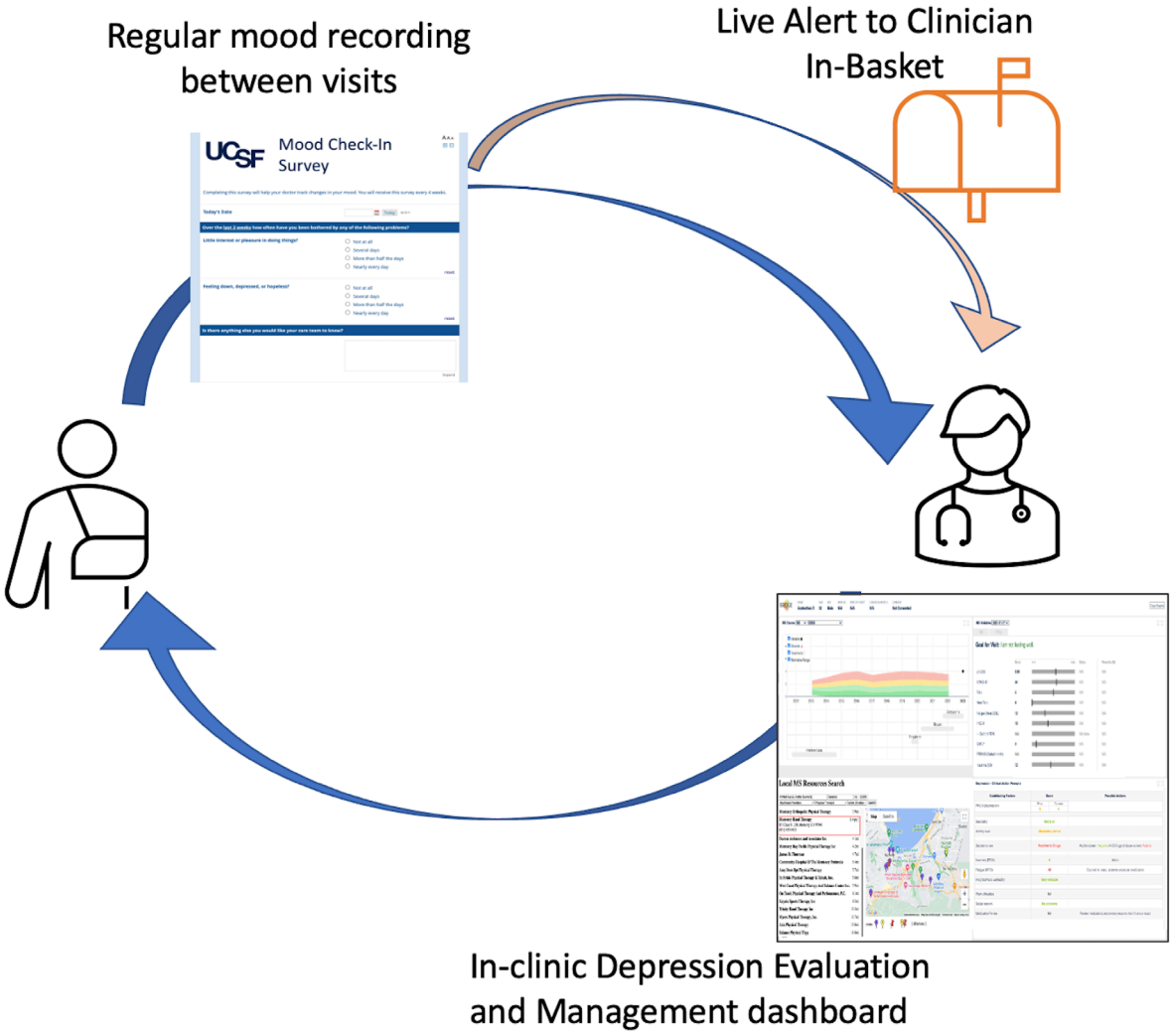

Figure 1 Care technology to Ascertain, Treat and engage the Community to Heal depression in patients with multiple sclerosis, a closed-loop depression management tool for use in multiple sclerosis. Between clinical visits, at 4-week intervals, participants complete a brief mood survey through the secure University of California, San Francisco (UCSF) REDCap portal. These surveys are emailed to ensure accessibility from any device at any time. Scores indicating severe mood trigger an alert directly to their physician's electronic health record (EHR) in-basket. During the medical encounter or between visits, the physician can access a dashboard via one click on the EHR that displays regarding a participant's current mood status. Clinicians can then open BRIDGE through the EHR to see treatment recommendations, participant health information, contributing risk factors and resources local to the patient.

Patient-facing tool

Monthly, participants receive a reminder to access their tool and complete the PHQ-9. This tool was chosen to be used as the screening tool for its adaptability and simplicity (can expand from two to nine questions), prior validation[19] including in MS, and to maintain consistency with the pilot study.[18] PHQ-9 survey responses then trigger real-time alerts delivered to the clinician's medical record inbox, who can then access a comprehensive dashboard. The patient-facing tool also displays simplified mood tracker each month and a list of resources relating to mental health. If a patient participant screens positive for suicidality or responds with a score above 15 on the PHQ-9, the participant receives an immediate alert through the tool that states 'If you are having thoughts of harming yourself or others, dial 988 to reach the National Suicide Prevention Lifeline, or seek immediate medical care. Please note that it can be up to 36 hours before your doctor receives your survey responses'.

Comprehensive clinical dashboard

The comprehensive dashboard accessed by the clinician displays the patient's longitudinal mood scores, multimodal risk factors and interventions to be considered, as well as mental health resources, including those supporting patient self-management and a map of clinicians providing comprehensive services for PwMS local to the patient, sourced from the National MS Society

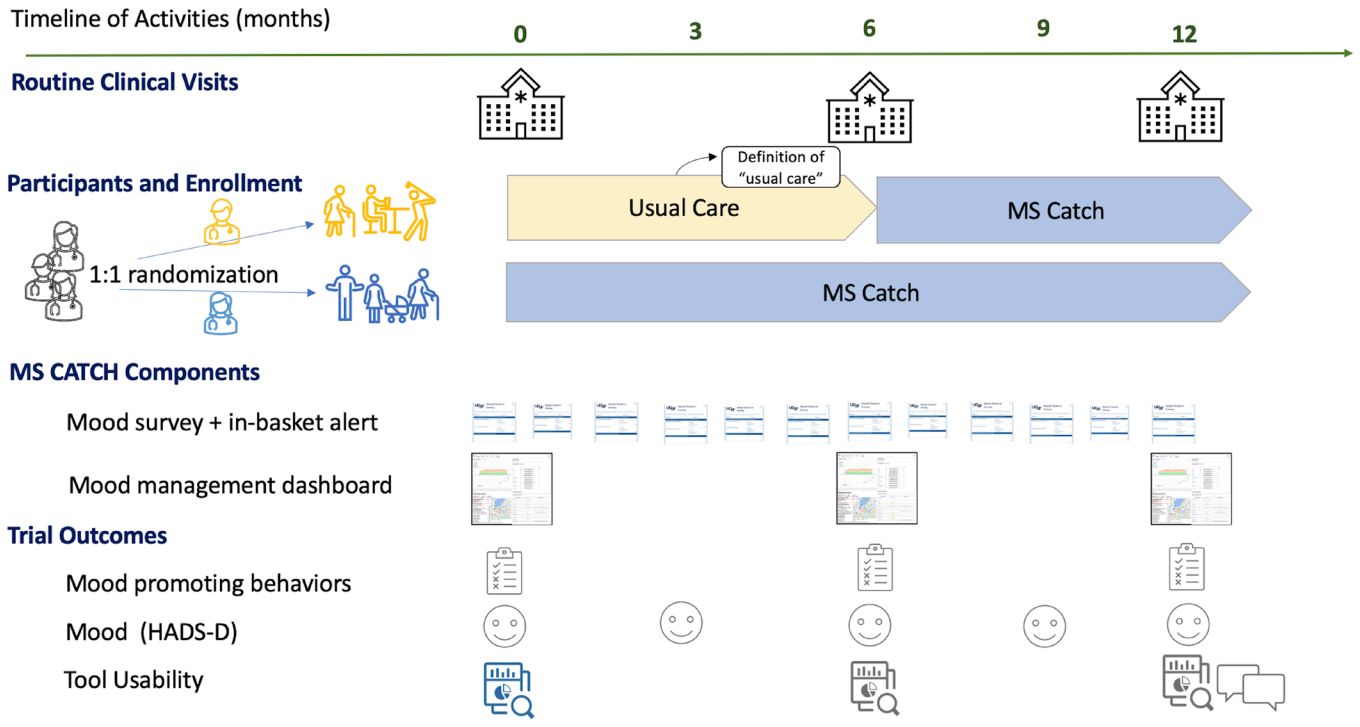

Figure 2 Overview of trial design and outcomes. A 12-month, delayed-start, randomised phase II clinical trial evaluating the efficacy (behaviours, mood) and usability of a closed loop tool for comprehensive real-time depression care in MS. HADS-D, Hospital Anxiety Depression Scale-Depression; MS, multiple sclerosis; MS CATCH, Care technology to Ascertain, Treat and engage the Community to Heal depression in patients with MS.

(figure 1). Since there can be some overlap between PHQ-9 items and other domains affected by MS such as fatigue or concentration, clinicians will also have access to the individual PHQ-9 item responses to gauge severity and best treatment options.

Design rationale

Patient participation in the trial will last 12 months, with each participant receiving at least 6 months of the intervention. Randomisation to arm 1 or arm 2 will occur by clinician, since (1) it is difficult to blind clinicians to tool use and (2) to prevent 'contamination' (ie, practice changes incurred as a result of using MS CATCH). The 6-month 'usual care' period for the delayed start group will allow us to define 'usual care' in a consistent way and to appropriately compare MS CATCH to usual care. This is important to do for a real-world clinic, given that the sparse guidelines for mood care in MS[6] are heterogeneously followed in typical clinical practice.[3] The 12-month trial will: (a) inform the feasibility of a larger randomised controlled clinical trial (RCT) and (b) allow us to account for a potential therapeutic lag between the time that interventions are recommended (eg, consult talk therapist) and a treatment duration sufficient for effect (eg, identity a therapist in-network, schedule intake, attend 12 weeks of talk therapy) has occurred. Further, (c) we will also be able to compare the effect of 6 versus 12 months of tool use on depression outcomes, that is, whether benefits incurred over months 6–12 exceed those over months 0–6 of tool use.

Patient recruitment, enrolment and screening

From the participating neurologists' practices, adult patients with MS—whether relapsing or progressive[20]—with upcoming clinical visits will be prescreened via chart review. Patients who fulfil the initial study inclusion criteria will be contacted by IRB-approved means (telephone or email). The study is also posted on clinicaltrials.gov. Interested patients will be scheduled for an enrolment visit with the study coordinator either in-person, or via telephone or Zoom (an institutionally approved, HIPAA-secure, televideo platform[9 10]). During this visit, they will complete a detailed informed consent (Docusign, paper on request) with the study team trained in ethical research conduct. In the rare event that there is concern, a potential participant's capacity will be evaluated using the University of California Brief Assessment of Capacity to Consent.[21 22] Patients who cannot or will not consent will *not be enrolled* as they would be unlikely to be able to adhere to the study protocol and personally respond to PROs. Any participant will be able to discontinue the study at any time.

Inclusion/exclusion criteria

Once they sign the informed consent, participants will be asked to complete a PHQ-9 questionnaire. This tool

Table 1 Major inclusion and exclusion criteria for patients

Inclusion criteria	Exclusion criteria
► Diagnosis of MS (relapsing or progressive) by 2017 McDonald criteria[45] ► Ages 18–80 (by going much beyond 80, you run the risk of including cognitively compromised people with another diagnosis ► PHQ-9 score of 5–19 ► Any MS therapy, or no treatment ► California resident to enable clinical telemedicine visits if warranted during the study visit	► Cognitive dexterity or visual impairment (typically defined as corrected acuity less than 20/70) that, in the opinion of the study neurologist (RB), would put the participant at risk or limit their ability to adhere to the study protocol ► Inability to provide informed consent ► Psychotic disorders: bipolar disorder, schizophrenia, schizoaffective disorder ► Substance use that in the treating neurologist's perspective could influence the patient's safety on study or adherence to study protocol ► Another comorbid central nervous system diagnosis, for example, traumatic brain injury

MS, multiple sclerosis; PHQ-9, Patient Health Questionnaire-9.

is a validated, self-administered survey used in the field of mental health that assists clinicians in objectively assessing the severity of depression. It has sensitivity and specificity values, and is comprised the nine criteria that comprise the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition depressive disorder diagnosis.[23] Given its validity and its brevity, this will be the screening tool to confirm inclusion criteria and monitor patients remotely. Patients with mild (scores: 5–9), moderate (scores: 10–14) or moderately severe (scores: 15–19), depression will be included. Given existing literature,[24 25] we anticipate that about half of the clinical patients will fall in these categories. Participants with severe depression (ie, PHQ-9 scores 20+) will result in the study team alerting their primary clinician to support safety; they will not be enrolled. Other inclusion/exclusion criteria are presented in table 1. Patients meeting these criteria will be formally enrolled in the study.

Plans to recruit a diverse research cohort

To ensure demographic representativeness, enrolment will be capped at 65% white non-Hispanic patients to ensure that trial participants reflect the composition of the diverse clinic population (30–35% of whom identify as Black, Hispanic, Asian or 'other'), and that the data are generalisable to the experiences of diverse PwMS. Enrolment is expected to be approximately 2:1 to 3:1 female to male, reflecting the clinical MS population.

Language. The REDCap mood survey can be viewed on the patient's device in the other major languages spoken by UCSF patients, namely Spanish or Cantones.

Device access. To reduce the likelihood of excluding participants based on socioeconomic status, participants who do not have a smartphone, laptop, tablet or computer in their homes will be provided with a device and/or Wi-Fi card for the duration of the study.

Study blinding

In this pragmatic trial, neither patients, clinicians nor patient-facing research coordinators will be blinded to tool use due to the difficulties of blinding, including lack of a comprehensive 'sham' tool. To minimise potential for bias, all patient-reported and physician-reported scores will be automated using REDCap to reduce any potential for subjective scoring. Further, the coordinator conducting chart review to extract data (such as depression mention in the clinical note, prescriptions filled) will be blinded to clinician group assignment. Finally, data analysis will be performed by our statistician who is blinded to treatment assignment until the dataset is locked and analysed.

Study procedures

Arm 1 participants will use the tool for the entire 12-month follow-up period, while arm 2 participants will do this for the latter 6 months. Clinical visits (whether in person or via telemedicine) will occur per usual routine at baseline, 6 months and 12 months. Over the course of the study, clinicians will have access to raw data scores of the PHQ-9 and will then provide care per their clinical judgement, using as indicated the resources available through MS CATCH, including treatment prescriptions, referrals and follow-up appointments. The nursing team will monitor the clinician's in-basket per clinical routine and alert the clinician to in-basket messages triggered by a patient's severe PHQ-9 scores. The clinic's nurses, social worker, neuropsychologist and other professionals will provide evaluation, treatment or follow-up assistance as per the treating neurologist's discretion. Patients may on their end, access any other available self-management tools of their choosing. PROs will be collected via REDCap at the baseline, 6M and 12M timepoints as well as 3M and 9M. Tool use will be monitored over the course of the study, including patient behaviour (response to prompts to the REDCap mood survey), and clinician behaviour (launching of the MS CATCH clinician-facing dashboard during or in between the clinical visits, tracked via Google Analytics). After the 12-month visit, an exit interview will be performed and participants will rate the tool. To ensure participant safety, patients with severe HADS-D scores at screening or participants with severe depression

or suicidality at any point during the trial will be referred to the study psychiatrist (CM) for expedited psychiatric care.

Study measures to be collected

Baseline covariates from EHR

Demographics

Age, sex, race, ethnicity, employment, education and zip code to estimate socioeconomic status indicators[26] and travel distance from clinic.

Clinical

MS history (symptom onset, date of diagnosis, MS type), comorbidities, smoking, cannabis use,[27] medications. Clinical disability measures will include the EDSS[28 29] (if not available within the past year, a validated e-PR-EDSS[30]) and Symbol Digit Modalities Test.

Psychiatric history

History of psychiatric and/or substance use disorder diagnoses including pseudobulbar affect, history of treatment (psychological and psychotropic medications).

Health literacy

eHealth literacy (eHEALS)[31] measures patients' combined knowledge, comfort and perceived skills at finding, evaluating and applying electronic health information to health problems.

Trial outcomes: effectiveness

Primary effectiveness outcome

Mood promoting behaviours (process of care)

At the baseline, 6 and 12-month visits, the study record will be updated for information on behaviours—both on the part of the patient and the clinician—that can promote mood evaluation and care. Data about care utilisation from the UCSF EHR and other institutional medical records will be complemented by the patient exit survey, since some aspects of clinical care (eg, out of pocket talk therapy) will not necessarily figure in the EHR. An overview of these outcome categories is provided in table 2 and detailed in online supplemental table 1. Clinician screening of depression as documented in the clinical note in the EHR will be performed at each (baseline, 6 and 12 months) visit, along with a comprehensive list of contributors to depression (substance use, inactivity, medications). Preventative care recommendations from the clinician will be noted, and patient follow-through will be calculated by study staff as a percentage of these recommendations that were completed by the next 6M follow-up visit.

Secondary effectiveness outcome

Mood

The Hospital Anxiety Depression Scale (HADS) is widely used to assess psychological distress in non-psychiatric

Table 2 Major categories of care utilisation and follow-through, including mood-promoting behaviours that are targeted by the intervention and that represent effectiveness outcomes for the trial

Source(s)	Data type	Target outcomes
Patient exit interview	**Behavioural outcomes:** referrals completed, prescriptions filled, activity level Missed work, work productivity	► Patient follow-through: 6M follow-through of mutually agreed on interventions (eg, start medication, schedule visit with local therapist) to improve mood, as per Block *et al*[10]
Medical record/ patient survey (q6M) prospectively collected by a data analyst blinded to clinician/patient randomisation	Hospitalisation, ER visits, outpatient visits Orders: medication prescriptions and refills, Referrals to psychiatrist, psychotherapist, social worker, neuropsychologist, other behavioural health professional; other medical services. Non-medical approaches recommended (join support group or club)	
	Behavioural outcomes: EHR neurologist notes: mood screening, mood evaluation, mood treatments	► Clinician screening: % visits where mood screening was documented by the clinician in the EHR
		► Clinician comprehensive depression evaluation: % depression risk factors evaluated at each patient visit (goal: 80%) ► Clinician preventive care recommendations: % visits in which applicable care was recommended (goal: 80% visits), and the % applicable interventions recommended per visit

The cells shaded in grey depict the primary behavioural change outcomes for the trial.
EHR, electronic health record.

patients. It consists of two subscales, measured via 14 items: seven items for the Anxiety subscale (HADS-Anxiety) and seven for the Depression (HADS-Depression) subscale.[32] It has been validated for use in MS demonstrated satisfactory psychometric properties in several different populations.[2] Each item is scored on a response scale with four alternatives ranging between 0 and 3. A higher score indicates greater anxiety or depression.

Thresholds

The clinically meaningful HADS-D cut-off in MS is ≥8.0.[2] The minimal clinically important difference (MCID) in HADS-D has not been defined in MS but in other chronic conditions ranges 1.5–17.[33 34] The SD of HADS-D scores in a recent survey of PwMS was 3.7.[35] The HADS *will be* collected at baseline and at 3, 6, 9 and 12 months.

Exploratory effectiveness outcomes: other self-reported measures

Self-efficacy

Patient self-efficacy at managing a chronic disease will be self-reported at the baseline, 6-month and 12-month visits.[36] Clinician self-efficacy at evaluating and managing mood will be evaluated at their enrolment and final patient out, on a 1–5 Likert scale as per.[37]

Mood

Mini-International Neuropsychiatric Interview (MINI): This is a 15 min structured assessment of depression that can be administered by a non-clinician. This will be an exploratory endpoint given project focus on the behavioural modifiers of this ultimate clinical outcome (depression).[4] Changes in PHQ-9 will also be evaluated.

Possible contributors to mood/covariates

Fatigue (Modified Fatigue Index[38]), sleep (Pittsburgh Sleep Quality Index),[39–41] social engagement (Short form Oxford Participation and Activities Questionnaire[42]) and work participation for employed participants (Impact on Participation and Autonomy[43]) will also be collected and analysed.

Trial outcomes: tool evaluation

To evaluate tool adoption, engagement, adherence and satisfaction, we will use an implementation science framework which is depicted in table 3. We will gauge *adoption* based on the percentage of patients using the tool during the first month of the study, *engagement* based on the percentage of patients who use the tool at least once every 3 months, and *adherence* based on the percentage of patients who respond to the mental health prompts at 12 months. Adherence will be seen as a participant responding to 75% or more of the depression prompts. We will also measure the engagement of patient-clinician encounters who use the tool during the 6 and 12-month visits. Trial retention will also be calculated at the 3, 6, 9 and 12-month timepoints.

Data collection and storage

Data will be collected through the institutionally secured HIPAA-compliant REDCap portal.[8] All information

Table 3 Implementation science-informed framework for evaluating the trial intervention: MS CATCH

	Patients	Clinicians
Adoption (uptake)	% who respond to the survey during the initial month of study (goal 80%)	% patient-clinician dyads who use the in-visit dashboard at the 6M follow
	% approached for the study who initially agreed to 'adopt' MS CATCH % patients screened for the study who were not able to participate due to cognitive, motor or visual impairments	% willing to participate in the study
Engagement (sustained use)	% who continued to use the patient-facing tool at least quarterly (goal 60%)	% clinician-patient dyads in arm 1 who use the in-visit dashboard at the 12M clinical visit (goal 60%)
Adherence	% depression-reporting prompts adhered to per participant on their patient-facing tool (goal 75%)	
	% adhering to >75% depression prompts (goal 75%)	
	Trial retention at 3, 6, 9 and 12 months	
Satisfaction	Net Promoter Score[46]	Net Promoter Score[46]
	12 item mHealth Satisfaction Questionnaire[47]	
	Qualitative (exit interview, or early exit): patterns of use, barriers to use, and ask for feedback about the tool's functionality; if applicable, dropout and to elicit any additional qualitative feedback on their depression care	Qualitative: satisfaction with the tool (also using perceived barriers to use, feedback about the tool's functionality, and to estimate the overall burden of use of MS CATCH

The primary usability outcomes are shaded in grey (tool adoption by patients and clinicians).
MS CATCH, Care technology to Ascertain, Treat and engage the Community to Heal depression in patients with multiple sclerosis.

relating to patients will be stored behind the UCSF firewall.

Statistical plan and data analysis

Effectiveness outcomes

The primary effectiveness outcome will be a behavioural change outcome. For the clinician, this will be descriptive only and will be the % visits in which they screened for depression. For the patient, this will be categorised as the % of mutually agreed on interventions (eg, start medication, schedule visit with local therapist) to improve mood (as per Block *et al*[10]) that were followed through by the patient at their 6-month visit. Additional behavioural outcomes are summarised in table 2 and detailed in online supplemental table 1. The secondary effectiveness outcome is mood improvement (change on the HADS-D). We will use 1.7 for an MCID and compare the % participants in each arm achieving improvement of at least 1.7 over the first 6 months of the trial. Additional patterns of HADS-D improvement will be evaluated, including mean change in HADS-D score over the course of the trial, and categorical improvement, such as reduction in HADS-D scale to <8; reduction in HADS-D scale by 2+ points). Exploratory outcomes will include MINI scores, self-efficacy, and other patient-reported outcomes.

The Full Analysis Set comprises all doctors to whom trial intervention has been randomly assigned 1:1 to arm 1 or arm 2 by block randomisation. According to the intent-to-treat principle, doctors will be analysed according to the intervention arm they were randomised to and strata (employment status: part-time (<50%) vs full-time (≥50%)) they have been assigned to during the randomisation procedure. The primary behavioural effectiveness outcome is depression screening at 6 months. The following statistical hypotheses will be tested to address the primary objective: H_o:θ_1=0 vs HA: $\theta_1 \neq 0$ where θ_1 is the difference in proportions between usual care and CATCH at 6 months. The primary analysis to test this hypothesis and compare the two treatment arms will use the adjusted χ^2 analysis via the Rao-Scott design that adjusts for clustering of doctors via SAS V.9.4 proc surveyfreq stratified by employment status (full-time vs part-time (<50%)). A similar approach will be undertaken examining other effectiveness outcomes.

We will also use mixed effects linear (for continuous outcomes) and logistic (for dichotomous secondary outcomes) regression analyses. We will flexibly model trajectories by testing whether including quadratic or cubic terms for time (up to five visits: baseline, 3 months, 6 months, 9 months, 12 months) or random slopes for individuals improve the model fit and include them if indicated by a significant (p<0.05) likelihood ratio test. The overall difference will be assessed using an F-test and post-estimation t-test in SAS V.9.4. Using this approach, we can assess whether intervention improvement occurs during different time-periods to assess duration of use of CATCH. We will present the results with and without a multiple testing adjustment when comparing across time periods. All of the secondary efficacy outcomes will be assessed using this mixed modelling approach. We will also adjust for important covariates (eg, PHQ-9 levels, age, sex, EDSS), as well as explore potential interactions between intervention arm and PHQ-9, and other covariates. We will also assess the model fit (eg, residuals) and assess whether transforming the outcomes (eg, log transformations) provides the best fit. We will also assess whether the baseline values are subject to confounding by isolating within person changes. The benefit of mixed-effects models is that they produce unbiased estimates even when some individuals have missing observations, adjust for differential loss to follow-up, accommodate irregular time measurements, and account for clustering of individuals, as required in this study. For all results, the estimates and associated 95% CIs will be reported.

Tool usability outcomes

Tool usability outcomes are detailed in table 3; the primary usability outcome will be adoption. To evaluate the tool itself, descriptive statistics (mean, SD, median and IQR) for numeric will be used.

Sample size considerations

Behavioural outcomes

125 patients will be enrolled in each arm to ensure at least 60 patients with PHQ-9>5 in each arm who can participate in the trial (ie, 250 total sign informed consent and 125 total participate), with 50 anticipated completers in each arm (82% retention at 6 months). In our experience, this sample size is ample to identify significant predictors of adherence to a digital health intervention and to ensure reasonable completers should there be unforeseen low study retention.[15 44] Based on our preliminary data, we expect a small design effect of 1.72 with an intraclass correlation coefficient of 0.45 (1+0.45×(17−1)). Using this design effect, our effective sample size is 118, which results in at least 80% statistical power to detect an improvement in our primary outcome (clinician depression screening) from 25% visits in the usual care arm to 50% in the MS CATCH arm at the 6-month visit using a χ^2 test assuming a two-sided a level of 0.05.

Mood outcomes

For our secondary outcome, change in HADS-D, an effective sample size of 114 (57 per group, assuming design effect of 1.72, see above) will provide approximately 80% statistical power to detect a clinically meaningful difference of 1.7 at 6 months, assuming a change from baseline in HADS-D of 1.7 for arm 1 (MS CATCH intervention), and no change with usual care (arm 2), with common SD of 3.2. This sample size calculation was based on a two-sample independent t-test using Stata V.15.1, assuming an alpha level of 0.05. We did not adjust for multiple tests given in the calculation given that HADS-D is a secondary outcome. However, we plan to present results adjusted for multiple testing using the Bonferroni correction to maintain the family-wise alpha level at 0.05.

Statistical analysis plan

The detailed statistical analysis plan (SAP) will include details about handling of potential missing data, including consideration of multiple imputation if more than 5% of the data is missing at random. It is possible that the data is not missing at random, but rather missing not at random, and we will explore this possibility. To account for possible effects of crossover and contamination, we will consider performing marginal structural modelling as a sensitivity analysis. Though, given that the primary endpoints of the study occur at 6 months before arm 2 crosses over to MS CATCH, we do not expect this to be a major limitation. The SAP will follow the guidelines provided by Consolidated Standards of Reporting Trials.

ETHICS AND DISSEMINATION

The trial received ethical approval from the UCSF Institutional Review Board (22-36620) and is registered to ClinicalTrials.gov. Both patient and clinician participants will provide informed consent, either electronically, through a secured online portal, or on paper, prior to study enrolment. Results from the study will only report aggregated deidentified patient information and will be published through peer-reviewed journals and presented at conferences. Deidentified data will be shared with qualified investigators on reasonable request. Adverse events will be reported through the UCSF IRB. Should the trial achieve its intended outcomes, then key features will be tested in a large, multi-site RCT.

Author affiliations
[1]UCSF Weill Institute for Neurosciences, Department of Neurology, University of California San Francisco, San Francisco, California, USA
[2]Patient Stakeholder, University of California San Francisco, San Francisco, California, USA
[3]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada
[4]Department of Psychiatry, University of California San Francisco, San Francisco, California, USA
[5]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA

Contributors RB designed the trial, obtained study funding and aided in drafting the manuscript. KH drafted the manuscript and is involved in coordinating the ongoing trial. KK and JW collected the preliminary data and revised the manuscript. JR was responsible for trial design and conducted focus group interviews, interpreted the results from these interviews, and helped to revise the manuscript. AF, CM, MD, AM, C-YG, LO and JS contributed expertise from their respective fields through stakeholder feedback sessions and other avenues and all took part in revising the manuscript. NS and NM oversaw the iterative development of the tool and contributed revisions to the manuscript. AL created the statistical analysis plan and revised the manuscript.

Funding This trial is funded by a Department of Defense USAMRAA Clinical Trials Award (Project #: W81XWH2210894, PI: Bove).

Competing interests The authors report no disclosures relevant to this clinical trial. KH, KK, JW, JR: none. RB reports: research support from NIH, DOD, National Multiple Sclerosis Harry Weaver Award, and from Biogen, Novartis and F Hoffman-LaRoche. RB has received personal advisory board or consulting fees from EMD Serono, Horizon, Janssen and TG Therapeutics. C-YG reports: personal advisory board or consulting fees from Genentech and Horizon. CM reports: supported by several grants from the National Institutes of Health (NIMH; NIMHD; NIAID; NIDA), Health Resources & Services Administration, the Doris Duke Charitable Foundation, the California Health Care Foundation, Genentech and United Health Group.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD
Kyra Henderson http://orcid.org/0009-0006-8350-0474

REFERENCES

1 Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* 1999;89:1322–7.
2 Honarmand K, Feinstein A. Validation of the hospital anxiety and depression scale for use with multiple sclerosis patients. *Mult Scler* 2009;15:1518–24.
3 Marrie RA, Horwitz R, Cutter G, *et al*. The burden of mental comorbidity in multiple sclerosis: frequent, underdiagnosed, and undertreated. *Mult Scler* 2009;15:385–92.
4 Sheehan DV, Lecrubier Y, Sheehan KH, *et al*. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59 Suppl 20:22–33;
5 Dzewaltowski DA, Glasgow RE, Klesges LM, *et al*. RE-AIM: evidence-based standards and a Web resource to improve translation of research into practice. *Ann Behav Med* 2004;28:75–80.
6 Rae-Grant A, Bennett A, Sanders AE, *et al*. Quality improvement in neurology: Multiple sclerosis quality measures: Executive summary. *Neurology* 2015;85:1904–8.
7 Kozubal DE, Samus QM, Bakare AA, *et al*. Separate may not be equal: a preliminary investigation of clinical correlates of electronic psychiatric record accessibility in academic medical centers. *Int J Med Inform* 2013;82:260–7.
8 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
9 Bove R, Schleimer E, Sukhanov P, *et al*. Building a precision medicine delivery platform for clinics: the University of California, San Francisco, BRIDGE experience. *J Med Internet Res* 2022;24:e34560.
10 Block VJ, Gopal A, Rowles W, *et al*. CoachMS, an innovative closed-loop, interdisciplinary platform to monitor and proactively treat MS symptoms: a pilot study. *Mult Scler J Exp Transl Clin* 2021;7:2055217321988937.
11 Block VJ, Koshal K, Wijangco J, *et al*. A closed-loop falls monitoring and prevention app for multiple sclerosis clinical practice: human-centered design of the multiple sclerosis falls insighttrack. JMIR Hum Factors 2023;11:e49331.
12 Bove R, Bruce CA, Lunders CK, *et al*. Electronic health record technology designed for the clinical encounter. *Neur Clin Pract* 2021;11:318–26.
13 Montalban X, Hauser SL, Kappos L, *et al*. Ocrelizumab versus placebo in primary progressive multiple sclerosis. *N Engl J Med* 2017;376:209–20.
14 Green AJ, Gelfand JM, Cree BA, *et al*. Clemastine fumarate as a remyelinating therapy for multiple sclerosis (ReBUILD): a randomised, controlled, double-blind, crossover trial. *Lancet* 2017;390:2481–9.
15 Block VJ, Bove R, Zhao C, *et al*. Association of continuous assessment of step count by remote monitoring with disability

progression among adults with multiple sclerosis. *JAMA Netw Open* 2019;2:e190570.
16 Block VJ, Lizée A, Crabtree-Hartman E, *et al*. Continuous daily assessment of multiple sclerosis disability using remote step count monitoring. *J Neurol* 2017;264:316–26.
17 Block VJ, Gopal A, Rowles W, *et al*. CoachMS, an innovative closed-loop, interdisciplinary platform to monitor and proactively treat MS symptoms: A pilot study. *Multiple Sclerosis Journal - Experimental, Translational and Clinical* 2021;7:205521732198893.
18 Kyra Henderson B, Kanishka Koshal M, Jaeleene Wijangco BS, *et al*. *A closed-loop digital health tool to improve depression care in multiple sclerosis: iterative design and crosssectional pilot RCT impact on depression care*. 2023.
19 Kroenke K, Wu J, Yu Z, *et al*. Patient health questionnaire anxiety and depression scale: initial validation in three clinical trials. *Psychosom Med* 2016;78:716–27.
20 Thompson AJ, Banwell BL, Barkhof F, *et al*. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet Neurology* 2018;17:162–73.
21 Jeste DV, Palmer BW, Appelbaum PS, *et al*. A new brief instrument for assessing decisional capacity for clinical research. *Arch Gen Psychiatry* 2007;64:966–74.
22 Gilbert T, Bosquet A, Thomas-Antérion C, *et al*. Assessing capacity to consent for research in cognitively impaired older patients. *Clin Interv Aging* 2017;12:1553–63.
23 Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16:606–13.
24 Gotoh M, Homma Y, Yokoyama O, *et al*. Responsiveness and minimal clinically important change in overactive bladder symptom score. *Urology* 2011;78:768–73.
25 Tajali S, Shaterzadeh-Yazdi M-J, Negahban H, *et al*. Predicting falls among patients with multiple sclerosis: comparison of patient-reported outcomes and performance-based measures of lower extremity functions. *Mult Scler Relat Disord* 2017;17:69–74.
26 Berkowitz SA, Traore CY, Singer DE, *et al*. Evaluating area-based socioeconomic status indicators for monitoring disparities within health care systems: results from a primary care network. *Health Serv Res* 2015;50:398–417.
27 Artigaud L, Fener C, Bisch M, *et al*. Screening tools for cannabis use disorders and their adaptation to DSM-5: A literature review. *Encephale* 2020;46:382–9.
28 Hobart J, Freeman J, Thompson A. Kurtzke scales revisited: the application of psychometric methods to clinical intuition. *Brain* 2000;123:1027–40.
29 Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;33:1444–52.
30 Romeo AR, Rowles WM, Schleimer ES, *et al*. An electronic, unsupervised patient-reported Expanded Disability Status Scale for multiple sclerosis. *Mult Scler* 2021;27:1432–41.
31 Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. *J Med Internet Res* 2006;8:e27e27.
32 Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361–70.
33 Puhan MA, Frey M, Büchi S, *et al*. The minimal important difference of the hospital anxiety and depression scale in patients with chronic obstructive pulmonary disease. *Health Qual Life Outcomes* 2008;6:46:18597689.:.
34 Lemay KR, Tulloch HE, Pipe AL, *et al*. Establishing the minimal clinically important difference for the hospital anxiety and depression scale in patients with cardiovascular disease. *J Cardiopulm Rehabil Prev* 2019;39:E6–11.
35 Chiaravalloti ND, Amato MP, Brichetto G, *et al*. The emotional impact of the COVID-19 pandemic on individuals with progressive multiple sclerosis. *J Neurol* 2021;268:1598–607.
36 Lorig KR, Sobel DS, Ritter PL, *et al*. Effect of a self-management program on patients with chronic disease. *Eff Clin Pract* 2001;4:256–62.
37 Salinas I, Sinha N, Sen A. Androgen-induced epigenetic modulations in the ovary. *J Endocrinol* 2021;249:R53–64.
38 Fisk JD, Ritvo PG, Ross L, *et al*. Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. *Clin Infect Dis* 1994;18 Suppl 1:S79–83.
39 Veauthier C, Gaede G, Radbruch H, *et al*. Poor sleep in multiple sclerosis correlates with beck depression inventory values, but not with polysomnographic data. *Sleep Disord* 2016;2016:8378423.
40 Buysse DJ, Reynolds CF III, Monk TH, *et al*. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research* 1989;28:2748771:193–213.:.
41 Utomo E, Korfage IJ, Wildhagen MF, *et al*. Validation of the urogenital distress inventory (udi-6) and incontinence impact questionnaire (IIQ-7) in a Dutch population. *Neurourol Urodyn* 2015;34:24–31.
42 Jenkinson C, Kelly L, Dummett S, *et al*. The oxford participation and activities questionnaire (Ox-PAQ): development of a short form and index measure. *Patient Relat Outcome Meas* 2019;10:227–32.
43 Vazirinejad R, Lilley JM, Ward CD. The “Impact on Participation and Autonomy”: acceptability of the English version in a multiple sclerosis outpatient setting. *Mult Scler* 2003;9:612–5.
44 Bove R, White CC, Giovannoni G, *et al*. Evaluating more naturalistic outcome measures: A 1-year smartphone study in multiple sclerosis. *Neurol Neuroimmunol Neuroinflamm* 2015;2:e162.
45 Thompson AJ, Banwell BL, Barkhof F, *et al*. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol* 2018;17:162–73.
46 Fontil V, Khoong EC, Hoskote M, *et al*. Evaluation of a health information technology-enabled collective intelligence platform to improve diagnosis in primary care and urgent care settings: protocol for a pragmatic randomized controlled trial. *JMIR Res Protoc* 2019;8:e13151.
47 Melin J, Bonn SE, Pendrill L, *et al*. A questionnaire for assessing user satisfaction with mobile health apps: development using rasch measurement theory. *JMIR Mhealth Uhealth* 2020;8:e15909.