Contemporary Clinical Trials Communications 16 (2019) 100435

Contents lists available at ScienceDirect

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc

Research paper

Comparison of surf and hike therapy for active duty service members with major depressive disorder: Study protocol for a randomized controlled trial of novel interventions in a naturalistic setting

Kristen H. Walter[a,b,*], Nicholas P. Otis[a,b], Lisa H. Glassman[a,b], Travis N. Ray[a,b], Betty Michalewicz-Kragh[c], Kim T. Kobayashi Elliott[c], Cynthia J. Thomsen[b]

[a] Leidos, 140 Sylvester Road, San Diego, CA, 92106-3521, United States
[b] Health and Behavioral Sciences Department, Naval Health Research Center, 140 Sylvester Road, San Diego, CA, 92106-3521, United States
[c] Directorate of Public Health, Naval Medical Center San Diego, San Diego, CA, United States

ARTICLE INFO

Keywords:
Major depressive disorder
Surf therapy
Hike therapy
Physical activity
Military

ABSTRACT

Many active duty service members suffer from major depressive disorder (MDD). Although traditional treatments exist, alternative approaches may also be effective in treating depressive symptoms. Previous research has shown that physical activity has significant positive effects on depression symptoms in individuals with MDD, and that these benefits may be enhanced when physical activity occurs in a natural environment. Even though physical activity (i.e., hiking, walking) in natural environments has been shown to reduce depressive symptoms, water-based activity occurring in a natural environment (e.g., surfing) may produce even greater improvements in depressive symptoms. We detail an ongoing randomized controlled trial (RCT) comparing the efficacy of surf therapy and hike therapy with respect to immediate and longer-term psychological, physical, and functional outcomes in active duty service members with MDD. We describe the methodological development of this RCT evaluating novel treatment approaches and discuss considerations for evaluating physical activity interventions in a naturalistic setting.

1. Introduction

Many U.S. service members suffer from major depressive disorder (MDD; [1,2], with prevalence estimated at 8% across services [3], comparable to the 8.6% estimate for the U.S. population [4]. Individuals with MDD are at increased risk of suicidal ideation, mortality, and medical/psychological comorbidities [5,6] as compared to the general population. In addition to the significant health consequences of MDD, there are substantial financial costs associated with related medical care and loss of productivity [2].

Traditional treatment approaches, such as cognitive behavioral therapy (CBT; [7–10] and antidepressant medication [11,12] are effective in treating MDD. However, there is conflicting evidence regarding the relative benefits of CBT for MDD compared with other active treatments [13]. Furthermore, nonadherence rates for antidepressant medication may be as high as 45% [14], which can reduce their effectiveness [15] and increase the risk for relapse (e.g., Ref. [16].

Due to these observed variations in the efficacy and tolerance of traditional treatments [17,18], there has been increasing interest in other interventions for MDD, including physical activity. Exercise has been theorized to lead to increases in neurotransmitters (e.g., serotonin) and neurotrophins (e.g., brain-derived neurotrophic factor), alleviate neuroendocrine abnormalities (e.g., hypothalamic-pituitary-adrenal axis), and enhance perceived self-efficacy—all of which could lead to feelings of well-being [19]. Indeed, certain types of physical activity have demonstrated large and significant effects on depression symptoms [20], including among individuals with MDD [21–25]. Additionally, exercise may be a viable adjunctive treatment to standard care [26,27].

Physical activity occurring in the natural environment may have an especially potent influence on mental health, providing greater benefits than physical activity alone. Meta-analytic findings show that engaging in physical activity outdoors (versus indoors) has greater effects on tension, anger, and depression [28]. Furthermore, water-based

* Corresponding author. Naval Health Research Center, 140 Sylvester Road, San Diego, CA, 92106-3521, United States.
E-mail addresses: kristen.h.walter.ctr@mail.mil (K.H. Walter), nicholas.p.otis.ctr@mail.mil (N.P. Otis), lisa.h.glassman.ctr@mail.mil (L.H. Glassman), travis.n.ray2.ctr@mail.mil (T.N. Ray), betty.michalewiczkragh.civ@mail.mil (B. Michalewicz-Kragh), kim.t.elliott2.civ@mail.mil (K.T. Kobayashi Elliott), cynthia.j.thomsen.civ@mail.mil (C.J. Thomsen).

https://doi.org/10.1016/j.conctc.2019.100435
Received 7 June 2019; Received in revised form 31 July 2019; Accepted 19 August 2019
Available online 21 August 2019

Abbreviations

AIMS	Athletic Identity Measurement Scale
CBT	cognitive behavioral therapy
CSQ-8	Client Satisfaction Questionnaire-8
DSM-5	Diagnostic and Statistical Manual of Mental Disorders (5th Edition)
GAD-7	Generalized Anxiety Disorder 7-item scale
IPAQ-SF	International Physical Activity Questionnaire – Short Form
ISI	Insomnia Severity Index
MADRS	Montgomery–Åsberg Depression Rating Scale
MDD	major depressive disorder
MET	metabolic equivalent
MINI-7	Mini International Neuropsychiatric Interview 7.0
MLM	multilevel modeling
NMCSD	Naval Medical Center San Diego
NPRS	Numerical Pain Rating Scale
OSU TBI-ID-SF	The Ohio State University Traumatic Brain Injury Identification Method – Short Form
PANAS	Positive and Negative Affect Schedule
PAS	Positive Affect Schedule
PCL-5	PTSD Checklist for DSM-5
PHQ-4	Patient Health Questionnaire-4
PHQ-9	Patient Health Questionnaire-9
PTSD	posttraumatic stress disorder
RCT	randomized controlled trial
RSES-4	Response to Stressful Events Scale-4
SF-36v2®	Short Form Health Survey – 36 Item, Version 2
TBI	traumatic brain injury
WII	Wounded, Ill, and Injured

activities in a natural environment may be particularly suited to improving mental health. For example, [29] demonstrated that exercise outdoors in the presence of water (versus without water) generated greater improvements in symptoms associated with depression. These results provide support for the beneficial effects of water-based environments, which is often referred to as 'blue space' (for a systematic review, see Ref. [30]. Theoretical proponents of blue space suggest the additive benefits of water may result from its unique sensory inputs [31], such as the relaxing sounds of waves or the smell of salt water. Based on the notion of blue space, surfing—a physical activity immersed in water—could be expected to produce enhanced psychological benefits when compared to physical activity occurring outside the presence of water.

The current study compares the efficacy of two outdoor physical activity-based interventions, surf therapy and hike therapy, for service members with MDD. The robust study methodology allows for an evaluation of both surf and hike therapy, providing valuable insight into the immediate and longer-term psychological, physical, and functional effects of each therapy on MDD symptomology. Further, the randomized controlled trial (RCT) design allows for a direct comparison of water-based and land-based activity, which may provide support for the additive effects of physical activity in the presence of water. The primary study objective is to evaluate whether surf therapy yields a greater reduction in depression symptoms than hike therapy, which may be due to the salutary effects of water, or "blue space." It is hypothesized that both programs will significantly reduce depression symptoms; however, due to the added benefits of exercising in water, surf therapy is expected to result in greater improvements in depressive symptomology than hike therapy. Secondary study aims include comparisons across participants in surf and hike therapy of MDD remission rates, changes in related symptoms both over the study period and within session, drop-out rates, and self-monitored physical activity.

2. Design and methods

This prospective, longitudinal study will screen up to 125 active duty service members seeking surf or hike therapy in order to identify at least 86 eligible study participants and account for potential loss of participants to attrition (see section 2.7). Recruitment began in January 2018 and is expected to be completed within 2 years of that date. Eligible participants are randomly assigned to surf or hike therapy. Participants complete clinical interviews and self-report measures prior to, following, and 3 months after completion of the intervention; change in depressive symptom severity will serve as the primary study outcome. Additionally, participants complete brief self-report measures before and after each of the physical activity sessions in order to evaluate immediate effects. Participants are also issued a wearable activity tracker (i.e., Fitbit Charge™ 2; Fitbit Inc., San Francisco, CA) to be worn as often as possible throughout the study period from pre-program to 3-month follow-up. All study procedures were approved by the Naval Medical Center San Diego (NMCSD) Institutional Review Board.

2.1. Participants

The final sample will include 86–110 active duty service members (i.e., 43–55 participants per condition) who meet current diagnostic criteria for MDD and are seeking care at the Wounded, Ill, and Injured (WII) Wellness Program at NMCSD. Service members are screened at the initial assessment and provide voluntary written informed consent prior to beginning any study-related procedures. If service members do not consent to take part in the study, they may still participate in the surf or hike therapy programs as standardly provided.

As the study intervention is incorporated into standard practice, the current study features broad eligibility criteria. Individuals are eligible for study participation if they (a) are active duty service members seeking care at the WII Wellness Program at NMCSD, and (b) have a current diagnosis of MDD based on the Diagnostic and Statistical Manual of Mental Disorders (5th Edition [DSM-5]; [32] criteria, as assessed by the Mini International Neuropsychiatric Interview 7.0 (MINI-7; [33]. Service members who have previously received or are currently receiving surf or hike therapy are ineligible for the study; this single exclusion criterion was selected to control for dose of the intervention received. For ethical care provision reasons, we do not exclude individuals who are engaged in psychotherapy or prescribed psychotropic medication; however, data regarding other treatments received are collected so that the use and effects of these treatments can be empirically evaluated.

2.2. Treatment

The WII Wellness Program in the Health and Wellness Department at NMCSD provides novel, 6-week surf and hike therapy programs as part of standard medical treatment (i.e., a scheduled medical appointment) for active duty service members seeking care. Both surf and hike therapy programs follow the same schedule, which consists of 6 consecutive weeks of the respective program, followed by a 2-week break from programming prior to the start of the next cohort (and during which pre- and post-program assessments are completed). Both programs are considered "progressive" in that individuals in each program are anticipated to build skill and ability over the course of the 6-week program. It is important to note that the programs do not include a structured psychotherapy/therapy component; rather, engagement in the activity within the natural environment and social interaction/support are considered the therapeutic elements (e.g., Refs. [34–37].

The two programs are conducted as usual (i.e., service members receive the same surf or hike intervention regardless of whether they participate in the study), consistent with standard operating procedures. Only three modifications are made to the programs' current procedures for the study. First, clinical interview and self-report assessments are completed before and after each program, as well as 3 months after each program's conclusion, to evaluate the effects of program participation on patient outcomes. Brief self-report assessments are also completed immediately before and after each activity session. Second, only patients who currently meet diagnostic criteria for MDD will qualify for participation in the study. Third, patients will be randomized to either surf or hike therapy rather than choosing the activity in which they will participate. However, all patients will have the opportunity to receive the alternate activity after completion of the program to which they were randomized (see section 2.4). Randomization of eligible study participants to either the surf or hike therapy condition is conducted using a blocked randomization scheme (in blocks of 10) designed to yield balanced groups across the recruitment period. Once randomized, participants receive either surf or hike therapy sessions once per week for 6 weeks.

2.3. Surf therapy

A master's degree-level exercise physiologist serves as the program manager and coordinates all aspects of the surf therapy program. The program is 6 weeks in duration; sessions occur on Thursday mornings at the same Southern California beach. The program follows a cohort format that accommodates about 20–25 service members per program cycle. Each service member is paired with a volunteer surf instructor who usually works with him or her each week for the duration of the program. Goals are individually tailored to the service member using a strengths-based approach that considers the service member's abilities, skills, and comfort with the water-based environment (e.g., Ref. [36]; they can incorporate both physical and psychological objectives. All volunteer surf instructors are required to be certified through the Armed Services Young Men's Christian Association, and to attend an initial clinic orientation given by the program manager and an annual refresher training on program policies and patient safety. Volunteers also attend a 15 min debrief prior to every surf therapy session. The ratio of volunteer surf instructors to patients is 1:1 or 1:2. In addition to the volunteer surf instructors, there are approximately 2–3 staff members (including the program manager), land volunteers, and local lifeguards on site to monitor safety. For each session, participants are provided with necessary equipment, including wet suits, surfboards, boogie boards, leashes, fins, and rash guards.

Before each surf therapy session, coffee and fruit are available to all individuals affiliated with the program. An elective hour of group yoga integrating Ashtanga and Vinyasa elements is also offered at the same beach location where the surf therapy occurs. At the start of each surf therapy session, a briefing is provided that includes program policies and procedures, as well current environmental conditions and safety considerations. Participants then surf with their instructor for approximately 3 h. Following each surf session, a lunch is provided for all patients, volunteers, and staff, and socialization is encouraged. The six sessions of surf therapy follow a similar format, but participants work to develop greater proficiency to surf safely and independently.

2.4. Hike therapy

The program manager of the hike therapy program is a Certified Therapeutic Recreation Specialist who manages all aspects of the program. Similar to surf therapy, the hike therapy program is 6 weeks in duration. Hike therapy sessions occur on Wednesday mornings and take place at various locations throughout a Southern California county. The hike program is also structured in a cohort format that can include up to 20 service members per program cycle. Approximately four staff members (including the program manager and student interns) and volunteers hike alongside patients to ensure safety, yielding a maximum 1:4 ratio of staff to patients. During hike therapy, service members may hike together or at a self-selected pace. Goals are also personalized for the service member using a strengths-based approach [36] and can include both physical and mental health aims to facilitate each participant's recovery (e.g., complete one hike independently, identify community hiking groups).

During the first week, participants attend an orientation where program policies, hiking etiquette, and safety information are provided. Participants are also asked to download a specific hiking app that provides extensive trail information to their cell phones. In the orientation meeting, participants are typically divided into small groups and work together to select a Southern California trail that program participants will hike for their designated week (i.e., within weeks 3–6). Following the orientation, participants hike an easy, flat urban trail so that the physical endurance and speed of the group can be assessed. The first two locations are consistent across program cycles, but the remaining four vary depending on the preferences of the small groups. Hike therapy sessions are approximately 3 h in duration, with trails

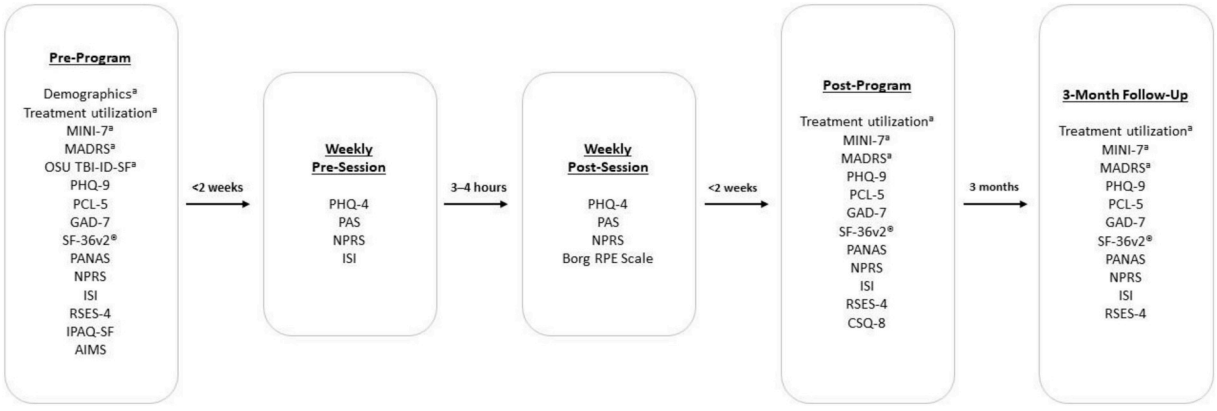

Fig. 1. Time course of data collection and measures used. Both surf and hike therapy programs are 6 weeks in duration. [a]Assessor-administered. AIMS = Athletic Identity Measurement Scale; Borg RPE Scale = Borg Rating of Perceived Exertion Scale; CSQ-8 = Client Satisfaction Questionnaire-8; GAD-7 = Generalized Anxiety Disorder 7-item scale; IPAQ-SF = International Physical Activity Questionnaire – Short Form; ISI = Insomnia Severity Index; MADRS = Montgomery–Åsberg Depression Rating Scale; MINI-7 = Mini International Neuropsychiatric Interview 7.0; NPRS = Numerical Pain Rating Scale; OSU TBI-ID-SF = The Ohio State University Traumatic Brain Injury Identification – Short Form; PANAS = Positive and Negative Affect Schedule; PAS = Positive Affect Schedule; PCL-5 = PTSD Checklist for DSM-5; PHQ-4 = Patient Health Questionnaire-4; PHQ-9 = Patient Health Questionnaire-9; RSES-4 = Response to Stressful Events Scale-4; SF-36v2® = Short Form Health Survey – 36-Item, Version 2.

becoming progressively more challenging across the 6 weeks. After the last hike at week 6, there is a potluck to celebrate program completion.

2.5. Assessments

Participants provide voluntary, written informed consent before engaging in any study assessments. Upon consent, participants complete a diagnostic, pre-program assessment (administered in person by an independent evaluator/study assessor) that includes both clinical interview measures and self-report questionnaires assessing MDD and other related symptoms. The study assessor is blind to participants' randomized condition. Participants are also provided with a wearable activity tracker (Fitbit Charge 2) along with instructions on how to use the device. They are informed that they will be permitted to keep the device if they use it for more than 50% of the days between the pre-program and 3-month follow-up assessments. Within 2 weeks of the pre-program assessment, patients begin the program in the modality to which they were randomly assigned (i.e., surf or hike therapy). Throughout these programs, participants complete brief, on-site self-report assessments before and after each of the six weekly sessions.

Within 2 weeks after the final session, participants complete an in-person post-program assessment. The final, in-person follow-up assessment occurs 3 months after the post-program assessment. Like the pre-test assessment, these assessments are administered by the study assessor and include both clinical interview measures and self-report questionnaires. If an individual discontinues participation, they are still contacted for assessment at post-treatment and follow-up time points unless they request to be removed from study participation. A fidelity assessor routinely reviews assessment recordings and provides quarterly feedback to the study assessor throughout the study.

2.5.1. MDD diagnosis and depressive symptoms

Change in depressive symptom severity serves as the primary study outcome. Depression symptoms are evaluated at all three main assessment time points using the Montgomery–Åsberg Depression Rating Scale (MADRS; [38]. The MADRS is a semi-structured, clinical interview for the assessment of depressive symptom severity and has been shown to be sensitive to change after an intervention [38].

The MINI-7 [33], an assessor-administered diagnostic interview, is used to evaluate MDD diagnostic status at main study assessment time points (see Fig. 1 for all assessment details). Additionally, at pre-program, it is used to determine participant eligibility for study participation. The MINI-7 is also presented in Section 2.5.2 because it is used to assess for other comorbid conditions. Self-reported severity of depression symptoms within the past 2 weeks is examined using the 9-item Patient Health Questionnaire (PHQ-9; [39]. Each item is rated on a scale (ranging from 0 to 3; higher scores reflect greater depression severity) that can either be summed to create a total severity score or used to evaluate diagnostic criteria. The PHQ-9 has demonstrated sound psychometric properties, including good internal consistency and test–retest reliability [39].

Current depression and anxiety symptomatology is assessed before and after each activity session with the 4-item Patient Health Questionnaire (PHQ-4; [40]. The PHQ-4 consists of two depression items from the PHQ-9 and two anxiety items from the Generalized Anxiety Disorder (GAD) 7-item scale (GAD-7; discussed in *2.3.2.*; [41]. Items are scored from 0 to 3 and summed to create a total score (with higher scores indicating greater severity). Scores on the anxiety and depression subscales of the PHQ-4 are highly correlated with scores on longer measures of anxiety and depression, making it an efficient and valid assessment instrument [40].

2.5.2. Other comorbid conditions

In addition to assessing MDD, the MINI-7 [33] is administered to evaluate the presence of other comorbid psychological disorders, including posttraumatic stress disorder (PTSD), anxiety disorders, other mood disorders, and psychotic symptoms. Self-reported PTSD symptomatology in the past month is assessed using the PTSD Checklist for the DSM-5 (PCL-5; [42]. This instrument contains the revised Life Events Checklist for DSM-5 [43] and an extended Criterion A assessment. Items on the PCL-5 directly correspond with the 20 DSM-5 diagnostic criteria for PTSD and are rated on a scale ranging from 0 to 4 (higher scores suggest greater symptom severity). The PCL-5 has high internal consistency ($\alpha = 0.94$) and good test–retest reliability ($r = 0.82$; [44].

Presence and severity of anxiety over the past 2 weeks is evaluated using the self-report GAD-7 [41]. Each symptom item is rated from 0 to 3 (higher scores indicate greater anxiety severity). The GAD-7 has been found to have an internal consistency of $\alpha = 0.92$ [41].

Traumatic brain injury (TBI) history is evaluated using The Ohio State University TBI Identification Method – Short Form (OSU TBI-ID-SF; [45]. The OSU TBI-ID-SF is a 7-item structured interview for the assessment of lifetime TBI history. The measure assesses for history of head and neck injuries due to motor vehicle accident, sports, and combat, as well as for the neurological aspects of each injury (e.g., loss of consciousness). Summary indices of lifetime injury are generated to reflect the likelihood that neurological consequences have resulted from the TBI history. The predictive validity of these indices is supported in the literature [45,46].

The nature, severity, and impact of insomnia over the past week is evaluated with the 7-item Insomnia Severity Index (ISI; [47]. Each item is rated from 0 to 4 and summed to yield a total score (higher scores reflect greater insomnia severity). The ISI has also demonstrated high internal consistency ($\alpha = 0.90$; [48].

Current level of pain is rated using the Numerical Pain Rating Scale (NPRS; [49]. The NPRS consists of a single item on which the participant rates their current level of pain on an 11-point scale (0 = *no pain*; 10 = *worst possible pain*).

2.5.3. Measures of affect, resilience, and function

Current emotions are assessed using the 20-item Positive and Negative Affect Schedule (PANAS; [50]. Each affect item is rated on a scale ranging from 0 to 4. The 10 positive affect items are summed to create a positive affect scale (higher scores reflect greater positive affect); the 10 negative affect items are added to yield a negative affect scale (lower scores indicate less negative affect). Internal consistencies for the PANAS subscales are acceptably high; for positive affect, they range from $\alpha = 0.86$ to 0.90; for negative affect, they range from $\alpha = 0.84$ to 0.87 [50].

Resilience to the most stressful events experienced by participants is measured by the 4-item Response to Stressful Events Scale (RSES-4; [51]. The RSES-4 is a brief self-report resilience measure that was developed from the 22-item Response to Stressful Events Scale (RSES-22; [52]. Construct validity is supported by a strong correlation with the RSES-22 ($r = 0.90$), as well as acceptable correlations ($r = 0.29$–0.39) with measures of burnout and distress [51].

Functional physical and mental health are determined using responses to the Short Form Health Survey – 36 Item, Version 2 (SF-36v2® [94]; QualityMetric Inc.). The SF-36 is a widely used questionnaire that allows the computation of eight summary health measures: physical functioning, role limitations due to physical health, pain, general health, energy/fatigue, social functioning, role limitations due to emotional problems, and emotional well-being. The SF-36, as well as the eight summary measures, have been found to have good criterion validity ($r = 0.51$ to 0.85; [53,54].

2.5.4. Demographic questionnaire

This questionnaire collects demographic information (e.g., age, years of education, race, military service branch), which will be used to describe the sample and to determine whether demographic factors influence outcomes (and therefore need to be included as covariates in study analyses).

2.5.5. Treatment use

The Treatment Utilization Questionnaire, created for this study, asks participants about their use of outpatient counseling, psychotropic medication, couples or family counseling, and support groups, as well as any psychiatric hospitalizations. At the pre-program assessment, information is collected over the participant's lifetime as well as within the past 3 months; at the post-program and 3-month follow-up assessments, information is only collected for the past 3 months.

2.5.6. Activity level and athletic identity

The subjective intensity of exertion a person experiences during physical activity is evaluated using a modified version of the single-item Borg Rating of Perceived Exertion Scale (Borg RPE Scale; [55]. The modified response scale ranges from 0 (*no exertion at all*) to 10 (*maximal exertion*) and has been used in prior research (e.g., Refs. [56–59]. Validity of the Borg RPE Scale is well-established against various physiological criteria (e.g., $r = 0.62$ with heart rate), and correlations are particularly high in activities involving swimming (e.g., $r = 0.83$ with heart rate; [60]. Athletic identity is determined by the Athletic Identity Measurement Scale (AIMS; [61]. The AIMS has displayed strong psychometric properties including construct validity, internal consistency ($\alpha = 0.93$), and test–retest reliability ($r = 0.89$; [61].

Baseline physical activity is assessed with the 9-item self-report International Physical Activity Questionnaire-Short Form (IPAQ-SF; [62]. The measure yields a summary score reflecting the frequency and intensity of activity over the last 7 days, as well as sedentary periods. Criterion validity of the IPAQ-SF in measuring total activity has been found to be acceptable ($rs = 0.29–0.30$) in relation to other established self-report activity measures [62–64]. Because the IPAQ-SF lacks sufficient specificity to identify physical activity change in small samples [64,65], it is only administered once at the pre-program assessment.

Ongoing physical activity and physiological data are collected via the Fitbit Charge 2 device. The Fitbit Charge 2 is a popular, consumer-based, wrist-worn activity tracker. In this study, the device provides additional information on physical activity. Heart rate is collected, and sleep data serves as secondary data to augment self-report measures of insomnia. The Fitbit Charge 2 displays good specificity (0.61–0.96) in detecting sleep–wake states and sleep stage composition, but is less accurate (0.49) at detecting deep sleep compared with polysomnography [66]. In research published since the start of the study, the device has shown mixed results regarding heart rate accuracy [67–69], active minutes [70], and energy expenditure [71,72]. Research on step count reports modest accuracy ($r = 0.58$; [70]. Heart rate, physical activity, and sleep pattern data are continuously synced and collected throughout the duration of the study. Due to data privacy issues and inability to reuse devices, and as an added incentive to adhere to the study protocol, participants who wear the Fitbit for more than 50% of the study days are given the device to keep; otherwise, devices are destroyed.

2.5.7. Program and surf/hike engagement

The extent of program participation at each session is assessed using a self-report activity participation questionnaire that was developed for this study. The measure asks participants to indicate the specific activities they engaged in during each surf or hike therapy session. At pre-program and 3-month follow-up, we also ask participants to report the number of days on which they surfed and hiked within the past month. This variable will be used to determine whether engagement in surf or hike therapy affects longer-term levels of engagement in these activities.

2.5.8. Program satisfaction

The 8-item Client Satisfaction Questionnaire (CSQ-8; [73] evaluates satisfaction with the therapy programs provided. Each of the eight items is rated on a 4-point scale (response anchors vary by question). The CSQ-8 has demonstrated strong psychometric properties, including high levels of internal consistency [73,74] and validity through higher satisfaction scores among individuals who completed treatment versus those who dropped out [74].

2.6. Design considerations

A number of issues were considered in developing this RCT in a naturalistic setting. A primary consideration was to develop study methodology that would allow for the main study hypothesis to be appropriately examined (i.e., to determine whether physical activity that occurs in water has enhanced benefits for reducing depression symptoms relative to physical activity that occurs on land). Additionally, the literature on surf therapy often lacks methodological rigor and we wanted to build not only upon our previous work [75], but also the state of the science within the field. Specifically, we wanted to address design limitations regarding incorporation of a comparison group, use of a longitudinal design, and inclusion of multiple methods of data collection.

2.6.1. Selection of a comparison group

To date, no study examining surf therapy has included a comparison group in its design, which is critically important for understanding the benefits that are specific to surf therapy. Not only was it imperative that the study include a comparison group, but we carefully evaluated several options. Data from the program evaluation phase of our surf therapy study (which used a single-group design) demonstrated significant improvements in psychological symptoms, as well as increases in positive affect and decreases in negative emotions [75]. Because significant changes in outcomes were observed for most of the study variables, we wanted to compare surf therapy and an active control condition rather than a minimal contact condition. Comparing two active conditions can produce results that begin to address the question of which interventions work for whom and under what circumstances.

When evaluating potential active control conditions, we considered rock climbing, which, like surfing can be considered an “extreme sport,” but that occurs on land. However, the WII Wellness Program at NMCSD did not offer a rock climbing program, and comparing the WII Wellness surf therapy program and a rock climbing program offered elsewhere could introduce a variety of confounds. Furthermore, we were concerned about potential differences in energy expenditure during rock climbing (5–8 metabolic equivalents [METs]) compared with surfing (3–6 METs; [76]; see also [77]. Cycling was another land activity offered through the WII Wellness Program that was considered but not selected for our comparison group. Similar to rock climbing, there were potential energy expenditure differences between cycling (3.5–10 METs; Ainsworth et al., n.d.; see also [77] and surfing. Finally, and perhaps most importantly, we were concerned about the accessibility of the comparison sport given that surf therapy is not widely accessible. Surf therapy requires a large, dynamic body of water (or a specially designed pool), so we wanted to select an activity that might be more easily accessible to a broader population than either rock climbing or cycling, which can require access to specific locations or gear that can be resource prohibitive.

Hike therapy, as provided through the WII Wellness Program, was considered and ultimately decided upon as the active comparison condition in this RCT. Hike therapy is an option available to many because the activity requires minimal gear and can take place on various types of terrain. As a comparative condition, hike therapy provides an excellent parallel to surfing in several ways. First, hike therapy is a physical activity where participants can progress at their own pace, and the energy expenditure (5.3 METs) is comparable to surfing (3–6 METs; Ainsworth et al., n.d.; see also [77]. Additionally, hike therapy is conducted in a group setting, which may help to control for social interaction. Lastly, hike therapy takes place outdoors and allows for an immersive experience in the natural environment. Taken together, these factors ultimately led to the selection of hike therapy as the active

comparison condition for this RCT. The surf–hike comparison will allow for the evaluation of any unique effects of water-based exercise while controlling for several important components (i.e., physical activity, social interaction, natural environment, level of exertion, and even scheduling and department policies) of both programs.

2.6.2. Longitudinal study design

Another methodological weakness of extant surf and hike therapy studies is the limited use of longitudinal study designs. Evaluating the outcomes of both therapies over time is important to determine the duration of the therapeutic effects and the dosing required to achieve them. Although there have been some hiking studies that report longitudinal findings (e.g., Refs. [78–81] conducted the only surf therapy study that reported follow-up assessment findings, and the follow-up occurred only 30 days after the end of participation in surf therapy. Including a follow-up assessment is critical to the current RCT because it provides essential information about the duration of therapeutic effects. However, it is important to take into account that study participants are active duty service members, a transient population due to temporary duty assignments, relocation, deployments, and transitions out of the military. In psychotherapy treatment outcome studies with military samples, follow-up periods typically range from 1 to 6 months. We chose to conduct follow-up assessments 3 months following treatment completion to longitudinally assess program effects while minimizing attrition at follow-up caused by participant deployment or relocation.

It is also important to examine the immediate outcomes of both surf and hike therapy. Some evidence suggests that physical activity in the natural environment produces positive effects on mental health immediately following participation in the activity (e.g., Ref. [28]. Results from our previous research showed significantly decreased depression and anxiety symptoms, as well as increased positive affect, immediately following engagement in a surf therapy session [75]. In order to assess the immediate effects of both surf and hike therapy, participants in the current RCT complete brief assessments immediately prior to and following their respective activity sessions. The brief assessments are abbreviated versions of the measures completed at the pre-program, post-program, and 3-month follow-up assessments.

As a part of the longitudinal study design, we employed a cross-over strategy in which participants can enroll in the other type of therapy after they complete the post-program assessment. For program evaluation purposes, participants would ideally receive the intervention (surf or hike therapy) to which they are randomized and complete any longitudinal follow-up assessments prior to receiving the other intervention. However, we determined that allowing participants to receive the other intervention after completion of the post-program assessment was more practical given the transient nature of our sample population. Since study participants may no longer be located in the area 3 months following program completion, we thought it was important to provide them with the opportunity to receive the other intervention after completing the first one in order to reduce bias in both recruitment and retention. To statistically evaluate the influence of the cross-over design on results, we code participants to indicate whether they participated in the other program (or elected to repeat the program to which they were randomized) during their follow-up period. Analyses will test whether program effects differ depending on whether participants received either intervention during the follow-up period. Furthermore, data are collected on the number of sessions attended by each participant to allow for a more nuanced understanding of dosing effects.

2.6.3. Multimodal assessment

An over-reliance on self-report data is another limitation of prior research on surf therapy outcomes. Four studies [34,75,81,82] have examined the psychological effects of surfing, of which three [75,81,82] collected quantitative data using validated measures. Furthermore, all four studies used only self-report measures, which are subject to biases due to variations in question interpretation, socially desirable responses, and the like. Similarly, although many hiking studies have used validated psychological self-report measures (e.g., Refs. [78,80,83], as well as physiological data (e.g., Refs. [84–86], none have utilized diagnostic clinical interviews to measure diagnoses or other primary outcome variables. While the assessments used in this study include validated self-report measures, ours is the first surf therapy study to also incorporate clinical interview data to strengthen conclusions about the effects of surf and hike therapy on relevant outcomes.

To augment information reported in both clinical interviews and on self-report instruments, physiological data are also collected throughout the program using the Fitbit Charge 2 device. The Fitbit Charge 2 was selected because of its ability to collect relevant data while still being cost effective. Additionally, Fitbit is a popular brand of wearable activity trackers, and could be issued to and kept by participants after completion of the study if they used the device as stated in the consent form. It should be mentioned that one limitation of the Fitbit Charge 2 for this particular study is that the device is not water resistant, so unlike participants randomized to hike therapy, those randomized to surf therapy cannot wear their devices during their activity sessions. Although Fitbit now has newer devices that are water resistant, these were not available at the outset of this study.

2.7. Assessment fidelity

The clinical interview portion of each assessment at pre-program, post-program, and 3-month follow-up is audio recorded. Participants provide written informed consent for audio recording as a part of the initial consent process. Twenty percent of participants ($n = 22$) will be randomly selected for review to determine assessment fidelity, and all assessment time points will be reviewed for each selected participant. A fidelity assessor, who holds a doctoral degree in clinical psychology and has expertise in the diagnosis of MDD, will be reviewing all cases selected for fidelity and providing ratings for the MINI-7 and the MADRS. Cohen's kappa statistic will be used to assess the interrater reliability of MDD diagnoses between the initial assessor and the fidelity rater; intraclass correlations will be used to evaluate interrater reliability for MADRS scores.

2.8. Data analysis

The study is a two (Group: surf therapy/hike therapy) by three (Assessment time point: pre-program/post-program/3-month follow-up) mixed effects design, where the first factor is between participants and the second is within participants. The minimum sample size of $N = 86$ (surf $n = 43$; hike $n = 43$) was determined based on the power to test the primary study research question (i.e., do changes in depression symptoms over time vary between participants in surf versus hike therapy?). We assumed an effect size of $d = 0.40$, with $\alpha = 0.05$; $1\text{-}\beta = 0.90$; and $r = 0.40$ between repeated measures. The effect size of $d = 0.40$ was derived from the program evaluation phase and reflects the most conservative value found for pre- to post-program changes in depression [75]. Using these parameters, the estimated required sample size was $N = 83$. Based on our previous study, 27 additional participants were included to account for attrition and ensure sufficient power, resulting in a final target sample size of at least 86.

Data will be analyzed as intent to treat, meaning that all eligible participants will be included in study analyses. If study participants do not complete the program or are discontinued from the study due to a physician-determined change in medical clearance, then they will be coded as such and analyses will be conducted to evaluate differences between program completers and non-completers. Chi-square tests of association (categorical variables) and independent samples t tests (continuous variables) will be used to test for differences between program completers and non-completers. Similarly, although the

purpose of randomization is to balance the effects of potentially confounding variables across treatment groups, chi-square tests of association and *t* tests will be used to explore whether any such confounds exist; if so, they will be included as covariates in subsequent analyses. Sensitivity analyses will also be conducted for variables that could influence the robustness of results (e.g., program completers; level of participation). For all analyses, effect size indices will be examined, and significance testing will be conducted.

The primary aims of this study are (1) to determine whether surf and hike therapy result in significantly reduced depression symptoms, and (2) to evaluate whether surf therapy results in greater symptom reduction than hike therapy. The second aim is based on prior data indicating a greater effect on mood for activities that occur near water compared with those that do not [29]. To examine these aims, multilevel modeling (MLM) will be used to examine depression symptom trajectories over time for both surf and hike therapy conditions. The same analytic strategy will be used to examine effects of both interventions – and any differential effects of each – on secondary outcomes including PTSD symptom severity, anxiety symptom severity, pain severity, insomnia severity, negative affect, positive affect, and functional impairment.

To evaluate both immediate and longer-term effects of activity sessions on psychological and physical symptoms, data from measures completed before and after each activity session will be analyzed using an MLM framework; each participant will have up to 12 data points for each measure (i.e., six pre- and six post-session assessments). MLM will be used to analyze the data from pre- and post-session assessments in order to assess changes in symptoms *within* as well as *across* sessions. The use of MLM is particularly advantageous for examining weekly session data because it can account for variability in the amount of time between measurements and can retain in analyses participants who have missing data for one or more sessions.

Rates of remission from MDD (i.e., not meeting diagnostic criteria at post-program or 3-month follow-up assessments) of service members who participate in surf therapy versus hike therapy will also be statistically compared using chi-square tests of association. The outcome measures for this aim will be based on MINI-7 data (i.e., whether diagnostic criteria for MDD are met or not) from the post-program and follow-up assessments (with separate analyses for each assessment). Chi-square test of association will also be used to compare the number of participants who did not complete each treatment condition (i.e., missed more than two sessions). Logistic regression will be used if confounds are identified and need to be controlled for in analyses; odds ratios will provide a measure of effect size.

Independent samples *t* tests will be used to compare patient satisfaction with the two interventions at post-program. The outcome measure for this aim will be the continuous CSQ-8 score derived from the post-program assessment.

Fitbit data for heart rate, physical activity, and sleep will be extracted from a non-identifiable data set and converted to an Excel file. Continuous scores (e.g., average heart rate, number of steps, hours of sleep) will be created and analyzed using *t* tests, and indices of effect size will be calculated.

3. Discussion

For active duty service members and civilians alike, it is imperative to identify alternative approaches to treating MDD and other related symptoms that do not rely on traditional psychotherapy or pharmacotherapy. Although there are efficacious, evidence-based psychotherapy and pharmacotherapy options for individuals with MDD, some individuals may not have access to, desire, or benefit from these treatments. Outdoor physical activity may address these issues given that it may be readily accessible and could supplement evidence-based treatments as an adjunctive intervention or prove beneficial as a standalone treatment. Additionally, outdoor physical activity may offer significant physical health benefits, which could address physical comorbidities associated with MDD [87] and other psychological disorders [88]. Interventions involving outdoor physical activity may be particularly appealing to service members since fitness is a core component of mission readiness [89].

Although outdoor recreation programs are commonly offered to service members and veterans in the United States, there have been few evaluations of such programs, and those that exist yield limited information due to methodological shortcomings. Recent meta-analyses [20,90] concluded that studies investigating the effect of exercise on depression must improve their methodological rigor across several domains. Those authors recommended several improvements to study assessments, including masking assessors to study conditions and increasing the use of validated and objective measures. Additionally, they advised that exercise studies report greater levels of methodological detail in order to enable replication of findings. Lastly, the authors also recommended that longitudinal data be collected, and that these data assess both activity engagement and symptom change over time.

Our study aims to provide a foundation for further research on outcomes of outdoor activity programs, and to begin to understand the specific outcomes of such programs and how they are best achieved. It incorporates all of these recommendations, and not only increases the scientific rigor of the extant body of research investigating psychological and physical outcomes of exercise and outdoor activity interventions, but also addresses whether water-based activities yield enhanced benefit over land-based ones. Hike therapy was carefully selected as a comparison condition because it occurs in a group setting within the natural environment while differing in the type of environment (land vs. water). Furthermore, participants in both activities are immersed in their respective environment settings, and both activities incur similar energy expenditures. This allows us to isolate the unique effect of water on participant outcomes. Second, the RCT is conducted in naturalistic settings (i.e., a military treatment facility; natural environment), which has the advantage of allowing the use of stringent scientific methods while maximizing generalizability. The use of minimal exclusion criteria to determine participant eligibility further bolsters generalizability. Third, data collected from the study use a variety of modalities, with complementary strengths and limitations, including structured clinical interviews, validated self-reports, and physiological/activity data. Although some prior research on hiking has utilized physiological measures, the literature on both hiking and surfing have, to date, been limited by reliance on self-report measures. The inclusion of structured clinical interviews and physiological/activity data in the current study serves to validate data collected through self-report and also allows for a more comprehensive examination of the range of outcomes that may be impacted by engaging in outdoor physical activity. Finally, the current study design incorporates a 3-month follow-up, which is the longest for any surf therapy study to date, and comparable to the longest that has been used for any hike therapy study [80]. In addition to the longer-term effects, the study methodology was intended to test the immediate effects of engaging in each activity. In combination, data collected in this study will provide evidence regarding immediate and longer-term effects of both surf and hike therapy on a broad range of relevant outcomes, including continued engagement in the focal physical activities.

Although this study improves considerably on the research methodology used in prior work in this area, it is not without limitations. The surf and hike therapy programs are both available through the WII Wellness Program at NMCSD, which provides some assurance that they follow similar policies, procedures, and schedules. However, there may be differences between the two programs due to inherent differences in their focal activities or to differences between instructors and program managers. Furthermore, optional yoga is provided immediately prior to the surf therapy program whereas it is not available immediately prior to the hike therapy program. The influence of optional yoga on surf outcomes will be statistically explored in a manner consistent with our

prior work [75], where it was found to have no associations with treatment outcomes for the overall sample or MDD subsample; this may be due to the fact that 40% of participants did not attend any of the optional yoga sessions, and the mean number of yoga sessions attended was 1.7 (out of a potential 6). Another limitation is that, although both activities occur in a group-based setting and are administered within the same department, the extent to which the programs deliver quantitatively comparable opportunities for social interaction are unclear. A third limitation is that the Fitbit Charge 2 devices used in the study are not water resistant. This creates a difference between conditions in that participants randomized to the hike condition can wear their device during their weekly program activity whereas those randomized to the surf condition cannot. This prohibits a direct comparison of activity data collected during the respective activity sessions. A fourth limitation derives from the cross-over design, which allows participants to participate in the other activity option following completion of the therapy activity to which they were randomized (i.e., to participate in the activity to which they were not originally randomized or re-enrollment in the activity they originally received). This means that some participants may be participating in the other activity—or additional "doses" of the original treatment—during the follow-up period. As mentioned previously, we elected to use this design to maximize recruitment and retention, given that study participants, as service members, are transient. To statistically address this limitation, participants who choose the cross-over option are coded as such and will be examined in sub-analyses. A fifth limitation is that we do not exclude participants who are receiving current psychotherapy or pharmacotherapy due to ethical considerations. Although existing research has supported exercise as an intervention for individuals with MDD [21–26,91,92], the data do not conclusively indicate that physical activity is superior to (or equally effective as) evidenced-based treatments (e.g., Refs. [90,93]; as such, we did not want to withhold evidence-based treatments from participants. Further, our previous research found that 75% of surf therapy study participants used surf therapy as an adjunctive treatment to psychotherapy or pharmacotherapy [75]; thus, outcome data for surf therapy as a primary treatment are limited. In order to investigate the relationship between intervention condition and concurrent treatments, we are collecting data regarding engagement in various forms of therapy as well as psychotropic medication use at pre-program, post-program, and 3-month follow-up assessments. Lastly, the study sample includes active duty service members, but results could be relevant to populations such as law enforcement, firefighters, and other first responders. However, it is important that any observed benefits of surf or hike therapy for individuals varying in age or fitness level be addressed in future research. Furthermore, activity-based interventions that take place outdoors may require specific environmental settings or equipment that may not be readily accessible, thus potentially limiting generalizability.

The current study explores novel interventions—surf and hike therapy—for treating depression and related symptoms among active duty service members with MDD. The use of an RCT evaluating these approaches in a naturalistic setting aims to provide efficacy data that are externally valid. The study methodology was developed to advance the current state of surf therapy and outdoor activity research by including a comparison group, random assignment, multimodal data, and a moderately long follow-up period, which sets this study apart from other research in this area. In addition to understanding the unique outcomes associated with surf and hike therapy, the study may provide a methodological framework for future research evaluating outcomes of outdoor physical activity in both military and civilian settings. Importantly, results from this RCT will provide information regarding the extent to which outdoor activities can serve as a potential intervention option for symptoms associated with MDD.

Funding

This work was supported by the U.S. Navy Bureau of Medicine and Surgery [work unit no. N1600].

Clinical trials identifier

NCT03302611.

Disclaimer

I am a military service member or employee of the U.S. Government. This work was prepared as part of my official duties. Title 17, U.S.C. §105 provides that copyright protection under this title is not available for any work of the U.S. Government. Title 17, U.S.C. §101 defines a U.S. Government work as work prepared by a military service member or employee of the U.S. Government as part of that person's official duties.

Report No. 19–41 was supported by the U.S. Navy Bureau of Medicine and Surgery under work unit no. N1600. The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government.

The study protocol was approved by the Naval Medical Center San Diego Institutional Review Board in compliance with all applicable Federal regulations governing the protection of human subjects. Research data were derived from an approved Naval Medical Center San Diego Institutional Review Board protocol number NMCSD.2017.0007.

Acknowledgments

The authors would like to acknowledge Jody DeLaPeña Murphy, Renée G. Dell'Aqua, Robyn M. Highfill-McRoy, Andrew M. Kewley, Casey B. Kohen, James E. LaMar, and Alexandra L. Powell for their significant efforts on this study. We are grateful to Naval Medical Center San Diego (NMCSD) and to the Wounded, Ill, and Injured Wellness Program at NMCSD for their openness to collaborate on this research study. We deeply thank all surf therapy volunteers, Del Mar lifeguards, yoga instructors, study participants, and student interns for their efforts and commitment to these programs. We are also appreciative for the support and hospitality of the Del Mar community. Additionally, the Challenged Athletes Foundation (including Nico Marcolango) and the Semper Fi Fund provide considerable support to the NMCSD Surf Therapy Program and we would like to recognize these beneficent organizations.

References

[1] K. Blakeley, D.J. Jansen, Post-traumatic Stress Disorder and Other Mental Health Problems in the Military: Oversight Issues for Congress, (2013) Retrieved from https://fas.org/sgp/crs/natsec/R43175.pdf.

[2] T. Tanielian, L.H. Jaycox (Eds.), Invisible Wounds of War: Psychological and Cognitive Injuries, Their Consequences, and Services to Assist Recovery, RAND Corporation, Santa Monica, CA, 2008.

[3] A.M. Gadermann, C.C.C. Engel, J.A. Naifeh, M.K. Nock, M. Petukhova, L.P.N. Santiago, R.C. Kessler, Prevalence of DSM-IV major depression among U.S. military personnel: meta-analysis and simulation, Mil. Med. 177 (Suppl. 8) (2012) 47–59, https://doi.org/10.7205/MILMED-D-12-00103.

[4] R.C. Kessler, M. Petukhova, N.A. Sampson, A.M. Zaslavsky, H. Wittchen, Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States, Int. J. Methods Psychiatr. Res. 21 (3) (2012) 169–184 https://doi.org/10.1002/mpr.1359.

[5] G.K. Brown, A.T. Beck, R.A. Steer, J.R. Grisham, Risk factors for suicide in psychiatric outpatients: a 20-year prospective study, J. Consult. Clin. Psychol. 68 (3) (2000) 371–377, https://doi.org/10.1037/0022-006X.68.3.371.

[6] P. Cuijpers, F. Smit, Excess mortality in depression: a meta-analysis of community studies, J. Affect. Disord. 72 (3) (2002) 227–236, https://doi.org/10.1016/S0165-0327(01)00413-X.

[7] A.C. Butler, J.E. Chapman, E.M. Forman, A.T. Beck, The empirical status of cognitive-behavioral therapy: a review of meta-analyses, Clin. Psychol. Rev. 26 (1)

(2006) 17–31, https://doi.org/10.1016/j.cpr.2005.07.003.
[8] P. Cuijpers, M. Berking, G. Andersson, L. Quigley, A. Kleiboer, K.S. Dobson, A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments, Can. J. Psychiatr. 58 (7) (2013) 376–385, https://doi.org/10.1177/070674371305800702.
[9] S.G. Hofmann, A. Asnaani, I.J. Vonk, A.T. Sawyer, A. Fang, The efficacy of cognitive behavioral therapy: a review of meta-analyses, Cogn. Ther. Res. 36 (5) (2012) 427–440, https://doi.org/10.1007/s10608-012-9476-1.
[10] S.D. Hollon, A.T. Beck, Cognitive and cognitive behavioral therapies, in: M.J. Lambert (Ed.), Bergin and Garfield's Handbook of Psychotherapy and Behavior Change, sixth ed., John Wiley & Sons, Hoboken, NJ, 2013, pp. 393–443.
[11] Jc C. Fournier, R.J. DeRubeis, S.D. Hollon, S. Dimidjian, J.D. Amsterdam, R.C. Shelton, J. Fawcett, Antidepressant drug effects and depression severity: a patient-level meta-analysis, J. Am. Med. Assoc. 303 (1) (2010) 47–53, https://doi.org/10.1001/jama.2009.1943.
[12] A. Khan, R.M. Leventhal, S.R. Khan, W.A. Brown, Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database, J. Clin. Psychopharmacol. 22 (1) (2002) 40–45.
[13] P. Cuijpers, A. van Straten, E. Bohlmeijer, S.D. Hollon, G. Andersson, The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size, Psychol. Med. 40 (2) (2010) 211–223, https://doi.org/10.1017/S0033291709006114.
[14] A.G. Bulloch, S.B. Patten, Non-adherence with psychotropic medications in the general population, Soc. Psychiatry Psychiatr. Epidemiol. 45 (1) (2010) 47–56, https://doi.org/10.1007/s00127-009-0041-5.
[15] G.I. Papakostas, Tolerability of modern antidepressants, J. Clin. Psychiatry 69 (Suppl. E1) (2008) 8–13.
[16] J.R. Geddes, S.M. Carney, C. Davies, T.A. Furukawa, D.J. Kupfer, E. Frank, G.M. Goodwin, Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review, The Lancet 361 (9358) (2003) 653–661, https://doi.org/10.1016/S0140-6736(03)12599-8.
[17] M. Fava, A.J. Rush, M.H. Trivedi, A.A. Nierenberg, M.E. Thase, H.A. Sackeim, D.J. Kupfer, Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study, Psychiatr. Clin. N. Am. 26 (2) (2003) 457–494, https://doi.org/10.1016/S0193-953X(02)00107-7.
[18] R. Kessler, P. Berglund, O. Demler, R. Jin, D. Koretz, K. Merikangas, P. Wang, The epidemiology of major depressive disorder: results from the national comorbidity Survey replication (NCS-r), J. Am. Med. Assoc. 289 (23) (2003) 3095–3105, https://doi.org/10.1001/jama.289.23.3095.
[19] M. aan het Rot, K.A. Collins, H.L. Fitterling, Physical exercise and depression, MSJM (Mt. Sinai J. Med.) 76 (2) (2009) 204–214, https://doi.org/10.1002/msj.20094.
[20] S. Rosenbaum, A. Tiedemann, C. Sherrington, J. Curtis, P.B. Ward, Physical activity interventions for people with mental illness: a systematic review and meta-analysis, J. Clin. Psychiatry 75 (9) (2014) 964–974, https://doi.org/10.4088/JCP.13r08765.
[21] M. Babyak, J.A. Blumenthal, S. Herman, P. Khatri, M. Doraiswamy, K. Moore, K.R. Krishnan, Exercise treatment for major depression: maintenance of therapeutic benefit at 10 months, Psychosom. Med. 62 (5) (2000) 633–638.
[22] L.L. Craft, D.M. Landers, The effect of exercise on clinical depression and depression resulting from mental illness: a meta-analysis, J. Sport Exerc. Psychol. 20 (4) (1998) 339–357, https://doi.org/10.1123/jsep.20.4.339.
[23] J. Mota-Pereira, J. Silverio, S. Carvalho, J.C. Ribeiro, D. Fonte, J. Ramos, Moderate exercise improves depression parameters in treatment-resistant patients with major depressive disorder, J. Psychiatr. Res. 45 (8) (2011) 1005–1011, https://doi.org/10.1016/j.jpsychires.2011.02.005.
[24] F.B. Schuch, D. Vancampfort, J. Richards, S. Rosenbaum, P.B. Ward, B. Stubbs, Exercise as a treatment for depression: a meta-analysis adjusting for publication bias, J. Psychiatr. Res. 77 (2016) 42–51, https://doi.org/10.1016/j.jpsychires.2016.02.023.
[25] H. Silveira, H. Moraes, N. Oliveira, E.S.F. Coutinho, J. Laks, A. Deslandes, Physical exercise and clinically depressed patients: a systematic review and meta-analysis, Neuropsychobiology 67 (2) (2013) 61–68, https://doi.org/10.1159/000345160.
[26] J. Gourgouvelis, P. Yielder, S.T. Clarke, H. Behbahani, B.A. Murphy, Exercise leads to better clinical outcomes in those receiving medication plus cognitive behavioral therapy for major depressive disorder, Front. Psychiatry 9 (2018) 37, https://doi.org/10.3389/fpsyt.2018.00037.
[27] J. Knapen, D. Vancampfort, Y. Moriën, Y. Marchal, Exercise therapy improves both mental and physical health in patients with major depression, Disabil. Rehabil. 37 (16) (2015) 1490–1495, https://doi.org/10.3109/09638288.2014.972579.
[28] J. Thompson Coon, K. Boddy, K. Stein, R. Whear, J. Barton, M.H. Depledge, Does participating in physical activity in outdoor natural environments have a greater effect on physical and mental wellbeing than physical activity indoors? A systematic review, Environ. Sci. Technol. 45 (5) (2011) 1761–1772, https://doi.org/10.1021/es102947t.
[29] J. Barton, J. Pretty, What is the best dose of nature and green exercise for improving mental health? A multi-study analysis, Environ. Sci. Technol. 44 (10) (2010) 3947–3955, https://doi.org/10.1021/es903183r.
[30] E. Britton, G. Kindermann, C. Domegan, C. Carlin, Blue care: a systematic review of blue space interventions for health and wellbeing, Health Promot. Int. (2018) 1–20, https://doi.org/10.1093/heapro/day103.
[31] M. Haeffner, D. Jackson-Smith, M. Buchert, J. Risley, Accessing blue spaces: social and geographic factors structuring familiarity with, use of, and appreciation of urban waterways, Landsc. Urban Plan. 167 (2017) 136–146, https://doi.org/10.1016/j.landurbplan.2017.06.008.
[32] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, fifth ed., American Psychiatric Association, Washington, DC, 2013.
[33] D.V. Sheehan, Mini International Neuropsychiatric Interview 7.0, Medical Outcomes Systems, Jacksonville, FL, 2015.
[34] N. Caddick, C. Phoenix, B. Smith, Collective stories and well-being: using a dialogical narrative approach to understand peer relationships among combat veterans experiencing post-traumatic stress disorder, J. Health Psychol. 20 (2015) 286–299, https://doi.org/10.1177/1359105314566612.
[35] N. Caddick, B. Smith, C. Phoenix, The effects of surfing and the natural environment on the well-being of combat veterans, Qual. Health Res. 25 (1) (2015) 76–86, https://doi.org/10.1177/1049732314549477.
[36] B. Hawkins, J. Townsend, B. Garst, Nature-based recreational therapy for military service members, Ther. Recreat. J. 50 (2016) 55–74, https://doi.org/10.18666/TRJ-2016-V50-I1-6793.
[37] J. Marshall, P. Kelly, A. Niven, "When I go there, I feel like I can be myself." Exploring programme theory within the Wave Project surf therapy intervention, Int. J. Environ. Res. Public Health 16 (2019) 2159, https://doi.org/10.3390/ijerph16122159.
[38] S.A. Montgomery, M. Åsberg, A new depression scale designed to be sensitive to change, Br. J. Psychiatry 134 (1979) 382–389, https://doi.org/10.1192/bjp.134.4.382.
[39] K. Kroenke, R. Spitzer, J. Williams, The PHQ-9: validity of a brief depression severity measure, J. Gen. Intern. Med. 16 (9) (2001) 606–613, https://doi.org/10.1046/j.1525-1497.2001.016009606.x.
[40] K. Kroenke, R.L. Spitzer, J.B. Williams, B. Löwe, An ultra-brief screening scale for anxiety and depression: the PHQ-4, Psychosomatics 50 (6) (2009) 613–621, https://doi.org/10.1016/S0033-3182(09)70864-3.
[41] R.L. Spitzer, K. Kroenke, J.B. Williams, B. Löwe, A brief measure for assessing generalized anxiety disorder: the GAD-7, Arch. Intern. Med. 166 (10) (2006) 1092–1097, https://doi.org/10.1001/archinte.166.10.1092.
[42] F.W. Weathers, B.T. Litz, T.M. Keane, P. Palmieri, B. Marx, P. Schnurr, The PTSD checklist for DSM-5 (PCL-5) [measurement instrument], Retrieved from, 2013. https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp#obtain.
[43] F.W. Weathers, D. Blake, P. Schnurr, D. Kaloupek, B. Marx, T. Keane, The Life events checklist for DSM-5 (LEC-5) [measurement instrument], Retrieved from, 2013. https://www.ptsd.va.gov/professional/assessment/te-measures/life_events_checklist.asp.
[44] C.A. Blevins, F.W. Weathers, M.T. Davis, T.K. Witte, J.L. Domino, The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation, J. Trauma. Stress 28 (2015) 489–498, https://doi.org/10.1002/jts.22059.
[45] J.D. Corrigan, J. Bogner, Initial reliability and validity of the Ohio state university TBI identification method, J. Head Trauma Rehabil. 22 (6) (2007) 318–329, https://doi.org/10.1097/01.HTR.0000300227.67748.77.
[46] J. Bogner, J.D. Corrigan, Reliability and predictive validity of the Ohio State University TBI Identification method with prisoners, J. Head Trauma Rehabil. 24 (4) (2009) 279–291, https://doi.org/10.1097/HTR.0b013e3181a66356.
[47] C.M. Morin, D.H. Barlow, Insomnia: Psychological Assessment and Management vol 104, Guilford Press, New York, NY, 1993.
[48] C.M. Morin, G. Belleville, L. Bélanger, H. Ivers, The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response, Sleep 34 (5) (2011) 601–608, https://doi.org/10.1093/sleep/34.5.601.
[49] M. McCaffery, A. Beebe, Pain: Clinical Manual for Nursing Practice, C.V. Mosby Company, St. Louis, MO, 1989.
[50] D. Watson, L.A. Clark, A. Tellegen, Development and validation of brief measures of positive and negative affect: the PANAS scales, J. Personal. Soc. Psychol. 54 (6) (1988) 1063–1070.
[51] G.M. De La Rosa, J.A. Webb-Murphy, S.L. Johnston, Development and validation of a brief measure of psychological resilience: an adaptation of the Response to Stressful Experiences Scale, Mil. Med. 181 (3) (2016) 202–208, https://doi.org/10.7205/MILMED-D-15-00037.
[52] D.C. Johnson, M.A. Polusny, C.R. Erbes, D. King, L. King, B.T. Litz, S.M. Southwick, Development and initial validation of the response to stressful experiences scale, Mil. Med. 176 (2) (2011) 161–169, https://doi.org/10.7205/MILMED-D-10-00258.
[53] I. McDowell, Measuring Health: A Guide to Rating Scales and Questionnaires, third ed., Oxford University Press, New York, NY, 2006.
[54] J.E. Ware Jr., SF-36 health Survey update, Spine 25 (24) (2000) 3130–3139.
[55] G.A. Borg, Psychophysical bases of perceived exertion, Med. Sci. Sport. Exerc. 14 (5) (1982) 377–381.
[56] M.L. Day, M.R. McGuigan, G. Brice, C. Foster, Monitoring exercise intensity during resistance training using the session RPE scale, J. Strength Cond. Res. 18 (2) (2004) 353–358, https://doi.org/10.1519/R-13113.1.
[57] K.R. Kendrick, S.C. Baxi, R.M. Smith, Usefulness of the modified 0-10 Borg scale in assessing the degree of dyspnea in patients with COPD and asthma, J. Emerg. Nurs. 26 (3) (2000) 216–222, https://doi.org/10.1016/S0099-1767(00)90093-X.
[58] R.C. Wilson, P.W. Jones, A comparison of the visual analogue scale and modified Borg scale for the measurement of dyspnoea during exercise, Clin. Sci. 76 (3) (1989) 277–282, https://doi.org/10.1042/cs0760277.
[59] R.C. Wilson, P.W. Jones, Long-term reproducibility of Borg scale estimates of breathlessness during exercise, Clin. Sci. 80 (4) (1991) 309–312, https://doi.org/10.1042/cs0800309.
[60] M.J. Chen, X. Fan, S.T. Moe, Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis, J. Sport. Sci. 20 (11) (2010) 873–899, https://doi.org/10.1080/026404102320761787.
[61] B. Brewer, J. Van Raalte, D. Linder, Athletic identity: hercules' muscles or Achilles heel? Int. J. Sport Psychol. 24 (1993) 237–254.
[62] C.L. Craig, A.L. Marshall, M. Sjöström, A.E. Bauman, M.L. Booth, B.E. Ainsworth, P. Oja, International physical activity questionnaire: 12-country reliability and

validity, Med. Sci. Sport. Exerc. 195 (9131/03) (2003), https://doi.org/10.1249/01.MSS.0000078924.61453.FB 3508–1381.
[63] M.K. Dinger, T.K. Behrens, J.L. Han, Validity and reliability of the international physical activity questionnaire in college students, Am. J. Health Educ. 37 (6) (2006) 337–343, https://doi.org/10.1080/19325037.2006.10598924.
[64] P.H. Lee, D.J. Macfarlane, T. Lam, S.M. Stewart, Validity of the international physical activity questionnaire short form (IPAQ-SF): a systematic review, Int. J. Behav. Nutr. Phys. Act. 8 (1) (2011) 115, https://doi.org/10.1186/1479-5868-8-115.
[65] A. Bauman, B.E. Ainsworth, F. Bull, C.L. Craig, M. Hagströmer, J.F. Sallis, M. Sjöström, Progress and pitfalls in the use of the International Physical Activity Questionnaire (IPAQ) for adult physical activity surveillance, J. Phys. Act. Health 6 (Suppl. 1) (2009) S5–S8, https://doi.org/10.1123/jpah.6.s1.s5.
[66] M. de Zambotti, A. Goldstone, S. Claudatos, I.M. Colrain, F.C. Baker, A validation study of Fitbit Charge 2™ compared with polysomnography in adults, Chronobiol. Int. 35 (4) (2018) 465–476, https://doi.org/10.1080/07420528.2017.1413578.
[67] S. Benedetto, C. Caldato, E. Bazzan, D.C. Greenwood, V. Pensabene, P. Actis, Assessment of the Fitbit Charge 2 for monitoring heart rate, PLoS One 13 (2) (2018) e0192691, , https://doi.org/10.1371/journal.pone.0192691.
[68] V.E. Salazar, N.D. Lucio, M.D. Funk, Accuracy of Fitbit Charge 2 worn at different wrist locations during exercise, Proc. Int. J. Exerc. Sci.: Conf. Proc. 2 (9) (2017).
[69] E. Thomson, Validity of heart rate measurements for the apple watch and Fitbit Charge HR 2, Med. Sci. Sport. Exerc. 50 (5S) (2018) 666, https://doi.org/10.1249/01.mss.0000538196.40788.41.
[70] I.A. Figueroa, N.D. Lucio, J.L. Gamez Jr., V.E. Salazar, M.D. Funk, Validity of daily physical activity measurements of Fitbit Charge 2, Int. J. Exerc. Sci.: Conf. Proc. 2 (10) (2018).
[71] K. Li, K. Nuss, E.A. Thomson, A. Comstock, S. Blake, S. Reinwald, B. Tracy, Validation of overall energy expenditure measurements in the Fitbit Charge HR 2 and apple watch, Med. Sci. Sport. Exerc. 50 (5S) (2018) 661, https://doi.org/10.1249/01.mss.0000538181.87422.7a.
[72] N.D. Lucio, E.V. Salazar, I.A. Figueroa, J.L. Gamez, R.D. Russell, M.D. Funk, Accuracy of Fitbit Charge 2 at estimating VO2max, calories, and steps on a treadmill, Int. J. Exerc. Sci.: Conf. Proc. 2 (10) (2018).
[73] D.L. Larsen, C.C. Attkisson, W.A. Hargreaves, T.D. Nguyen, Assessment of client/patient satisfaction: development of a general scale, Eval. Program Plann. 2 (1979) 197–207, https://doi.org/10.1016/0149-7189(79)90094-6.
[74] C.C. Attkisson, R.J. Zwick, The Client Satisfaction Questionnaire: psychometric properties and correlation with service utilization and psychotherapy outcome, Eval. Program Plann. 5 (1982) 233–237, https://doi.org/10.1016/0149-7189(82)90074-X.
[75] K.H. Walter, N.P. Otis, T.N. Ray, A.L. Powell, L.H. Glassman, B. Michalewicz-Kragh, C.J. Thomsen, Psychol. Sport Exerc. (2019), https://doi.org/10.1016/j.psychsport.2019.101551.
[76] B.E. Ainsworth, W.L. Haskell, S.D. Herrmann, N. Meckes, D.R. Bassett Jr., C. Tudor-Locke, A.S. Leon, Compendium of Physical Activities Tracking Guide (n.d), Retrieved from, https://sites.google.com/site/compendiumofphysicalactivities/.
[77] B.E. Ainsworth, W.L. Haskell, S.D. Herrmann, N. Meckes, D.R. Bassett Jr., C. Tudor-Locke, A.S. Leon, 2011 Compendium of physical activities: a second update of codes and MET values, Med. Sci. Sport. Exerc. 43 (8) (2011) 1575–1581, https://doi.org/10.1249/MSS.0b013e31821ece12.
[78] T. Hartig, M. Mang, G.W. Evans, Restorative effects of natural environment experiences, Environ. Behav. 23 (1) (1991) 3–26, https://doi.org/10.1177/0013916591231001.
[79] Q. Li, K. Morimoto, M. Kobayashi, H. Inagaki, M. Katsumata, Y. Hirata, T. Kawada, Visiting a forest, but not a city, increases human natural killer activity and expression of anti-cancer proteins, Int. J. Immunopathol. Pharmacol. 21 (1) (2008) 117–127, https://doi.org/10.1177/039463200802100113.
[80] M.R. Marselle, K.N. Irvine, S.L. Warber, Examining group walks in nature and multiple aspects of well-being: a large-scale study, Ecopsychology 6 (3) (2014) 134–147, https://doi.org/10.1089/eco.2014.0027.
[81] R.T. Crawford, The Impact of Ocean Therapy on Veterans with Posttraumatic Stress Disorder (Doctoral dissertation). Retrieved f rom, ProQuest, 2016 (10252268).
[82] C.M. Rogers, T. Mallinson, D. Peppers, High-intensity sports for posttraumatic stress disorder and depression: feasibility study of ocean therapy with veterans of Operation Enduring Freedom and Operation Iraqi Freedom, Am. J. Occup. Ther. 68 (4) (2014) 395–404, https://doi.org/10.5014/ajot.2014.011221.
[83] E. Morita, S. Fukuda, J. Nagano, N. Hamajima, H. Yamamoto, Y. Iwai, ... T. Shirakawa, Psychological effects of forest environments on healthy adults: shinrin-yoku (forest-air bathing, walking) as a possible method of stress reduction, Public Health 121 (1) (2007) 54–63, https://doi.org/10.1016/j.puhe.2006.05.024.
[84] Q. Li, T. Otsuka, M. Kobayashi, Y. Wakayama, H. Inagaki, M. Katsumata, H. Suzuki, Acute effects of walking in forest environments on cardiovascular and metabolic parameters, Eur. J. Appl. Physiol. 111 (11) (2011) 2845–2853, https://doi.org/10.1007/s00421-011-1918-z.
[85] G.X. Mao, X.G. Lan, Y.B. Cao, Z.M. Chen, Z.H. He, Y.D. Lv, Y.A.N. Jing, Effects of short-term forest bathing on human health in a broad-leaved evergreen forest in Zhejiang Province, China, Biomed. Environ. Sci. 25 (3) (2012) 317–324, https://doi.org/10.3967/0895-3988.2012.03.010.
[86] B.J. Park, Y. Tsunetsugu, T. Kasetani, T. Morikawa, T. Kagawa, Y. Miyazaki, Physiological effects of forest recreation in a young conifer forest in Hinokage Town, Japan, Silva Fenn. 43 (2) (2009) 291–301.
[87] S. Moussavi, S. Chatterji, E. Verdes, A. Tandon, V. Patel, B. Ustun, Depression, chronic diseases, and decrements in health: results from the world health surveys, Lancet 370 (2007) 851–858, https://doi.org/10.1016/S0140-6736(07)61415-9.
[88] S. Rosenbaum, B. Stubbs, P.B. Ward, Z. Steel, O. Lederman, D. Vancampfort, The prevalence and risk of metabolic syndrome and its components among people with posttraumatic stress disorder: a systematic review and meta-analysis, Metabolism 64 (8) (2015) 926–933, https://doi.org/10.1016/j.metabol.2015.04.009.
[89] Chief of Naval Operations, *Physical Readiness Program Policy Changes* (OPNAVINST 6110.1), (2011) Retrieved from https://navadmin.dodreads.com/2018/03/03/opnavinst-6110-1j-physical-readiness-program-policy-changes/.
[90] G. Cooney, K. Dwan, C. Greig, D. Lawlor, J. Rimer, F. Waugh, G. Mead, Exercise for depression, Cochrane Database Syst. Rev. (9) (2013) 1–157, https://doi.org/10.1002/14651858.CD004366.pub6 2013.
[91] J.A. Blumenthal, M.A. Babyak, K.A. Moore, W.E. Craighead, S. Herman, P. Khatri, P.M. Doraiswamy, Effects of exercise training on older patients with major depression, Arch. Intern. Med. 159 (19) (1999) 2349–2356, https://doi.org/10.1001/archinte.159.19.2349.
[92] J.A. Blumenthal, M.A. Babyak, P.M. Doraiswamy, L. Watkins, B.M. Hoffman, K.A. Barbour, A. Hinderliter, Exercise and pharmacotherapy in the treatment of major depressive disorder, Psychosom. Med. 69 (7) (2007) 587–596, https://doi.org/10.1097/PSY.0b013e318148c19a.
[93] P. Ekkekakis, Honey, I shrunk the pooled SMD! Guide to critical appraisal of systematic reviews and meta-analyses using the Cochrane review on exercise for depression as example, Ment. Health Phys. Act. 8 (2015) 21–36, https://doi.org/10.1016/j.mhpa.2014.12.001.
[94] QualityMetric Inc. SF-36v2: Short Form Health Survey Instrument - 36 Item Version (Version 2). Lincoln, RI.