European Journal of Applied Physiology
https://doi.org/10.1007/s00421-026-06481-6

ORIGINAL ARTICLE

Creatine supplementation effects on cognition, sleep, and menopause symptoms in response to six weeks of high-intensity resistance training in perimenopausal females

Paul A. Baker[1,2] · Gabrielle M. DelBiondo[1,3] · Caroline G. Kelly[1] · William M. Johnson[1] · Gregory J. Grosicki[4] · Abbie E. Smith-Ryan[1,3,5]

Received: 24 July 2026 / Accepted: 9 October 2026
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2026

Abstract

Introduction Perimenopause (PERI) women experience cognitive dysfunction, sleep disturbances, and menopause-related symptoms, yet few accessible and time-efficient interventions have been evaluated. The purpose of the current study was to determine the impact of creatine monohydrate (CM) during six weeks of high-intensity resistance training (HIRT) on cognition, sleep, and menopause symptoms in PERI females.

Methods PERI females ($n=51$; Age: 46 ± 5 yrs) were randomized to CM ($n=18$), placebo (PLA) ($n=20$), or a non-training control (CON) group ($n=13$). CM and PLA groups completed a 5-day loading phase (20 g·day^{-1}) followed by six weeks of maintenance (~4.5 g·day^{-1}) combined with twice-weekly progressive HIRT. Cognitive performance (Stroop, Rey 15-item memory test) and prefrontal oxygenation via functional near-infrared spectroscopy were assessed at baseline (BASE), pre- (PRE), and post-training (POST). Sleep and self-reported menopause symptoms (temperature sensitivity, fatigue, soreness, brain fog, and bloating) were monitored daily using a WHOOP wearable wristband.

Results Stroop reaction time improved over time (congruent: $p<0.001$; incongruent: $p=0.004$) with no group differences. Prefrontal oxygenation during the Stroop test increased with CM at POST (+0.32 ± 2.01%) compared to a decrease with PLA (−1.62 ± 2.01%; $p=0.003$). Sleep consistency showed a time × treatment interaction ($p<0.001$), with higher consistency in CM than PLA at week 6 (baseline-adjusted mean difference: +7.4% points, $p=0.015$).

Conclusions In this exploratory analysis, CM combined with HIRT was associated with greater prefrontal oxygenation, higher sleep consistency, and fewer self-reported temperature sensitivity days relative to PLA in PERI females.

Keywords Cerebral oxygenation · Midlife women · Menopause transition · Exercise

Communicated by Mark Elisabeth Willems

✉ Abbie E. Smith-Ryan
abbsmith@email.unc.edu

[1] Department of Exercise and Sport Science, University of North Carolina, 209 Fetzer Hall, CB# 8700, Chapel Hill, NC 27599-8700, USA

[2] Department of Applied Exercise Science, College of Charleston, Charleston, SC, USA

[3] Department of Allied Health Sciences, Human Movement Science Curriculum, University of North Carolina, Chapel Hill, NC, USA

[4] Department of Performance Science, WHOOP, Inc, Boston, MA, USA

[5] Department of Nutrition, Gillings School of Public Health, University of North Carolina, Chapel Hill, NC, USA

Abbreviations

CM	Creatine monohydrate
HIRT	High-intensity resistance training
PERI	Perimenopause
PLA	Placebo
CON	Control group
fNIRS	Functional near-infrared spectroscopy
NIRS	Near-infrared spectroscopy
TSI	Tissue saturation index
REM	Rapid eye movement
SWS	Slow wave sleep
BASE	Baseline testing visit
PRE	Post-loading/pre-training testing visit
POST	Post-intervention testing visit
1RM	One-repetition maximum

Published online: 24 October 2026

6RM	Six-repetition maximum
BMI	Body mass index
FSH	Follicle-stimulating hormone
MD	Mean difference
OR	Odds ratio
CI	Confidence interval
SD	Standard deviation

Introduction

Each year, over a million women in the United States go through the menopause transition and are often met with significant physiological changes (El Khoudary et al. 2019; Khoudary et al. 2020; Talaulikar 2022; O'Reilly et al. 2023). Specifically, the perimenopausal (PERI) period is marked by reduced cognition (Metcalf et al. 2023), sleep disruptions (Troìa et al. 2025), and a host of menopause symptoms such as hot flashes, brain fog, and elevated fatigue (Monteleone et al. 2018). In a recent community cohort of 14,234 women aged 45–55 years, including 4,044 PERI, those in PERI had higher odds of reporting poor memory, brain fog, and sleep difficulties than premenopausal females (Naysmith et al. 2026).

Evidence regarding the effects of exercise on these outcomes in PERI females remains limited and has focused primarily on aerobic exercise. A meta-analysis of 12 randomized trials involving 1,493 PERI females found a small-to-moderate improvement in self-reported sleep outcomes, although the included interventions primarily involved aerobic exercise, walking, and yoga (Zhao et al. 2023). Acute aerobic exercise has improved Stroop reaction time (Kotopoulea-Nikolaidi et al. 2019; Wu et al. 2024). Evidence for sleep is similar with eight weeks of aerobic exercise reduced Pittsburgh Sleep Quality Index scores from 19.99 ± 3.69 to 10.23 ± 4.94 in PERI females with sleep disturbances, although the study lacked a control group (Zhao et al. 2022).

Resistance training is an effective strategy for improving strength, muscle, and physical function (Currier et al. 2026), cognition (Dhahbi et al. 2025; Oporto-Colicoi et al. 2025), sleep (Kovacevic et al. 2018; Cunha et al. 2025), and even a reduction of hot flashes in post-menopausal females (Berin et al. 2019; Capel-Alcaraz et al. 2023; Nilsson et al. 2024). However, most middle aged women note a lack of time as a major barrier for regular engagement in exercise (Kimberly et al. 2018; Vasudevan and Ford 2022; Stimson et al. 2024). High intensity resistance training (HIRT) is an evidence-based time-efficient exercise modality for improving lean body mass, strength, and cardiometabolic health (Wingfield et al. 2015; Pihoker et al. 2019; Moro et al. 2020; Cabre et al. 2022). Importantly, exercise paradigms incorporating moderate-to-high intensity resistance exercise have been associated with improvements in executive function, mood, sleep quality, and fatigue, potentially through enhanced cerebral blood flow, neurotrophic signaling, and metabolic regulation (Sudo et al. 2022; Buzdagli et al. 2024).

Creatine monohydrate (CM) supplementation has been associated with improvements in cognition and sleep, potentially through its role in phosphocreatine buffering and adenosine triphosphate maintenance, with prior research demonstrating increased cerebral high-energy phosphates following supplementation (Candow et al. 2023). In a randomized crossover study of 15 young adults undergoing sleep deprivation, a single 0.35 $g \cdot kg^{-1}$ dose of CM increased cerebral total creatine by 5.3% and improved processing speed by approximately 10.6–18.4% across language, logic, and numerical tasks (Gordji-Nejad et al. 2024). In 21 naturally menstruating females completing twice-weekly resistance training, six weeks of CM supplementation resulted in approximately 48 min more sleep on training days than placebo (PLA), although habitual sleep duration and self-reported sleep quality did not improve (Aguiar Bonfim Cruz et al. 2024). Direct evidence during the menopause transition remains limited; however, a small trial of 36 PERI and postmenopausal females found that eight weeks of creatine hydrochloride improved reaction time by 6.6% compared with 1.2% with PLA and increased frontal brain creatine by 16.4% compared with 0.9% with PLA (Korovljev et al. 2026). In addition, CM supplementation may support thermoregulation by enhancing cellular hydration and fluid balance, a mechanism that could be relevant to menopause-related temperature sensitivity, although this relationship has not been directly established (Deminice et al. 2015; Moore et al. 2023). The effects of CM may be particularly relevant during the menopause transition, where hormonal fluctuations contribute to impaired cognition, disrupted sleep, altered thermoregulation, and neurological symptoms such as brain fog and fatigue (Weber et al. 2014; Troìa et al. 2025).

Although HIRT and CM supplementation may offer meaningful benefits during this midlife transition, no studies to date have specifically examined the effects of CM supplementation alongside HIRT on cognition, sleep, and menopause symptoms in PERI females. Therefore, the purpose of this study was to determine the effects of CM supplementation (loading and maintenance) during a six-week HIRT program on cognitive performance (scores & prefrontal oxygenation), wearable-derived sleep characteristics, and self-reported menopause symptoms in PERI females.

Materials and methods

Participants

Perimenopausal females ($n=51$) were enrolled in the study (Fig. 1; Table 1). Eligibility was confirmed via an initial screening survey followed by a telephone screening and a health/medical history questionnaire. Inclusion criteria included females aged 38–60 yrs who were transitioning through perimenopause (defined by the North American Menopause Society), alongside confirmation of either (1)≥2 months of self-reported menstrual cycle irregularity or (2) the presence of ≥2 perimenopause-related symptoms as defined by the questionnaire (Table 1).

Participants were recreationally active with at least 30 min of moderate intensity activity per week, but less than 3 days of progressive resistance training per week. Exclusion criteria included recent musculoskeletal injury limiting exercise participation, current use of CM or performance-enhancing supplements, pregnancy, and the presence of depression, anxiety, sleep-disorders or taking sleep-medication medical conditions or medications. All study procedures were approved by the University's Institutional Review Board, and participants provided written informed consent prior to participation. This trial was registered at ClinicalTrials.gov (NCT07027800).

Experimental design

This study used a randomized, double-blind, PLA-controlled, parallel-group design with an additional non-blinded control (CON) group. Following enrollment, participants completed an initial testing session (BASE) in which a WHOOP 5.0 wearable device (WHOOP, Inc., Boston, MA) was supplied, and cognitive assessments were performed while wearing an functional near infrared spectroscopy (fNIRS) device (PortaLite MKII, Artinis Medical Systems, Elst, The Netherlands) (Fig. 2). Participants were asked to wear the WHOOP band continuously throughout the study to track their sleep. Following the BASE visit, participants were randomized using a computer-generated block allocation scheme in a 2:2:1 ratio to CM, PLA, or CON. Participants assigned to CM or PLA completed a 5-day loading phase (20 g per day) and came back for a second visit (PRE). Because the WHOOP device was issued at BASE, wearable sleep and journal data began with supplementation and no pre-supplementation wearable data exist. The 7-day period preceding HIRT (hereafter Pre-HIRT), which encompassed the 5-day loading phase, therefore served as the reference period for the sleep and symptom analyses. This period was anchored to the initiation of HIRT so that training began after loading was complete, when muscle creatine stores in the CM group were expected to be saturated, and so that WHOOP sleep metrics that require several nights of data (see Sleep) were available before HIRT. Pre-HIRT is an on-supplement reference period and not an untreated baseline; it is distinct from the laboratory BASE visit (see Limitations). For the PRE visit, the same tests were completed as the BASE visit. Participants in the CM or PLA group continued with the assigned treatment a maintenance dose (~4.5 g CM per day in gummy form; Create Wellness Inc®) combined with a twice weekly 6-week progressive HIRT. Following the six weeks, all groups completed POST testing. The CON group participated in all testing visits (BASE and POST) and were asked to maintain their current exercise routine throughout the 6-weeks but did not come in for training sessions. For all laboratory testing visits, participants arrived after an overnight fast and were instructed to refrain from vigorous exercise (24 h) and caffeine/alcohol (12 h) prior to testing. Hydration status was assessed via

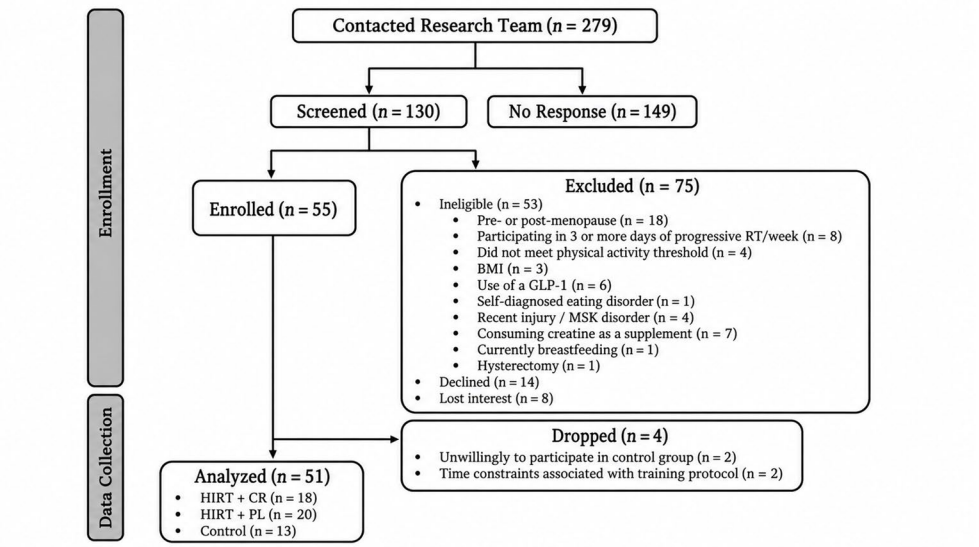


Fig. 1 Consort Diagram

Table 1 Baseline Participant Characteristics

	Total ($n=51$)
Demographics	
Age (years)	46 ± 5
Height (cm)	164 ± 7
Weight (kg)	69 ± 11
Body Mass Index (kg/m^2)	25.5 ± 4.0
White race, n (%)	42 (82)
Hispanic ethnicity, n (%)	4 (8)
Menopausal Transition Characteristics	
Self-identified as menopausal transition, n (%)	51 (100)
Reported ≥ 5 menopause-related symptoms, n (%)	51 (100)
Menstrual cycles irregular, n (%)	18 (35)
Menstrual periods less than once/month, n (%)	8 (16)
Menstrual periods more than once/month, n (%)	4 (8)
Recent change in bleeding duration, n (%)	28 (55)
Intermenstrual spotting/bleeding, n (%)	17 (33)
Recently heavier menstrual bleeding, n (%)	8 (16)
Days since last menstrual period, days	35 ± 52 days
Amenorrhea ≥ 60 days, n (%)	7 (14)
Hormonal Status	
Menopausal hormone therapy use, n (%)	15 (29)
Hormonal contraceptive use, n (%)	8 (16)
Intrauterine device use (hormonal, non-hormonal)	13 (25%)
Serum FSH, mIU/mL	23.6 ± 29.5
Serum FSH > 25 mIU/mL, n (%)	15 (29)

Values are presented as mean ± standard deviation. Participants ($n=51$) were randomized to creatine monohydrate, placebo, or control groups in a 2:2:1 allocation. Baseline (BASE) characteristics were assessed prior to the loading phase and HIRT intervention. There were no significant differences between groups at BASE. Perimenopausal status was determined from self-reported menstrual history and menopause-related symptom burden rather than physician diagnosis. All participants reported at least five menopause-related symptoms. Due to hormonal contraceptive use (39%) and menopausal hormone therapy use (25%), definitive STRAW + 10 staging was not possible in all participants. Menstrual characteristics, symptom burden, and hormonal measures are therefore presented to characterize reproductive-aging status

urine specific gravity using a refractometer and participants had to have a urine specific gravity of ≤ 1.025 A.U. These data were collected as part of a larger pilot randomized trial examining the effects of CM supplementation and six weeks of HIRT, for which muscle quality and strength were the registered primary outcomes (ClinicalTrials.gov: NCT07027800).

Exposures

Creatine supplementation

Participants assigned to the CM group completed a 5-day loading phase of 20 g·day^{-1} divided into four 5 g doses (5 g Creapure® [AlzChem Trostberg GmbH, Trostberg, Germany] + 2 g sugar-free Crystal Light® [Kraft Heinz, Chicago, IL, USA] per serving) mixed with 4–6 oz of water, followed by a maintenance dose of CM in gummy form (3 gummies: ~4.5 g·day^{-1}; 33 kcal; 8 g carbohydrate; 5 g sugar; Create Wellness Inc®, New York, NY, USA) daily for six weeks. Participants assigned to the PLA group completed a matched 5-day loading phase (~ 14 g·day^{-1} divided into four 3.5 g doses of sugar-free Crystal Light® [Kraft Heinz, Chicago, IL, USA]) mixed with 4–6 oz of water, followed by PLA gummies (Welch's® Fruit Snacks; Promotion in Motion Companies, Inc., Allendale, NJ, USA; 3 fruit snacks; 30 kcal; 7 g carbohydrate; 4.3 g sugar) daily for six weeks. Supplements were packaged in opaque containers by a study team member not involved in data collection to maintain blinding, and both participants and investigators were blinded to CM versus PLA assignment. During the 5-day loading phase, supplements were provided in premeasured opaque cups that participants were required to return, and during the daily supplementation phase, supplements were distributed in individual opaque bags that were also returned to monitor compliance. The CON group was not blinded and completed BASE and POST testing only without supplementation or participation in HIRT.

High-intensity resistance training

The HIRT intervention consisted of two supervised sessions per week for six weeks. Each session began with a self-selected warm-up (5 min of treadmill walking at a self-selected pace), followed by leg press, bench press, lunges, shoulder press, biceps curl, and triceps extension (Wingfield et al. 2015; Pihoker et al. 2019; Cabre et al. 2022). Participants were instructed to complete three sets of 6–10 repetitions per exercise with 30 s of rest between sets and 120 s between exercises. Training loads were individualized based on baseline strength testing, with initial loads set at 75% of one-repetition maximum (1RM) for leg press and bench press and 75% of 6-repetition maximum (6RM) for accessory exercises (lunges, shoulder press, biceps curl, and triceps extension). Load progression occurred when participants completed at least two sets of 10 repetitions at the prescribed weight; during the subsequent session, lower-body exercises (leg press and lunges) were increased by 10% and upper-body exercises (bench press, shoulder press, biceps curl, and triceps extension) were increased by 5%. For the accessory exercises (lunges, shoulder press, biceps curl, and triceps extension), initial loads were set at 75% of the estimated 6RM. Both training groups completed the same HIRT program using identical intensity and progression criteria. Although individual loads progressed according to performance, total training volume did not significantly differ between CM and PLA.

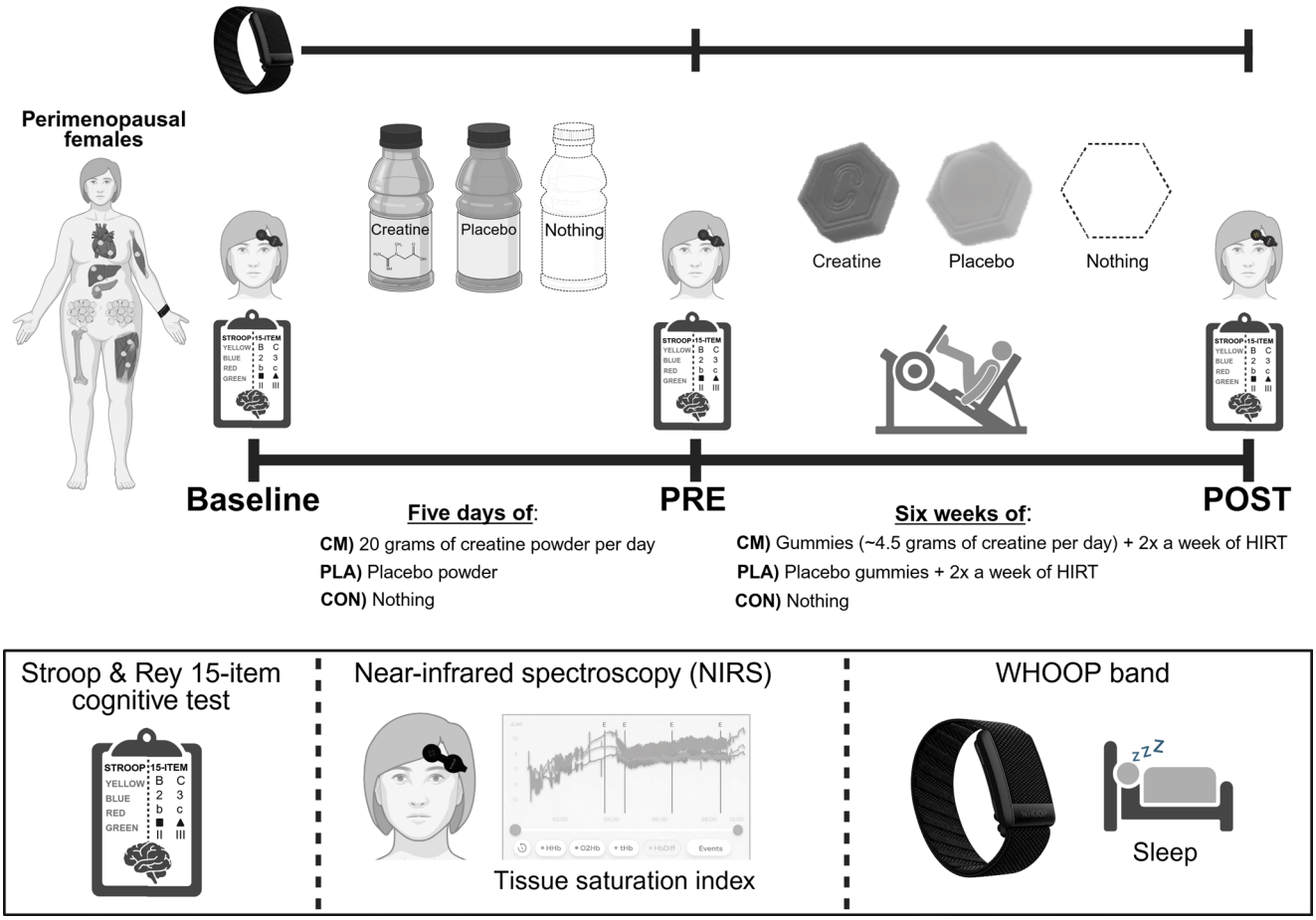


Fig. 2 Experimental overview. Perimenopausal females completed baseline testing (BASE), including cognitive assessments and functional near-infrared spectroscopy (fNIRS) measurements while wearing a WHOOP 5.0 device. Participants were randomized to creatine monohydrate (CM), placebo (PLA), or control (CON). Following BASE testing, participants completed a 5-day loading phase consisting of CM powder (20 g·day^{-1}), placebo powder, or no supplementation, followed by a pre-training visit (PRE). Participants then completed a 6-week high-intensity resistance training (HIRT) intervention with CM gummies (~4.5 g·day^{-1}), placebo gummies, or no intervention. Wearable outcome windows were categorized as Pre-HIRT (the 7 days preceding the start of HIRT, encompassing the supplement loading period), Early (Week 2 of HIRT), Middle (Week 4 of HIRT), and Late (Week 6 of HIRT), with final testing conducted at POST. Outcome assessments included Stroop and Rey 15-item memory test, prefrontal oxygenation measured using functional near-infrared spectroscopy (fNIRS) via tissue saturation index, and continuous sleep and recovery monitoring using the WHOOP 5.0 wearable device

Outcomes

Cognition

Cognition was assessed using two standardized tasks: the 2-min demo version of Stroop test (executive function/inhibitory control) administered through the PsyToolkit (Scarpina and Tagini 2017) and the Rey 15-item memory test (short-term visual memory) (Goldberg and Miller 1986). The 2-min PsyToolkit Stroop demonstration consisted of 40 trials, with participants using the r, g, b, and y keyboard keys to identify the displayed font color. Congruent, incongruent, and incongruent-congruent reaction times were recorded from the output. Cognitive testing was administered at BASE, following the 5-day loading phase and before training (PRE), and following the 6-week HIRT intervention (POST) using standardized instructions and testing conditions in the morning. The same standardized PsyToolkit Stroop task was administered at each testing visit, with the trial sequence randomized at each administration. The BASE assessment, conducted before supplementation loading, served as familiarization with the cognitive tasks, whereas the post-loading PRE assessment served as the baseline for evaluating changes during HIRT. Investigators responsible for processing the cognitive data remained blinded to CM versus PLA assignment. Oxygenation during cognitive testing was assessed using a continuous-wave fNIRS device (PortaLite MKII, Artinis Medical Systems, Elst, The Netherlands) positioned over the left frontal region. The device uses near-infrared wavelengths of approximately 760 and 850 nm and three long source-detector separations of approximately 29, 35, and 41 mm.

Tissue saturation index (TSI) was calculated by the device using spatially resolved spectroscopy across these source-detector distances. Data were acquired at a sampling frequency of 100 Hz using OxySoft (Artinis Medical Systems, Elst, The Netherlands). Probe placement was standardized by measuring a fixed distance above the eyebrow (3 cm) and to the left (3 cm) of the midline of the forehead (Boushel and Piantadosi 2000; Jones et al. 2016). The device was secured with tape and an opaque headband to minimize ambient light. Participants rested for five minutes before the commencement of each cognitive test to have a standardized baseline for fNIRS recordings. Outcomes derived from fNIRS included tissue saturation index (TSI) peak and average during the Stroop and the Rey 15-item memory test tasks.

Sleep

Sleep was assessed using a wrist-worn WHOOP 5.0 device (WHOOP, Inc., Boston, MA) worn continuously throughout the study period. Sleep staging was performed by WHOOP's proprietary algorithm, and nightly summary data were exported from WHOOP's research data platform on 26 January 2026. Only the primary nocturnal sleep episode of each day was analyzed; naps were not included. A night was considered valid if a primary sleep episode was detected; nights without a detected sleep episode (device not worn or not charged) were treated as missing. Participants could view their sleep and recovery feedback in the WHOOP application and could edit the start and end bounds of a detected sleep period; sleep and wake classification within the sleep period was determined solely by the WHOOP algorithm and could not be edited. The WHOOP strap combines accelerometry and photoplethysmography to detect major sleep periods of at least one hour in duration. Within each detected sleep period, WHOOP classified each 30-second interval into one of four sleep stages: wake, light, slow wave sleep (SWS), and rapid eye movement (REM) sleep. Sleep duration was defined as total sleep time (light, SWS, REM). Sleep disturbances were defined by WHOOP as the number of discrete periods of wakefulness detected within the sleep period.

The device also reports sleep consistency, a measure adapted from the Sleep Regularity Index (Phillips et al. 2017), which evaluates sleep onset and offset times over a rolling window, comparing the most recent 24 h to the preceding four days, with comparisons of days further apart being assigned progressively lower weights. This approach acknowledges that variations in sleep schedules have an impact on circadian rhythms beyond 24 h due to physiological lag. The WHOOP sleep consistency measure excludes brief wake events within sleep periods, focusing specifically on sleep onset and offset patterns relevant to circadian alignment. The final metric is provided on a 100-point scale, with 0 representing complete misalignment between the most recent 24 h and the preceding four days and 100 representing perfect alignment. Because the score compared each night with the preceding four nights, it is not generated for the first few nights of wear. The 7-day Pre-HIRT period therefore allowed the metric to stabilize before HIRT began. Adjacent daily scores share input nights by construction, which was addressed analytically using participant-level window summaries (see Statistical Analysis). Sleep consistency is a proprietary score adapted from, but not equivalent to, the Sleep Regularity Index (Phillips et al. 2017). WHOOP-derived sleep measures have been validated against polysomnography for two-stage (sleep/wake) and four-stage classification in earlier hardware generations (86–89% agreement for two-stage classification) (Miller et al. 2020, 2022). Independent validation of the specific hardware and algorithm version used in the present study has not been published, and sleep-stage and consistency outputs should be interpreted with this in mind.

Menopause symptoms

Participants logged menopause symptoms daily through the WHOOP journal feature. Five PERI/menopause symptoms were selected as a priori for analysis: temperature sensitivity, fatigue, soreness, brain fog, and bloating. Each was treated as binary (present/absent) on any given day for analysis. Journal items are fixed yes/no prompts answered each morning in the WHOOP application (e.g., "Experienced bloating?"; "Felt unusually sensitive to temperature changes?"); the items have not been psychometrically validated in perimenopausal women. Modeled percentages therefore represent the probability that a given day was symptom-positive, not the percentage of women who experienced the symptom. Temperature sensitivity refers to increased perceived sensitivity to heat or difficulty tolerating temperature changes. Fatigue refers to persistent physical or mental exhaustion, reduced energy levels, and diminished capacity for daily activities or exercise. Cognitive difficulties, commonly referred to as "brain fog," include perceived impairments in concentration, memory, mental clarity, multitasking ability, or processing speed. Bloating describes feelings of abdominal fullness, swelling, tightness, or gastrointestinal discomfort often associated with fluid retention or digestive changes during the menopause transition. To ensure adequate data density for symptom models, participants were required to have at least 3 days of symptom data (yes or no) in at least three of the four-time windows.

Statistical analysis

BASE characteristics were compared using one-way analysis of variance (ANOVA). For cognitive outcomes, a two-way repeated measures mixed-effects model using restricted maximum likelihood was performed for time×treatment interaction, with fixed effects of group (CM, PLA, CON), and time, with participant included as a random effect with data missing at random. When significant interaction or main effects were observed, Tukey's multiple comparisons tests were performed. Given the 2:2:1 randomization scheme, two-way repeated measures mixed-effects analyses were conducted including only the CM and PLA. The three-group analysis included CM, PLA, and CON. In addition, a planned CM-versus-PLA analysis directly compared the equally allocated groups that completed the same HIRT intervention and differed in supplementation. This comparison was specified and was not conducted in response to the three-group result. Statistical significance was set a priori at $p<0.05$. BASE cognitive, and NIRS outcomes analyses were conducted in GraphPad Prism (GraphPad Software, San Diego, CA, USA).

Sleep and menopause symptom outcomes were analyzed using daily-level WHOOP data across four predefined time windows anchored to each participant's PRE visit, which immediately preceded the first HIRT session: Pre-HIRT (the 7 days before PRE, encompassing the supplement loading phase), Early (days 7–13 after PRE; HIRT week 2), Middle (days 21–27; week 4), and Late (days 35–41; week 6). Each window was a fixed 7-day block, so no observation could fall in two windows and no interpolation at window boundaries was required. Sleep outcomes included sleep duration, sleep consistency, sleep disturbances, slow wave sleep percentage, and REM sleep percentage. Because WHOOP devices were issued at BASE, no wearable data preceded supplementation and the Pre-HIRT window is an on-supplement reference period rather than an untreated baseline.

The primary analysis was a covariate-adjusted mixed-effects model (analysis-of-covariance form) fit to the Early, Middle, and Late observations. For sleep outcome Y of participant i on night j: $Y_{ij} = \beta_0 + \beta_1 Window_{ij} + \beta_2 Group_i + \beta_3 (Window \times Group)_{ij} + \beta_4 Age_i + \beta_5 BMI_i + \beta_6 \bar{Y}_i,Pre + u_i + \varepsilon_{ij}$, where $\bar{Y}_i,Pre$ is the participant's mean of Y over the Pre-HIRT window, $u_i \sim N(0, \sigma^2_u)$ is a participant random intercept, and $\varepsilon_{ij} \sim N(0, \sigma^2_e)$. Pre-HIRT nights entered the model only through $\bar{Y}_i,Pre$ and were not part of the outcome vector, so the covariate and the modeled observations did not overlap. Menopause symptoms were modeled with the same linear predictor on the logit scale, $logit[Pr(Y_{ij} = 1)] = \beta_0 + \ldots + \beta_6 \bar{p}_i,Pre + u_i$, where $\bar{p}_i,Pre$ is the proportion of symptom-positive days during Pre-HIRT. Age and BMI were included a priori because they may influence sleep and menopause-related symptoms; the Pre-HIRT mean was included a priori to adjust for differences in starting values between groups irrespective of their statistical significance.

Sleep models were fit with lmerTest (R version 4.4.2), using restricted maximum likelihood for estimation and Satterthwaite degrees of freedom for contrasts and symptom models were fit by maximum likelihood with glmmTMB. Interaction and main effects were tested with likelihood ratio tests between nested models refit by maximum likelihood.

Primary analyses compared CM versus PLA; secondary analyses included all three groups. When the Window × Group interaction was significant, estimated marginal means were compared between groups within each window (emmeans; Tukey adjustment for the three-group model); sleep contrasts are reported as adjusted mean differences and symptom contrasts as odds ratios, each with 95% confidence intervals. Convergence was verified for every model, and participant random-intercept standard deviations are reported in the supplementary table footnotes.

Two supporting analyses were added. First, to describe change from the Pre-HIRT window, a time-level mixed model with the same fixed effects but without the Pre-HIRT covariate was fit to all four windows with Pre-HIRT as the reference level; when a main effect of time was present without an interaction, Early, Middle, and Late were compared with Pre-HIRT using Dunnett adjustment. Second, because daily observations within a participant are serially correlated, and adjacent sleep consistency scores share input nights by construction, the primary model was refit to participant-level window summaries (the mean per participant per window for sleep outcomes and the proportion of symptom-positive days for symptoms). For outcomes with a significant interaction, leave-one-participant-out analyses were performed, and symptom models were inspected for complete or quasi-complete separation (group-by-window cells with no symptom-positive days).

Statistical significance was set a priori at $p<0.05$. Given the number of exploratory outcomes, no adjustment for multiplicity across outcome families was applied, and isolated p values below 0.05 are not interpreted as confirmatory.

An a priori sample-size calculation was conducted using for the parent trial's primary outcomes of muscle size, muscle quality, and muscle strength. Assuming an effect size of 0.48, 80% power, $\alpha=0.05$, a repeated-measures correlation of 0.50, and a nonsphericity correction of 1.0, 42 participants were required. Target enrollment was increased to 48 to account for approximately 15% attrition. Cognition, fNIRS, sleep, and symptom outcomes were planned secondary outcomes and were not separately used to determine the sample size. Only participants that had ≥80% compliance with both the supplementation and training protocol were analyzed.

Results

Cognition

Stroop test & fNIRS

For congruent reaction time (ms) there was no time×treatment interaction with HIRT ($p=0.853$), but a lower reaction time was observed over the six weeks (time $p<0.001$); Table 2. For incongruent reaction time (ms), there was no difference between the CM, PLA, or CON groups (time×treatment $p=0.890$), but a lower reaction time was observed over the six weeks (time $p=0.004$). For the incongruent – congruent difference, there were no interactions or main effects (time×treatment $p=0.834$, time $p=0.340$, treatment $p=0.568$).

During the Stroop test, TSI peak difference did not differ between CM, PLA, or CON groups across visits (time×treatment $p=0.083$), but there was a main effect of time (time $p=0.015$); Fig. 3A. TSI Stroop peak difference at POST was greater in CM (mean ± SD: +0.32 ± 2.01%) compared with PLA (−1.62 ± 2.01%; time × treatment $p=0.013$; CM vs. PLA at POST: $p=0.003$; Fig. 3B). There was no interaction in TSI Stroop average difference between groups over time (time×treatment $p=0.659$), but a main effect of time ($p=0.038$) and treatment ($p=0.049$) was found.

Rey 15-item memory test & fNIRS

For the Rey 15-item memory test, there was no difference in the number of items remembered between the CM, PLA, or CON groups during the loading phase or after the six weeks of supplementation with HIRT; Table 3. During the 10-second memorization phase of the Rey 15-item memory test, there were no differences in TSI peak difference (time×treatment $p=0.204$) or average difference (time×treatment $p=0.131$) between CM, PLA, and CON groups. Similarly, during the recall phase of the Rey 15-item memory test, there was no difference in TSI peak difference (time×treatment $p=0.150$) or average difference (time×treatment $p=0.555$) between CM, PLA, and CON.

Sleep outcomes

Sleep variables are presented in Fig. 4 and Supplementary Table 1 (CM vs. PLA) and **2** (all three groups). Of the 38 participants randomized to CM or PLA, 37 had WHOOP data and 30 (CM: $n=15$; PLA: $n=15$) had data in the Pre-HIRT window and were included in the primary analysis; the seven excluded participants (CM: 2; PLA: 5) began wearing the device after the PRE visit and therefore had no Pre-HIRT reference data. They did not differ from included

Table 2 Stroop outcome values and frontal lobe oxygenation responses assessed via near-infrared spectroscopy

	Creatine			Placebo			Control			*p*-value		
	BASE ($n=18$)	PRE ($n=18$)	POST ($n=18$)	BASE ($n=20$)	PRE ($n=20$)	POST ($n=20$)	BASE ($n=13$)	PRE ($n=13$)	POST ($n=13$)	Time×Treatment	Treatment	Time
Congruent (ms)	990 ± 204	900 ± 149	864 ± 127	1014 ± 201	930 ± 182	904 ± 195	1058 ± 164	968 ± 163	910 ± 152	0.853	0.665	<0.001
Incongruent (ms)	1083 ± 178	1047 ± 231	968 ± 156	1077 ± 311	1027 ± 194	1010 ± 157	1124 ± 170	1048 ± 192	1026 ± 161	0.890	0.893	0.004
Incongruent – Congruent (ms)	94 ± 103	147 ± 214	105 ± 103	63 ± 234	99 ± 81	106 ± 96	66 ± 122	78 ± 115	116 ± 71	0.834	0.568	0.340
TSI Peak (%)	-0.06 ± 0.67	-0.36 ± 0.97	0.32 ± 2.01	-0.29 ± 0.79	-0.87 ± 1.28	-1.62 ± 2.01	0.36 ± 0.78	-1.38 ± 3.01	-2.80 ± 5.87	0.083	0.057	0.015
TSI Average (%)	0.30 ± 0.83	0.08 ± 0.90	0.35 ± 0.99	0.25 ± 0.79	-0.31 ± 1.22	-0.22 ± 1.20	0.92 ± 1.48	0.55 ± 1.05	0.78 ± 1.14	0.659	0.049	0.038

Values are presented as mean ± standard deviation. Participants were assigned to creatine (CM), placebo (PLA), or control (CON) groups and assessed at three time points: baseline (BASE), post-loading/pre-training (PRE), and post-intervention (POST). Cognitive performance was evaluated using the Stroop test, with outcomes reflecting reaction time (ms) and/or accuracy for congruent and incongruent conditions. Frontal lobe oxygenation during Stroop task performance was assessed using near-infrared spectroscopy (NIRS), with tissue saturation index (TSI) reported as both peak and average responses (change from baseline during task)

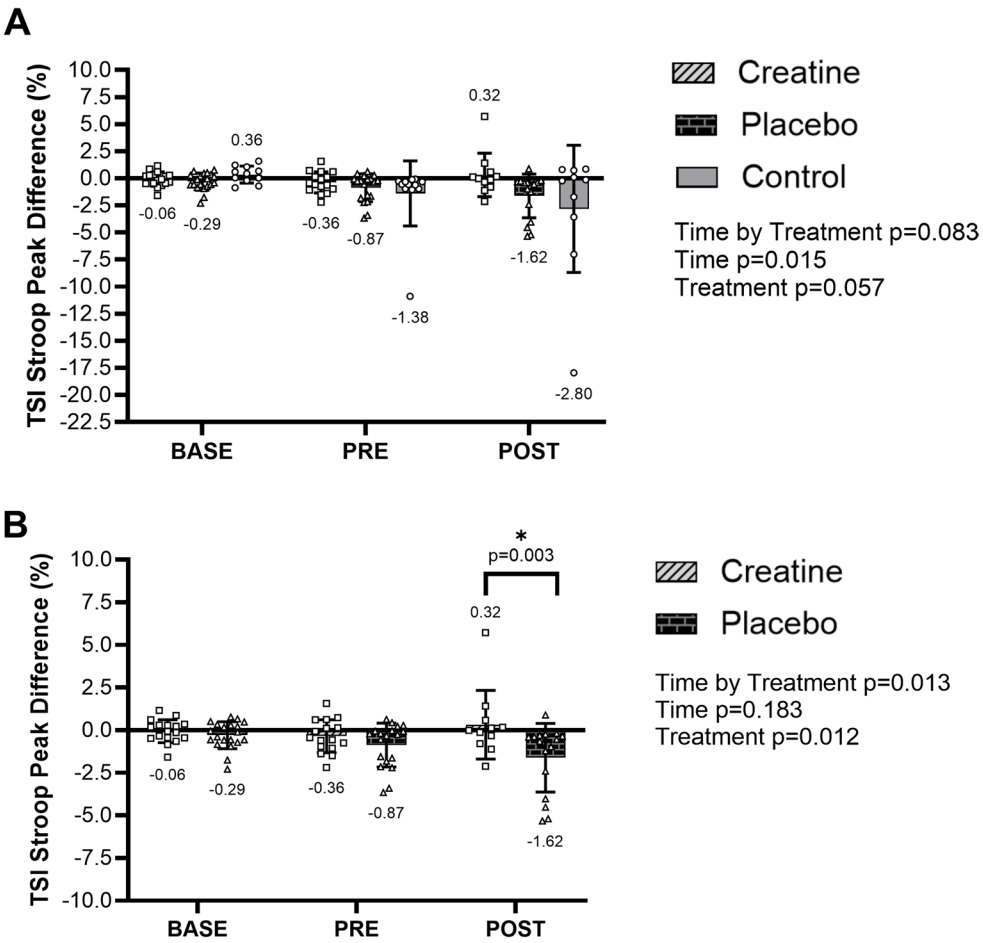


Fig. 3 Change in prefrontal tissue saturation index (TSI) during the Stroop task before and after six weeks of HIRT with creatine or placebo supplementation in peri-menopausal females. TSI Stroop Peak Difference was calculated as peak Stroop TSI minus baseline TSI at BASE, PRE (before six weeks of HIRT), and POST (after six weeks of HIRT). (A) Creatine, placebo, and a non-training control group. (B) Creatine and placebo groups only. Data are presented as individual values with means. Abbreviations: TSI, tissue saturation index. (*$p<0.05$)

participants in age ($p=0.68$) or BMI ($p=0.29$). The three-group analysis included 37 participants (CM: $n=15$; PLA: $n=15$; CON: $n=7$). The primary model comprised 596–599 nights across the Early, Middle, and Late windows (median 7 nights per participant per window; range 3–7), with 178 Pre-HIRT nights (median 6 per participant; range 3–7) used to derive the baseline covariate. Nights per participant and window are provided in Supplementary Table S5.

Sleep consistency was the only sleep outcome for which a significant time × treatment interaction was observed in the primary analysis ($p<0.001$; Fig. 4B; Table S1). CM and PLA did not differ at Early (adjusted mean difference [MD]: -2.0% points; 95% CI -7.7, 3.7; $p=0.485$) or Middle (MD: +1.5; 95% CI -4.3, 7.2; $p=0.612$), but diverged at Late, with higher consistency in CM than PLA (MD: +7.4; 95% CI 1.5, 13.3; $p=0.015$). Descriptively, PLA sleep consistency declined from Pre-HIRT (74.8 ± 10.4%) to Late (69.4 ± 17.1%), whereas CM was stable (Pre-HIRT: 75.8 ± 12.4%; Late: 77.2 ± 14.9%); the groups did not differ at Pre-HIRT (time-level model: MD + 0.2; $p=0.958$). The result was robust to the analytic approach: in the participant-level sensitivity analysis the Late difference was +7.8% points (95% CI 0.6, 14.9; $p=0.034$; interaction $p=0.067$), and in leave-one-participant-out analyses the interaction remained significant in all 30 refits ($p\leq0.010$) with Late differences of +5.8 to +9.6% points (all $p<0.05$). The pattern was consistent in the three-group model (interaction $p=0.002$; Table S2), in which CM was higher than PLA at Late (MD: +7.4; 95% CI 0.6, 14.2; $p=0.029$) and PLA was lower than CON at Late (MD: -9.7; 95% CI -18.2, -1.1; $p=0.023$), whereas CM and CON did not differ (MD: -2.3; 95% CI -11.2, 6.6; $p=0.815$).

No time × treatment interactions or main effects of time were observed for sleep duration (interaction $p=0.662$; time $p=0.734$; Fig. 4A), SWS% (interaction $p=0.467$; time $p=0.223$), REM% (interaction $p=0.341$; time $p=0.086$; Fig. 4C-D), or sleep disturbances (interaction $p=0.197$; time $p=0.856$; Table S1). A main effect of treatment group was observed for sleep duration ($p=0.044$), reflecting shorter sleep in CM than PLA that was already present during Pre-HIRT (CM: 409 ± 69 min; PLA: 442 ± 78 min) and persisted after adjustment for the Pre-HIRT mean (MD: -17.5 min; 95% CI -36.6, 1.6; participant-level analysis: -17.3 min; 95% CI -36.3, 1.6). In the three-group model, a time × treatment interaction was observed for REM% ($p=0.036$; Table

Table 3 Rey 15-Item Memory Test performance and frontal lobe oxygenation responses assessed via near-infrared spectroscopy

	Creatine			Placebo			Control			Interaction	Treatment	Time
	BASE	PRE	POST	BASE	PRE	POST	BASE	PRE	POST	*p*-value	*p*-value	*p*-value
Items recalled	14.70 ± 0.92	15.00 ± 0.00	14.78 ± 0.94	15.00 ± 0.00	14.85 ± 0.67	15.00 ± 0.00	14.93 ± 0.26	15.00 ± 0.00	15.00 ± 0.00	0.397	0.293	0.776
TSI Peak Memory	-2.82 ± 6.63	-2.17 ± 3.97	-1.48 ± 1.05	-1.46 ± 1.30	-1.78 ± 1.44	-2.73 ± 3.50	-1.79 ± 1.21	-1.60 ± 1.23	-0.73 ± 0.86	0.204	0.813	0.634
TSI Average Memory	0.65 ± 0.98	0.66 ± 0.71	0.47 ± 094	0.36 ± 0.88	0.25 ± 1.01	0.20 ± 0.78	0.15 ± 0.79	0.69 ± 1.48	1.18 ± 1.52	0.131	0.381	0.602
TSI Peak Recall	-2.83 ± 6.73	-2.24 ± 4.22	-1.51 ± 1.05	-1.25 ± 1.32	-1.81 ± 1.77	-2.72 ± 3.47	-1.24 ± 0.78	-1.54 ± 1.25	-0.66 ± 1.03	0.150	0.644	0.720
TSI Average Recall	0.63 ± 0.99	0.59 ± 0.86	0.44 ± 0.79	0.57 ± 0.85	0.22 ± 1.33	0.22 ± 1.23	0.70 ± 1.19	0.75 ± 1.43	1.26 ± 1.59	0.555	0.371	0.775

Values are presented as mean ± standard deviation. Participants were assigned to creatine (CM), placebo (PLA), or control (CON) groups and assessed at three time points: baseline (BASE), post-loading/pre-training (PRE), and post-intervention (POST). Cognitive performance was evaluated using the Rey 15-Item Memory Test, with "items recalled" representing the total number of correctly recalled items. Frontal lobe oxygenation during the memory and recall phases was assessed using NIRS, with tissue saturation index (TSI) reported as both peak and average responses (change from baseline during task)

S2), but no pairwise contrast reached significance after Tukey adjustment.

Menopause symptoms

Menopause symptom data are presented in Fig. 5 and Supplementary Table 3 (CM vs. PLA) and **S4** (all three groups). After applying inclusion thresholds, 18 to 24 participants were retained per outcome in the primary analysis (CM: $n = 7–11$; PLA: $n = 10–13$), contributing 330 to 429 daily observations across the Early, Middle, and Late windows (median 7 days per participant per window; range 1–7), with 72 to 92 Pre-HIRT days used to derive the baseline covariate. Numbers of symptom-positive and observed days by group and window are provided in Supplementary Table S6. Percentages in the text and tables represent the proportion of days on which a symptom was reported, not the proportion of women reporting it. All models converged, and no group-by-window cell in the CM versus PLA models was completely separated (each contained at least one symptom-positive day).

Temperature sensitivity showed a significant time × treatment interaction ($p < 0.001$; Fig. 5A; Table S3). CM and PLA did not differ at Early (OR = 0.55; 95% CI 0.06, 5.33; $p = 0.607$), but CM had lower odds of a temperature-sensitivity day at Middle (OR = 0.013; 95% CI 0.0004, 0.40; $p = 0.013$) and Late (OR = 0.010; 95% CI 0.0004, 0.22; $p = 0.004$). The observed prevalence of temperature-sensitivity days in CM declined from 14.6% (6 of 41 days) at Pre-HIRT to 3.5% at Late, whereas PLA was 35.9% (14 of 39 days) at Pre-HIRT and 34.9% at Late. Main effects of time ($p = 0.049$) and treatment group ($p = 0.010$) were also observed. This finding was less robust than the sleep consistency result. In the participant-level sensitivity analysis, the difference in the proportion of symptom-positive days at Late was −0.23 (95% CI -0.47, 0.01; $p = 0.064$) and the interaction was not significant ($p = 0.28$). In leave-one-participant-out analyses the interaction remained significant in 18 of 19 refits, but omitting the only CM participant who reported temperature sensitivity at Middle or Late left no symptom-positive days in CM at those windows, and the contrast could not be estimated. The interaction was also present in the three-group model ($p < 0.001$; Table S4), although CON cells at Middle and Late contained no symptom-positive days, so contrasts involving CON are not interpretable.

Bloating also demonstrated a significant time × treatment interaction ($p < 0.001$; Table S3). CM and PLA did not differ at Early (OR = 0.99; 95% CI 0.29, 3.38; $p = 0.988$), but CM had higher odds of a bloating day at Middle (OR = 16.6; 95% CI 4.3, 63.8; $p < 0.001$) and Late (OR = 7.2; 95% CI 1.6, 31.8; $p = 0.009$), with an overall group effect (OR = 4.3;

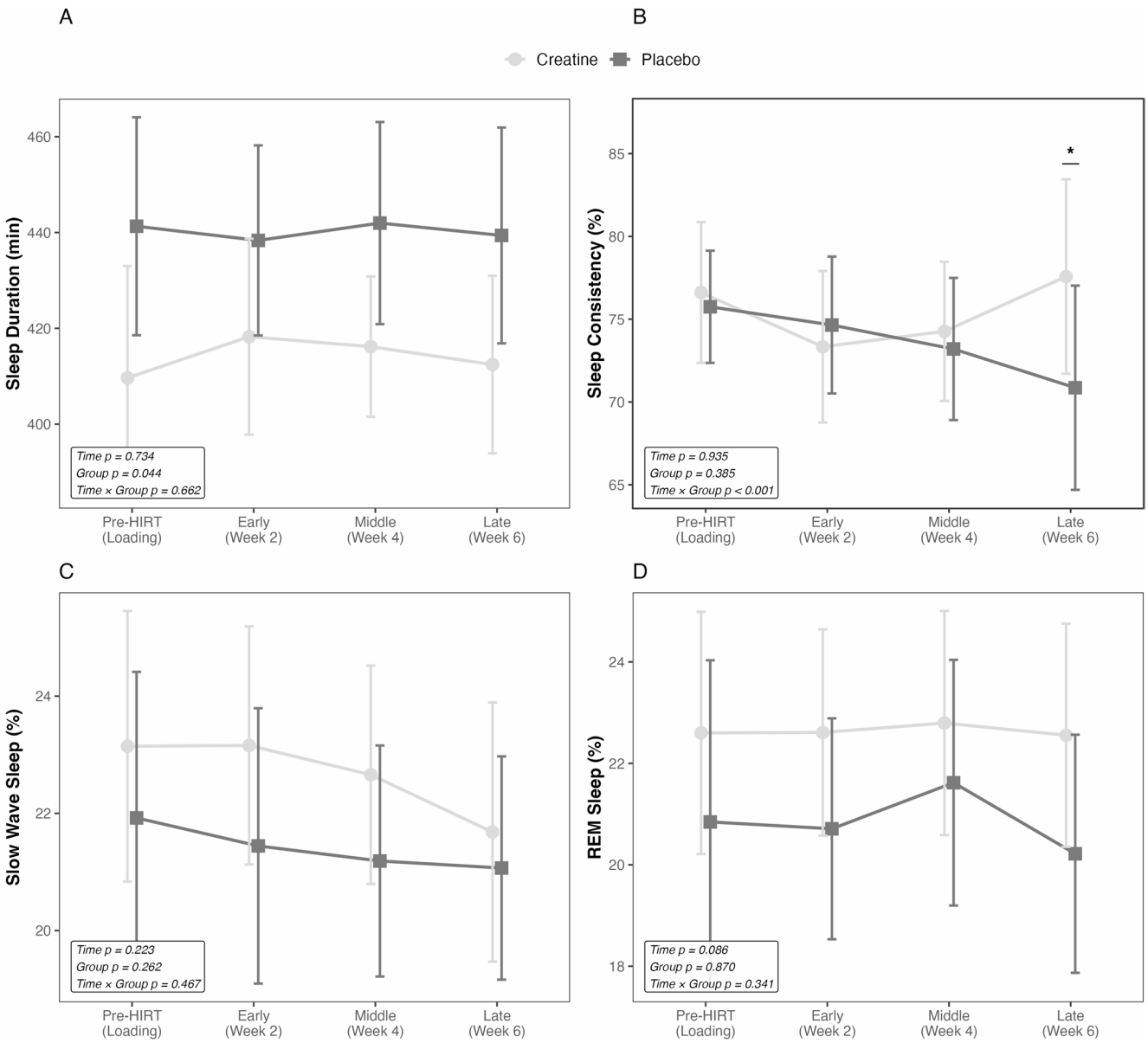


Fig. 4 Sleep outcomes during six weeks of high-intensity resistance training (HIRT) in perimenopausal females. Sleep outcomes including (A) sleep duration (min), (B) sleep consistency (%), (C) slow wave sleep (SWS; %), and (D) rapid eye movement (REM) sleep (%) were assessed using a WHOOP wearable device. Data are group means of participant-level window means with 95% confidence intervals for the creatine and placebo groups across four time windows: Pre-HIRT (the 7 days before HIRT, encompassing the loading phase), Early (week 2), Middle (week 4), and Late (week 6) of the intervention. *$p<0.05$ for the creatine versus placebo contrast from the covariate-adjusted mixed model

95% CI 1.6, 11.2; $p=0.005$). The observed prevalence of bloating days was 64.3% (27 of 42 days) in CM versus 17.1% (12 of 70 days) in PLA at Middle, and 32.6% (14 of 43) versus 10.6% (7 of 66) at Late, compared with 37.9% and 39.5% at Pre-HIRT.

No time × treatment interactions were observed for fatigue ($p=0.555$), soreness ($p=0.259$), or brain fog ($p=0.533$; Table S3). Main effects of time were detected for fatigue ($p=0.050$; Fig. 5B), soreness ($p<0.001$; Fig. 5C), and brain fog ($p=0.006$; Fig. 5D). In the time-level model, fatigue was lower than Pre-HIRT at Middle (OR = 0.40; 95% CI 0.18, 0.88; $p=0.017$) and Late (OR = 0.28; 95% CI 0.13, 0.62; $p<0.001$) but not at Early (OR = 0.55; 95% CI 0.25, 1.21; $p=0.180$), with similar declines in both groups (observed prevalence of fatigue days, CM: 63.8% to 47.9%; PLA: 66.7% to 48.1%). Soreness was higher than Pre-HIRT at Early (OR = 4.38; 95% CI 1.88, 10.24; $p<0.001$), returned toward Pre-HIRT at Middle (OR = 1.30; 95% CI 0.55, 3.04; $p=0.786$), and remained modestly elevated at Late (OR = 2.54; 95% CI 1.10, 5.87; $p=0.025$). For brain

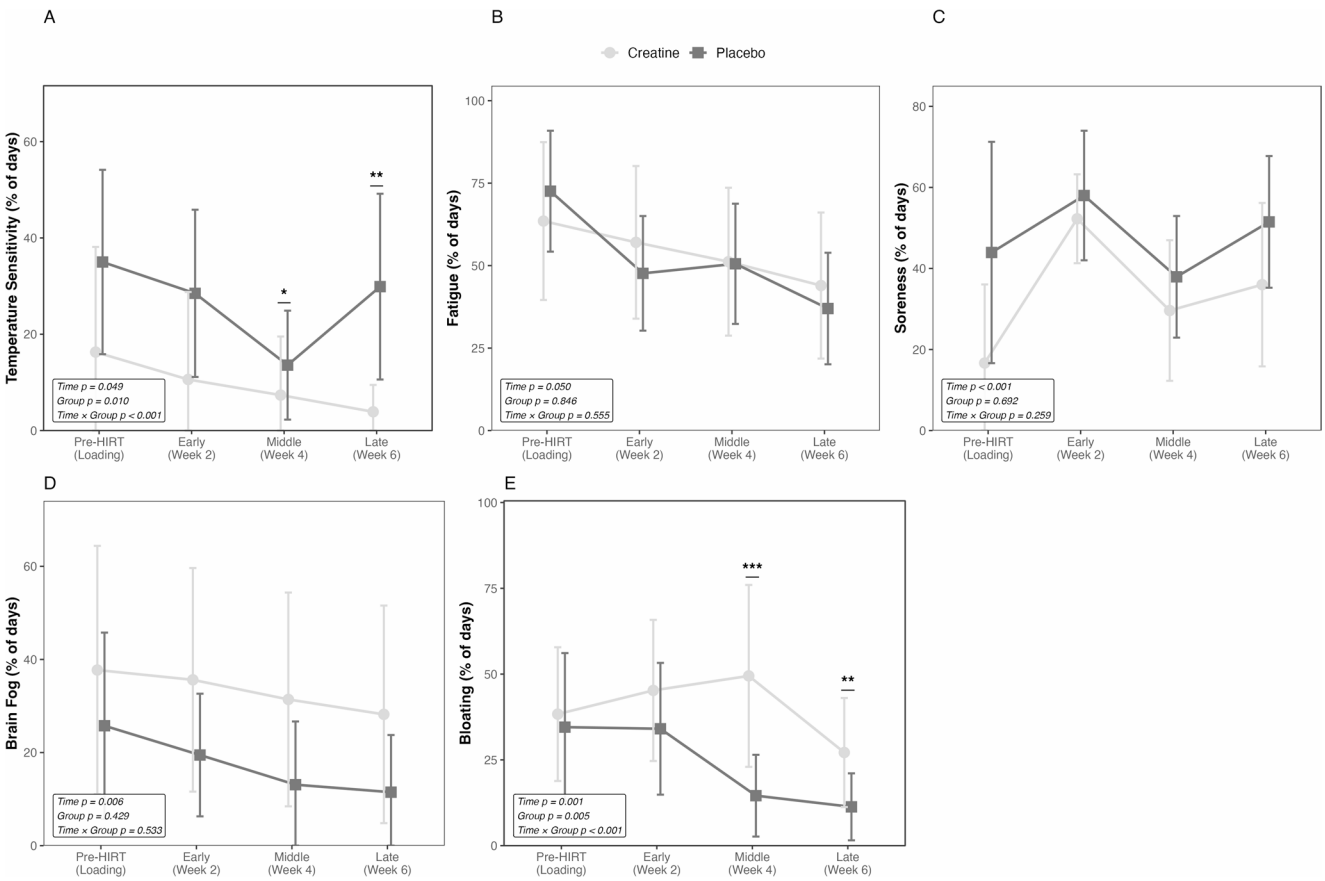


Fig. 5 Menopause symptom responses during six weeks of HIRT in PERI females. Menopause-related symptoms including (A) temperature sensitivity (%), (B) fatigue (%), (C) soreness (%), (D) brain fog (%), and (E) bloating (%) were assessed via daily self-report using the WHOOP journal feature. Data are the percentage of symptom-positive days (group means of participant-level proportions) with 95% confidence intervals for the creatine and placebo groups across four time windows: Pre-HIRT (the 7 days before HIRT, encompassing the loading phase), Early (week 2), Middle (week 4), and Late (week 6) of the intervention. Percentages represent the probability of a symptom-positive day, not the proportion of participants reporting the symptom. (*$p<0.05$, **$p<0.01$, ***$p<0.001$)

fog, no individual window differed from Pre-HIRT (Early $p=0.789$; Middle $p=0.144$; Late $p=0.429$). In the three-group models, interactions were observed for fatigue ($p=0.018$), soreness ($p<0.001$), brain fog ($p<0.001$), and bloating ($p=0.002$; Table S4); however, the CON group contributed few participants and, for soreness, contained a window with no symptom-positive days, so these interactions should be interpreted cautiously.

Discussion

The current study examined the effects of CM supplementation during six weeks of HIRT on cognition, sleep, and menopause symptoms in PERI females. The primary findings were: **(1)** Cognitive performance improved over time regardless of group, while CM supplementation with HIRT enhanced prefrontal oxygenation during the Stroop task, as evidenced by greater increases in tissue saturation index at POST compared to PLA; (2) sleep consistency was higher with CM than PLA at Late (week 6), reflecting a decline in PLA that was not observed with CM; (3) CM was associated with fewer self-reported temperature sensitivity days at Middle (week 4) and Late (week 6) but more self-reported bloating days at the same windows compared to PLA, with no group differences in fatigue, soreness, or brain fog. Together, these exploratory findings suggest that CM supplementation with HIRT may support cerebral oxygenation during cognitive tasks and help preserve sleep consistency in PERI females, with mixed effects on self-reported symptoms.

Cognition

In the present study, cognitive performance during the Stroop but not the Rey 15-item memory test improved over time, independent of CM supplementation or HIRT training (congruent: time $p<0.001$, incongruent: time $p=0.004$).

The abbreviated two-minute Stroop test may not have provided sufficient cognitive demand or sensitivity to detect subtle treatment-related differences, particularly as CM supplementation appears to enhance cognition primarily under conditions of mental fatigue or metabolic stress (Gordji-Nejad et al. 2024), which were likely not elicited during the short-duration (~2 min) task used in the current study (Fabiano and Candow 2025). This is consistent with null findings on cognitive performance following CM supplementation and resistance training in postmenopausal females (Alves et al. 2013). Further, Stroop congruent and incongruent reaction times decreased over time. Due to the task being repeated at all three testing visits, the possible influence of repeated practice cannot be ruled out. In addition, Rey scores were near the maximum at all testing visits, limiting the ability of this measure to detect changes in memory and thus a more difficult memory recall was likely needed to observe differences.

Most literature supporting CM yielding cognitive benefits is in older individuals (60 + years) (Marshall et al. 2026), it is likely that the younger age group and normal baseline cognitive status were not sufficient to elicit measurable changes to cognitive performance (Pratt et al. 2026). Despite a lack of improvements in cognitive performance, six weeks of CM+HIRT supplementation enhanced prefrontal oxygenation during the Stroop task as measured via fNIRS. CM is known to improve brain energetics through phosphocreatine buffering and ATP availability maintenance, suggesting altered cerebral hemodynamic responses during the Stroop task without changes in cognitive performance (Candow et al. 2023). The current findings may suggest a CM-related difference in cerebral oxygenation relative to PLA supplementation, broadly aligning with prior work indicating that CM may influence cerebral bioenergetics, while cognitive performance benefits have been more consistently observed under fatiguing conditions (Gordji-Nejad et al. 2024). These changes may be particularly relevant in this cohort given the neurological basis of many PERI symptoms (temperature sensitivity, brain fog) (Lu et al. 2023). Overall, these findings suggest that CM supplementation may influence underlying brain energetics in PERI females, as reflected by increased prefrontal oxygenation, without translating to measurable improvements in cognitive performance under the conditions tested. Longer supplementation periods and/or higher doses of creatine should be evaluated to further understand the effects on brain energetics and symptoms.

Sleep measures

In PERI females undergoing six weeks of HIRT, the CM group had higher sleep consistency at Late than PLA, with sleep consistency preserved in CM but declining in PLA over the six-week intervention, suggesting a potential moderating effect of CM supplementation under conditions of elevated physiological stress. Notably, in the three-group model CM did not differ from the non-training CON group at Late, whereas PLA was lower than CON, consistent with CM attenuating a training-associated decline in sleep regularity rather than improving it above usual levels. The clinical relevance of a 7-percentage-point difference in this metric is unknown. For context, the between-participant standard deviation of sleep consistency in this sample was 10–17 points, so the difference corresponds to roughly half a standard deviation, and sleep consistency is a proprietary score adapted from, but not equivalent to, the Sleep Regularity Index.

High-intensity exercise is known to increase sympathetic activity and metabolic demand, which may disrupt sleep patterns; thus, the addition of CM may help attenuate these perturbations (Frisch et al. 2024; Zhang et al. 2026). Acute within-subject comparisons in premenopausal females suggest that CM supplementation increases total sleep duration on training days (Aguiar Bonfim Cruz et al. 2024). The present study used between-group contrasts and observed no within-group change in sleep duration in either group. A main effect of group reflected shorter sleep duration with CM than PLA; however, this difference was already present during the Pre-HIRT window. Taken together with sleep consistency results, these findings suggest that CM supplementation with HIRT may help preserve sleep consistency under conditions of increased metabolic and physiological demand, rather than directly improving sleep duration. This distinction may be particularly relevant in PERI females, where hormonal fluctuations, altered thermoregulation, and increased sleep fragmentation already contribute to reduced sleep stability (Troìa et al. 2025).

Menopause symptoms

In the present study, CM was associated with fewer self-reported temperature sensitivity days in PERI females at Middle (week 4) and Late (week 6) compared to PLA. This exploratory finding rests on a small analyzed sample (9 CM and 10 PLA participants) in which the CM group reported the symptom on only 3 of 109 days at Middle and Late, which produces extreme odds ratios with wide confidence intervals, and it was attenuated in the participant-level sensitivity analysis ($p = 0.064$). It should be regarded as hypothesis-generating. It nonetheless extends the limited literature on CM supplementation in PERI populations (Smith-Ryan et al. 2025; Korovljev et al. 2026), and warrants further investigation using validated measures of vasomotor symptoms and thermoregulation. Although these physiological mechanisms provide a potential rationale

for future investigation, their relevance to the present finding remains uncertain because temperature sensitivity was assessed using a self-reported journal item and direct measures of vasomotor symptoms or thermoregulation were not obtained. The CM group also had greater odds of reporting bloating than PLA during the Middle and Late periods, a result that was consistent in participant-level and leave-one-participant-out analyses. Bloating was assessed using a binary yes/no journal entry, which did not quantify severity or establish its cause. This exploratory finding warrants consideration as a potential tolerability signal, although it does not establish that a regular dose of CM itself caused bloating. Eight weeks of creatine hydrochloride supplementation in peri- and postmenopausal women improved reaction time, increased frontal brain CM concentrations, and reduced the severity of mood swings compared to PLA, supporting the potential for CM based interventions to positively influence some menopause-related symptoms and neurophysiological outcomes (Korovljev et al. 2026). CM has been shown to enhance thermoregulatory efficiency during exercise in the heat, including reductions in core temperature, heart rate, and sweat rate (Kilduff et al. 2004; Twycross-Lewis et al. 2016). These effects are often attributed to CM-induced increases in total body water and improved fluid balance (Deminice et al. 2015; Antonio et al. 2021). Supporting this mechanism, Moore and colleagues (Moore et al. 2023) reported that 5 days of CM loading increased total body water, extracellular fluid, and intracellular fluid during the luteal phase of the menstrual cycle in recreationally active females, without changes in body mass. More recently, CM supplementation improved phase angle and maintained countermovement jump performance across menstrual cycle phases, suggesting improved cellular hydration and cellular integrity during periods characterized by greater hormonal fluctuation (Moore et al. 2025). However, findings are not entirely consistent, as some studies report enhanced plasma volume maintenance without corresponding changes in core temperature, hemodynamic responses, or perceptual outcomes (Mendel et al. 2005; Branch et al. 2007). Although these investigations were conducted in male athletes, the underlying physiological mechanisms may translate to improved temperature regulation in PERI females experiencing vasomotor symptoms. Nevertheless, these findings should be interpreted with caution, as temperature sensitivity was assessed via self-report and direct measures of thermoregulation were not obtained.

Limitations & future directions

Important limitations to acknowledge for the current study are that cognitive outcomes were likely influenced by limited task familiarization, the narrow assessment battery (Stroop and Rey 15-item memory test) and lack of extreme deficits in cognitive health, which may not have captured broader cognitive changes. Future research should investigate the effects of higher CM dosing strategies and include individuals with greater baseline cognitive stress or impairment, as creatine-related cognitive benefits may be more apparent in populations with reduced cognitive reserve or heightened metabolic demand. The sleep and symptom analyses were based on available data and did not include all randomized participants. Because participants with missing WHOOP or journal data may have differed from those included, the smaller outcome-specific samples may have affected the group comparisons. Seven participants randomized to CM or PLA began wearing the WHOOP device after HIRT started and had no Pre-HIRT reference data; they did not differ from included participants in age or BMI, but adherence to device wear and journaling may be related to symptom burden or treatment tolerability, so the analyzed samples cannot be assumed to be missing at random. In addition, incorporating more comprehensive and sensitive cognitive batteries, alongside tasks designed to induce mental fatigue, may better reveal the cognitive effects of CM in PERI populations. Lastly, although STRAW + 10 provides the gold-standard framework for reproductive-aging staging, classification was limited in some participants due to hormonal contraceptive use and menopausal hormone therapy, both of which can modify bleeding patterns. Nonetheless, participants exhibited multiple converging indicators of the menopausal transition, including substantial symptom burden (> 5 symptoms), menstrual irregularity, altered bleeding characteristics, prolonged intervals between menses, and elevated or variable follicle stimulating hormone (FSH) concentrations. These characteristics are representative of women commonly seeking treatment during the PERI years and support the ecological validity of the study population.

Several limitations relate specifically to the wearable data. First, WHOOP devices were issued at BASE, so no pre-supplementation sleep or symptom data exist; the Pre-HIRT reference period coincided with creatine loading, and any acute effect of loading on sleep or symptoms cannot be evaluated and is absorbed into the reference values. Statements that CM preserved or maintained an outcome therefore refer to the six-week HIRT period relative to an on-supplement reference, not to change from an untreated state. Second, sleep staging and sleep consistency are proprietary WHOOP outputs and validation against polysomnography is available for earlier hardware generations but not for the specific device and algorithm version used here, and sleep consistency is not equivalent to the Sleep Regularity Index. Third, the daily journal items are unvalidated binary self-report. Temperature sensitivity is not a validated

measure of vasomotor symptoms, and the small number of participants contributing to the symptom models limits the precision of the group-by-time estimates. Finally, although the participant-level sensitivity analyses addressed serial correlation among daily observations, residual autocorrelation was not modeled explicitly in the primary analysis, and no adjustment for multiplicity was applied across the exploratory sleep and symptom outcomes.

Conclusions

In this exploratory analysis of a pilot randomized trial, CM supplementation combined with HIRT was associated with greater task-evoked prefrontal oxygenation, higher wearable-derived sleep consistency, and fewer self-reported temperature sensitivity days compared with PLA, whereas cognitive performance did not differ between groups. These preliminary findings support further evaluation of CM as an adjunct to HIRT during perimenopause in an adequately powered trial using validated measures of sleep and vasomotor symptoms.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00421-026-06481-6.

Author contributions Funding acquisition: ASR, PAB; Conceptualization: ASR; Methodological development: PAB, GMD, ASR; Data collection: PAB, GMD, WMJ, CGK; Data analysis and interpretation: PAB, GJG; ASR; Manuscript preparation and writing: PAB, GJG, CGK, ASR; Figure preparation: PAB, GMD; GJG; Manuscript Review and editing: PAB, GMD, WMJ, CGK, GJG, ASR;

Funding AlzChem provided the creatine monohydrate powder and Create provided the creatine gummies for this study but had no input on the design, conduct, or interpretation of the current data. WHOOP, Inc. provided the wearable devices used in this study.

Data availability Deidentified individual data that supports the results will be shared beginning 9 to 36 months following publication provided the investigator who proposes to use the data has approval from an Institutional Review Board (IRB), Independent Ethics Committee (IEC), or Research Ethics Board (REB), as applicable, and executes a data use/sharing agreement with the University of North Carolina. The WHOOP-derived sleep and journal data cannot be shared publicly because of intellectual property considerations of WHOOP, Inc.; deidentified WHOOP data and a data dictionary may be made available upon reasonable request to WHOOP (science@whoop.com), subject to a methodologically sound proposal and a signed data use agreement. The complete analysis code for the sleep and symptom outcomes is publicly available at https://github.com/greggrosicki/Creatine_HIRT_Perimenopause.

Declarations

Author disclosures ASR serves as a scientific advisor for Alzchem, Create Wellness, and NeoLife, companies that manufacture or sell creatine-containing products. ASR also holds equity shares in Create Wellness and has received research funding from industry sponsors to conduct research on dietary supplements, including creatine. GJG is an employee of and holds stock options in WHOOP, Inc. GJG exported the WHOOP data and performed the sleep and symptom analyses while blinded to the treatment groups; these analyses were not independently reproduced by an investigator without a WHOOP affiliation, and the analysis code is publicly available to permit verification. GJG participated in interpretation and manuscript preparation as a coauthor. WHOOP, Inc. as a corporate entity had no role in the study design, interpretation, decision to publish, or approval of the manuscript, and did not review the manuscript.

Clinical trial This trial was registered at ClinicalTrials.gov (NCT07027800).

References

Aguiar Bonfim Cruz AJ, Brooks SJ, Kleinkopf K et al (2024) Creatine Improves Total Sleep Duration Following Resistance Training Days versus Non-Resistance Training Days among Naturally Menstruating Females. Nutrients 16:2772. https://doi.org/10.3390/nu16162772

Alves CRR, Merege Filho CAA, Benatti FB et al (2013) Creatine supplementation associated or not with strength training upon emotional and cognitive measures in older women: a randomized double-blind study. PLoS ONE 8:e76301. https://doi.org/10.1371/journal.pone.0076301

Antonio J, Candow DG, Forbes SC et al (2021) Common questions and misconceptions about creatine supplementation: what does the scientific evidence really show? J Int Soc Sports Nutr 18:13. https://doi.org/10.1186/s12970-021-00412-w

Berin E, Hammar M, Lindblom H et al (2019) Resistance training for hot flushes in postmenopausal women: A randomised controlled trial. Maturitas 126:55–60. https://doi.org/10.1016/j.maturitas.2019.05.005

Boushel R, Piantadosi CA (2000) Near-infrared spectroscopy for monitoring muscle oxygenation. Acta Physiol Scand 168:615–622. https://doi.org/10.1046/j.1365-201x.2000.00713.x

Branch JD, Schwarz WD, Van Lunen B, EFFECT OF CREATINE SUPPLEMENTATION ON CYCLE ERGOMETER EXERCISE IN A HYPERTHERMIC ENVIRONMENT (2007) J Strength Cond Res 21:57–61. https://doi.org/10.1519/00124278-200702000-00011

Buzdagli Y, Ozan M, Baygutalp N et al (2024) The effect of high-intensity intermittent and moderate-intensity continuous exercises on neurobiological markers and cognitive performance. BMC Sports Sci Med Rehabil 16:39. https://doi.org/10.1186/s13102-024-00831-7

Cabre HE, Gordon AN, Patterson ND, Smith-Ryan AE (2022) Evaluation of pre-workout and recovery formulations on body composition and performance after a 6-week high-intensity training program. Front Nutr. https://doi.org/10.3389/fnut.2022.1016310

Candow DG, Forbes SC, Ostojic SM et al (2023) Heads Up for Creatine Supplementation and its Potential Applications for Brain Health and Function. Sports Med 53:49–65. https://doi.org/10.1007/s40279-023-01870-9

Capel-Alcaraz AM, García-López H, Castro-Sánchez AM et al (2023) The Efficacy of Strength Exercises for Reducing the Symptoms of Menopause: A Systematic Review. J Clin Med 12:548. https://doi.org/10.3390/jcm12020548

Cunha PM, Santos LD, Herold F et al (2025) Impact of resistance training on sleep quality, mental health, and functional capacity in older women with varying baseline sleep quality: A randomized

controlled trial. Psychol Sport Exerc 81:102968. https://doi.org/10.1016/j.psychsport.2025.102968

Currier BS, D'Souza AC, Singh MAF et al (2026) American College of Sports Medicine Position Stand. Resistance Training Prescription for Muscle Function, Hypertrophy, and Physical Performance in Healthy Adults: An Overview of Reviews. Med Sci Sports Exerc 58:851–872. https://doi.org/10.1249/MSS.0000000000003897

Deminice R, Rosa F, Pfrimer K et al (2015) Creatine Supplementation Increases Total Body Water in Soccer Players: a Deuterium Oxide Dilution Study. Int J Sports Med 37:149–153. https://doi.org/10.1055/s-0035-1559690

Dhahbi W, Briki W, Heissel A et al (2025) Physical Activity to Counter Age-Related Cognitive Decline: Benefits of Aerobic, Resistance, and Combined Training—A Narrative Review. Sports Med - Open 11:56. https://doi.org/10.1186/s40798-025-00857-2

El Khoudary SR, Greendale G, Crawford SL et al (2019) The menopause transition and women's health at midlife: a progress report from the Study of Women's Health Across the Nation (SWAN). Menopause 26:1213. https://doi.org/10.1097/GME.0000000000001424

El Khoudary SR, Aggarwal B, Beckie TM et al (2020) Menopause Transition and Cardiovascular Disease Risk: Implications for Timing of Early Prevention: A Scientific Statement From the American Heart Association. Circulation 142:e506–e532. https://doi.org/10.1161/CIR.0000000000000912

Fabiano N, Candow D (2025) Creatine Supplementation: More Is Likely Better for Brain Bioenergetics, Health and Function. J Psychiatry Brain Sci. https://doi.org/10.20900/jpbs.20250006

Frisch N, Heischel L, Wanner P et al (2024) An acute bout of high-intensity exercise affects nocturnal sleep and sleep-dependent memory consolidation. J Sleep Res 33:e14126. https://doi.org/10.1111/jsr.14126

Goldberg JO, Miller HR (1986) Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. J Clin Psychol 42:792–795. https://doi.org/10.1002/1097-4679(198609)42:5%3C792::AID-JCLP2270420519%3E3.0.CO;2-8

Gordji-Nejad A, Matusch A, Kleedörfer S et al (2024) Single dose creatine improves cognitive performance and induces changes in cerebral high energy phosphates during sleep deprivation. Sci Rep 14:4937. https://doi.org/10.1038/s41598-024-54249-9

Jones S, Chiesa ST, Chaturvedi N, Hughes AD (2016) Recent developments in near-infrared spectroscopy (NIRS) for the assessment of local skeletal muscle microvascular function and capacity to utilise oxygen. Artery Res 16:25. https://doi.org/10.1016/j.artres.2016.09.001

Kilduff LP, Georgiades E, James N et al (2004) The Effects of Creatine Supplementation on Cardiovascular, Metabolic, and Thermoregulatory Responses during Exercise in the Heat in Endurance-Trained Humans. Int J Sport Nutr Exerc Metab 14:443–460. https://doi.org/10.1123/ijsnem.14.4.443

Kimberly SH, Flippin KJ, Blom LC et al (2018) Practices, Perceived Benefits, and Barriers to Strength Training Among Women Enrolled in College. Int J Exerc Sci 11:226–238. https://doi.org/10.70252/ZRMT3507

Korovljev D, Ostojic J, Panic J et al (2026) The Effects of 8-Week Creatine Hydrochloride and Creatine Ethyl Ester Supplementation on Cognition, Clinical Outcomes, and Brain Creatine Levels in Perimenopausal and Menopausal Women (CONCRET-MENOPA): A Randomized Controlled Trial. J Am Nutr Assoc 45:199–210. https://doi.org/10.1080/27697061.2025.2551184

Kotopoulea-Nikolaidi M, Watkins E, Giannopoulou I (2019) Effects of High Carbohydrate vs. High Protein Pre-exercise Feedings on Psychophysiological Responses to High Intensity Interval Exercise in Overweight Perimenopausal Women. Front Nutr 5:141. https://doi.org/10.3389/fnut.2018.00141

Kovacevic A, Mavros Y, Heisz JJ, Fiatarone Singh MA (2018) The effect of resistance exercise on sleep: A systematic review of randomized controlled trials. Sleep Med Rev 39:52–68. https://doi.org/10.1016/j.smrv.2017.07.002

Lu W, Sun Y, Gao H, Qiu J (2023) A review of multi-modal magnetic resonance imaging studies on perimenopausal brain: a hint towards neural heterogeneity. Eur Radiol 33:5282–5297. https://doi.org/10.1007/s00330-023-09549-5

Marshall S, Kitzan A, Wright J et al (2026) Creatine and Cognition in Aging: A Systematic Review of Evidence in Older Adults. Nutr Rev 84:333–344. https://doi.org/10.1093/nutrit/nuaf135

Mendel RW, Blegen M, Cheatham C et al (2005) Effects of creatine on thermoregulatory responses while exercising in the heat. Nutrition 21:301–307. https://doi.org/10.1016/j.nut.2004.06.024

Metcalf CA, Duffy KA, Page CE, Novick AM (2023) Cognitive Problems in Perimenopause: A Review of Recent Evidence. Curr Psychiatry Rep 25:501–511. https://doi.org/10.1007/s11920-023-01447-3

Miller DJ, Lastella M, Scanlan AT et al (2020) A validation study of the WHOOP strap against polysomnography to assess sleep. J Sports Sci 38:2631–2636. https://doi.org/10.1080/02640414.2020.1797448

Miller DJ, Sargent C, Roach GD (2022) A Validation of Six Wearable Devices for Estimating Sleep, Heart Rate and Heart Rate Variability in Healthy Adults. Sensors 22:6317. https://doi.org/10.3390/s22166317

Monteleone P, Mascagni G, Giannini A et al (2018) Symptoms of menopause — global prevalence, physiology and implications. Nat Rev Endocrinol 14:199–215. https://doi.org/10.1038/nrendo.2017.180

Moore SR, Gordon AN, Cabre HE et al (2023) A Randomized Controlled Trial of Changes in Fluid Distribution across Menstrual Phases with Creatine Supplementation. Nutrients 15:429. https://doi.org/10.3390/nu15020429

Moore SR, Mutuwa AG, Brown A, Smith-Ryan AE (2025) The effects of creatine loading on cellular integrity and jump performance across the menstrual cycle. J Int Soc Sports Nutr 22:2533651. https://doi.org/10.1080/15502783.2025.2533651

Moro T, Marcolin G, Bianco A et al (2020) Effects of 6 Weeks of Traditional Resistance Training or High Intensity Interval Resistance Training on Body Composition, Aerobic Power and Strength in Healthy Young Subjects: A Randomized Parallel Trial. Int J Environ Res Public Health 17:4093. https://doi.org/10.3390/ijerph17114093

Naysmith LF, Ward H, Elliott P et al (2026) Cognition and the menopause transition: cross-sectional evidence from a large community cohort. Npj Womens Health 4:14. https://doi.org/10.1038/s44294-026-00132-z

Nilsson S, Henriksson M, Hammar M et al (2024) A 2-year follow-up to a randomized controlled trial on resistance training in postmenopausal women: vasomotor symptoms, quality of life and cardiovascular risk markers. BMC Womens Health 24:511. https://doi.org/10.1186/s12905-024-03351-1

O'Reilly K, McDermid F, McInnes S, Peters K (2023) An exploration of women's knowledge and experience of perimenopause and menopause: An integrative literature review. J Clin Nurs 32:4528–4540. https://doi.org/10.1111/jocn.16568

Oporto-Colicoi V, Sepúlveda-Lara A, Marzuca-Nassr GN, Sepúlveda-Figueroa P (2025) Mild Cognitive Impairment and Sarcopenia: Effects of Resistance Exercise Training on Neuroinflammation, Cognitive Performance, and Structural Brain Changes. Int J Mol Sci 26:11036. https://doi.org/10.3390/ijms262211036

Phillips AJK, Clerx WM, O'Brien CS et al (2017) Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing. Sci Rep 7:3216. https://doi.org/10.1038/s41598-017-03171-4

Pihoker AA, Peterjohn AM, Trexler ET et al (2019) The effects of nutrient timing on training adaptations in resistance-trained females. J Sci Med Sport 22:472–477. https://doi.org/10.1016/j.jsams.2018.09.236

Pratt J, Fuchs CJ, van Loon LJC et al (2026) Creatine supplementation and brain health: Methodological challenges, current evidence, and translational perspectives. J Nutr Physiol 5:100017. https://doi.org/10.1016/j.jnphys.2026.100017

Scarpina F, Tagini S (2017) The Stroop Color and Word Test. Front Psychol. https://doi.org/10.3389/fpsyg.2017.00557

Smith-Ryan AE, DelBiondo GM, Brown AF et al (2025) Creatine in women's health: bridging the gap from menstruation through pregnancy to menopause. J Int Soc Sports Nutr 22:2502094. https://doi.org/10.1080/15502783.2025.2502094

Stimson AM, Anderson C, Holt A-M, Henderson AJ (2024) Why don't women engage in muscle strength exercise? An integrative review. Health Promot J Austr 35:911–923. https://doi.org/10.1002/hpja.857

Sudo M, Costello JT, McMorris T, Ando S (2022) The effects of acute high-intensity aerobic exercise on cognitive performance: A structured narrative review. Front Behav Neurosci 16:957677. https://doi.org/10.3389/fnbeh.2022.957677

Talaulikar V (2022) Menopause transition: Physiology and symptoms. Best Pract Res Clin Obstet Gynaecol 81:3–7. https://doi.org/10.1016/j.bpobgyn.2022.03.003

Troìa L, Garassino M, Volpicelli AI et al (2025) Sleep Disturbance and Perimenopause: A Narrative Review. J Clin Med 14:1479. https://doi.org/10.3390/jcm14051479

Twycross-Lewis R, Kilduff LP, Wang G, Pitsiladis YP (2016) The effects of creatine supplementation on thermoregulation and physical (cognitive) performance: a review and future prospects. Amino Acids 48:1843–1855. https://doi.org/10.1007/s00726-016-2237-9

Vasudevan A, Ford E (2022) Motivational Factors and Barriers Towards Initiating and Maintaining Strength Training in Women: a Systematic Review and Meta-synthesis. Prev Sci 23:674–695. https://doi.org/10.1007/s11121-021-01328-2

Weber MT, Maki PM, McDermott MP (2014) Cognition and mood in perimenopause: A systematic review and meta-analysis. J Steroid Biochem Mol Biol 142:90–98. https://doi.org/10.1016/j.jsbmb.2013.06.001

Wingfield HL, Smith-Ryan AE, Melvin MN et al (2015) The acute effect of exercise modality and nutrition manipulations on post-exercise resting energy expenditure and respiratory exchange ratio in women: a randomized trial. Sports Med - Open 1:11. https://doi.org/10.1186/s40798-015-0010-3

Wu L-Y, Hsu H-C, Ni L-F et al (2024) Effect of Physical Exercise on Executive Functions Using the Emotional Stroop Task in Perimenopausal Women: A Pilot Study. Behav Sci 14:338. https://doi.org/10.3390/bs14040338

Zhang B, Zheng C, Liao Q et al (2026) High-Intensity Circuit Training Plus Sleep Health Intervention for Sleep Improvement: A Randomized Clinical Trial. JAMA Netw Open 9:e2556927. https://doi.org/10.1001/jamanetworkopen.2025.56927

Zhao Y, Niu H, Liu S (2022) Effects of aerobics training on anxiety, depression and sleep quality in perimenopausal women. Front Psychiatry 13:1025682. https://doi.org/10.3389/fpsyt.2022.1025682

Zhao M, Sun M, Zhao R et al (2023) Effects of exercise on sleep in perimenopausal women: A meta-analysis of randomized controlled trials. EXPLORE 19:636–645. https://doi.org/10.1016/j.explore.2023.02.001

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.