



Rationale and study design of an intervention of increased energy intake in women with exercise-associated menstrual disturbances to improve menstrual function and bone health: The REFUEL study

Nancy I. Williams*, Rebecca J. Mallinson, Mary Jane De Souza

Pennsylvania State University, Department of Kinesiology, Noll Laboratory, Women's Health and Exercise Laboratory, University Park, PA, 16802, USA

ARTICLE INFO

Keywords:

Clinical trial
Female athlete triad
Amenorrhea
Bone health
Nutrition

ABSTRACT

Purpose: Exercising women who consume inadequate energy relative to expenditure are at risk for downstream health consequences, such as menstrual cycle disturbances and poor bone health. Collectively, these conditions are known as the Female Athlete Triad (Triad). Clinicians often prescribe hormonal contraceptives to address this issue; however, the recommended treatment is reversal of the energy deficit. This paper describes the design of the REFUEL study, a randomized controlled trial (RCT) that explored the effectiveness of a 12-month intervention of increased energy intake on the reversal of an unhealthy energetic status and menstrual dysfunction and subsequent improvements in bone health in exercising women with severe menstrual cycle disturbances. **Methods:** Women between the ages of 18–35 years and participating in at least 2 h/week of purposeful exercise were recruited. Those who reported irregular or absent menstrual cycles and were determined to have an exercise-associated menstrual disturbance (EAMD) were randomized into either the treatment group (EAMD + Cal), which was instructed to increase caloric intake throughout the intervention, or a control group (EAMD Control). Women who reported eumenorrhea were eligible for the ovulatory (OV) Control group. Repeated measures of energetic and metabolic status, reproductive status, and skeletal health were obtained. **Discussion:** The REFUEL study is the first RCT to explore a non-pharmacological treatment approach among exercising women with the Triad. 118 women were randomized, and 55 women completed the entire study. The findings of this study have the potential to inform and alter clinical practice for exercising young women who present with this condition.

1. Introduction

Inadequate energy intake relative to energy expenditure among exercising women contributes to low energy availability (EA). Defined as insufficient energy available for metabolic processes after subtracting energy used for exercise [1], low EA has negative effects, including menstrual cycle disturbances and poor bone health. Low EA, menstrual dysfunction, and low bone mass constitute the medical condition known as the Female Athlete Triad [2,3].

Low EA can present with or without disordered eating [4] and leads to the repartitioning of energy away from non-vital physiological functions, including reproduction and growth [5]. Related changes in metabolic hormones and neuropeptides then act upon the hypothalamus to suppress the reproductive axis. Low EA-related menstrual cycle disturbances that result from low EA range from subtle to severe

[6,7]. Subtle menstrual disturbances, including luteal phase defects and anovulation, occur in the face of regular menses and thus are often undetected. Notably, these subtle disturbances occur in approximately 50% of cycles in exercising women, highlighting the magnitude of this health problem [8]. Severe menstrual disturbances include amenorrhea, defined as the absence of menses for at least 90 days, and oligomenorrhea, defined as long, inconsistent menstrual cycles (36–90 days), and have been reported to occur in approximately 30% of cycles in exercising women [8].

Low EA and menstrual cycle disturbances work in tandem to compromise bone health by suppressing bone formation [9] and upregulating bone resorption [10]. This uncoupling of bone turnover leads to low bone mass and poor bone geometry. Because the adolescent years are critical for building bone mass [11,12], low EA and menstrual disturbances at this time can lead to lifelong skeletal health

* Corresponding author. Department of Kinesiology, Women's Health and Exercise Laboratory, 276 Recreation Building, Penn State University, University Park, PA, 16802, USA.

E-mail address: niw1@psu.edu (N.I. Williams).

<https://doi.org/10.1016/j.conctc.2019.100325>

Received 24 August 2018; Received in revised form 4 January 2019; Accepted 11 January 2019

Available online 18 January 2019

2451-8654/ © 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

consequences [13].

Accordingly, the first treatment goal is to increase food intake and/or decrease exercise training to increase EA [3]. Using a monkey model, a randomized prospective experiment demonstrated that increasing food intake in the face of daily exercise can reverse reproductive suppression caused by low EA [14]. In humans, current published reports of the efficacy of increasing energy intake to reverse severe menstrual disturbances are limited to case studies [15–19], one retrospective chart review [20], and one prospective, uncontrolled intervention [21,22]. These human studies have provided observational evidence that an increase in energy intake and modest weight gain are effective for recovery of menstrual function and improvements in bone health in exercising women. To date, however, there have been no randomized controlled trials exploring the effectiveness of this approach on the three components of the Female Athlete Triad in human females.

The purpose of this paper is to describe the design of the REFUEL study, a randomized controlled trial (RCT) that explored the effectiveness of a 12-month intervention of increased energy intake on recovery of a healthy energetic status, menstrual function and bone health in women with oligomenorrhea and amenorrhea. The significance of this study is its potential to inform and alter clinical practice for exercising young women who present with the Female Athlete Triad.

2. Specific aims & hypotheses

The overall purpose of this study was to evaluate the effects of 12 months of increased energy intake on energetic status, menstrual status, and bone health among exercising women with exercise-associated menstrual disturbances (EAMD), including oligomenorrhea and amenorrhea.

The first primary outcome focused on energy status. This study tested the effectiveness of 12 months of increased energy intake on reversal of a chronic energy deficit in women with severe EAMD by promoting significant changes in physiological, metabolic and hormonal indices of energy conservation. We hypothesized that increased energy intake during the 12 month intervention will be associated with significant increases in resting energy expenditure (REE), body weight and fat mass, and body mass index (BMI) in exercising women with EAMD with final values similar to those of exercising ovulatory controls. Additionally, we hypothesized that increased energy intake during the 12-month intervention will be associated with favorable changes in metabolic hormones among exercising women with EAMD with final values similar to those of exercising ovulatory controls.

The second primary outcome focused on recovery of menstrual function. We hypothesized that increased energy intake during the 12-month intervention will lead to the resumption of menstrual function in amenorrheic, exercising women and increase the frequency of menses in oligomenorrheic, exercising women. Further, we hypothesized that recovery of menses in EAMD women will be preceded by reversal of chronic energy deficiency, as indicated by changes in one or more key metabolic indicators, i.e., circulating total triiodothyronine, insulin-like growth-factor-1 (IGF-1), leptin, ghrelin, peptide YY (PYY), and increases in REE.

The third primary outcome focused on bone health. We hypothesized that bone mineral density (BMD) at the total body, lumbar spine, and/or hip will improve concomitant with improvements in bone remodeling as indicated by an increase in markers of bone formation (osteocalcin and N-terminal propeptide of Type I collagen (P1NP)) and a decrease or no change in markers of bone resorption (urinary C-Terminal telopeptide (CTX)) relative to baseline. We anticipated that the improvements in bone remodeling and BMD will be preceded by significant improvements in metabolic and hormonal indices of energy status, especially those factors important in the regulation of bone cell metabolism, i.e., IGF-1 and leptin. We also hypothesized that improvement in estrogen status as indicated by increases in urinary excretion of estrone-1-glucuronide (E1G) would be predictive of

improvements in bone parameters.

Overall, this study aimed to determine the association among each of the aforementioned primary outcomes, i.e., improved metabolic status, resumption of menses, and improved bone health in women with severe EAMD. We expect that women with EAMD exhibiting the largest magnitude of improvement in metabolic hormones and markers of energy status and the earliest resumption of menses will experience the largest gains in BMD.

3. Methods

3.1. Study design

The REFUEL study is a RCT designed to determine the effect of a 12-month intervention of increased energy intake on recovery of energetic status, menstrual function, and bone health in exercising women with severe exercise-associated menstrual disturbances (EAMD), including oligomenorrhea (long and inconsistent menstrual cycles of 36–90 days) and functional hypothalamic amenorrhea (FHA, the absence of menses for > 90 days in women who previously menstruated). Women with EAMD were randomized into one of two groups: 1) a group that increased energy intake during the intervention (EAMD + Cal) and 2) a control group that was instructed to maintain current dietary and exercise behaviors (EAMD Control). The EAMD + Cal and EAMD Control groups were compared to an exercising ovulatory control group (OV Control), consisting of exercising women who presented with a 6-month history of regular menstrual cycles and an ovulatory menstrual cycle during baseline monitoring (see Fig. 1). To be considered “exercising,” participants were required to participate in purposeful physical activity for at least 2 h/week (OV Control group) or 3 h/week (EAMD groups).

Participants were recruited on a rolling basis, and repeated measures of energetic and metabolic status, reproductive status, and

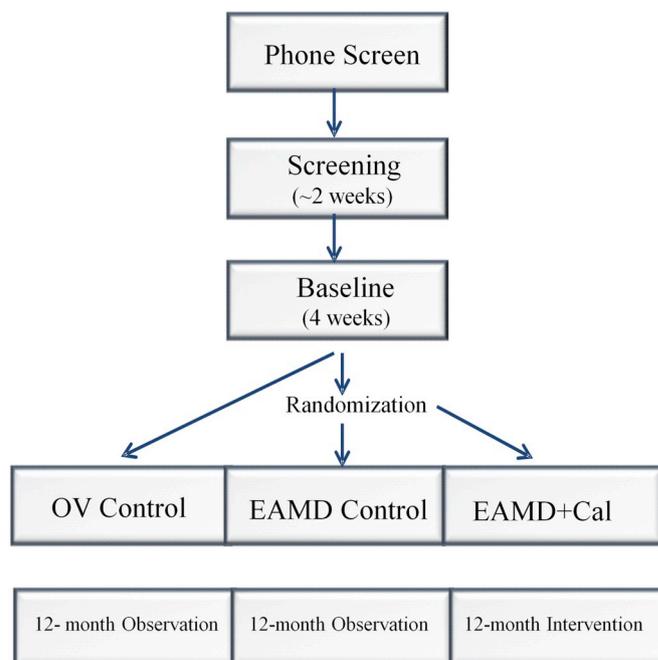


Fig. 1. Study Flow and Groups. After an initial screening questionnaire, eligible participants completed a screening phase to determine final eligibility. Those who met all inclusion criteria completed a baseline phase. After baseline, women with exercise-associated menstrual disturbances (EAMD) were randomized into the treatment group (EAMD + Cal) or the control group (EAMD Control). Women with regular menstrual cycles were assigned to the Ovulatory (OV) Control group. Participants then either increased caloric intake for 12 months (EAMD + Cal group) or were observed for 12 months (EAMD Control and OV Control groups).

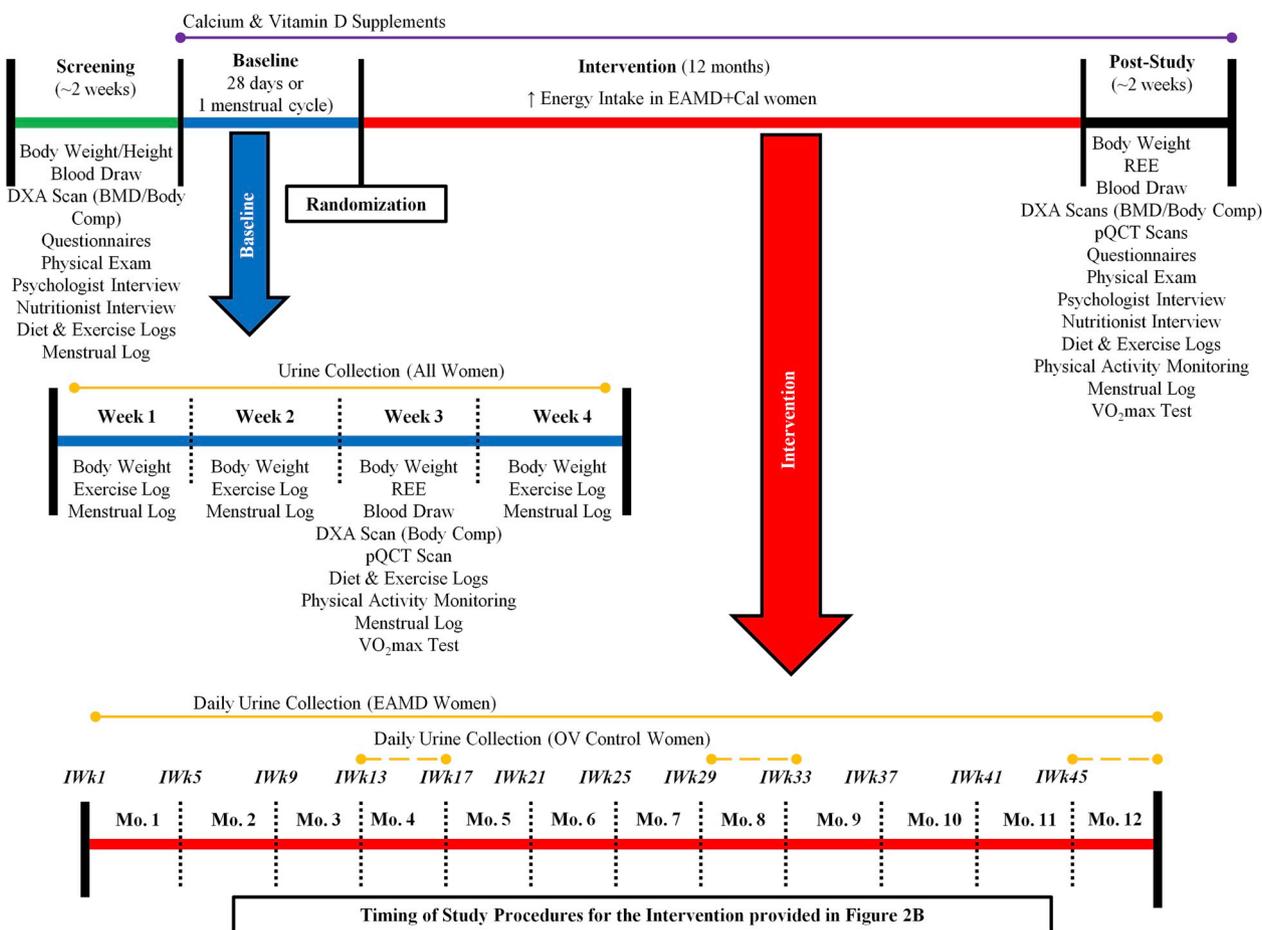


Fig. 2. Study Design. This figure provides an overview of the type and timing of measurements collected during this randomized controlled trial. Fig. 2A is the overall study design. Fig. 2B provides a detailed study design, including the type and timing of study procedures, during the intervention phase of the study. *Body Composition Scan Only. **BMD & Body Composition Scans. REE: resting energy expenditure; BD: blood draw; DXA: dual-energy x-ray absorptiometry; pQCT: peripheral quantitative computed tomography; PE: physical exam.

skeletal health were obtained throughout the approximately 14-month study, which included a 2–4 week screening period, 4-week baseline period, 12-month intervention, and 2-week post-study period. The study design is depicted in Fig. 2.

Data collection for the study was conducted over 8 years (2006–2014) at two sites: University of Toronto (UT) and the Pennsylvania State University (PSU). The study was approved by the Research Ethics Board of UT and the Institutional Review Board of PSU, and participants signed an approved informed consent prior to study initiation.

3.2. Recruitment strategies

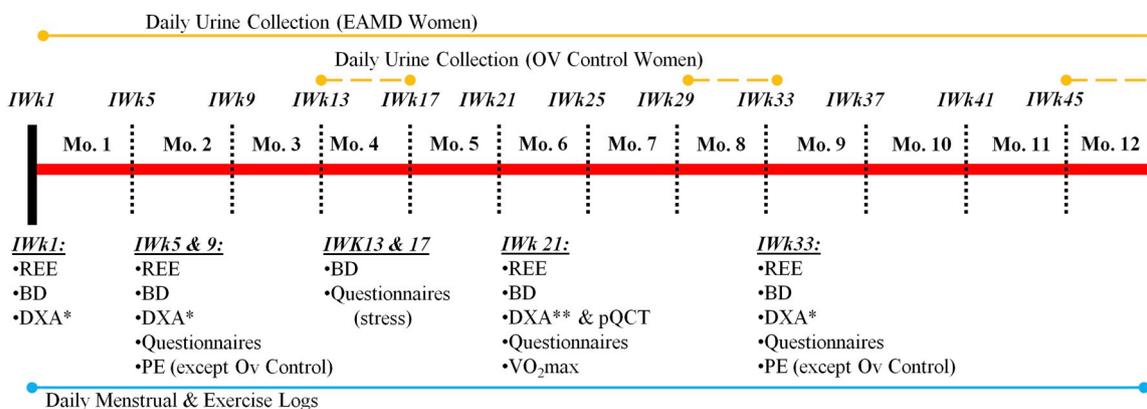
Recruitment was performed on the two university campuses and in the surrounding communities. Recruitment methods included flyers posted on bulletin boards and in student mail boxes, newspaper ads, emails sent to college/department listservs, announcements in classrooms and at club meetings, announcements posted on the PSU news email and on the PSU website for research volunteers and the use of a website that described the study.

Recruitment occurred from August 2006 to November 2013, and data collection ended in December 2014. The progression of participants through the study is displayed in Fig. 3. Two hundred thirty-three women signed an informed consent and were assessed for eligibility during the screening phase of the study. Ninety-one women were excluded or withdrew during screening; therefore, 142 women entered the baseline period. One hundred and eighteen women were

randomized and entered the intervention, including 42 women in the EAMD + Cal group, 36 women in the EAMD Control group, and 40 women in the OV Control group. Eighty-two women completed IWK9 (month 3) of the intervention (30 EAMD + Cal, 24 EAMD Control, and 28 OV Control). Seventy-one women completed IWK21 (month 6) (26 EAMD + Cal, 19 EAMD Control, and 26 OV Control). Sixty-three women completed IWK 33 (month 9) (24 EAMD + Cal, 16 EAMD Control, and 23 OV Control), and 55 women completed all 12 months of the intervention (19 EAMD + Cal, 16 EAMD Control, and 20 OV Control). The screening dropout rate was 39%, and the failure rate (those who did not qualify for the study) was 16%. Primary reasons for participant withdrawal/exclusion included loss of interest in the study, not responding to communication from study team (lost to follow-up), and not meeting the study criteria. Overall dropout/early termination rate among women who were randomized was 53%. The dropout/early termination rate for each group did not differ (OV Control, 50%; EAMD Control, 56%; EAMD + Cal, 55%). During the baseline and intervention phases of the study, primary reasons why participants were screened out, dropped out or were terminated early included time commitment, menstrual status that was incompatible with study criteria, and non-compliance with study protocol. Detailed reasons for participant exclusion or drop out are provided in Fig. 3.

3.3. Screening & eligibility of participants

When volunteers expressed interest in the study, a screening questionnaire was completed to determine the volunteer's initial eligibility.



Additional Measures:

Body Weight: EAMD – Every 2 weeks

OV Control – Once per month

Nutritionist & Psychologist Interview: EAMD – every 2 weeks for first 3 months then monthly (IWk1, 3, 5, 7, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45)
 OV Control – every 3 months (IWk9, 21, 33)

Diet Logs & Physical Activity Monitoring: EAMD – once per month (IWk1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41, 45)
 OV Control – every 3 months (IWk1, 9, 21, 33)

*Body Composition Scan Only

**BMD & Body Composition Scans

REE – resting energy expenditure; BD: blood draw; DXA: dual-energy x-ray absorptiometry; pQCT: peripheral quantitative computed tomography; PE: physical exam

Fig. 2. (continued)

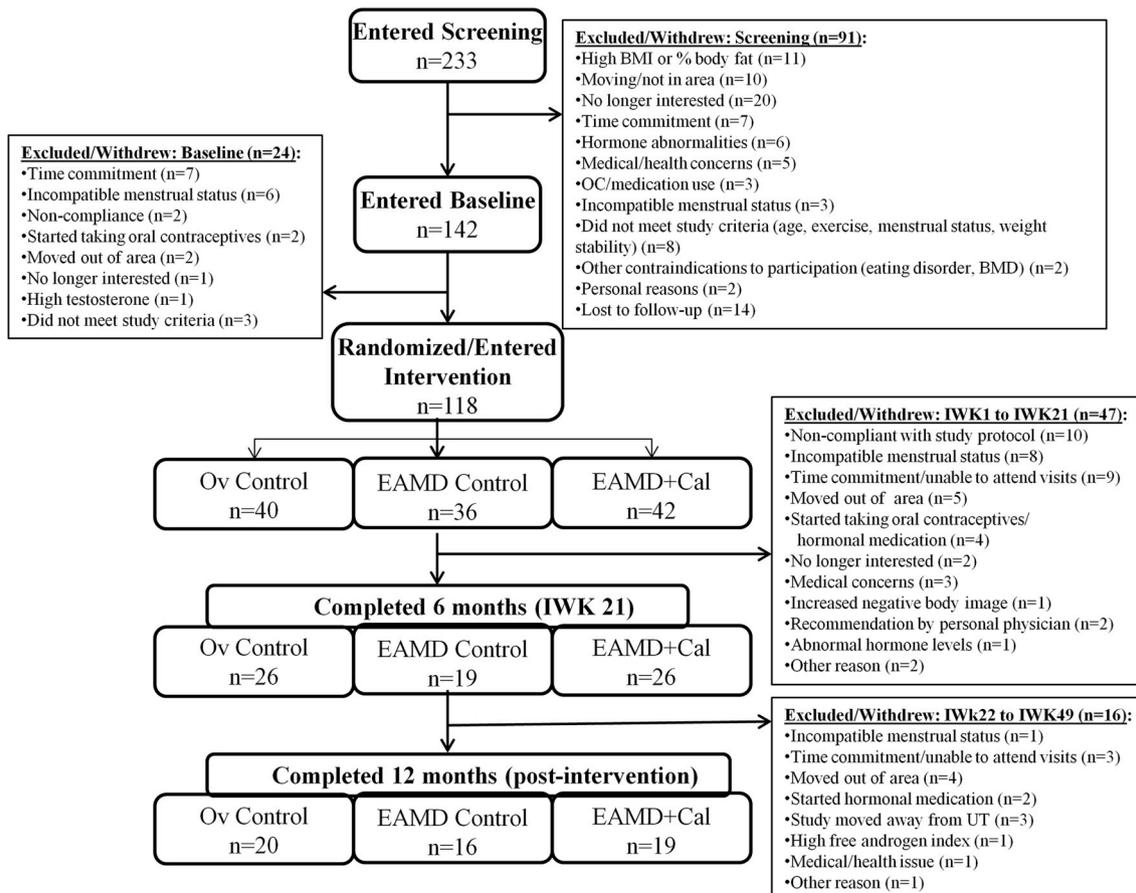


Fig. 3. Study enrollment and dropout. This figure provides the number of participants enrolled at each study phase and the reasons for dropout.

If a volunteer met the initial criteria, the first visit to the lab was scheduled during which the volunteer was informed of the purpose, procedures, and potential benefits/risks of study participation prior to signing an approved informed consent document. Once consent was obtained, height (cm) and weight (kg) were measured, and additional questionnaires were completed to assess demographic information, medical and menstrual history, eating attitudes and behaviors [23,24], exercise participation, bone health, and psychological health [25–28]. A physical exam was performed to determine health status and to rule out any physical signs or symptoms of polycystic ovarian syndrome (i.e., acne, hirsutism) or eating disorders. To rule out endocrine and metabolic disease, a fasting blood sample was analyzed for complete blood count, basic chemistry panel, and an endocrine panel which included follicular stimulating hormone, luteinizing hormone (LH), estradiol, prolactin, thyroid stimulating hormone, thyroxine, total and free testosterone, and dehydroepiandrosterone sulfate. Participants completed a 3-day diet log for assessment of dietary energy intake and a 7-day exercise log for assessment of purposeful exercise energy expenditure. Further, a research psychologist with trained expertise in clinical eating disorders completed a semi-structured interview with each participant to rule out current clinical eating disorders and other psychiatric disorders. Participants met with a registered dietitian for a structured interview designed to assess the participants' eating patterns, relationship with food, and food preferences in an effort to determine participants' willingness to comply with the study protocol. A dual-energy x-ray absorptiometry (DXA) scan was performed to assess body composition and bone mineral density (BMD) at the total body, lumbar spine, and dual femur.

Eligibility criteria for this study were as follows: (1) women aged 18–35 years, (2) body mass index (BMI) between 16 and 25 kg/m²; (3) good health as determined by a medical exam, (4) no chronic illness, including hyperprolactinemia, thyroid disease, and bone disease, (5) currently participating in at least 2 h/week of purposeful exercise, (6) non-smoker, (7) not currently dieting, (8) not taking any hormonal therapy for at least six months, (9) no current clinical diagnosis of eating or psychiatric disorders, (10) not pregnant or lactating or planning a pregnancy, (11) no medication use that would alter metabolic or reproductive hormone concentrations, and (12) no other contraindications that would preclude participation in the study. Women reporting regular menstrual cycles of 26–35 days for the past 6 months were eligible for the OV Control group; whereas, women who reported no menses in the past three months or ≤6 cycles in the past 12 months were eligible for the EAMD groups.

3.4. Baseline & randomization

The baseline phase was 4 weeks (28 days) in duration for the women with EAMD and the length of one menstrual cycle for women in the OV Control group. During this phase, participants collected daily urine samples and recorded menses on menstrual calendars. The urine samples were used to measure urinary metabolites of reproductive hormones, including estrone-1-glucuronide (E1G), pregnandiol glucuronide (PdG), and luteinizing hormone (LH). These metabolites allow for an assessment of the daily fluctuation of the ovarian hormones, estrogen and progesterone, and the pituitary hormone, LH [29,30]. As such, these metabolites are used to assess phases of the menstrual cycle, ovulation occurrence, and hormonal exposure [8,31]. Representative cycles of a woman in the EAMD group and a woman in the OV Control group are provided in Fig. 4.

Participants began a calcium and vitamin D run-in period on day 1 of the first week of baseline which coincided with the first day of menses for the OV Control group and a random day for the EAMD groups. All participants received oral calcium and vitamin D₃ supplements to ensure that they consumed the adequate intake (AI) of 1200 mg/per day of calcium and 400 IU of vitamin D (usual dietary intake was considered in achieving this goal and supplemented when

necessary) [32,33]. Calcium and vitamin D₃ were used as standard of care similar to other studies of bone health [34–37].

During the third week of baseline, participants arrived at the laboratory between 600 and 830hr (fasted and having refrained from exercise and caffeine for the prior 24hr and alcohol for the prior 12hr) and completed the following: (1) body weight, (2) resting energy expenditure (REE) assessment, and (3) blood sampling for the determination of metabolic hormones. During the same appointment or on a different day that week, participants also completed 1) a body composition assessment (DXA) and 2) peak oxygen uptake (VO_{2peak}) test to evaluate aerobic fitness. Participant were given a 3-day diet log and a 7-day exercise and activity log to complete during the following week. In a subset of participants (n = 35), bone geometry and structure were assessed using peripheral quantitative computed tomography (pQCT). Details about how the measurements were obtained are provided below.

3.4.1. Anthropometric assessment

Total body weight was measured to the nearest 0.1 kg each week during baseline; height was measured in centimeters using a stadiometer. BMI was calculated as the body mass divided by height squared (kg/m²).

3.4.2. Resting energy expenditure test

REE was determined by indirect calorimetry using a ventilated hood system (SensorMedics Vmax Series, Yorba Linda, CA, USA). RMR measurements were performed between 0630 and 1000 h in a dimly lit room at a comfortable temperature setting (20–24C). After volunteers lay quietly for 45 min, a transparent canopy was placed over their head. Volunteers were instructed to lie flat on their back and remain awake during the 30- to 45-min measurement period. Oxygen consumption (VO₂; mL/min) and carbon dioxide production (VCO₂; mL/min) were measured every 20 s. Following the REE test, the VO₂ and VCO₂ measures were assessed for steady state, which was achieved when the volume of expired air (VO₂) and respiratory quotient values were not varying by more than 10% or 5%, respectively. Only steady state data was used to calculate REE with the Weir equation [38]: REE (kcal/day) = [3.94(VO₂) + 1.11(VCO₂)] * 1.44. REE was used to calculate baseline energy expenditure needs and to provide an indicator of energetic status.

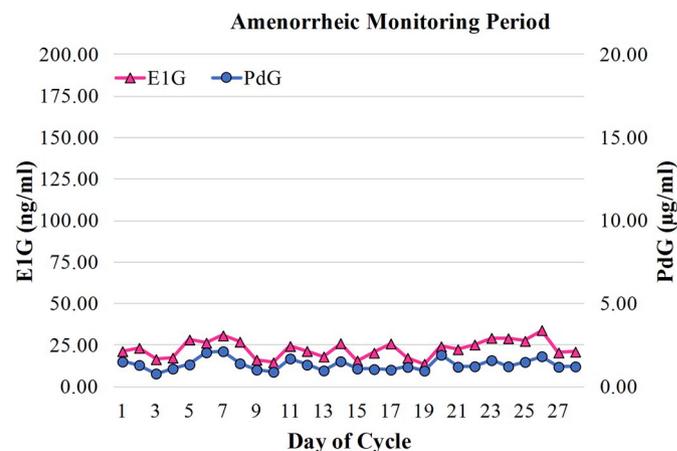


Fig. 4. Representative Profiles of Reproductive Hormones. Fig. 4A is the profile of urinary metabolites of reproductive hormones for an amenorrheic monitoring period of a woman in the EAMD group. The suppressed hormone concentrations are characteristic of exercising women with amenorrhea. Fig. 4B is the profile of urinary metabolites of reproductive hormones for a healthy ovulatory menstrual cycle of a woman in the Ovulatory Control group. The normal peaks in estrogen, progesterone, and luteinizing hormone that signify ovulation are evident. E1G: estrone-1-glucuronide (urinary metabolite of estrogen); PdG: pregnandiol glucuronide (urinary metabolite of progesterone); LH: luteinizing hormone.

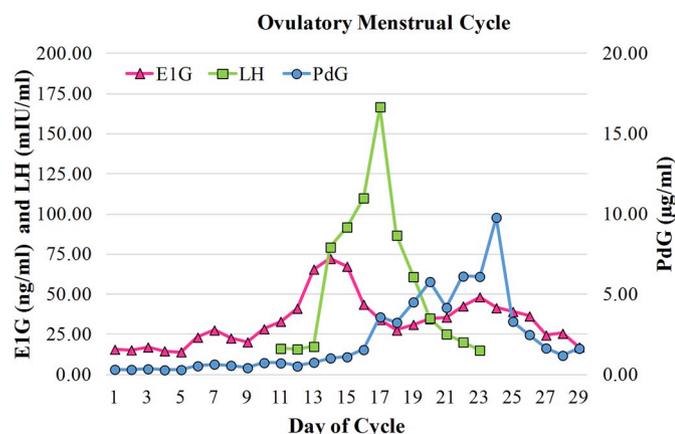


Fig. 4. (continued)

3.4.3. Dietary energy intake

Energy intake ($\text{kcal}\cdot\text{day}^{-1}$) was assessed using three-day diet logs recorded for two week days and one weekend day. On-site registered dietitians met with the participants to instruct them on how to accurately record food intake on the 3-day diet log. Participants were instructed to measure (using standard measuring cups/tools) and record all food and beverages consumed in detail. The nutrient data from the 3-day diet logs were coded and analyzed for total kilocalories using the Nutrition Data System for Research (NDSR 2008 Version; University of Minnesota; Minneapolis, MN, USA). Daily kilocalories consumed over the 3-day recording period were averaged. Three-day diet logs recording energy intake have been shown to provide comparable data to 7-day logs in women who may under-report their energy intake, including lean women [39]. Additionally, 3-day diet logs have been shown to reduce participant burden and improve compliance [40].

3.4.4. Purposeful exercise energy expenditure

Participants kept logs of their purposeful exercise each week during baseline. Participants were instructed to record the type of exercise and duration of each purposeful exercise session; as such, these logs provided a measurement of exercise volume over a 7-day period (min/wk). Participants were also provided with a Polar heart rate monitor to wear during each exercise session for one 7-day period; data collected from the monitor was used to estimate purposeful exercise energy expenditure (EEE). The OwnCal feature of the Polar S610 or RS400 heart rate monitors (Polar Electro Oy, Kempele, Finland), has been validated for using heart rate to calculate EEE [41–43]. The Polar S601 and RS400 heart rate monitors include rest in their estimation of energy expenditure. To estimate only net exercise energy expenditure, we subtracted the measured REE (kilocalories/min) from the Polar heart rate monitors' estimation of energy expenditure during exercise bouts. In the event that the Polar heart rate monitor was not worn during an exercise session, the metabolic equivalent (MET) level for the specific activity was determined using the Ainsworth et al. [44,45] compendium of physical activities [46]. The MET level of the activity was multiplied by the duration (min) of the exercise session and the measured REE (kcal/min) to estimate the energy expended during that session; because MET data include REE, the measured REE (kcal/min) was subtracted from this calculated EEE value.

3.4.5. Non-exercise activity thermogenesis

Energy expended during activities of daily living (non-exercise activity thermogenesis (NEAT)) was assessed using the RT3 triaxial accelerometer (Stayhealthy, Inc., Monrovia, CA) [47,48]. Participants wore the accelerometer on the anterior aspect of their right hip for 7 days. Accelerometers were removed during sleep and water activities (bathing, swimming) and when participating in purposeful exercise (at which time the Polar heart rate monitor was worn). Participants

completed a daily activity diary for 7 consecutive days during which all activities and their duration were recorded (i.e., all time during the 24-h day was accounted for). The accelerometer provided calories expended during each minute of the day; from these data, NEAT (kcal/day) was estimated as previously published [49].

3.4.6. Total daily energy expenditure

Baseline energy requirements were operationally defined as 24-h total daily energy expenditure (TDEE), which is the sum of laboratory-measured REE, purposeful EEE, NEAT, and the thermic effect of food (TEF). TEF was calculated as 10% of the sum of REE, EEE, and NEAT. The baseline energy requirement for each woman was used to determine individual "energy prescriptions," or the quantity of additional kilocalories to be consumed throughout the intervention, as explained below, using previously published methods [7].

3.4.7. Serum hormone measurement

Fasting blood samples were collected by a trained nurse or phlebotomist via venipuncture between 0700 and 1000h during week 3 of baseline. Another sample was collected at the end of the baseline period. The two samples were pooled for all baseline hormone analyses. After collection, samples were stored and processed as explained in the "Laboratory Analyses" section below. These samples were used to measure metabolic hormones, bone markers, and gut peptides.

3.4.8. Body composition assessment

A total body DXA scan was performed to assess body composition. Participants were scanned on either a GE Lunar Prodigy DXA scanner (GE Lunar Corporation, Madison, WI, enCORE 2002 software, version 6.50.069) or a GE Lunar iDXA scanner (GE Lunar Corporation, Madison, WI, enCORE 2008 software version 12.10.113). Consistent with the International Society for Clinical Densitometry guidelines, cross calibration studies were performed to remove systematic bias between the systems. Equations were derived using simple linear regression to remove biases and BMD and body composition absolute values obtained from the Lunar Prodigy were calibrated to the Lunar iDXA. More information about the cross calibration is provided below. Three DXA technicians, certified by the ISCD, performed and analyzed the scans.

3.4.9. VO_2 peak test

VO_2 peak ($\text{mL}/\text{kg}/\text{min}$) was measured during a progressive treadmill test using the Modified Astrand protocol [50] on a single occasion during the baseline period. The test was performed to volitional exhaustion [51] using breath-by-breath indirect calorimetry (SensorMedics Vmax metabolic cart, Yorba Linda, Ca). After a 3–5 min warm-up, the VO_2 peak test was initiated at a participant-selected comfortable running speed at a 0.0% grade. Every 2 min the grade was increased by 2.0% until reaching 6.0%; thereafter, the incline was increased 1.0% every minute until volitional exhaustion. The goals for achieving peak VO_2 criteria included a perceived exertion ≥ 18 , age predicted maximal heart rate, respiratory exchange ratio ≥ 1.1 , and an observed plateau in oxygen consumption despite an increase in exercise workload [52].

3.4.10. Classification of Menstrual Status

Initial classification of menstrual status prior to the intervention was based on self-reported menstrual history and urinary E1G, PdG, and LH profiles. Participants were assigned a "screening menstrual status" based on the self-reported number of menstrual cycles in the past 3–12 months. Prior to the baseline phase, women completed a questionnaire which asked them how many periods they had experienced in the past 3, 6, 9, and 12 months. Women were defined as "eumenorrheic" if they reported 2 or 3 menses in the past 3 months and at least 5 menses in the past 6 months. Women were defined as "amenorrheic" if they reported no menses in the past 3 months. Women were defined as "oligomenorrheic" if they reported 1 or 2 menses in the past 3 months and < 7 menses in the past 12 months.

Table 1
Definitions of menstrual cycle categories.

Menstrual Cycle Category	Definition ^a	Cycle & Phase Lengths
Optimal Ovulatory	<ul style="list-style-type: none"> ● E1G: > 35 ng/mL ● PdG: peak > 5 µg/mL AND ± 1d sum of > 10µg/mL ● LH: > 25 mIU/mL <u>within 5 days</u> after the E1G peak 	<ul style="list-style-type: none"> ● Cycle: 21–35 days ● Luteal phase: ≥ 10 days
Ovulatory LPD Inadequate	<ul style="list-style-type: none"> ● E1G: > 35 ng/mL ● PdG: peak > 2.5 µg/mL AND < 5 µg/mL AND/OR ± 1d sum of < 10µg/mL ● LH: > 25 mIU/mL <u>within 5 days</u> after the E1G peak 	<ul style="list-style-type: none"> ● Cycle: 21–35 days ● Luteal phase: ≥ 10 days
Ovulatory LPD short	<ul style="list-style-type: none"> ● E1G: > 35 ng/mL ● PdG: peak > 5 µg/mL AND/OR ± 1d sum of > 10µg/mL ● LH: > 25 mIU/mL <u>within 5 days</u> after the E1G peak 	<ul style="list-style-type: none"> ● Cycle: 21–35 days ● Luteal phase: < 10 days
Ovulatory LPD inadequate and short	<ul style="list-style-type: none"> ● E1G: > 35 ng/mL ● PdG: peak > 2.5 µg/mL AND < 5 µg/mL AND/OR ± 1d sum of < 10µg/mL ● LH: > 25 mIU/mL <u>within 5 days</u> after the E1G peak 	<ul style="list-style-type: none"> ● Cycle: 21–35 days ● Luteal phase: < 10 days
Anovulatory	<ul style="list-style-type: none"> ● E1G: suppressed (< 35 ng/mL) ● PdG: peak < 2.5 µg/mL ● LH: no peak (< 25 mIU/mL) 	<ul style="list-style-type: none"> ● Cycle: 21–35 days
Oligomenorrheic Ovulatory	<ul style="list-style-type: none"> ● E1G: > 35 ng/mL ● PdG: peak > 2.5 µg/mL ● LH: > 25 mIU/mL <u>within 5 days</u> after the E1G peak 	<ul style="list-style-type: none"> ● Cycle: 36–89 days
Oligomenorrheic Anovulatory	<ul style="list-style-type: none"> ● E1G: suppressed (< 35 ng/mL) ● PdG: peak < 2.5 µg/mL ● LH: no peak (< 25 mIU/mL) 	<ul style="list-style-type: none"> ● Cycle: 36–89 days
Amenorrheic	<ul style="list-style-type: none"> ● E1G: suppressed (< 35 ng/mL) ● PdG: suppressed (< 2.5 µg/mL) ● LH: no peak (< 25 mIU/mL) 	<ul style="list-style-type: none"> ● ≥ 90 days
Short Ovulatory	<ul style="list-style-type: none"> ● E1G: > 35 ng/mL ● PdG: peak > 2.5 µg/mL ● LH: > 25 mIU/mL <u>within 5 days</u> after the E1G peak 	<ul style="list-style-type: none"> ● Cycle: < 21 days
Short Anovulatory	<ul style="list-style-type: none"> ● E1G: suppressed (< 35 ng/mL) ● PdG: peak < 2.5 µg/mL ● LH: no peak (< 25 mIU/mL) 	<ul style="list-style-type: none"> ● Cycle: < 21 days

E1G: estrone-1-glucuronide; PdG: pregnanediol glucuronide; LH: luteinizing hormone; LPD: luteal phase defect.

^a PdG criteria applied to the luteal phase.

Participants' "baseline menstrual status" was also assessed using profiles of E1G, PdG, and LH in urine samples collected daily during the baseline phase. During the baseline phase, eumenorrheic women collected daily urine samples for one menstrual cycle, oligomenorrheic women for no more than 90 days, and women with amenorrhea for one 28-day monitoring period. Participants were instructed to collect first morning void urine samples each day of the baseline period and to record menses on daily menstrual calendars.

Baseline cycles were classified in 10 different categories based on the hormonal profile of the cycle. See Table 1 for the definitions of these menstrual cycle classifications. Cycles that were incomplete or missing substantial data, thereby precluding the ability to accurately classify the cycle, were categorized as incomplete, and excluded from the study or the baseline cycle was repeated. Menstrual cycle length was defined as the number of days from the first day of menses up to the day preceding the next menses [8,31]. The follicular phase was defined as the first day of menses up to and including the day of the urinary LH peak, identified on the day of or within a few days after the mid-cycle E1G peak [8,31]; whereas, the luteal phase was defined as the day after the LH peak to the last day preceding the onset of the next menses. Ovulation was confirmed by presence of a urinary LH peak preceded by an E1G peak. Specific hormonal criteria for E1G, PdG, and LH are identified in Table 1 [8,10,31,53,54].

Participants were also assigned an "overall menstrual status" which incorporated both the self-reported "screening menstrual status" and the hormonal profile "baseline menstrual status." The rules applied to determine "overall menstrual status" are provided in Table 2.

A participant was assigned to the EAMD group if her overall menstrual status was amenorrheic or oligomenorrheic. After meeting study criteria in the screening phase and completing a 4-week baseline phase, women with EAMD were randomly assigned into one of two groups, the EAMD + Cal group or the EAMD Control group. A participant was assigned to the OV Control group if her overall menstrual status

indicated eumenorrhea (self-reported eumenorrhea) with an ovulatory baseline cycle.

For all participants, we calculated the free androgen index (FAI) to rule out the possibility that hyperandrogenism was contributing to the menstrual irregularities. For those who presented with a high FAI ($n = 5$), it was suspected that the menstrual irregularity *may* be from a cause other than functional hypothalamic amenorrhea (FHA) [55–57]; thus, these subjects were removed from data analyses. To determine the cutoff for FAI, we calculated the 95% confidence interval of FAI for all women in our study who completed baseline ($n = 133$). Any woman in the study with a FAI above the upper limit of the 95% confidence interval ($FAI \geq 6.6$) was categorized as "high FAI." FAI was calculated as follows: $(Total\ Testosterone/Sex\ Hormone\ Binding\ Globulin) * 100$.

3.5. Intervention

3.5.1. Energy prescription

The women in the intervention group (EAMD + Cal) were provided an energy prescription of increased energy intake 20–40% above baseline energy requirements and asked to maintain their usual exercise training regimen for the intervention phase of the study. Participants in the EAMD + Cal group were counseled by a nutritionist and given energy bars that contained approximately 220–300 calories as well as pre-measured servings of nuts. These women were also counseled by a dietitian about how to increase calories using the foods that they typically ate. Participants in the EAMD Control and OV Control groups were asked to maintain baseline physical activity level and energy intake but were monitored in a similar manner to the EAMD + Cal group.

3.5.2. Primary outcome measures

During the intervention phase of the study, study visits occurred either twice each month or monthly, depending on the group assignment and time point in the intervention. An overview of which

Table 2
Determination of overall menstrual status.

Screening Menstrual Status ^a	Baseline Menstrual Status ^b	Overall Menstrual Status
Eumenorrhea	Ovulatory Cycle < 36 days in length (Eumenorrheic)	Eumenorrhea
Amenorrhea	Suppressed E1G, PdG, LH; No menses (Amenorrheic)	Amenorrhea
Oligomenorrhea	Menstrual Cycle 36–89 days in length (Oligomenorrheic)	Oligomenorrhea
Oligomenorrhea	Suppressed E1G, PdG, LH; No menses (Amenorrheic)	Oligomenorrhea
Oligomenorrhea	Ovulatory or Anovulatory Menstrual Cycle < 36 days in length (Eumenorrheic)	Oligomenorrhea
Oligomenorrhea	Evidence of menses but missing collection days	Oligomenorrhea
Oligomenorrhea	Evidence of ovulation	Oligomenorrhea
Amenorrhea	Menses occurred	Oligomenorrhea
Amenorrhea	Menstrual Cycle 36–89 days in length (Oligomenorrheic)	Oligomenorrhea
Amenorrhea	Anovulatory Cycle	Oligomenorrhea
Amenorrhea	Ovulatory Cycle < 36 days in length (Eumenorrheic)	Baseline Failure ^c
Amenorrhea	Incomplete; unable to be characterized	Baseline Failure ^c
Eumenorrhea	Menstrual Cycle 36–89 days in length (Oligomenorrheic)	Baseline Failure ^c
Eumenorrhea	Anovulatory Cycle	Baseline Failure ^c

E1G: estrone-1-glucuronide; PdG: pregnanediol glucuronide; LH: luteinizing hormone.

^a Self-reported: Eumenorrhea: 2 or 3 menses in the past 3 months and at least 5 menses in the past 6 months. Amenorrhea: no menses in the past 3 months. Oligomenorrhea: 1 or 2 menses in the past 3 months and < 7 menses in the past 12 months.

^b See Table 1.

^c Not used in statistical analyses because they did not meet established baseline criteria for study inclusion. Note: There were 2 women whose *data are not usable* for the following reasons: a) bleeding/spotting occurred continuously throughout the monitoring period without clear menses (n = 1), and b) data could not be trusted because hormonal profile indicated ovulatory cycles; however, participant indicated that menses never occurred (n = 1).

measures were performed at the various time points is provided in the study flow chart (Fig. 2A and B). The primary outcome measures for this RCT included 1) change in energy intake and body mass, and other measures of energy status, 2) change in menstrual function, and 3) change in bone mineral density and secondary indices of bone health. Specific methods of measuring each outcome measure are described below.

3.6. Primary outcome I) energy status: REE, energy intake & expenditure, body weight & composition, metabolic hormones & gut peptides

3.6.1. Energy intake

The EAMD + Cal women met with a registered dietitian at screening and baseline then bi-weekly for the first 3 months and monthly for the remainder of the study. Throughout the intervention, participants in the EAMD and OV Control groups met with the registered dietitian at monthly and 3-month intervals, respectively (see Fig. 2). Participants were monitored by the dietitian for changes in nutritional and eating behavior characteristics and compliance to energy prescription, i.e., the dietitian reviewed participants' diet logs and provided strategies to achieve prescribed energy intake.

3.6.2. Energy expenditure (REE and EEE)

Both REE and EEE were measured as described in the "Baseline" section. For all groups, REE measurements were completed monthly for the first 3 months of the intervention (IWk1, 5, 9) then every 3 months for the remainder of the intervention (IWk21/month 6 and IWk33/month 9). This physical activity monitoring was completed monthly for women in the EAMD groups and every 3 months for women in the OV Control group. Participants were also instructed to keep a daily exercise log that recorded the type and duration of all purposeful exercise sessions completed during the study.

3.6.3. Body weight & composition

Body weight was measured on a digital scale every 2 weeks in EAMD women and monthly in OV Control women. Participants were asked to wear light clothing and remove their shoes for each measurement. Body composition was assessed using a total body DXA scan. Scans were performed and analyzed by 3 certified DXA technicians. Body composition assessment occurred monthly for the first 3 months

of the intervention (IWk1, 5, 9) then every 3 months for the duration of the intervention (IWk 21, 33). Variables obtained included fat mass, fat free mass, lean body mass, and bone mass of the total body as well as that of specific regions, such as the android, gynoid, trunk, legs, and arms.

3.6.4. Metabolic hormones & gut peptides - blood samples

Fasting blood samples were collected by a trained nurse or phlebotomist via venipuncture between 0700 and 1000h monthly for the first 6 months of the intervention (IWk1, 5, 9, 13, 17, 21) and then every 3 months for the remainder of the intervention (IWk33). Participants were instructed to fast overnight and laid supine for at least 15 min prior to the blood draw. After collection, samples were stored and processed as explained in the "Laboratory Analyses" section below. These samples were used to measure metabolic hormones, bone markers, and gut peptides.

3.7. Primary outcome II) menstrual function

3.7.1. Urine collection

Women in the EAMD groups collected first morning void urine samples every day for the duration of the intervention. These collections were divided into 28-day monitoring periods. OV Control women collected first morning void urine samples every day for one menstrual cycle at 3 different time points during the intervention; these collection periods coincided closely with months 4, 8, and 12 of the intervention. Urine samples were assessed for E1G, PdG, and LH as described in the "Baseline" and "Laboratory Analyses" sections. Further, each monitoring period and menstrual cycle was classified into 10 different categories based on the hormonal profile of the monitoring period or cycle, as described in Table 1 and the "Baseline" section.

3.7.2. Menstrual logs

Participants were instructed to record menses and other menstrual symptoms on monthly calendars and return them at the end of each month.

3.8. Primary outcome III) BMD and bone health

3.8.1. DXA

DXA scans of the total body, lumbar spine, and dual femur were performed to assess areal BMD at screening, IWK21 (month 6), and Post-study (month 13). In addition, femoral neck geometry and strength were estimated from dual femur scans by Hip Strength Analysis (HSA) as developed by Yoshikawa et al. [58]. HSA is a feature of the GE Lunar software that is used to estimate the structural properties of the hip, and the measurements were obtained from the automatic analysis. Specific measurements obtained include femoral neck cross-sectional moment of inertia (CSMI), cross-sectional area (CSA), strength index (SI), and diameter.

The majority of participants were scanned on either a GE Lunar Prodigy DXA scanner (n = 50) (GE Lunar Corporation, Madison, WI, enCORE 2002 software, version 6.50.069) or a GE Lunar iDXA scanner (n = 100) (GE Lunar Corporation, Madison, WI, enCORE 2008 software version 12.10.113). Three certified DXA technicians performed the scans for BMD. Consistent with the International Society for Clinical Densitometry guidelines, cross calibration studies were performed to remove systematic bias between the systems. For the cross calibration study between the Lunar Prodigy and Lunar iDXA, fourteen subjects were scanned in triplicate on both machines. The majority (n = 8) were scanned on both machines within 5 days; however, there was approximately one month between scans for some subjects (n = 6). Body composition, BMD, and femoral neck geometry measurements on the Lunar Prodigy and iDXA systems were found to be highly correlated with $r \geq 0.930$ for body composition and $r \geq 0.983$ for BMD and femoral neck geometry. Equations were derived using simple linear regression to remove biases, and BMD and body composition absolute values obtained from the Lunar Prodigy were calibrated to the Lunar iDXA. The same software version was used to obtain all HSA measurements. A small subset of participants (n = 17) were scanned on a Hologic QDR 4500W (Hologic Inc., Bedford, MA) prior to the availability of the GE Lunar iDXA. A similar cross calibration procedure was also performed using duplicate scans of n = 32 participants, and Hologic data were converted to the Lunar iDXA.

3.8.2. pQCT

As an optional procedure, peripheral quantitative computed tomography (pQCT) scans (Stratec XCT-3000, software version 6.00B, Stratec Medical, Pforzheim, Germany) of the distal (4%) and proximal (66%) radius and tibia were performed at baseline, IWK21 (month 6), and post-study (month 13) in a subset of participants (n = 40). The pQCT scans evaluated volumetric BMD, bone geometry, and estimated bone strength at these sites.

As a general rule, the non-dominant radius and opposite tibia were scanned; however, if a participant had a history of fracture in a bone generally scanned, the opposite limb was scanned. Prior to scanning, bone length was measured with a measuring tape. The tibia was measured from the tibial plateau to the base of the medial malleolus, and the radius was measured from the olecranon process to the head of the ulnar styloid process. pQCT scans were performed at 4% and 66% of the bone length, proximal to the tibial and radial endplates. A scout view scan was performed prior to each scan to position the reference in the endplates according to manufacturer guidelines.

3.9. Secondary outcome measures: psychological status & eating behavior, aerobic fitness

3.9.1. Secondary outcome I) psychological status/eating behavior (questionnaires)

The EAMD + Cal and EAMD Control participants met with a clinical psychologist twice each month for the first three months and then monthly for the remainder of the study to monitor general psychological and eating behavior status and provide assistance in implementing

the energy prescription and other lifestyle changes to ensure compliance to the intervention. The participants on the OV Control group met with the clinical psychologist at 3-month intervals throughout the study. Surveys regarding eating behavior and psychological stress were completed at regular intervals throughout the study. The questionnaires that were administered were as follows: 1) an overall health questionnaire (demographics, general health, menstrual and bone health, medications and supplements), 2) Three-Factor Eating Questionnaire (TFEQ) [23], 3) Eating-Disorder Inventory II (EDI-II) [24], 4) Brief Resilient Coping Scale [59], 5) Profile of Mood States (POMS) [60], 6) Importance of Change Scale, 7) Beck Depression Inventory [61], and 8) stress questionnaires, including Perceived Stress Scale [27], Daily Stress Inventory [25], and Dysfunctional Attitude Scale [62,63]. The Health Questionnaire was only completed at screening. All other questionnaires were completed at IWk 5, 9, 21, 33 (months 2, 3, 6, 9) of the intervention. At IWk 13 and 17 (months 4 and 5), only the stress questionnaires were completed (Please refer to Fig. 2.).

3.9.2. Post-intervention

At the conclusion of the 12-month intervention or observation period, participants underwent post-intervention assessments. These measurements included 1) body weight, 2) REE assessment, 3) DXA and pQCT scans, 4) eating behavior and stress questionnaires as described in the "Intervention" section, 5) psychologist and nutritionist interviews, 6) diet and exercise logs and monitoring of physical activity, and 7) menstrual log. All procedures were completed as described in the "Baseline" or "Intervention" sections above. The DXA scans included scans of the total body for body composition and BMD, the lumbar spine for BMD, and the dual hip for BMD and HSA measurements.

3.10. Laboratory Analyses

3.10.1. Fasting blood samples

As described above, fasting blood samples were collected at specified time points throughout the study between 0700 and 1000h. After collection, the samples were permitted to clot and were spun in a Eppendorf centrifuge. The serum was aliquoted into 300 μ l samples and stored in a -80°C freezer until analysis. These samples were analyzed for a variety of metabolic hormones, gut peptides, and bone markers. All samples from a given participant were analyzed in duplicate. The assays were performed in our lab at Penn State University; details for the analysis of each hormone, peptide or bone marker will be provided in forthcoming original research publications.

3.10.2. Urinary reproductive hormone measurements: calculation of hormone exposure

To determine estrogen and progesterone exposure, area under the curve (AUC) for E1G and PdG was calculated for a complete menstrual cycle (if eumenorrheic or oligomenorrheic) or 28-day monitoring period (if amenorrheic) using Kaleidagraph Software (Synergy Software, Reading, PA, USA). Mean E1G and PdG concentrations across the cycle were also calculated. E1G and PdG AUC and mean were calculated for the following variables: a) entire cycle, b) follicular phase (in ovulatory cycles), c) luteal phase (in ovulatory cycles), d) days 2–5 of cycle/monitoring period (E1G only), and e) days 2–12 of cycle/monitoring period (E1G only). The following data were also collected in ovulatory cycles: a) E1G peak concentration during the follicular phase and the day of the cycle that it occurred, b) PdG peak concentration during the luteal phase and the day of the cycle that it occurred, c) LH peak concentration and the day of the cycle that it occurred.

3.10.3. Urinary reproductive hormone measurements: Laboratory Analyses

All urine samples were corrected for specific gravity using a hand refractometer (NSG Precision Cells; Farmingdale, NY) to account for hydration status [64], which has been reported to perform as well as creatinine correction for adjusting urinary hormone concentrations

[64]. The secretion of E1G and PdG metabolites in the urine parallels serum concentrations of the parent hormones, estrogen and progesterone, respectively [29].

Microtiter plate competitive enzyme immunoassays were used to measure the urinary metabolites E1G and PdG. The E1G (R522-2) and PdG (R13904) assays use a polyclonal capture antibody supplied by Coralie Munro University of California (Davis, CA). The inter-assay coefficients of variation for high and low internal controls for the E1G assay are 12.2% and 14.0% respectively. The PdG intra- and inter-assay variability was determined in-house as 13.6% and 18.7%, respectively [8,65].

Urinary LH was determined by coat-a-count immunoradiometric assay (Siemens Healthcare Diagnostics, Deerfield, IL). The sensitivity of the LH assay is 0.15 mIU/mL. The intra- and inter-assay coefficients of variation were 1.6% and 7.1%, respectively.

3.11. Sample size calculations

The appropriate sample size was determined using two primary outcomes of the study, 1) recovery of menstrual function and 2) bone health. Based on the calculations explained below for these two outcomes, a target enrollment of 30 subjects per group, for a total of $n = 90$ volunteers, should yield adequate power to detect differences among groups. All sample size calculations were performed with $\alpha = 0.05$ using two tailed tests. All sample size calculations were performed using Sample Power v 2.0 (SPSS, Chicago, IL).

3.11.1. Sample size calculation for resumption of menses

The sample size required to detect that a significant proportion of volunteers in the treatment group will resume menses as a result of the diet intervention was based on reported results in four studies [16,17,66,67]. Three of these studies [16,17,66] represent the only studies found in the literature where volunteers with severe EAMD were prospectively studied before and after an alteration in energy balance that resulted in weight gain, thus paralleling the EAMD + Cal group in the study described herein. In a study by Drinkwater et al. [66], 7 of 9 amenorrheic volunteers who were followed over time experienced a return of menses that was associated with decreases in training, weight gain, and dietary modifications. In a study by Koppe-Woodruffe et al. [17], amenorrheic volunteers underwent an intervention designed to increase dietary intake and decrease exercise volume. They reported ovulation in approximately 15–20 weeks in 3 of 5 volunteers. In an earlier report of a case study, Dueck et al. [16] reported that ovulation was successfully induced in one amenorrheic individual with a similar intervention, i.e., 1 of 1. When all three studies are considered together, 11 of 15 or 73% of volunteers resumed menses, indicated by at least a single occurrence of menses, as a result of a presumed change in energy balance. For comparison purposes we used a conservative (high) estimate of 0.40 as the proportion of volunteers who might experience a spontaneous resumption of menses. As a reference for this assumption, we rely on the report by Gibson et al. [67] that 33% of volunteers with severe EAMD who were assigned to a no treatment group experienced a spontaneous resumption of menses. Using a chi-square test to compare these proportions of women who may resume as a result of an intervention vs. those who may spontaneously resume without treatment (0.40 vs. 0.33), a proposed sample size of 17 will have 80% power to yield a statistically significant result. In this way, this calculation parallels a comparison between the EAMD + Cal group and EAMD Control in the REFUEL study. It must be noted that the definition of menstrual resumption is not standardized and varies across studies.

3.11.2. Sample size calculation for BMD

The primary outcome variable used to calculate sample size was BMD at the lumbar spine (L1-L4). The literature provides no studies wherein a dietary intervention of any type was employed to reverse bone loss in women with severe EAMD who are not considered to have

clinical eating disorders. We therefore used a report by Drinkwater et al. [66] as a basis for sample size calculations. In the study by Drinkwater et al., a subset of amenorrheic athletes volunteered for follow-up BMD testing an average of 15.5 months after a previous baseline study was performed [66]. An average 6.3% increase in L1-L4 BMD was observed in 7 of 9 athletes. Each of the 7 athletes had resumed menses after an average of 4.7 months since baseline testing. In 2 athletes who did not resume menses and were otherwise untreated, BMD decreased by 3.4%. In the 7 formerly amenorrheic athletes, Drinkwater et al. [66] documented several factors that might have contributed to the resumption of menses and subsequent observation of increased BMD, including increased consumption of dairy products and or calcium supplements, reductions in training, and an increase in body weight. Since the latter changes were likely to have been associated with a positive change in energy balance, similar to what we expect in our EAMD + Cal, we based our estimated difference on the changes observed in these volunteers. From the differences in L1-L4 illustrated in Drinkwater et al. [66], we calculated an annualized percentage increase of +4.87%, then based our sample size calculation on the expected difference in BMD that a 4.87% increase represents, i.e., +0.0536 gm/cm² (which represents 4.87% of the baseline value of 1.127 g/cm²). For the standard deviation to be used in the sample size calculation, we used the estimated SD of the individual change scores for BMD at 15.5 months (estimated from Fig. 1 of Drinkwater et al. [66] that included individual plots of changes in BMD), i.e., 0.086 g/cm². To estimate the % change in the EAMD Control, we used data reported by Gibson et al. [67], who tested whether the addition of calcium and vitamin D could increase bone mass in amenorrheic athletes. In the latter study, there was no significant difference in BMD from pre-to post-study after 11.2 months, and we therefore estimated 0% change for this group. As per Sowers et al. [68], who studied annual changes in BMD for three years in Caucasian women aged 24–44 yrs, we also estimated 0% for the expected change over 12 months in L1-L4 for the OV Control group. Thus, our expected difference, when our EAMD + Cal intervention group is compared against the OV Control (expected difference = 0%) and the EAMD Control (expected difference = 0%), is 4.87%, or 0.0536 gm/cm². Using the SD of 0.086 gm/cm², the sample size was calculated using a single sample *t*-test = 0 test (Sample Power, SPSS, Inc). The sample size required to detect an expected difference of 4.87% with $\alpha = 0.05$ and 80% power is 23 completed participants per group. This number was increased by 30% to account for drop outs. Therefore, targeted enrollment was 30 volunteers in each group. The least significant change (LSC) for lumbar spine BMD in our lab on the GE Lunar iDXA is 1.45%, indicating the smallest amount of change needed for a clinically meaningful improvement.

3.11.3. Sample size calculations for bone turnover

No studies were identified that prospectively examined changes in bone markers with a dietary intervention in volunteers with severe EAMD. Although numerous studies exist that examine weight gain in women with anorexia nervosa, most of these involve adolescents under the targeted age for the current study, and many include concurrent treatment with hormonal therapy. However, Caillot-Augusseau et al. [69] studied changes in bone markers with weight gain and no hormonal therapy in 9 women with anorexia nervosa that were 16–30 years (mean of 21 yrs) and amenorrheic. The duration of refeeding was 8 ± 5.3 months during which time BMI changed 27% from 13.8 ± 1.6 to 17.5 ± 2.3 kg/m². Thus, for the purpose of the calculation these women represent the EAMD + Cal group in the current study; however, we do not anticipate a BMI change of that magnitude. A reasonable expectation for an increase in BMI in the current study would be equivalent to the difference between amenorrheic athletes (BMI = 17.5 ± 0.2 kg/m²) and eumenorrheic athletes (19.6 ± 0.1 kg/m²), as reported by Zanker and Cooke [70], or 2.1 kg/m², which represents an approximate 12% increase. We therefore based our sample size calculation on an expected difference in bone markers

equivalent to 60% of what was observed by Caillot-Augusseau et al. [69], since they reported a much larger change in BMI than we anticipate in the current study. We estimated sample size based on expected changes in our experimental groups; we anticipate 0% changes in the OV Control group and the EAMD Control group. We used an expected difference in uCTX of 54 $\mu\text{g}/\text{mmol creat}$ (60% of that reported in Caillot-Augusseau et al. [69]), along with an estimated standard deviation of 91.8 $\mu\text{g}/\text{mmol creat}$, estimated by assuming a 0.5 correlation between the standard deviations of the pre- and post-scores for uCTX. A proposed sample of 23 volunteers per group yields 77% power to detect a statistically significant result for the EAMD + Cal group. We also estimated sample size based on results reported in Caillot-Augusseau et al. [69], for serum intact osteocalcin. We used an expected difference of 60% of that reported (reported = 11.8 ng/mL) and a standard deviation of 5.03 ng/mL, conservatively estimated by assuming a low ($r = 0.2$) correlation between reported standard deviations for pre and post data for osteocalcin. A proposed sample of 23 volunteers per group will yield 100% power to detect a significant result.

4. Discussion

The REFUEL RCT explored the influence of 12 months of increased caloric intake on energetic status, menstrual status, and bone health among exercising women who presented with menstrual disturbances, such as amenorrhea or oligomenorrhea, secondary to an energy deficit. To our knowledge, this is the first RCT that assessed the aforementioned outcomes in detail for a 12-month period, using appropriate control groups. It has been recommended that the mainstay of treatment for menstrual disturbances among female athletes who may be suffering from low energy availability associated with the Female Athlete Triad is to increase energy intake [3,71]. To date, there are no published RCTs that provide scientific evidence that this non-pharmaceutical approach is effective. Moreover, there are no experimental studies that quantify the magnitude of energy surplus required to reverse the Female Athlete Triad or document the underlying metabolic, endocrine, and behavioral mechanisms whereby an intervention of increased energy intake is successful. Prior observational studies [20] and uncontrolled interventions [16,17] are informative, but the possibility of bias introduced by individual susceptibility cannot be ruled out and these studies are small in size and do not document underlying mechanisms. Because menstrual disturbances and poor bone health among exercising women are often downstream effects of an energy deficit, it is logical that addressing the root of the problem through lifestyle changes, i.e. an increase in energy intake, may be an effective and sustainable treatment, especially when energy intake is increased in a gradual manner that is initiated with only 200–400 kcal increments. Notably, improvements in menstrual cyclicity have been observed with a moderate increase of approximately 300 kcal/day [15–17,22]; since this increase in energy intake is equivalent to one additional snack or food item each day, it appears to be well-received and manageable for exercising young women.

However, to address menstrual disturbances among exercising women, many clinicians resort to pharmaceutical approaches, such as the administration of oral contraceptives, to address the symptoms, i.e., the irregular or absent menstrual cycles, and/or low bone mass [72]. Oral contraceptives, however, do not address the underlying problem of low EA/energy deficit and, therefore, may not benefit the long-term health of the athletes. Moreover, oral contraceptives suppress key anabolic hormones such as IGF-1 due to first pass effects through the liver and have downstream effects on bone metabolism [73–75]. As such, the results of this study have the potential to inform and alter clinical practice for the long-term benefit of female athletes and recreationally-exercising women.

A primary strength and unique aspect of this study is the thorough monitoring of menstrual status and detailed analysis of reproductive

hormones throughout the entire study; consequently, we were able to create a comprehensive picture of the menstrual environment at baseline and throughout the intervention as subjects experienced changes in their menstrual status, including recovery of menstrual function. With the detailed analyses, we were able to complete a careful characterization of each menstrual cycle, which allowed for determining the presence or absence of ovulation and the identification of subtle menstrual disturbances, i.e., luteal phase defects and anovulation. In this way, we were able to observe the spectrum of menstrual function and ascertain varying degrees of menstrual recovery as a result of the intervention. Furthermore, to our knowledge, this study is the only RCT that explores the three interrelated components of the Female Athlete Triad, i.e., energetic status, menstrual status, and bone health, in great detail for a period of 12 months.

Although 12 months is a longer period of time for an intervention of its type compared to currently published studies, it still remains a relatively short period of time to observe significant changes in bone health when clinical outcomes such as BMD are used. As such, the duration of the study may mask important physiological and skeletal changes occurring as a result of the intervention. However, it is difficult to balance subject compliance and retention with optimal study duration. The average dropout rate during the intervention exceeded 50%, primarily due to time commitment and non-compliance; however, the dropout rate was similar among all groups (range: 50–56%), and group assignment did not statistically influence time to dropout during the intervention ($p > 0.10$). The number of physiological data points and the 12-month duration resulted in high subject burden for this population, which was primarily composed of college students. The high dropout rate reflects the challenges associated with an adequate assessment of the study's primary outcome measures. Lastly, this study relied on self-reported measures of energy intake and some self-reported measures of energy expenditure in a free-living situation. We acknowledge that the subjective nature of these measurements impacts the accuracy of the energetic data; however, we also collected a wide array of objective laboratory measurements that will be used to provide an accurate assessment of energetic and metabolic changes, such as REE, energy expenditure via an accelerometer (non-exercise) and heart rate monitor (exercise), and serum concentrations of metabolic hormones. Therefore, this array of measurements provide a comprehensive assessment of energy status while simultaneously evaluating eating behavior and psychological status via validated questionnaires.

The REFUEL study provides the first RCT data for the non-pharmaceutical treatment approach to the three interrelated conditions present in the Female Athlete Triad: low EA, menstrual disturbances, and poor bone health. Despite the aforementioned limitations, we believe the results of this study will have important clinical implications, including the potential to alter clinical practice to benefit the health of exercising girls and women.

Funding

This work was supported by the United States Department of Defense Congressionally Directed Medical Research Programs (CDMRP) (grant number PR05431).

References

- [1] A.B. Loucks, Low energy availability in the marathon and other endurance sports, *Sports Med.* 37 (2007) 348–352.
- [2] A. Nattiv, A.B. Loucks, M.M. Manore, C.F. Sanborn, J. Sundgot-Borgen, M.P. Warren, American College of Sports Medicine position stand. The female athlete triad, *Med. Sci. Sports Exerc.* 39 (2007) 1867–1882.
- [3] M.J. De Souza, A. Nattiv, E. Joy, M. Misra, N.I. Williams, R.J. Mallinson, J.C. Gibbs, M. Olmsted, M. Goolsby, G. Matheson, 2014 female athlete triad coalition consensus statement on treatment and return to play of the female athlete triad: 1st international conference held in san Francisco, California, may 2012 and 2nd international conference held in indianapolis, Indiana, may 2013, *Br. J. Sports Med.* 48 (2014) 289.

- [4] A.B. Loucks, B. Kiens, H.H. Wright, Energy availability in athletes, *J. Sports Sci.* 29 (Suppl 1) (2011) S7–S15.
- [5] G.N. Wade, J.E. Schneider, H.Y. Li, Control of fertility by metabolic cues, *Am. J. Physiol.* 270 (1996) E1–E19.
- [6] J.L. Lieberman, D.E.S. MJ, D.A. Wagstaff, N.I. Williams, Menstrual disruption with exercise is not linked to an energy availability threshold, *Med. Sci. Sports Exerc.* 50 (2018) 551–561.
- [7] N.I. Williams, H.J. Leidy, B.R. Hill, J.L. Lieberman, R.S. Legro, M.J. De Souza, Magnitude of daily energy deficit predicts frequency but not severity of menstrual disturbances associated with exercise and caloric restriction, *Am. J. Physiol. Endocrinol. Metab.* 308 (2015) E29–E39.
- [8] M.J. De Souza, R.J. Toombs, J.L. Scheid, E. O'Donnell, S.L. West, N.I. Williams, High prevalence of subtle and severe menstrual disturbances in exercising women: confirmation using daily hormone measures, *Hum. Reprod.* 25 (2010) 491–503.
- [9] M. Russell, M. Misra, Influence of ghrelin and adipocytokines on bone mineral density in adolescent female athletes with amenorrhea and eumenorrheic athletes, *Med. Sport Sci.* 55 (2010) 103–113.
- [10] M.J. De Souza, S.L. West, S.A. Jamal, G.A. Hawker, C.M. Gundberg, N.I. Williams, The presence of both an energy deficiency and estrogen deficiency exacerbate alterations of bone metabolism in exercising women, *Bone* 43 (2008) 140–148.
- [11] L.K. Bachrach, Acquisition of optimal bone mass in childhood and adolescence, *Trends Endocrinol. Metabol.* 12 (2001) 22–28.
- [12] G. Theintz, B. Buchs, R. Rizzoli, D. Slosman, H. Clavien, P.C. Sizonenko, J.P. Bonjour, Longitudinal monitoring of bone mass accumulation in healthy adolescents: evidence for a marked reduction after 16 years of age at the levels of lumbar spine and femoral neck in female subjects, *J. Clin. Endocrinol. Metab.* 75 (1992) 1060–1065.
- [13] T. Csermely, L. Halvax, M. Vizer, I. Drozdyk, P. Tamas, P. Gocze, I. Szabo, S. Jeges, A. Szilagyi, Relationship between adolescent amenorrhea and climacteric osteoporosis, *Maturitas* 56 (2007) 368–374.
- [14] N.I. Williams, D.L. Helmreich, D.B. Parfitt, A. Caston-Balderrama, J.L. Cameron, Evidence for a causal role of low energy availability in the induction of menstrual cycle disturbances during strenuous exercise training, *J. Clin. Endocrinol. Metab.* 86 (2001) 5184–5193.
- [15] R.J. Mallinson, N.I. Williams, M.P. Olmsted, J.L. Scheid, E.S. Riddle, M.J. De Souza, A case report of recovery of menstrual function following a nutritional intervention in two exercising women with amenorrhea of varying duration, *J Int Soc Sports Nutr* 10 (2013) 34.
- [16] C.A. Dueck, K.S. Matt, M.M. Manore, J.S. Skinner, Treatment of athletic amenorrhea with a diet and training intervention program, *Int. J. Sport Nutr.* 6 (1996) 24–40.
- [17] S.A. Kopp-Woodroffe, M.M. Manore, C.A. Dueck, J.S. Skinner, K.S. Matt, Energy and nutrient status of amenorrheic athletes participating in a diet and exercise training intervention program, *Int. J. Sport Nutr.* 9 (1999) 70–88.
- [18] M. Fredericson, K. Kent, Normalization of bone density in a previously amenorrheic runner with osteoporosis, *Med. Sci. Sports Exerc.* 37 (2005) 1481–1486.
- [19] C.L. Zanker, C.B. Cooke, J.G. Truscott, B. Oldroyd, H.S. Jacobs, Annual changes of bone density over 12 years in an amenorrheic athlete, *Med. Sci. Sports Exerc.* 36 (2004) 137–142.
- [20] J.C. Arends, M.Y. Cheung, M.T. Barrack, A. Nattiv, Restoration of menses with nonpharmacologic therapy in college athletes with menstrual disturbances: a 5-year retrospective study, *Int. J. Sport Nutr. Exerc. Metabol.* 22 (2012) 98–108.
- [21] C.P. Guebels, L.C. Kam, G.F. Maddalozzo, M.M. Manore, Active women before/after an intervention designed to restore menstrual function: resting metabolic rate and comparison of four methods to quantify energy expenditure and energy availability, *Int. J. Sport Nutr. Exerc. Metabol.* 24 (2014) 37–46.
- [22] L. Cialdella-Kam, C.P. Guebels, G.F. Maddalozzo, M.M. Manore, Dietary intervention restored menses in female athletes with exercise-associated menstrual dysfunction with limited impact on bone and muscle health, *Nutrients* 6 (2014) 3018–3039.
- [23] A.J. Stunkard, S. Messick, The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger, *J. Psychosom. Res.* 29 (1985) 71–83.
- [24] D.M. Garner, M.P. Olmsted, *Eating Disorder Inventory Manual*, Psychological Assessment Resources, Odessa, FL, 1991.
- [25] P.J. Brantley, C.D. Waggoner, G.N. Jones, N.B. Rappaport, A Daily Stress Inventory: development, reliability, and validity, *J. Behav. Med.* 10 (1987) 61–74.
- [26] D.J. Dozois, K.S. Dobson, J.L. Ahnberg, A psychometric evaluation of the Beck depression inventory-II, *Psychol. Assess.* 10 (1998) 83–89.
- [27] S. Cohen, T. Kamarck, R. Mermelstein, A global measure of perceived stress, *J. Health Soc. Behav.* 24 (1983) 385–396.
- [28] P.C. Terry, A.M. Lane, Normative values for profiles of Mood states for use in athletic samples, *J. Appl. Sport Psychol.* 12 (2000) 93–109.
- [29] C.J. Munro, G.H. Stabenfeldt, J.R. Cragun, L.A. Addiego, J.W. Overstreet, B.L. Lasley, Relationship of serum estradiol and progesterone concentrations to the excretion profiles of their major urinary metabolites as measured by enzyme immunoassay and radioimmunoassay, *Clin. Chem.* 37 (1991) 838–844.
- [30] M. Guida, G.A. Tommaselli, S. Palomba, M. Pellicano, G. Moccia, C. Di Carlo, C. Nappi, Efficacy of methods for determining ovulation in a natural family planning program, *Fertil. Steril.* 72 (1999) 900–904.
- [31] M.J. De Souza, B.E. Miller, A.B. Loucks, A.A. Luciano, L.S. Pescatello, C.G. Campbell, B.L. Lasley, High frequency of luteal phase deficiency and anovulation in recreational women runners: blunted elevation in follicle-stimulating hormone observed during luteal-follicular transition, *J. Clin. Endocrinol. Metab.* 83 (1998) 4220–4232.
- [32] Institute of Medicine, *Dietary Reference Intakes for Calcium and Vitamin D*, The National Academy Press, Washington DC, 2011.
- [33] Institute of Medicine, Food and Nutrition Board: dietary reference intakes for calcium, phosphorus, magnesium, vitamin D, and fluoride, National Academy of Sciences SCOTSeoDRI, The National Academy Press, Washington DC, 1997.
- [34] D. Baran, A. Sorensen, J. Grimes, R. Lew, A. Karellas, B. Johnson, J. Roche, Dietary modification with dairy products for preventing vertebral bone loss in premenopausal women: a three-year prospective study, *J. Clin. Endocrinol. Metab.* 70 (1990) 264–270.
- [35] K.T. Borer, Physical activity in the prevention and amelioration of osteoporosis in women - interaction of mechanical, hormonal and dietary factors, *Sports Med.* 35 (2005) 779–830.
- [36] A.L. Friedlander, H.K. Genant, S. Sadowsky, N.N. Byl, C.C. Gluer, A 2-year program of aerobics and weight training enhances bone-mineral density of young-women, *J. Bone Miner. Res.* 10 (1995) 574–585.
- [37] R.B. Mazess, H.S. Barden, Bone-density in premenopausal women - effects of age, dietary-intake, physical-activity, smoking, and birth-control pills, *Am. J. Clin. Nutr.* 53 (1991) 132–142.
- [38] J.B. Weir, New methods for calculating metabolic rate with special reference to protein metabolism. 1949, *Nutrition* 6 (1990) 213–221.
- [39] A.H. Goris, K.R. Westerterp, Underreporting of habitual food intake is explained by underreporting in highly motivated lean women, *J. Nutr.* 129 (1999) 878–882.
- [40] F. Magkos, M. Yannakoulia, Methodology of dietary assessment in athletes: concepts and pitfalls, *Curr. Opin. Clin. Nutr. Metab. Care* 6 (2003) 539–549.
- [41] S.E. Crouter, C. Albright, D.R. Bassett Jr., Accuracy of polar S410 heart rate monitor to estimate energy cost of exercise, *Med. Sci. Sports Exerc.* 36 (2004) 1433–1439.
- [42] H.K. Hiilloskorpi, M.E. Pasanen, M.G. Fogelholm, R.M. Laukkanen, A.T. Manttari, Use of heart rate to predict energy expenditure from low to high activity levels, *Int. J. Sports Med.* 24 (2003) 332–336.
- [43] L.R. Keytel, J.H. Goedecke, T.D. Noakes, H. Hiilloskorpi, R. Laukkanen, L. van der Merwe, E.V. Lambert, Prediction of energy expenditure from heart rate monitoring during submaximal exercise, *J. Sports Sci.* 23 (2005) 289–297.
- [44] B.E. Ainsworth, W.L. Haskell, M.C. Whitt, M.L. Irwin, A.M. Swartz, S.J. Strath, W.L. O'Brien, D.R. Bassett Jr., K.H. Schmitz, P.O. Emplancourt, D.R. Jacobs Jr., A.S. Leon, Compendium of physical activities: an update of activity codes and MET intensities, *Med. Sci. Sports Exerc.* 32 (2000) S498–S504.
- [45] B.E. Ainsworth, W.L. Haskell, S.D. Herrmann, N. Meckes, D.R. Bassett Jr., C. Tudor-Locke, J.L. Greer, J. Vezina, M.C. Whitt-Glover, A.S. Leon, 2011 Compendium of Physical Activities: a second update of codes and MET values, *Med. Sci. Sports Exerc.* 43 (2011) 1575–1581.
- [46] B.E. Ainsworth, D.R. Bassett Jr., S.J. Strath, A.M. Swartz, W.L. O'Brien, R.W. Thompson, D.A. Jones, C.A. Macera, C.D. Kimsley, Comparison of three methods for measuring the time spent in physical activity, *Med. Sci. Sports Exerc.* 32 (2000) S457–S464.
- [47] R. Maddison, Y. Jiang, S.V. Hoorn, C.N. Mhurchu, C.M. Lawes, A. Rodgers, E. Rush, Estimating energy expenditure with the RT3 triaxial accelerometer, *Res. Q. Exerc. Sport* 80 (2009) 249–256.
- [48] A.V. Rowlands, P.W. Thomas, R.G. Eston, R. Topping, Validation of the RT3 triaxial accelerometer for the assessment of physical activity, *Med. Sci. Sports Exerc.* 36 (2004) 518–524.
- [49] H.J. Leidy, J.K. Gardner, B.R. Frye, M.L. Snook, M.K. Schuchert, E.L. Richard, N.I. Williams, Circulating ghrelin is sensitive to changes in body weight during a diet and exercise program in normal-weight young women, *J. Clin. Endocrinol. Metab.* 89 (2004) 2659–2664.
- [50] M.L. Pollock, J.H. Wilmore, S.M. Fox, *Exercise in Health and Disease: Evaluation and Prescription for Prevention and Rehabilitation*, W.B. Saunders, Philadelphia, PA, 1984.
- [51] P. Astrand, R.K. Conlee, *Textbook of Work Physiology*, second ed., McGraw Hill, New York, 1977.
- [52] B. Saltin, P.O. Astrand, Maximal oxygen uptake in athletes, *J. Appl. Physiol.* 23 (1967) 353–358.
- [53] J.S. Kesner, D.M. Wright, S.M. Schrader, N.W. Chin, E.F. Krieg Jr., Methods of monitoring menstrual function in field studies: efficacy of methods, *Reprod. Toxicol.* 6 (1992) 385–400.
- [54] N. Santoro, S.L. Crawford, J.E. Allsworth, E.B. Gold, G.A. Greendale, S. Korenman, B.L. Lasley, D. McConnell, P. McGaffigan, R. Midgely, M. Schocken, M. Sowers, G. Weiss, Assessing menstrual cycles with urinary hormone assays, *Am. J. Physiol. Endocrinol. Metab.* 284 (2003) E521–E530.
- [55] A. Rickenlund, K. Carlstrom, B. Ekblom, T.B. Brismar, B. von Schoultz, A.L. Hirschberg, Hyperandrogenicity is an alternative mechanism underlying oligomenorrhea or amenorrhea in female athletes and may improve physical performance, *Fertil. Steril.* 79 (2003) 947–955.
- [56] A. Rickenlund, M. Thoren, K. Carlstrom, B. von Schoultz, A.L. Hirschberg, Diurnal profiles of testosterone and pituitary hormones suggest different mechanisms for menstrual disturbances in endurance athletes, *J. Clin. Endocrinol. Metab.* 89 (2004) 702–707.
- [57] M. Hagmar, B. Berglund, K. Brismar, A.L. Hirschberg, Hyperandrogenism may explain reproductive dysfunction in olympic athletes, *Med. Sci. Sports Exerc.* 41 (2009) 1241–1248.
- [58] T. Yoshikawa, C.H. Turner, M. Peacock, C.W. Slemenda, C.M. Weaver, D. Teegarden, P. Markwardt, D.B. Burr, Geometric structure of the femoral neck measured using dual-energy x-ray absorptiometry, *J. Bone Miner. Res.* 9 (1994) 1053–1064.
- [59] V.G. Sinclair, K.A. Wallston, The development and psychometric evaluation of the Brief resilient coping scale, *Assessment* 11 (2004) 94–101.
- [60] D.M. McNair, M. Lorr, L.F. Droppleman, *Manual for the Profile of Mood States*, Educational and Industrial Testing Services, San Diego, CA, 1971.
- [61] A.T. Beck, C.H. Ward, M. Mendelson, J. Mock, J. Erbaugh, An inventory for

- measuring depression, *Arch. Gen. Psychiatr.* 4 (1961) 561–571.
- [62] A. Weissman, The dysfunctional attitude scale: a validation study, Publicly Accessible Penn Dissertations, 1979.
- [63] A. Weissman, A. Beck, Development and validation of the dysfunctional attitude scale: a preliminary investigation, Paper Presented at the Annual Meeting of the American Educational Research Association. Toronto, Ontario, 1978.
- [64] R.C. Miller, E. Brindle, D.J. Holman, J. Shofer, N.A. Klein, M.R. Soules, K.A. O'Connor, Comparison of specific gravity and creatinine for normalizing urinary reproductive hormone concentrations, *Clin. Chem.* 50 (2004) 924–932.
- [65] M.J. De Souza, B.E. Miller, A.B. Loucks, A.A. Luciano, L.S. Pescatello, C.G. Campbell, B.L. Lasley, High frequency of luteal phase deficiency and anovulation in recreational women runners: blunted elevation in follicle-stimulating hormone observed during luteal-follicular transition, *J. Clin. Endocrinol. Metab.* 83 (1998) 4220–4232.
- [66] B.L. Drinkwater, K. Nilson, S. Ott, C.H. Chesnut 3rd, Bone mineral density after resumption of menses in amenorrheic athletes, *J. Am. Med. Assoc.* 256 (1986) 380–382.
- [67] J.H. Gibson, A. Mitchell, J. Reeve, M.G. Harries, Treatment of reduced bone mineral density in athletic amenorrhea: a pilot study, *Osteoporos. Int.* 10 (1999) 284–289.
- [68] M. Sowers, M. Crutchfield, R. Bandekar, J.F. Randolph, B. Shapiro, M.A. Schork, M. Jannausch, Bone mineral density and its change in pre-and perimenopausal white women: the Michigan Bone Health Study, *J. Bone Miner. Res.* 13 (1998) 1134–1140.
- [69] A. Caillot-Augusseau, M.H. Lafage-Proust, P. Margailan, N. Vergely, S. Faure, S. Paillet, F. Lang, C. Alexandre, B. Estour, Weight gain reverses bone turnover and restores circadian variation of bone resorption in anorexic patients, *Clin. Endocrinol.* 52 (2000) 113–121.
- [70] C.L. Zanker, C.B. Cooke, Energy balance, bone turnover, and skeletal health in physically active individuals, *Med. Sci. Sports Exerc.* 36 (2004) 1372–1381.
- [71] M.J. De Souza, A. Nattiv, E. Joy, M. Misra, N.I. Williams, R.J. Mallinson, J.C. Gibbs, M. Olmsted, M. Goolsby, G. Matheson, Female athlete triad C, American college of sports M, American medical society for sports M, American bone health A, 2014 Female Athlete Triad Coalition Consensus Statement on Treatment and Return to Play of the Female Athlete Triad: 1st International Conference Held in San Francisco, CA, May 2012, and 2nd International Conference Held in Indianapolis, IN, May 2013. *Clin J Sport Med*, vol. 24, 2014, pp. 96–119.
- [72] K.J. Pantano, Strategies used by physical therapists in the U.S. for treatment and prevention of the female athlete triad, *Phys. Ther. Sport* 10 (2009) 3–11.
- [73] B. Eden Engstrom, P. Burman, A.G. Johansson, L. Wide, F.A. Karlsson, Effects of short-term administration of growth hormone in healthy young men, women, and women taking oral contraceptives, *J. Intern. Med.* 247 (2000) 570–578.
- [74] J.D. Vescovi, J.L. VanHeest, M.J. De Souza, Short-term response of bone turnover to low-dose oral contraceptives in exercising women with hypothalamic amenorrhea, *Contraception* 77 (2008) 97–104.
- [75] M. Hansen, B.F. Miller, L. Holm, S. Doessing, S.G. Petersen, D. Skovgaard, J. Frystyk, A. Plyvbjerg, S. Koskinen, J. Pingel, M. Kjaer, H. Langberg, Effect of administration of oral contraceptives in vivo on collagen synthesis in tendon and muscle connective tissue in young women, *J. Appl. Physiol.* 106 (2009) 1435–1443.