Measuring the Exercise Component of Energy Availability during Arduous Training in Women

Robert M. Gifford[1,2]*, Julie P. Greeves[3,4], Sophie L. Wardle[3], Thomas J. O'Leary[3], Rebecca L. Double[3,4], Michelle Venables[5], Christopher Boos[6], Joss Langford[7,8], David R. Woods[2,6,9]*, Rebecca M. Reynolds[1]*

[1]University/British Heart Foundation Centre for Cardiovascular Science, Queen's Medical Research Institute, University of Edinburgh, Edinburgh, United Kingdom; [2]Research and Clinical Innovation, Royal Centre of Defence Medicine, Birmingham, United Kingdom; [3]Army Health and Physical Performance Research, Andover, United Kingdom; [4]Norwich Medical School, University of East Anglia, United Kingdom; [5]Medical Research Council Elsie Widdowson Laboratory, Cambridge, United Kingdom; [6]Research Institute for Sport, Physical Activity and Leisure, Leeds Beckett University, Leeds, United Kingdom; [7]College of Life and Environmental Sciences, University of Exeter, United Kingdom; [8]Activinsights Ltd, Cambridgeshire, United Kingdom; [9]Northumbria and Newcastle NHS Trusts, Wansbeck General and Royal Victoria Infirmary, Newcastle, United Kingdom

*These authors contributed equally as senior author

Accepted for Publication: 23 September 2020

Abstract

Introduction Low energy availability (EA) may impede adaptation to exercise, suppressing reproductive function and bone turnover. Exercise energy expenditure (EEE) measurements lack definition and consistency. This study aimed to compare EA measured from moderate and vigorous physical activity from accelerometry (EEE_{mpva}) with EA from total physical activity (EEE_{tpa}) from doubly-labelled water in women. The secondary aim was to determine the relationship of EA with physical fitness, body composition by DXA, heartrate variability (HRV) and eating behavior (brief eating disorder in athletes-questionnaire, BEDA-Q). **Methods** Prospective, repeated measures study, assessing EA measures and training adaptation during 11-month basic military training. 47 women (23.9 ±2.6 years) completed 3 consecutive 10-d assessments of EEE_{mvpa}, EEE_{tpa} and energy intake (EI). EA measures were compared using linear regression and Bland-Altman analyses; relationships of EA with fat mass, heartrate variability, 1.5-mile run times and BEDA-Q were evaluated using partial correlations. **Results** EA from EEE_{mvpa} demonstrated strong agreement with EA from EEE_{tpa} across the measurement range (R^2=0.76, r=0.87, p<0.001) and was higher by 10 kcal/kg FFM/d. However, EA was low in absolute terms due to underreported EI. Higher EA was associated with improved 1.5 mile run time (r=0.28, p<0.001) fat mass loss (r=0.38, p<0.001) and lower BEDA-Q score (r=–0.37, p<0.001) but not HRV (all p>0.10). **Conclusion** Accelerometry-based EEE demonstrated validity against DLW during multi-stressor training, the difference representing 10 kcal/kg FFM/d EEE from non-exercise activity. Beneficial physical but not autonomic adaptations were associated with higher EA. EA_{mvpa} and BEDA-Q warrant consideration for low EA assessment and screening. **Key words:** Relative energy deficiency, Exercise energy expenditure, Wearable technology, physical adaptation, Women, Female athlete triad, heart-rate variability

Introduction

Low energy availability (EA) (insufficient dietary intake (EI) in relation to exercise energy expenditure (EEE)) has important physiological and performance ramifications for male and female athletes (1). Measurement of EA is challenging and varies between studies. Measuring EI is notoriously challenging, being hampered by systematic under-reporting (2), while EEE is defined and measured inconsistently (3). Some have extrapolated the EEE of total physical activity (EEE_{tpa},) from calorimetry or doubly-labelled water (DLW) (4-6), while work demonstrating the importance of low EA (7), and subsequent studies (8-10), measured EEE directly from moderate or vigorous physical activity (EEE_{mvpa}). In a prospective study of 35 women, Lieberman et al. demonstrated low EA measured from purposeful EEE_{mvpa} was linearly related to ovulatory dysfunction (11). Thus EEE_{mvpa} is advantageous in that it focusses on specific activity addressed by the relative energy deficiency in sport (RED-S) or female athlete triad paradigms; however, measurement is often hampered by reporting bias and direct measurement is normally impossible in the field. As the importance of low EA becomes increasingly apparent, there is a pressing need to develop reliable and feasible methods for real-world measurement (3).

The context of basic military training is germane for developing EA measurement, since it is characterized with high physical demands and multiple stressors, in free-living but well circumscribed field environment. It is repeatable (12) and provides routine measures of training adaptation (1). Furthermore, since a ban on women joining the infantry has recently been lifted, women may be required to train more arduously which could put them at increased risk of conditions associated with low EA, like premature osteoporosis, increased cardiovascular risk or reproductive dysfunction (1, 13, 14).

Low EA may impair cardiovascular adaptation to exercise (1), regulated by the parasympathetic and sympathetic nervous systems (PNS and SNS). Increased resting PNS activity, considered a beneficial effect of exercise (15), is manifested by higher heart rate variability (HRV). Conversely, when SNS activity predominates, lower HRV is found and may accompany overtraining, psychological stress and restricted sleep (16, 17).

This study aimed to compare EA measured from EEE_{mvpa}, measured using an open-source accelerometry technique, with EA based on EEE_{tpa} using DLW, in women during an 11-month basic military training programme. The secondary aim was to determine the relationship of EA_{mvpa} and EA_{tpa} with putative benefits of training, namely physical adaptations (improved 1.5 mile run time and body composition changes) and autonomic adaptation (increased resting PNS activity), together with evidence of disordered eating behavior. These were assessed, respectively, by fitness test scores and body composition, HRV, and the brief eating disorder in athletes questionnaire (BEDA-Q). We hypothesized that EEE_{mvpa} as measured using accelerometry would correlate significantly with EEE_{tpa}, as measured using DLW, and EA calculated from EEE_{mvpa} would lead to a relative over-estimation of EA. We hypothesized EA_{mvpa} and EA_{tpa} would be associated with concordant changes in physical and autonomic adaptation, and eating behavior.

Methods

Participants and setting

Women commencing the British Army Officer Commissioning Course at the Royal Military Academy, Sandhurst, (the Course) were invited to participate. Participants underwent a routine detailed medical screen which included a full history, physical examination and an electrocardiograph (ECG) before participation, to meet exacting medical standards mandated prior to employment in the Army (18). The entry medical included review of pre-existing medical records for a multiplicity of conditions, prior to enrolment in the Army, including diagnosed thyrotoxicosis, eating disorder, malabsorption or food intolerance. The study was approved by the UK Ministry of Defence Research Ethics Committee (790/MoDREC/16) and was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent.

The study followed a repeated measures design summarized in Figure 1. Height was measured at visit 1 (Seca stadiometer model 217, Birmingham, UK) and weight at every study visit (Seca scales model 874), wearing T-shirt and combat trousers or shorts. Eating behavior and body composition were measured at the beginning and end of each term. Body composition was measured by dual-energy X-ray absorptiometry (DXA; GE Lunar iDXA, GE Healthcare Systems, Chalfont St Giles, UK) at study visits 1 (week 1), 2 (week 14), 4 (week 29) and 6 (week 43), wearing t-shirt and shorts. A self-constructed physical activity and diet questionnaire was completed at visit 1 with reference to the preceding 6 months, comprising 16 questions on exercise and diet (see Table, Supplemental Digital Content 1, Exercise and diet at the commencement of the study, http://links.lww.com/MSS/C158).

Energy availability assessment phases (EAPs)

Once per term, TEE, EEE and EI were measured over a 10-day EA assessment phases (EAPs, denoted 1 to 3). EAPs were selected in consort with training staff to be representative of the entire Course. Physical activity was a prominent feature of the Course and working days lasted mean 13.0 (SD ±4.0) hours per day (see Document, Supplemental Digital Content 2, Description of the Commissioning Course, http://links.lww.com/MSS/C159).

Exercise energy expenditure and EA calculation

We measured EEE_{mvpa} using wrist-worn GENEActiv Original tri-axial accelerometer (Activinsights, Cambridgeshire, UK). The device was worn for 24 hours per day throughout each EAP, sampling at 75 Hz. Data were processed with the GENEAread R package from CRAN (19) using a customized, openly available script (20). The data were calibrated and days with more than 7 hours non-wear were excluded: 1322 days (93%) of valid wear were included (21, 22). The mean absolute gravity-subtracted acceleration was calculated for each 1-minute epoch within a 24 hour period, per participant per EAP. Acceleration accumulated in sedentary activities was separated from acceleration accumulated in MVPA using a cut-point of 0.09g (23). Moderate and vigorous activities during each EAP were expressed as metabolic equivalents (METs), based on programmed activities (24). Exercise energy expenditure was calculated from the accumulated duration spent undertaking moderate and vigorous activity for each EAP as follows:

EEE

where t_{mvpa} is mean daily duration (minutes) of moderate and vigorous physical activity, MET is the mean daily metabolic equivalent of activity, 3.5 is the assumed oxygen cost for one MET (ml/kg/min), 0.0049 is the calorific value (kcal) of 1ml oxygen and weight (kg) before the EAP.

During each EAP, TEE was measured using DLW. In brief, a baseline urine sample was provided, before 174 mg/kg body weight $H_2{}^{18}O$ and 70 mg/kg body weight ${}^{2}H_2O$ were ingested. Ten consecutive daily urine samples were then obtained from which isotopes were measured (see Document, Supplemental Digital Content 3, Doubly Labelled Water Method, http://links.lww.com/MSS/C160). Analytical precision was 0.3 ppm for ${}^{2}H$ and 0.5 ppm for ${}^{18}O$, and a method precision of 1.2 %. Due to the compressed nature of the Course, it wasn't feasible to conduct indirect calorimetry. Therefore, resting metabolic rate (RMR) was estimated from fat free mass (FFM) according to the equation of Cunningham *et al.* (25) and the energy expenditure of all physical activity (EEE_{tpa}) was calculated by subtracting RMR from TEE.

$$\lfloor 370 + (21.6 \times \text{FFM}) \rfloor$$

where RMR is resting metabolic rate, FFM is fat-free mass, EEE_{tpa} is exercise energy expenditure of total physical activity and TEE is total energy expenditure.

Energy intake was measured using a 24-hour food diary, aided by interview with researchers at the end of each day to prompt missed items. Meals served in the canteen were weighed in a validatory cohort (see Document, Supplemental Digital Content 4, Energy Intake Assessment, http://links.lww.com/MSS/C161). All food diary data were entered into the Nutritics database (Dublin, Ireland) by the same member of the research team (RLD) to calculate EI, before EA was calculated as follows:

where EA_{mvpa} is energy availability after moderate and vigorous physical activity, EI is energy intake, EEE_{mvpa} is exercise energy expenditure of moderate and vigorous physical activity, FFM is fat free mass, EA_{tpa} is energy availability after total physical activity and EEE_{tpa} is exercise energy expenditure of total physical activity.

Physical and autonomic training adaptation and eating behaviour

A best-effort 1.5 mile (2.4 km) run test was undertaken during the same week as, but on a different day to, study visits 1, 2, 3, 5 and 6 (Figure 1). This test is a good indicator of cardiorespiratory fitness and correlates strongly with maximal rate of oxygen uptake ($\dot{V}O_{2max}$; $r = 0.79$; 95% CI 0.73 to 0.85) (26). Heart rate variability was measured at study visits 1, 2, 4 and 6, as described previously (27). A 5-minute single-lead ECG was measured using CheckMyHeart™ devices (DailyCare Biomedical, Taiwan), according to manufacturer's instructions, yielding time domain, frequency domain and nonlinear metrics of PNS and SNS balance (see Document, Supplemental Digital Content 5, Heartrate Variability Measurement, http://links.lww.com/MSS/C162).

Eating attitudes were assessed using the Brief Eating Disorders in Athletes Questionnaire (BEDA-Q), a sensitive screen for low EA (28). The binary item 'are you dieting?' was scored at each study visit, and 'have you ever dieted?' at study visit 1 only. Questionnaires were completed on a web-based application (SmartSurvey, Tewkesbury, UK).

Statistical analysis

Data were analyzed using IBM SPSS for Macintosh (Version 24.0, Armonk, NY). Baseline characteristics of participants who completed the study were compared with those who did not using independent samples t-tests. For each EAP, TEE was compared with EI, EEE_{tpa} with EEE_{mvpa}, and pre- with post-EAP weight using paired samples t-tests; EAPs were compared using repeated measures ANOVAs. Partial correlations and linear regression were used to describe the relationship between EEE_{tpa} and EEE_{mvpa}. Systematic bias between EA_{mvpa} and EA_{tpa} was assessed by the methods of Bland & Altman (29). Where EA_{mpva} data were missing due to technical issues (loss of, or failure to deploy, accelerometers) EA_{mpva} was imputed from EA_{tpa} via the regression equation for that EAP (41 EAP exposures, 24%) prior to determining the relationship with adaptation.

Repeated measures ANOVAs were used to compare weight, fat mass, fat-free mass, 1.5 mile run time, HRV indices and BEDA-Q scores across visits. Spearman's rank correlation (r_s) was used to assess dieting status over time. Physical adaptations were taken as the difference between post- and pre-EAP measurements: loss of FM, gain of FFM and improvement of 1.5 mile run time. Partial correlations were used to assess associations between EA measures and physical adaptations, HRV (measured at the study visit after each EAP) and BEDA-Q score (measured at the study visit before each EAP). Point-biserial correlation (r_{pb}) assessed relationships between EA measures and dieting status (whether dieting or not). Independent samples t-tests were used to compare EA measures between participants who reported ever having dieting was participants who did not. Alpha was set at $p<0.05$, except for multiple correlations of training adaptations with EA, where Bonferroni adjustment was made (HRV was treated as one adaptation, adjusted alpha $p<0.0083$).

Results

Participants

Recruitment and loss to follow up are illustrated in Figure 2. Fifty-nine women attended study visit 1 and 47 women completed three EAPs (aged 23.9 (SD ±2.6) years, mean baseline BMI 23.3 (SD ±2.1) kg/m^2). Age, height, BMI and body composition did not differ between participants who the completed study and those who withdrew, although more women who completed the study reported ever dieting than those who withdrew, as described in the table in Supplemental Digital Content 6 (see Table, Supplemental Digital Content 6, Baseline evaluation of participants at baseline who completed all study measures with those who did not, http://links.lww.com/MSS/C163).

Participants reported exercising more than most women their own age before the study, particularly running and weight training. Thirty-five participants (51%) reported skipping meals beforehand, but only 2 (3%) reported doing so after starting the Course. Six participants (10%) were vegetarian (4 lacto-ovo vegetarians and 2 vegan). Detailed diet and exercise findings are shown in the table in Supplemental Digital Content 1 (see Table, Supplemental Digital Content 1, Exercise and diet at the commencement of the study, http://links.lww.com/MSS/C158).

Energy availability assessment phases (EAPs)

As expected, EEE_{tpa} was higher than EEE_{mvpa} during EAPs 1, 2 and 3 (mean difference 452 (SD ±358) kcal/d, 504 (SD ±544) kcal/d and 506 (SD ±413) kcal/d, respectively, all $p < 0.001$) (Table 1). TEE was higher than EI during EAPs 1, 2 and 3 (energy balance –654 (SD ±558) kcal/d, –1573 (SD ±578) kcal/d, and –673 (SD 663) kcal/d, respectively, all $p<0.001$).

Comparisons of EA from accelerometry with EA from doubly-labelled water

Partial correlations between EA_{mvpa} and EA_{tpa} were strong at all EAPs (Figure 3 A and B); the linear regression equations are shown Supplemental Digital Content 7 (see Table, Supplemental Digital Content 7, Linear regression equations of energy availability measured by accelerometry with energy availability measured by doubly labelled water, http://links.lww.com/MSS/C164). Across the range of measurement, EEE_{tpa} was higher than EEE_{mvpa} by 10.2 (SD ±8.3) kcal/kg FFM/d therefore EA was lower when using EEE_{tpa} (Figure 3 C).

Training adaptation and eating behaviour

Participants gained weight during EAP 1, lost a similar amount of weight during EAP 2, and EAP 3 was weight neutral (Table 1). Overall, term 1 was weight neutral, but participants demonstrated modest gains in FFM and loss in FM. These beneficial changes were then reversed: weight and FM were higher than study baseline by the end of term 2 with no change in FFM. During term 3, weight and body composition regressed to baseline levels. Throughout the study, HRV metrics demonstrated beneficial adaptations, in particular from time domain, PNS and SNS indices and sample entropy. Table 2 shows physical, autonomic and eating behaviour scores. A narrative detailing body composition changes during training can be found in

Supplemental Digital Content 8 (see Document, Supplemental Digital Content 8, body composition changes, http://links.lww.com/MSS/C165).

Correlations of EA with training adaptations and eating behaviour

As demonstrated in Table 3, physical training adaptations correlated weakly with one another. FM loss and improvement in 1.5 mile run time correlated inversely with BEDA-Q score. There was no association between BEDA-Q scores or physical adaptations and autonomic adaptation. Increasing EA was associated with increased 1.5 mile run performance and fat mass loss. There was no correlation between EA and FFM gain or autonomic adaptation. As expected, EA demonstrated a modest negative association with BEDA-Q score and dieting status. Average EA_{tpa} and EA_{mvpa} were lower among participants who reported ever dieting, compared with those did not (-0.18 (SD ±11.7) kcal/kg FFM/d versus 5.5 (SD ±16.1) kcal/kg FFM/d, p=0.016; and 22.0 (SD ±13.0) kcal/kg FFM/d versus 29.0 (SD ±15.6) kcal/kg FFM/d, p=0.004, respectively).

Discussion

Energy availability based on estimation of moderate and vigorous physical activity using accelerometry demonstrated a strong agreement with EA based on total physical activity from the gold-standard technique across the measurement range in this setting of multi-stressor military training. EA_{mvpa} was higher than EA_{tpa} by 10 kcal/kg FFM/d, which likely represents the energy expenditure difference between ‘total physical activity’ and ‘moderate and vigorous exercise’, i.e. non-exercise activity.

In this study, all EA values were ostensibly well below the purported threshold of low EA for which has been previously been mooted and then refuted (7, 11). We found increased EA was associated with improved run times and body composition during the entire military training. TEE and EEE were commensurate with reports in athletes (2). On average, EI was 26 (SD ±8) % (958 (SD ±732) kcal/d) lower than TEE. Based on an average tissue density of 7,000 kcal/kg for adult women (30) such an energy deficit would be associated with an average weight loss of 1.06 kg per EAP. Instead, we observed no significant weight change during EAPs on average, implying crude energy balance. Applying the 95% confidence intervals of EI:EE ratio plausibility, derived from the Goldberg cutoff (0.82 to 1.18 (31)), 73% of our measurements fell below this level (mean ratio 0.73 (SD 0.19).We therefore surmise EI was underestimated, possibly due to participant motivation, fatigue during prolonged measurement durations, competing pressures from the Course itself or a combination of these. Such underreporting has been identified widely elsewhere (2, 32) including in similar settings of military training (33). Thus, while we have demonstrated the validity of the EEE component of an EA assessment, the usual limitations to EI assessment apply in this population.

Both EA_{mvpa} and EA_{tpa} were positively associated with loss in FM and improved 1.5 mile run time. These findings are particularly noteworthy for the very low measured EA at which they occurred, underlining the linearity of EA's effects on physical adaptation and performance, rather than a threshold below which its effects are seen (11). The adaptations we observed could be interpreted as relating to 'energy compensation' following lower EA experienced during EAPs, i.e. compensatory increases in EI and reduction in non-exercise activity 5 to 9 weeks between the EAP and the following study visit (34). Military training involves rapid changes in the volume and nature of physical activity, and although our participants reported active lifestyles beforehand, they found the training intensity highly challenging, as reported in a linked manuscript on stress responses in these participants (35). Changing feeding habits takes several years (36), so adjustment of habitual EI following the abrupt and immersive onset of initial military training was likely to have been delayed (37). Studies of controlled exercise protocols have found interventions were followed by compensatory EI increases (38), increased FM and reduced FFM (39). A synthesis of two randomized, controlled trials of exercise interventions for obesity showed such energy compensation was inversely associated with peak oxygen uptake ($\dot{V}O_{2peak}$)(40). Military training is not an exercise intervention *per se,* but our study suggests similar energy compensation took place in response to a multi-stressor environment, which could be relevant for women undertaking a wide range of physically demanding employment. This context makes it more remarkable that performance improvements were observed overall.

We found positive autonomic adaptations throughout the study, but these were not correlated with EA. Increased parasympathetic and decreased sympathetic activity were observed in time domain measures consistently throughout the Course, especially during term 1. Time and frequency domain measures were slightly below means for athletes reported elsewhere (41). Heart rate variability has been measured in studies of psychological stress as well as exercise and may decrease following negative effects of psychological stress (17) and increase with improved aerobic capacity (42). Although psychological stress was experienced throughout the course (35), we demonstrated autonomic benefits, independent of improvement in cardiovascular fitness or EA. In a study of six military women undertaking an arduous Antarctic crossing, HRV demonstrated a latent increase in non-linear, time frequency domains 2 weeks after the expedition (27), suggesting that beneficial autonomic adaptation of exercise occurred independently to the marked energy deficit seen during the expedition (43).

The BEDA-Q has received increasing recognition (1, 28). We found BEDA-Q scores were associated inversely with EA, FM loss and 1.5-mile improvement, participants reporting ever trying to lose weight demonstrated slightly reduced EA. Our findings support this tool's potential for ongoing use in stressful real-world settings.

Multiple stressors induced by the Course provided a challenging context for the measurement of EA. Assessing EA is often harder in real-world than experimental settings, owing to competing interests of data collection with other priorities (37). A key strength of our study was therefore to demonstrate the potential of a novel real-world measure of EEE_{mvpa}, a relevant metric to the EA paradigm in many occupations. Other strengths include its comparatively long duration, use of a gold-standard referent of TEE and measurement of concurrent training adaptations. This tool would warrant further validation alongside concurrent DLW or indirect calorimetry for use in other contexts.

Our study has several limitations. Accelerometry tends to underestimate TEE and EEE in free-living environments and may yield mean negative bias of 8% TEE compared with gold standard techniques, with significant inter-individual variation (44). Importantly for our population, currently available accelerometry platforms do not capture the energy cost of load carriage, which would be likely to increase estimates of EA. The wrist site is associated with less negative bias than the hip, although both demonstrate significant inter-individual variation (44, 45). On the other hand, our approach of calculating EA_{mvpa} from movement above the nonpurposeful activity cutoff (0.09g), differed from that of Loucks and Thuma (7), who added background TEE (from accelerometry) to EEE (from indirect calorimetry) to calculate EEE_{mvpa}. This may have led to modest relative underestimation of EA_{mvpa} during the prolonged bouts of activity we observed. However, overall, EA was grossly underestimated likely due to underreporting of EI. We sought to overcome this limitation by the use of weighed food analysis in a validatory cohort (46). Underestimation of EI is endemic in self-reports, applies to traditional diaries as much as mobile technology and varies widely between individuals (32). We were unable to carry out measurement of VO_{2peak}, VO_{2max} or indirect calorimetry due to constraints imposed by the Course timetable.

We conclude the simple accelerometry-based measure of moderate and vigorous physical activity may be recommended for women undertaking complex, multi-stressor training. Since purposeful exercise activity is a more useful concept for trainers and athletes than total physical activity, EEE_{mvpa} could be specified within EA definitions in future. Yet until the perpetual barrier of EI underreporting is overcome, it is difficult to rely heavily on field measures of EA. Instead, screening tools like the BEDA-Q (as in this study) and biomarkers (47) demonstrate promise. Low EA correlated with performance is clearly an important concept for men and women in sports and physical occupations; our findings underline the importance of addressing low EA to optimize performance in military training.

Acknowledgements, funding, and conflicts of interest

The authors would like to acknowledge the support of Staff at the Royal Military Academy, Sandhurst for generously accommodating the study and the Officer Cadets for enthusiastic participation. We are grateful to research assistants Miss Sally Handford, Miss Jennifer Wright, Mr Shaun Chapman, Miss Rachael Bradley, Dr Sam Saunders, Dr Susan Dewhurst, Miss Mollie Drew, Mr Fionn Sullivan, Miss Jessica Dearman, Mr Alfie Gordon, Miss Louise Corfield, Miss Amy Ranford and Miss Vicky Edwards. We acknowledge the support of the Wellcome Trust Clinical Research Facility led by Jo Singleton, Finny Patterson and Steve McSwiggan. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation, and statement that results of the present study do not constitute endorsement by ACSM.

The Study was funded by the Ministry of Defence, UK. The authors acknowledge the support of the British Heart Foundation Centre for Research Excellence Award III (RE/18/5/34216). None of the funders had any role in study design or completion.

J Langford works as a Technical Director of Activinsights Ltd, which manufactures the GENEActiv accelerometer. None of the other authors has any conflict of interest to declare.

References

1. Mountjoy M, Sundgot-Borgen JK, Burke LM et al. IOC consensus statement on relative energy deficiency in sport (RED-S): 2018 update. *British journal of sports medicine*. 2018;52(11):687-97.

2. Capling L, Beck K, Gifford J, Slater G, Flood V, O'Connor H. Validity of dietary assessment in athletes: A systematic review. *Nutrients*. 2017;9(12):1313.

3. Burke LM, Lundy B, Fahrenholtz IL, Melin AK. Pitfalls of Conducting and Interpreting Estimates of Energy Availability in Free-Living Athletes. *International journal of sport nutrition and exercise metabolism*. 2018;28(4):350-63.

4. Brown MA, Howatson G, Quin E, Redding E, Stevenson EJ. Energy intake and energy expenditure of pre-professional female contemporary dancers. *PloS one*. 2017;12(2):e0171998.

5. Silva AM, Matias CN, Santos DA et al. Compensatory Changes in Energy Balance Regulation over One Athletic Season. *Medicine and science in sports and exercise*. 2017;49(6):1229-35.

6. Levine JA. Non-exercise activity thermogenesis (NEAT). *Best practice & research. Clinical endocrinology & metabolism*. 2002;16(4):679-702.

7. Loucks AB, Thuma JR. Luteinizing hormone pulsatility is disrupted at a threshold of energy availability in regularly menstruating women. *The Journal of clinical endocrinology and metabolism*. 2003;88(1):297-311.

8. VanHeest JL, Rodgers CD, Mahoney CE, De MS. Ovarian suppression impairs sport performance in junior elite female swimmers. *Medicine and science in sports and exercise*. 2014;46(1):156-66.

9. Torstveit MK, Fahrenholtz I, Stenqvist TB, Sylta Ø, Melin A. Within-day energy deficiency and metabolic perturbation in male endurance athletes. *International journal of sport nutrition and exercise metabolism*. 2018;28(4):419-27.

10. Silva M-RG, Silva H-H, Paiva T. Sleep duration, body composition, dietary profile and eating behaviours among children and adolescents: a comparison between Portuguese acrobatic gymnasts. *European journal of pediatrics*. 2018;177(6):815-25.

11. Lieberman JL, MJ DES, Wagstaff DA, Williams NI. Menstrual Disruption with Exercise Is Not Linked to an Energy Availability Threshold. *Medicine and science in sports and exercise*. 2018;50(3):551-61.

12. Siddall AG, Powell SD, Needham-Beck SC et al. Validity of energy expenditure estimation methods during 10 days of military training. *Scandinavian journal of medicine & science in sports*. 2019;29(9):1313-21.

13. Friedl KE. Biomedical research on health and performance of military women: accomplishments of the Defense Women's Health Research Program (DWHRP). *Journal of women's health (2002)*. 2005;14(9):764-802.

14. Gifford RM, Reynolds RM, Greeves J, Anderson RA, Woods DR. Reproductive dysfunction and associated pathology in women undergoing military training. *Journal of the Royal Army Medical Corps*. 2017;163(5):301-10.

15. Bellenger CR, Fuller JT, Thomson RL, Davison K, Robertson EY, Buckley JD. Monitoring athletic training status through autonomic heart rate regulation: a systematic review and meta-analysis. *Sports medicine*. 2016;46(10):1461-86.

16. Booth CK, Probert B, Forbes-Ewan C, Coad RA. Australian army recruits in training display symptoms of overtraining. *Military medicine*. 2006;171(11):1059-64.

17. Kim HG, Cheon EJ, Bai DS, Lee YH, Koo BH. Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. *Psychiatry investigation*. 2018;15(3):235-45.

18. Ministry of Defence. Joint Service Publication 950. In. London, UK2018, pp. Lft 6-7-.

19. [Internet]. Available from: https://cran.r-project.org/.

20. Activinsights [Internet]. Available from: https://open.geneactiv.org/.

21. van Hees VT, Renstrom F, Wright A et al. Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. *PLoS One*. 2011;6(7):e22922.

22. van Hees VT, Fang Z, Langford J et al. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. *Journal of applied physiology (Bethesda, Md. : 1985)*. 2014;117(7):738-44.

23. Esliger DW, Rowlands AV, Hurst TL, Catt M, Murray P, Eston RG. Validation of the GENEA Accelerometer. *Medicine and science in sports and exercise*. 2011;43(6):1085-93.

24. Jette M, Sidney K, Blumchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. *Clin Cardiol*. 1990;13(8):555-65.

25. Cunningham JJ. Body composition as a determinant of energy expenditure: a synthetic review and a proposed general prediction equation. *The American journal of clinical nutrition*. 1991;54(6):963-9.

26. Mayorga-Vega D, Bocanegra-Parrilla R, Ornelas M, Viciana J. Criterion-Related Validity of the Distance- and Time-Based Walk/Run Field Tests for Estimating Cardiorespiratory Fitness: A Systematic Review and Meta-Analysis. *PLoS One*. 2016;11(3):e0151671.

27. Gifford RM, Boos CJ, Reynolds RM, Woods DR. Recovery time and heart rate variability following extreme endurance exercise in healthy women. *Physiological reports*. 2018;6(21):e13905-e.

28. Gordon CM, Ackerman KE, Berga SL et al. Functional Hypothalamic Amenorrhea: An Endocrine Society Clinical Practice Guideline. *The Journal of clinical endocrinology and metabolism*. 2017;102(5):1413-39.

29. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Statistical methods in medical research*. 1999;8(2):135-60.

30. Black AE, Prentice AM, Coward WA. Use of food quotients to predict respiratory quotients for the doubly-labelled water method of measuring energy expenditure. *Human nutrition. Clinical nutrition*. 1986;40(5):381-91.

31. Livingstone MB, Black AE. Markers of the validity of reported energy intake. *The Journal of nutrition*. 2003;133 Suppl 3:895s-920s.

32. Burrows TL, Ho YY, Rollo ME, Collins CE. Validity of Dietary Assessment Methods When Compared to the Method of Doubly Labeled Water: A Systematic Review in Adults. *Front Endocrinol (Lausanne)*. 2019;10:850.

33. Tharion WJ, Lieberman HR, Montain SJ et al. Energy requirements of military personnel. *Appetite*. 2005;44(1):47-65.

34. Riou ME, Jomphe-Tremblay S, Lamothe G, Stacey D, Szczotka A, Doucet E. Predictors of Energy Compensation during Exercise Interventions: A Systematic Review. *Nutrients*. 2015;7(5):3677-704.

35. Gifford RM, O'Leary TJ, Double RL et al. Positive adaptation of HPA axis function in women during 44 weeks of infantry-based military training. *Psychoneuroendocrinology*. 2019;110:104432.

36. Blundell JE, King NA. Physical activity and regulation of food intake: current evidence. *Medicine and science in sports and exercise*. 1999;31(11 Suppl):S573-83.

37. Ackerman KE, Holtzman B, Cooper KM et al. Low energy availability surrogates correlate with health and performance consequences of Relative Energy Deficiency in Sport. *British journal of sports medicine*. 2019;53(10):628-33.

38. Whybrow S, Hughes DA, Ritz P et al. The effect of an incremental increase in exercise on appetite, eating behaviour and energy balance in lean men and women feeding ad libitum. *The British journal of nutrition*. 2008;100(5):1109-15.

39. Schubert MM, Palumbo E, Seay RF, Spain KK, Clarke HE. Energy compensation after sprint- and high-intensity interval training. *PLoS One*. 2017;12(12):e0189590.

40. McNeil J, Brenner DR, Courneya KS, Friedenreich CM. Dose-response effects of aerobic exercise on energy compensation in postmenopausal women: combined results from two randomized controlled trials. *International journal of obesity (2005)*. 2017;41(8):1196-202.

41. Shaffer F, Ginsberg JP. An Overview of Heart Rate Variability Metrics and Norms. *Front Public Health*. 2017;5:258.

42. Task Force. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Eur Heart J*. 1996;17(3):354-81.

43. Gifford RM, O'Leary T, Cobb R et al. Female Reproductive, Adrenal, and Metabolic Changes during an Antarctic Traverse. *Medicine and science in sports and exercise*. 2019;51(3):556-67.

44. Murakami H, Kawakami R, Nakae S et al. Accuracy of wearable devices for estimating total energy expenditure: comparison with metabolic chamber and doubly labeled water method. *JAMA internal medicine*. 2016;176(5):702-3.

45. Crouter SE, Flynn JI, Bassett DR, Jr. Estimating physical activity in youth using a wrist accelerometer. *Medicine and science in sports and exercise*. 2015;47(5):944-51.

46. Gemming L, Utter J, Ni Mhurchu C. Image-assisted dietary assessment: a systematic review of the evidence. *Journal of the Academy of Nutrition and Dietetics*. 2015;115(1):64-77.

47. Elliott-Sale KJ, Tenforde AS, Parziale AL, Holtzman B, Ackerman KE. Endocrine Effects of Relative Energy Deficiency in Sport. *International journal of sport nutrition and exercise metabolism*. 2018;28(4):335-49.

Figure captions

Figure 1. Scheme of study visits and energy availability assessment phases (EAPs). ① recruitment, ② height,③ weight, brief eating disorder questionnaire (BEDA-Q) ④ dual-energy x-ray absorptiometry,(DXA) and heart rate variability (HRV), ⑤ 10-day energy availability assessment phase (EAP). PCBC, Pre-Course briefing course, 6 to 20 weeks before start of term 1

Figure 2. Recruitment and follow up. EAP, energy availability assessment phase, where energy requirement was measured using multi-point doubly labelled water and energy intake and exercise energy expenditure were estimated. At 'study visits', weight, HRV, and body composition were measured, and questionnaires were completed. * 2 participants declined and 2 provided insufficient urine samples. † 2 declined and 4 provided insufficient urine samples

Figure 3. Comparisons of EA_{tpa} and EA_{mvpa} A: Scatter plot of all paired EA_{tpa} against paired EA_{mvpa} values with overall linear regression equation, B: EAPs plotted separately. C: Bland-Altman Plot, demonstrating difference between EA_{mvpa} and EA_{tpa} at the range of values measured. D and E: change in weight during 10-d assessment EAPs (weight post – weight pre) plotted against EA_{tpa} and EA_{mvpa}, respectively. In panels B to E, blue circle represents EAP 1, unfilled red circles EAP 2 and green triangles EAP 3. EA_{tpa}: energy availability from total physical activity (measurement based on total energy expenditure from doubly-labelled water); EA_{mvpa}: energy availability from moderate and vigorous physical activity (measurement based on accelerometry).

Supplemental Digital Content

Supplemental Digital Content 1. Table: Exercise and diet at the commencement of the study

Supplemental Digital Content 2. Description of the Commissioning Course

Supplemental Digital Content 3. Doubly-Labelled Water Method

Supplemental Digital Content 4. Energy Intake Assessment

Supplemental Digital Content 5. Heartrate Variability Measurement

Supplemental Digital Content 6. Baseline evaluation of participants at baseline who completed all study measures with those who did not.

Supplemental Digital Content 7. Table: Linear regression equations of energy availability measured by accelerometry with energy availability measured by doubly labelled water.

Supplemental Digital Content 8. Narrative: body composition changes.

Figure 1

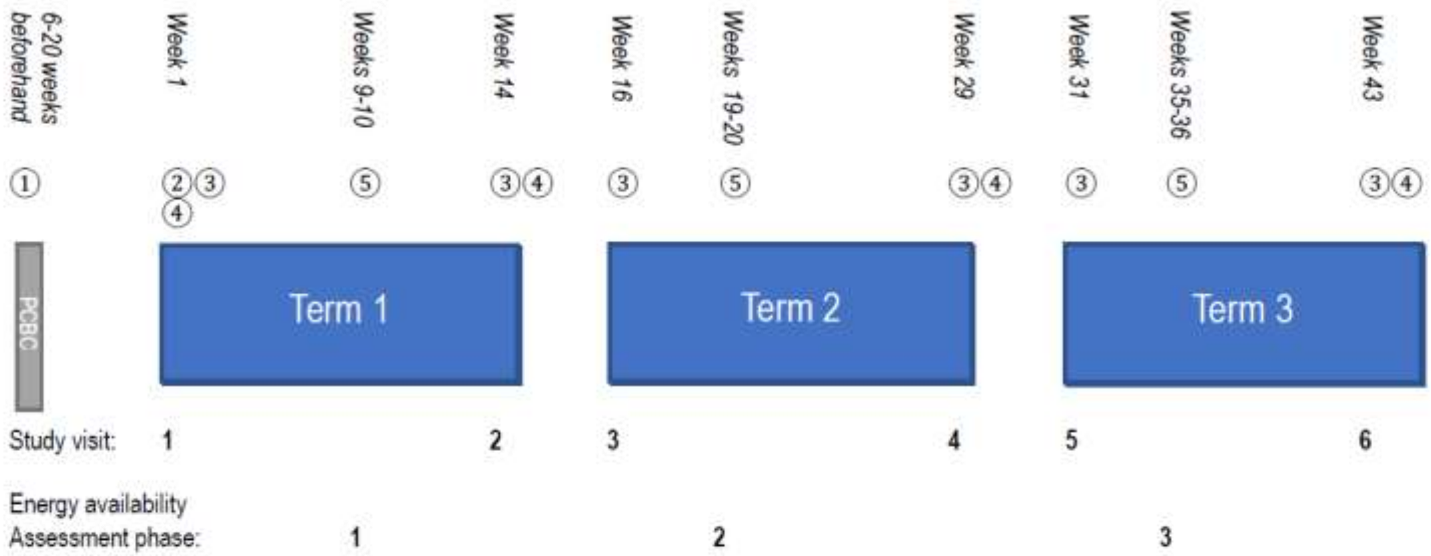

6-20 weeks beforehand
Week 1
Weeks 9-10
Week 14
Week 16
Weeks 19-20
Week 29
Week 31
Weeks 35-36
Week 43
①
②③④
⑤
③④
③
⑤
③④
③
⑤
③④
PCBC
Term 1
Term 2
Term 3
Study visit: 1 2 3 4 5 6
Energy availability Assessment phase: 1 2 3

Figure 2

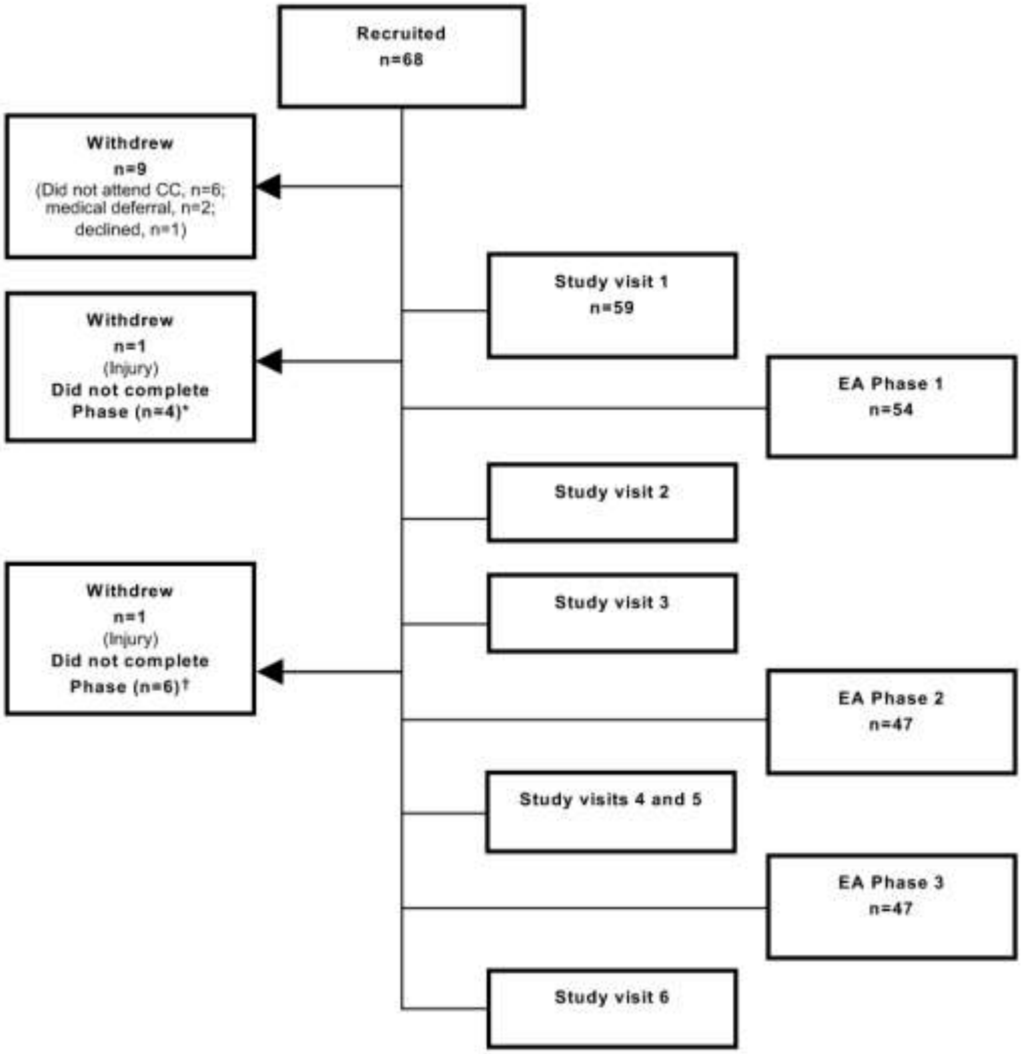
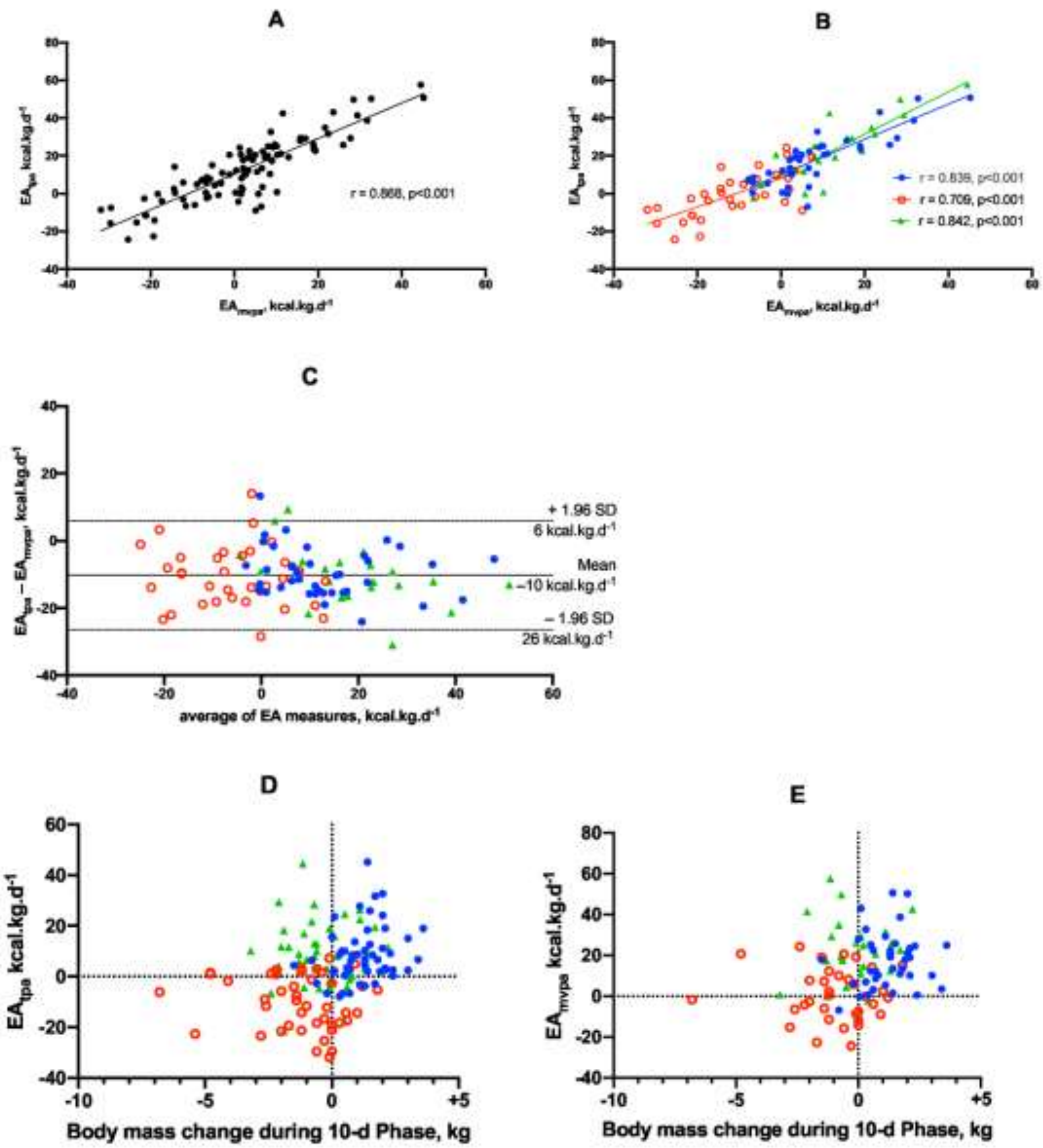

Recruited
n=68
Withdrew
n=9
(Did not attend CC, n=6; medical deferral, n=2; declined, n=1)
Study visit 1
n=59
Withdrew
n=1
(Injury)
Did not complete Phase (n=4)*
EA Phase 1
n=54
Study visit 2
Withdrew
n=1
(Injury)
Did not complete Phase (n=6)†
Study visit 3
EA Phase 2
n=47
Study visits 4 and 5
EA Phase 3
n=47
Study visit 6

Figure 3



A
r = 0.868, p<0.001
B
r = 0.839, p<0.001
r = 0.709, p<0.001
r = 0.842, p<0.001
C
+ 1.96 SD
6 kcal.kg.d-1
Mean
−10 kcal.kg.d-1
− 1.96 SD
26 kcal.kg.d-1
average of EA measures, kcal.kg.d-1
D
E
Body mass change during 10-d Phase, kg
Body mass change during 10-d Phase, kg

	Phase 1		Phase 2		Phase 3		ηp²	p
Weight pre / post, kg	64.4 (7.7)	65.6 (7.7)	65.0 (7.6)	63.7 (7.4)	65.8 (7.7)	65.3 (7.6)		
Difference, kg	+1.2 (1.1), p<0.001		–1.3 (3.0), p=0.003		–0.3 (1.3), p=0.11		0.471	<0.001
EI, kcal/d	2667 (696)		2320 (574)*		2358 (422)*		0.115	0.003
TEE, kcal/d	3332 (424)		3849 (363)*		3041 (286)*		0.321	<0.001
EEE_{tpa}, kcal/d	2228 (355)		2811 (455)*		1963 (325)* †		0.695	<0.001
EEE_{mvpa}, kcal/d	1865 (312)		2253 (536)*		1513 (336)* †		0.752	<0.001
EA_{tpa}, kcal/kg/d	8 (11)		–10 (11)*		9 (12)* †		0.511	<0.001
EA_{mvpa} kcal/kg/d	18 (13)		1 (13)*		23 (15)* †		0.419	0.001

Table 1. Energy availability measurements. Values are mean (SD). p for repeated measures ANOVA (main effect of time), ηp²: partial Eta squared, *significant (p<0.05) difference versus Phase 1. † significant vs Phase 2. TEE, total energy expenditure; EI energy intake; EEE exercise energy expenditure – either from total physical activity (tpa; measured by doubly labelled water) or from measured moderate and vigorous physical activity (mpva; measured by accelerometry); EA energy availability for each measure of EEE, expressed as kcal per kg fat–free mass per day.

	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6	ηp^2	p
Weight, kg	64.1 (7.9)	63.7 (7.9)	64.5 (7.9)	65.0 (7.7)*	65.2 (7.9)*	63.9 (7.8)	0.073	0.006
Fat free mass, kg	49.6 (5.3)	50.2 (5.4)*		49.4 (5.0)		49.4 (5.0)	0.072	0.041
Fat mass, kg	15.6 (4.0)	14.4 (3.9)**		16.1 (4.0)*		15.6 (4.2)	0.245	<0.001
1.5 mile run time, mm:ss	10:41 (1:02)	10:08 (0:55)**	10:20 (0:57)**		10:38 (0:58)	10:29 (1:00)*	0.399	<0.001
Heart rate	76.4 ±11.9	69.7 ±9.6*		71.2 ±8.9*		68.2 ±10.3**	0.122	<0.001
HRV								
Time domain								
RMSSD median (IQR)	35.7 [23.5, 56.4]	51.4 [33.6, 66.81]*		45.9 [37.6 ,55.29]*		47.2 [30.6 ,66.6]*	0.075	0.036
pNN50,%, median (IQR)	14.7 [4.9, 31.6]	28.4 [11.1, 43.8]*		24.4 [16, 35.7]*		27.6 [9.6, 44.9]*	0.079	0.003
Frequency domain (fast–Fourier transformed)								
LF (log)	6.87 ±0.88	7.09 ±0.97		6.84 ±0.77		7.00 ±1.07	0.002	0.70
HF (log)	6.25 ±1.19	6.70 ±1.17*		6.48 ±0.93		6.59 ±1.28	0.028	0.11
LF:HF median (IQR)	1.8 [1.2, 3.3]	1.4 [0.8, 3.1]		1.8 [1.1, 2.5]		1.6 [0.9, 2.5]	0.021	0.17
Non–linear								
Sample entropy	1.61 ±0.26	1.69 ±0.24		1.72 ±0.25*		1.75 ±0.26*	0.149	0.003
Autonomic nervous system indices								
PNS Index median (IQR)	–0.7 [–1.4, 0.2]	0.1 [–0.7, 0.6*]		–0.2 [–0.6, 0.1]*		0.3 [–0.7, 1.2]**	0.090	0.001
SNS Index	0.86 ±1.35	0.14 ±1.09*		0.30 ±0.89*		0.12 ±1.26**	0.083	0.002
BEDA-Q score median [IQR]	3 [1, 4]	3 [1, 5]	3 [2, 5]	4 [2, 6]	3 [1, 5]	3 [2, 5]	0.034	0.70
BEDA-Q dieting (“yes”) n (%)	11 (18.6%)	7 (13.0%)	5 (15.6%)	14 (29.2%)	16 (35.6%)	16 (30.8%)	0.139[a]	0.010[a]

Table 2. Physical, autonomic, eating behaviour changes. Data are mean (SD) unless otherwise stated. BEDA-Q Brief Eating Disorders in Athletes Questionnaire (score was log transformed prior to analysis). IQR, inter-quartile range. For continuous variables, p values refer to repeated measures ANOVA (main effect of time). ηp^2 Partial Eta squared, [a] Spearman’s correlation for BEDA-Q dieting (dichotomous) with visit week. RMSSD: root mean square of successive differences, pNN50: percentage of successive normal R-R intervals above 50 ms, IQR: inter quartile range, LF: low frequency power, HF: high frequency power, log: transformed by natural logarithm, PNS: parasympathetic nervous system, SNS: sympathetic nervous system.* pairwise difference with visit 1 (p<0.05), ** pairwise difference with visit 1 (p<0.001)

	EA_{tpa}	EA_{mvpa}	FM loss	FFM gain	1.5 mile run improvement	RMSSD (log)	pNN50 (log)	LF (log)	HF (log)	LF:HF (log)	Sample entropy	PNS Index (log)	SNS Index	BEDA-Q score
FM loss	.376**	.373**												
FFM gain	.161	.213	.222**											
1.5 mile run improvement	.284**	.252**	.482**	.082										
RMSSD (log)	.203	.208	.007	.154	.087									
pNN50 (log)	.182	.193	–.037	–.23	–.165	.838**								
LF (log)	–.118	–.097	–.018	–.005	–.227	.730**	.484**							
HF (log)	.252	.227	–.006	.201	–.115	.918**	.847**	.678**						
LF:HF (log)	–0.197	–0.182	.044	–.204	–0.067	–.271	–.472**	.276*	–.445**					
Sample entropy	.171	.204	.047	.059	–.048	.2938	.054	–.302	.286	–.391**				
PNS Index (log)	.168	.188	.068	0.028	–.042	.806**	.823**	.564**	.761**	–.348**	–.018			
SNS Index	–.163	–.184	–.035	.05	.074	–.799**	–.762**	–.672**	–.679	.413**	–.321**	–.880**		
BEDA–Q score (log)	–.367**	–.321*	–.285*	.036	–.302	–.318	–.19	–.246	–.299	.133	.029	–.157	.038	
BEDA–Q – “are you trying to lose weight?”[pb]	–.211*	–.206*	.037	–.057	.037	.002	.007	.024	–.010	.062	.024	–.088	.120	.343**

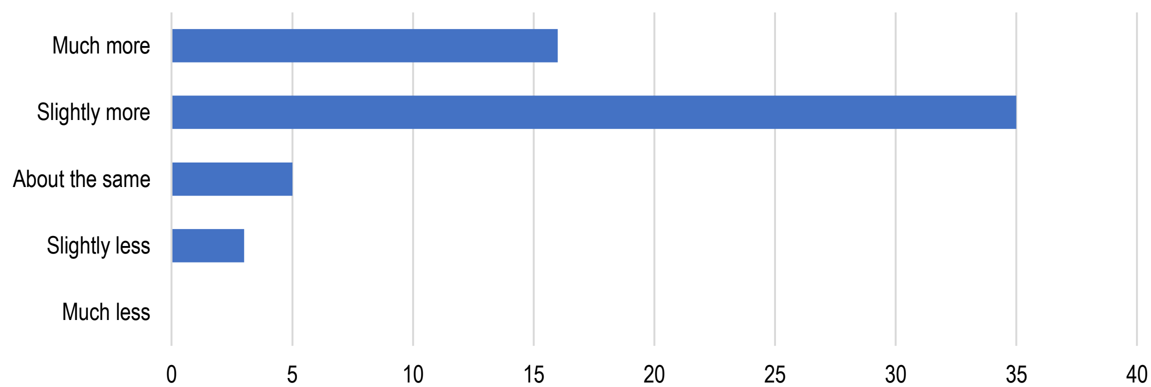
Table 3. Correlations between EA measures and training adaptation. Correlations are between EAtpa, EAmvpa measurements during Energy Assessment Phases, and Pre to post-Phase training adaptation, or between concurrent training adaptation measures. All are partial correlations (taking account of repeated measures in individuals) except marked [pb] (point biserial non-parametric correlation). Significant correlations after Bonferroni adjustment: ** p<0.0001; * p<0.008. EA: energy availability measured by $_{tpa}$ (total physical activity, from total energy expenditure) or $_{mvpa}$ (from moderate and vigorous physical activity, from accelerometry), FM and FFM loss and 1.5 mile run improvement: difference in fat mass, fat free mass and 1.5 mile best-effort run time, respectively, from pre to post-EA measurement (pre minus post). Heart-rate variability (HRV) and BEDA-Q

measured before each EA measurement. Associations of HRV variables are italicized; these are expected to correlate strongly with one other. RMSSD: root mean square of successive differences, log: transformed by natural logarithm, pNN50: percentage of successive normal R-R intervals above 50 ms, IQR: inter quartile range, LF: low frequency power, HF: high frequency power, PNS: parasympathetic nervous system, SNS: sympathetic nervous system, BEDA-Q: Brief eating disorder in athletes questionnaire.

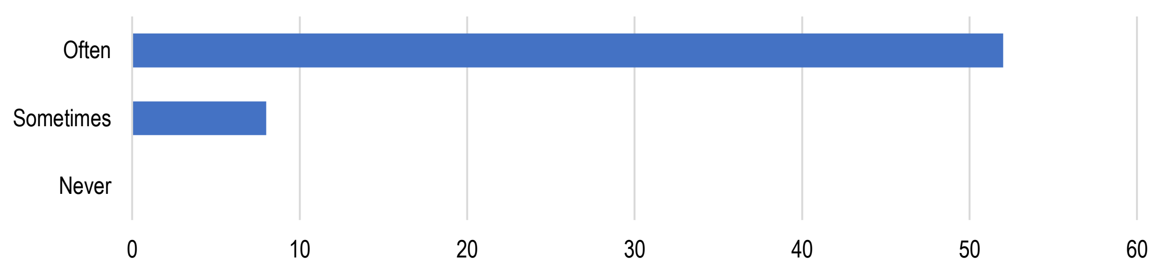
Supplemental Digital Content to Gifford et al.

Supplemental Digital Content 1. Table: Exercise and diet at the commencement of the study.

During the six months prior to starting the Commissioning Course, how active do you think you have been compared to other men or women your age?



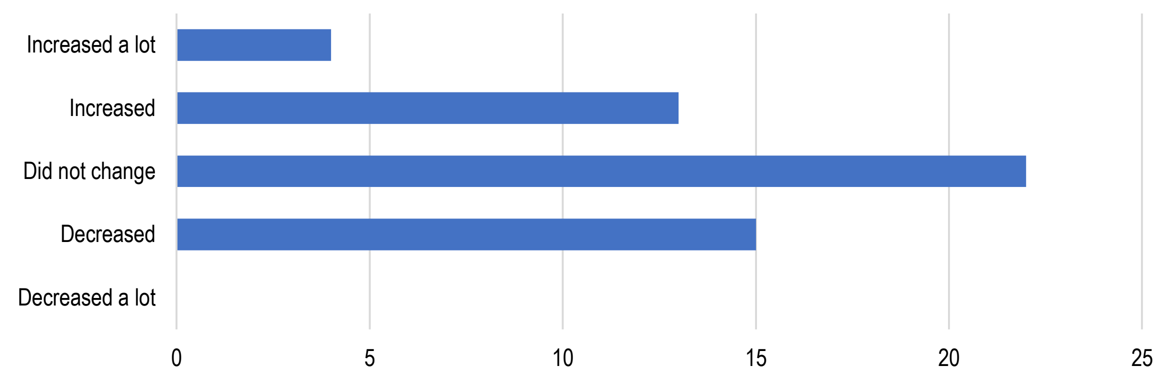
Before you arrived at Sandhurst, how frequently did you do any activity long enough to work up a sweat (heart beats quickly)?



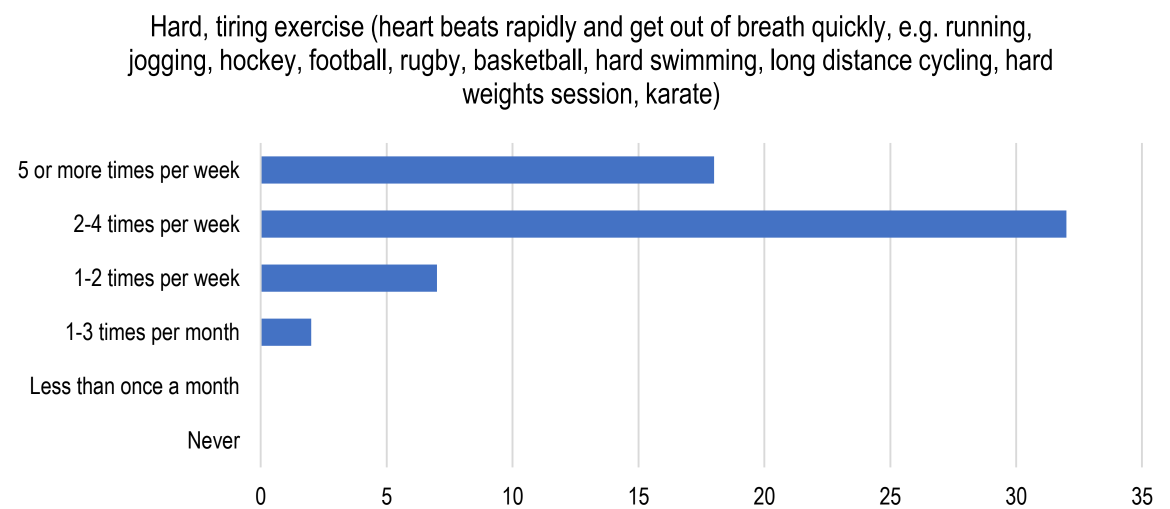
What type of exercise did you do most often before you arrived at Sandhurst?

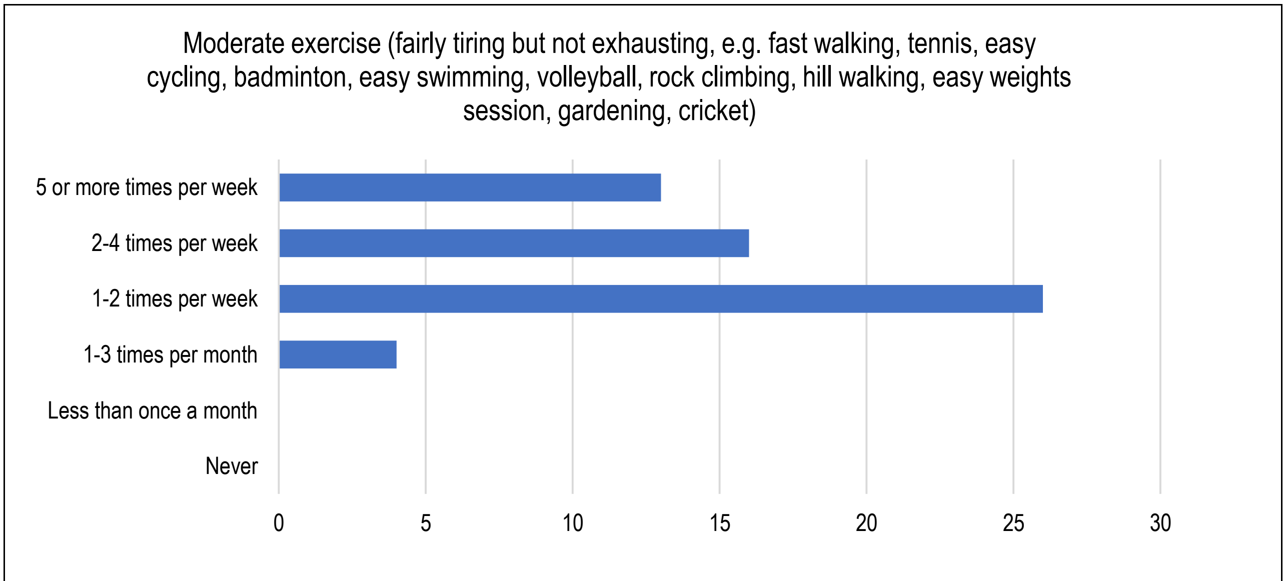
	I didn't exercise	Running	Team sport (football, hockey, rugby, etc.)	Swimming	Cycling	Weight training
	0 (0%)	49 (80%)	16 (26%)	10 (16%)	12 (20%)	38 (62%)

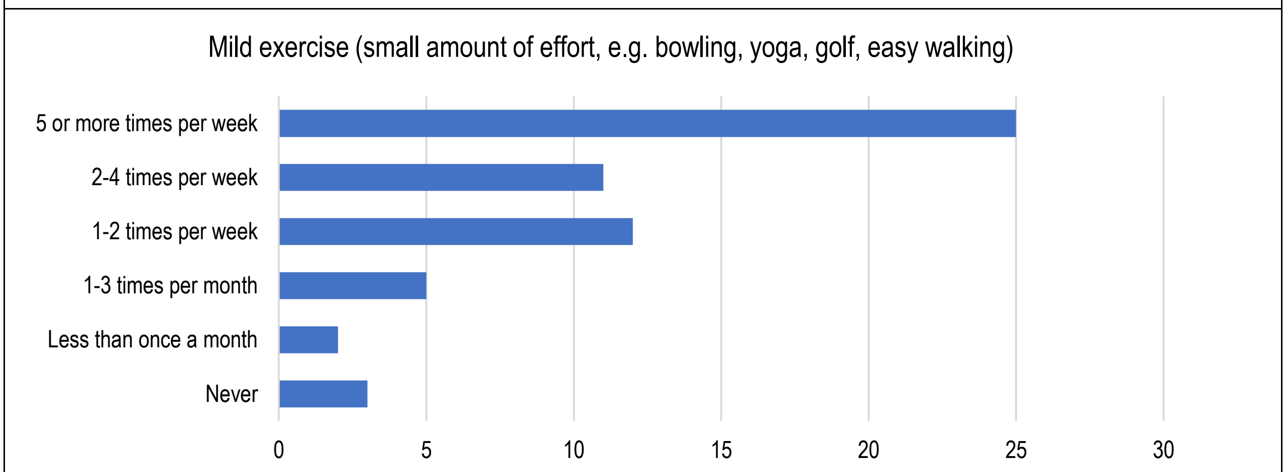
Between PCCBC and starting the Commissioning Course how much did you change the amount of exercise you did compared to the 6 months before that?



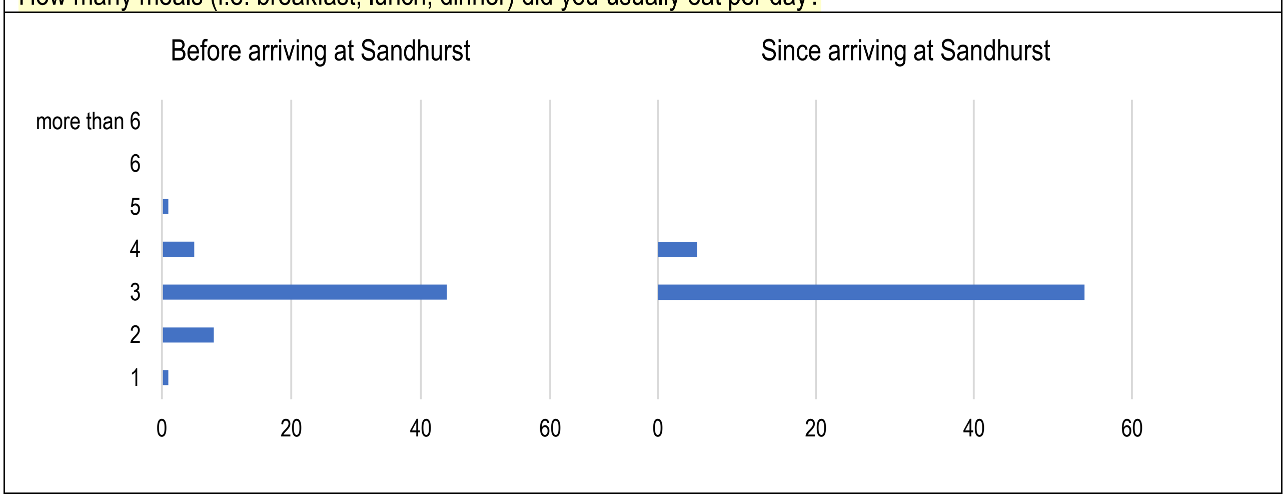
In a typical week before you arrived at Sandhurst, how many times did you do the following kinds of exercise for more than 15 minutes?



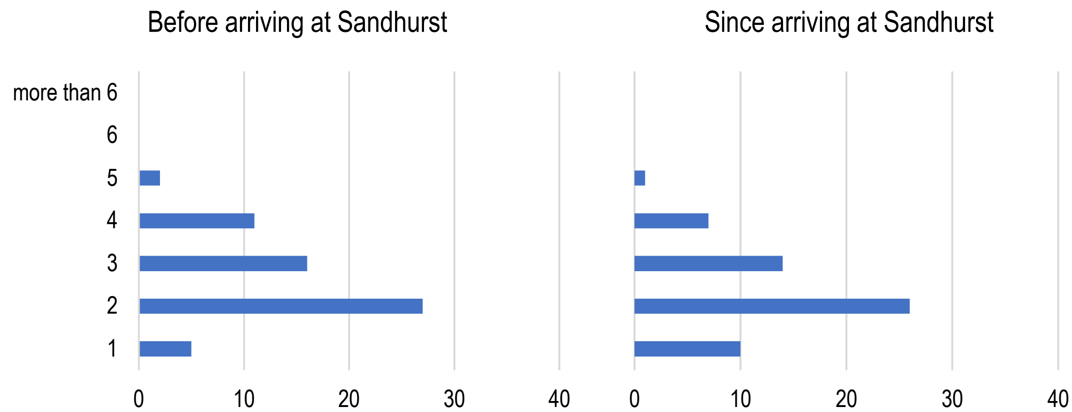




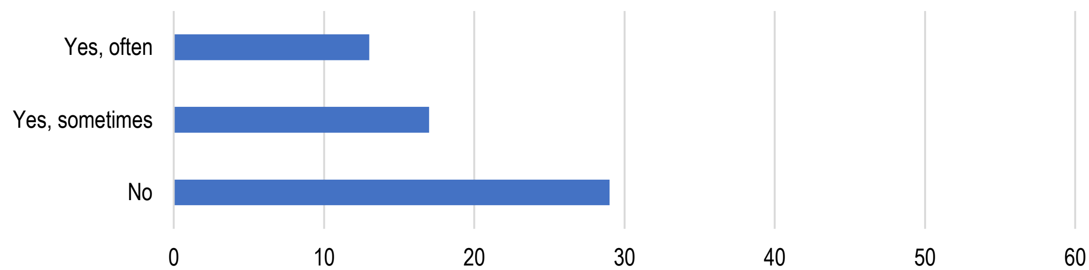
How many meals (i.e. breakfast, lunch, dinner) did you usually eat per day?



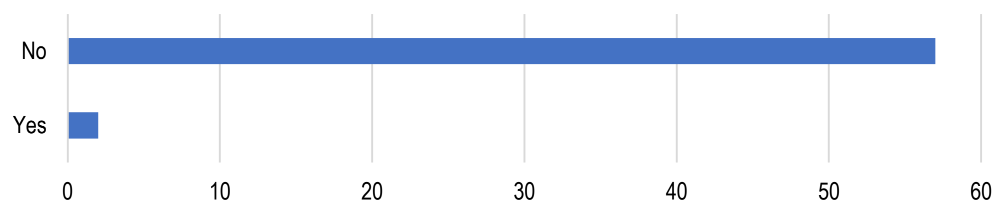
How many snacks (i.e. chocolate bar, sports bar, piece of fruit) did you usually eat per day?



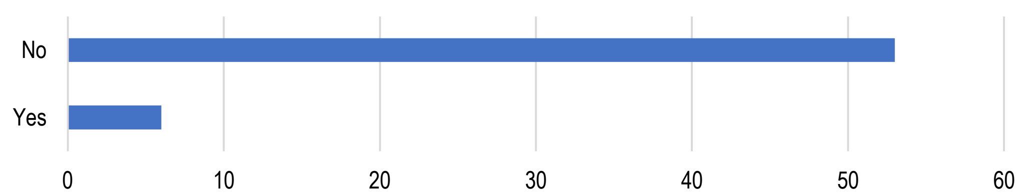
Before arriving at Sandhurst have you skipped meals?



Have you skipped any meals since arriving at Sandhurst?



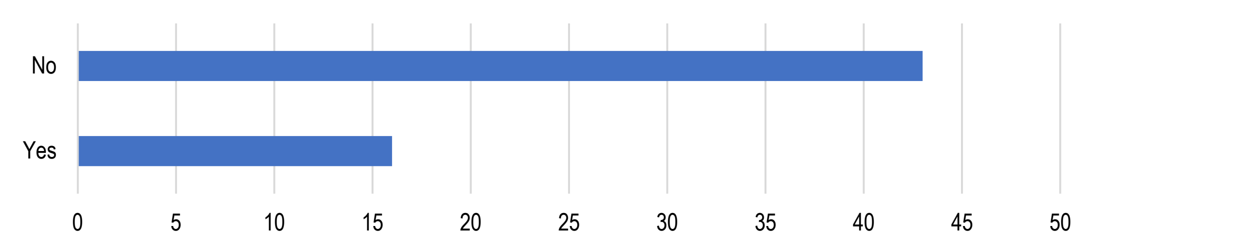
Are you vegetarian?



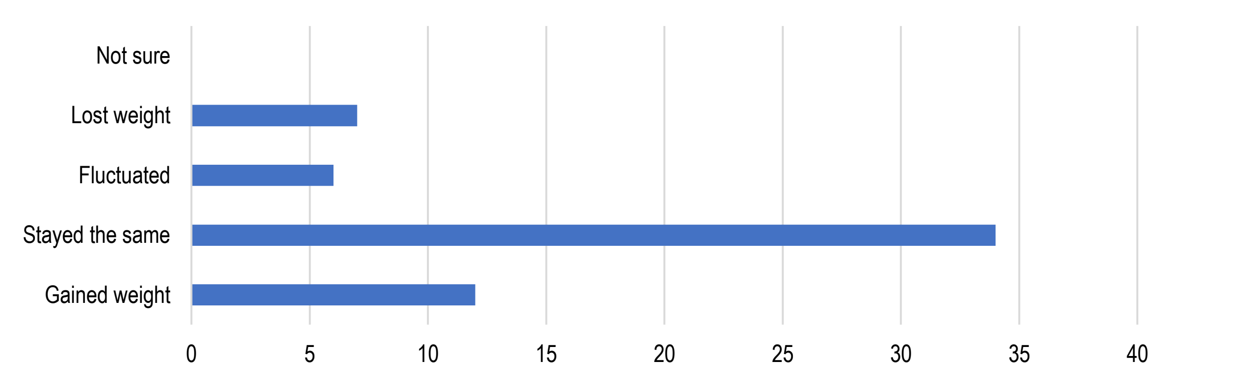
Please indicate which type of vegetarian

Vegan (does not eat meat, eggs or dairy) - 2
Lacto-ovo (does not eat meat, but does eat eggs and dairy) - 4

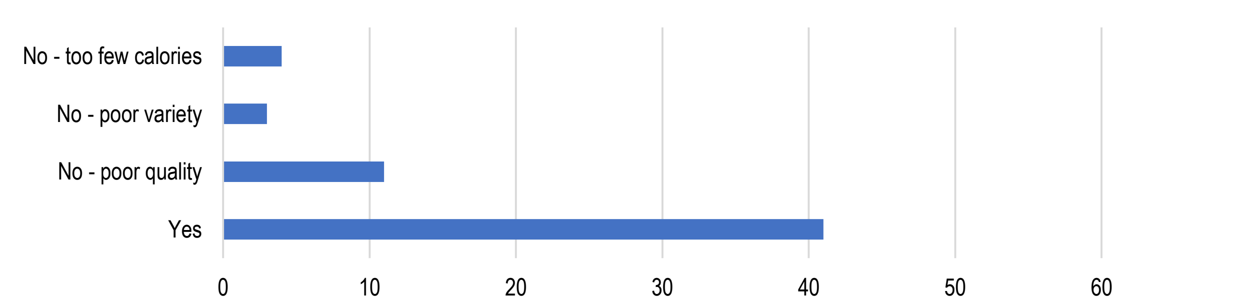
In the past 6 months have you limited / restricted the amount or type of food you eat to control your weight?



Since PCCBC what has happened to your weight overall?



Do you think your diet is nutritionally adequate (you have a good diet, eat enough and get what you need from the food you eat)?



Exercise and diet at the commencement of the study. PCCBC: Pre-Commissioning Course Briefing Course (5 to 13 weeks beforehand).

Supplemental Digital Content 2. Description of the Commissioning Course

The Course consists of three, 14-week terms, each separated by 2-3 weeks of leave, and an additional 2-weeks adventurous training (e.g. mountaineering, skiing or paddle sports). During the 44 training weeks, Officer Cadets undergo rigorous infantry-based training with physical, academic and leadership elements. The Course is designed to be immersive and intense; working days usually last over 16 hours. Inclusion criteria were commencing the Course, female sex, aged 18 to 30 years at start of the Course.

During Phase 1, participants resided in barrack accommodation and underwent programmed drill, physical exercise, field-based study and tactical leadership assessments. Programmed activities (excluding meals) took place on all ten days; median (range) duration was 12.75 (5.50 to 14.25) hours/day. During Phase 2, participants undertook a four-day field exercise that involved high physical demands and assessed leadership, strength, stamina and reaction to pressure. Participants rested where possible during the exercise, typically for 4-6 hours/day. The remainder of the Phase comprised weapons training and classroom activities with programmed activities lasting median (range) 12.50 (11.50 to 14.50) hours/day. During Phase 3, participants resided in barracks and underwent predominantly classroom-based lessons. Participants were expected to undertake daily physical exercise outside programmed activities (programmed for nine of the ten days, lasting median (range) 14.25 (11.50 to 16.25) hours/day).

Supplemental Digital Content 3. Doubly-labelled water method.

The evening before each 10-day EA assessment Phase, a baseline urine sample was collected, followed by administration of a single DLW dose containing 174 mg/kg BW $H_2{}^{18}O$ and 70 mg/kg BW ^2H_2O. Ten consecutive daily urine samples were then collected. Urine was stored at 5ºC for up to 10 days before being returned to MRC Elsie Widdowson Laboratory where they were stored at −20ºC until analysis. Urine samples were analysed for ^{18}O enrichment using the CO_2 equilibration method of Roether (1). Briefly, 0.5 ml of sample was transferred into 12 ml vials (Labco Ltd., Lampeter, UK), flush-filled with 5% CO_2 in N_2 gas and equilibrated overnight whilst agitated on rotators (Stuart, Bibby Scientific). Headspace of the samples was then analysed using a continuous flow isotope ratio mass spectrometer (IRMS) (AP2003, Analytical Precision Ltd, Northwich, Cheshire, UK). For ^2H enrichment, 0.4 mL of sample was flush-filled with H_2 gas and equilibrated over 6 hours in the presence of a platinum catalyst. Headspace of the samples was then analysed using a dual-inlet IRMS (Isoprime, GV Instruments Ltd, Wythenshawe, Manchester, UK). All samples were measured alongside secondary reference standards previously calibrated against the primary international standards Vienna-Standard Mean Ocean Water (vSMOW) and Vienna-Standard Light Antarctic Precipitate (International Atomic Energy Agency, Vienna, Austria). Sample enrichments were corrected for interference according to Craig (2) and expressed relative to vSMOW. Analytical precision was 0.3 ppm for ^2H and 0.5 ppm for ^{18}O. Total production of CO_2 was estimated using the multipoint method of Coward (3) and converted to TEE using the equations of Elia and Livesey (4) with an assumed RQ of 0.85.

1. Roether W. Water-CO_2 exchange set-up for the routine 18oxygen assay of natural waters. *The International journal of applied radiation and isotopes*. 1970;21(7):379-87.
2. Craig H. Isotopic standards for carbon and oxygen and correction factors for mass-spectrometric analysis of carbon dioxide. *Geochimica et Cosmochimica Acta*. 1957;12(1):133-49.
3. Coward W. The doubly-labelled-water (2 H 2 18 O) method: principles and practice. *Proceedings of the Nutrition Society*. 1988;47(3):209-18.
4. Elia M, Livesey G. Theory and validity of indirect calorimetry during net lipid synthesis. *The American journal of clinical nutrition*. 1988;47(4):591-607.

Supplemental Digital Content 4. Energy Intake Assessment.

On the first day of each Phase participants were given a full briefing detailing the correct way to populate the food diary and were given an example diary to refer to throughout the Phase. Participants were asked to list all food and drink consumed during the day, along with the brand, method of cooking and estimated portion size. At each evening visit to the lab, individual food diaries were reviewed by the research team and participants were asked to confirm the food diary entries and recall any items that may have been missed. Researchers used questioning to prompt participants to remember any missed items, for example, '*Did you have any dessert after dinner?*' or '*Did you eat any snacks this morning?*' In addition, standardized text messages were sent at 10:00 AM and 15:00 PM each day by the same member of the research team to remind participants to continue filling in food diaries (RLD). Every canteen meal for 16 women were weighed across all three Phases to create a database of average portion sizes served. Average portions, along with the nutritional content provided by the Royal Military Academy, Sandhurst, were entered into dietary analysis software (Nutritics Ltd., Dublin, Ireland). Using these portion sizes, a large, normal and small portion were entered as 1.5, 1.0 and 0.5 of these weighed average portions. For branded snack food, weight and nutritional content provided by the manufacturer was used.

Supplemental Digital Content 5. Heartrate Variability Measurement.

Participants were asked to avoid caffeine for 8 hours beforehand, were sitting upright in a quiet environment, and were asked to keep talking or movement to a minimum during measurements. Due to constraints placed by the Course, it was necessary to measure HRV prior to blood sampling in study visits 1, 4 and 6. Beat-to-beat time series were produced using proprietary software (CheckMyHeart software version 2.2) and inspected manually to ensure appropriate identification of normal-normal intervals. R-R intervals were exported and analyzed using Kubios® HRV Premium version 3.2.0 (http://www.kubios.com). We examined mean heart rate, traditional markers of time domain (root mean square of successive differences (RMSSD), percentage of successive normal RR intervals greater than 50 ms, (pNN50)), and frequency domain (fast-Fourier transformed logarithms of low-frequency (0.04–0.15 Hz) and high frequency (0.15–0.40 Hz) power, LnLF and LnHF, respectively, and their ratio, LF:HF) (1). Sample entropy, a non-linear measure of chaos within the HRV signal, and indices of PNS and SNS activity were also assessed. The parasympathetic index represents a synthesis of mean heart rate, RMSSD and the standard deviation of short term HRV (SD1), while the parasympathetic index represents heart rate, stress index (as per Baevsky and Berseneva (2)) and mean standard deviation of long-term HRV (SD2), both reported to reflect the mean deviation from normal values (3). Parasympathetic and sympathetic index values of zero mean that the parameters are on average equal to their normal values, while positive and negative values reflect a relative increase or decrease, respectively.

1. Task Force. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Eur Heart J*. 1996;17(3):354-81.
2. Baevsky R, Berseneva A [Internet]. Available from: https://www.academia.edu/35296847/Methodical_recommendations_USE_KARDiVAR_SYSTEM_FOR_DETERMINATION_OF_THE_STRESS_LEVEL_AND_ESTIMATION_OF_THE_BODY_ADAPTABILITY_Standards_of_measurements_and_physiological_interpretation_Moscow_-Prague_2008.
3. Nunan D, Sandercock GR, Brodie DA. A quantitative systematic review of normal values for short-term heart rate variability in healthy adults. *Pacing and clinical electrophysiology : PACE*. 2010;33(11):1407-17.

Supplemental Digital Content 6. Table: Baseline evaluation of participants at baseline who completed all study measures with those who did not.

	Completed study, n=47	Withdrew, n = 12	P value
Age, y	23.9 (2.6)	24.9 (2.5)	0.60
Height, cm	169.0 (8.6)	167.8 (3.2)	0.40
BMI, kg/m^2	23.36 (2.12)	22.77 (2.68)	0.92
Fractional fat, %	25.0 (5.1)	25.9 (4.9)	0.58
1.5 mile run time, mm:ss	10:54 (0:54)	10:36 (0:58)	0.89
BEDAQ score, median (IQR)	4 (1,6)	5 (0, 7)	0.61
BEDAQ ever dieted, 'yes', n (%)	27 (52)	6 (22)	0.64

Values are Mean (SD) unless otherwise stated. BMI: body mass index, BEDAQ: brief eating disorders in athletes questionnaire, IQR: interquartile range. P values are for independent samples t-test (participants who withdrew versus those who completed the study) except for BEDAQ ever dieted, which is for Chi squared test.

Supplemental Digital Content 7. Table: Linear regression equations of energy availability measured by accelerometry with energy availability measured by doubly labelled water.

	Equation	R^2	P
All Phases	Y = 0.94*X + 10.36	0.76	<0.001
Phase 1	Y = 0.93*X + 10.10	0.70	<0.001
Phase 2	Y = 0.76*X + 8.09	0.47	<0.001
Phase 3	Y = 1.12*X + 8.92	0.71	<0.001

Linear regression equations of energy availability measured by accelerometry (Energy availability from moderate and vigorous physical activity, EA_{mvpa}) and by doubly labelled water (energy availability from total physical activity, EA_{tpa}). Where $Y = EA_{tpa}$ and $X = EA_{mvpa}$. R^2: coefficient of determination.

Supplemental Digital Content 8: Narrative: body composition changes.

Modest fluctuations in weight were demonstrated with pairwise increases from visit 1 to visits 4 and 5 (+0.81 (SD ±2.65) kg, p = 0.020 and +0.82 (SD ±2.70) kg, p = 0.031, respectively) but no difference between visits 1 and 6 (–0.25 kg (SD ±3.03), p=0.60). Fat-free mass increased modestly from visits 1 to 2 (+0.47 (SD ±1.52) kg, p=0.032) but did not differ from visit 1 at visits 4 or 6 (–0.01 (SD ±1.26) kg, p=0.90 and +0.23 (SD ±2.59) kg, p=0.30, respectively). Fat mass decreased from visits 1 to 2 but increased to visit 4 (–0.89 (SD ±1.92) kg, p=0.001, and +0.85 (SD ±2.28) kg, p=0.003, respectively) but did was no different between visits 1 and 6 (+0.04 (SD ±2.30) kg, p=0.90). 1.5 mile run time was improved at visits 2, 3 and 6 compared with visit 1 (–0:30 (SD ±0:30) min, p<0.001, –0:20 (SD ±0:29) min, p<0.001 and –0:15 (SD ±0:40) min, p=0.022, respectively). Heart rate variability demonstrated beneficial adaptation during training (Table 3), particularly for time domain, parasympathetic and sympathetic indices (small effect sizes), and sample entropy (moderate effect size). Time domain measures (pNN50% and RMSSD) demonstrated a significant rise from visit 1 to 2 followed by a modest decline from visit 2 to 3, but remaining higher than visit 1. Frequency domain measures also suggested an improvement with a decrease in LF:HF power ratio from visit 1 to visits 2 and 3, driven by an increase in HF power. Sample entropy increased at visits 3 and 4 compared with visit 1, indicating increased chaotic variability. The PNS and SNS indices, representing a synthesis of time and domain variables, showed an increased parasympathetic and decreased sympathetic activity, respectively. The BEDA-Q score was low and did not change during the study (Table 3).