Melin, A., Tornberg, Å., Skouby, S., Faber, J. O., Ritz, C., Sjödin, A., Sundgot-Borgen, J. (2014). The LEAF questionnaire: a screening tool for the identification of female athletes at risk for the female athlete triad. *British Journal of Sports Medicine, 48*, 540-545.

The LEAF questionnaire: A screening tool for the identification of female athletes at risk for the female athlete triad

Anna Melin[1], Åsa B Tornberg [2, 3], Sven Skouby[4], Jens Faber[5], Christian Ritz[1], Anders Sjödin[1] & Jorunn Sundgot-Borgen[6]

[1] Department of Nutrition, Exercise and Sports, University of Copenhagen, Denmark,

[2] Physiotherapy, Department of Health Sciences, Lund University, Sweden; [3]Genetic & Molecular Epidemiology (GAME) Unit, Lund University Diabetes Center (LUDC) Clinical Research Center, Skåne University Hospital, Malmö, Sweden, [4] Endocrinological and Reproductive Unit, Department of Ob/Gyn. Herlev Hospital, Faculty of Health and Medical Sciences University of Copenhagen, Denmark, [5]Medical and Endocrinological Unit, Herlev Hospital, Faculty of Health and Medical Sciences, University of Copenhagen, Herlev, Denmark

[6] Norwegian School of Sport Sciences, Oslo, Norway

Corresponding author: Anna Melin, e-mail: aot@life.ku.dk, anna.katarina.melin@hotmail.com, cell phone: +46 732629714

Word count: 3956

Abstract

Background: Low energy availability (EA) in female athletes with or without an eating disorder (ED) increases the risk of oligomenorrhea/functional hypothalamic amenorrhea (FHA) and impaired bone health, a syndrome called the female athlete triad (*Triad*). There are validated psychometric instruments developed to detect disordered eating behaviour (DE), but no validated screening tool to detect persistent low EA and *Triad* conditions, with or without DE/ED, is available. Aim: The aim of this observational study was to develop and test a screening tool designed to identify female athletes at risk for the *Triad*. Methods: Female athletes (n=84) 18–39 years of age and training ≥5 times/week filled out the Low Energy Availability in Females Questionnaire (LEAF-Q), which is comprised of questions regarding injuries and gastrointestinal and reproductive function. Reliability and internal consistency were evaluated in a subsample of female dancers and endurance athletes (n=37). Discriminant as well as concurrent validity were evaluated by testing self-reported data against measured current EA, menstrual function and bone health in endurance athletes from sports such as long distance running and triathlon (n=45). Results: The 25-item LEAF-Q produced an acceptable sensitivity (78%) and specificity (90%) in order to correctly classify current EA and/or reproductive function and/or bone health. Conclusion: The LEAF-Q is brief and easy to administer, and relevant as a complement to existing validated DE screening instruments, when screening female athletes at risk for the *Triad*, in order to enable early detection and intervention.

New findings

- Validated screening tool
- Identified female athletes at risk for the *Triad*

How might it impact on clinical practice?

- Could enable early detection and intervention

Introduction

When energy intake is restricted or inadequate, the amount of energy available for basic physiological functions, such as reproductive function, becomes insufficient (1;2). Therefore, persistent low energy availability (EA), with or without an eating disorder (ED) present, may pose a significant health risk to female athletes. There are validated screening tools for the detection of disordered eating behaviour (DE) in athletes including the *Athletic Milieu Direct Questionnaire* (AMDQ) (3), the *Female Athlete Screening Tool* (FAST) (4) and the *American Physiological Screening Test for eating disorders among Female College Athletes* (PST) (5). The prevalence of low EA is assumed to be high in female athletes (1), but there are no screening tools based on self-reported physiological symptoms of low EA. It is, therefore, relevant to develop an instrument that can be routinely and widely used for screening female athletes to identify individuals at risk of the *Triad*.

Energy availability ≤125 kJ/kg fat-free mass (FFM) over more than five days has been shown to reduce blood glucose and leptin levels, to suppress the pulsatility of gonadotropin-releasing hormone (GnRH), and hypothalamic-pituitary-axis hormones, like luteal hormone (LH) and triiodothyronine (T_3) and to elevate cortisol as well as increase bone resorption markers in eumenorrheic sedentary women (6;7). If maintained for a longer period, low EA can cause functional hypothalamic amenorrhea (FHA) (2;8), and FHA has been shown to have a high predictive value when screening for ED among female athletes (9). Low levels of oestrogen increase the risk for stress fractures and osteoporosis (1;10), and results from a study on young female athletes indicate that restricted eating behaviour and FHA may also increase the risk for muscular and joint overload (11). Orthostatic hypotension is common among patients with ED

(10),(5), and the sensation of dizziness when rising from supine to standing position could, therefore, be a symptom of low EA. Reduced T_3 and T_4 have been reported in athletes with FHA (12;13) and in patients with ED (14). Hypothyroidism can result in an increased cold sensitivity and could therefore be a symptom of low EA. Persistent low EA, could have immunosuppressive consequences, and therefore increase the risk for infections in athletes (15). Irregular meal patterns and high fibre intake are commonly reported among females athletes dieting or with DE (1) and can cause a variety of different gastrointestinal symptoms like bloating and constipation (5;16).

The aim of this study was to develop a screening tool, the Low Energy Availability in Females questionnaire (LEAF-Q) designed to identify female athletes at risk for the *Triad* and evaluate reliability, internal consistency and discriminants as well as concurrent validity.

Method

A total of 84 female athletes from Sweden and Denmark, recruited through the national federations of endurance sport, competitive endurance sport clubs and professional dancers were included in the validation of the LEAF-Q. Subjects included were women 18–39 years of age, who trained ≥5 times/week. Dancing and endurance sports such as long-distance running and triathlon were chosen as representatives for leanness demanding sports, with an increased risk for the *Triad* (17;18). The study had two parts. In part one, the reliability of the LEAF-Q was assessed using endurance athletes and dancers (n=37). In part two, the self-reported symptoms, reported on the LEAF-Q were verified in another group of endurance athletes (n=45) using clinical assessments. The exclusion criteria were pregnancy, chronic illness, use of forms of

contraceptives other than oral, e.g. hormonal coil, not being willing to stop with oral contraceptives (OC) for at least six weeks prior to investigation, or injuries preventing the athlete from training ≥2 weeks. A total of 47 subjects participated in the clinical verification of self-reported symptoms; two dropped out and 45 were included in the analysis. Permission to undertake the study was provided by relevant regional ethics committees both in Sweden and Denmark (no 2011/576 and H-4-2011-096). Physiological symptoms in the literature frequently associated with long-term low EA and/or other *Triad* conditions, were included as variables in the LEAF-Q after being verified as especially relevant by a collective clinical expertise in endocrinology, sports nutrition, medicine and gastroenterology.

Internal consistency and reliability

The first version of the LEAF-Q included 29 items distributed on 5 main variables (Injuries, Dizziness, Cold sensitivity, Gastrointestinal function and menstrual dysfunction (MD), including questions regarding the use of OC). Since these endurance athletes and dancers (n=37), were of English, Swedish and Danish origin, the LEAF-Q was from the start translated by two bilingual employees at the University of Copenhagen. A test-retest was performed in order to assess item performance and estimate reliability. The subjects received the LEAF-Q twice within a two-week period and a letter of information to which the subjects were asked to reply within two days. After filling out the LEAF-Q the second time, the subjects were interviewed in order to assess response bias. The subjects were asked to comment if questions or response categories were not clearly stated, if they perceived any question as irrelevant, if response options were inadequate and whether the volume of the questionnaire was appropriate. One variable, illness, was added.

Verification of self-reported symptoms

The version of the LEAF-Q used in the next part of the validation process included 14 items in the introduction including occupation, type of sport, age, height, weight and training, followed by 30 items divided into 6 variables: Dizziness; Gastrointestinal function; Cold sensitivity; Illnesses (during the last year); Injuries (during the last year); Menstrual function (present and in the past). The variables for dizziness and cold sensitivity included one item assessed by self-reporting the frequency of the symptom on a Likert-type ordinal scale: 1) Yes, several times a day, 2) Yes, several times a week, 3) Yes, 1–2 times a week, 4) Rarely or never. The injury and illness variable contained three items for which the frequency and duration were assessed by ordinal scales and an open category to specify the types of injuries/illness. The open responses were for the variable of injuries divided in two main types of injuries—accidental or overload. The variable for gastrointestinal function contained four items and was assessed by self-reporting gastrointestinal symptoms such as pain/cramps/bloating on a Likert-type ordinal scale. Nominal scales assessed stool frequency and consistency. Menstrual function included 12 items and was measured on dichotomous as well as ordinal scales. Age of menarche was measured on an ordinal scale and previous MD on a nominal scale. Bleeding pattern related to exercise was measured on both dichotomous and nominal scales. The use of hormonal contraceptives contained six items and was assessed on dichotomous and nominal scales. After filling out the LEAF-Q, the endurance athletes (n=47) were invited to take part in a standardized verification program in order to verify self-reported symptoms by assessment of EA, clinical verification of DE/ED, orthostatic blood pressure, hypothyroidism, reproductive function and bone health.

Dietary intake and training intensity were recorded by the subjects for seven consecutive days in order to assess EA (19). Subjects were instructed to maintain and follow their normal eating pattern and training regime. Energy intake was calculated from weighed food records using a nutrient analysis program, Dankost 2000 (Dankost, Copenhagen, Denmark), for Danish foods, and Dietist XP (Kost och Näringsdata AB, Bromma, Sweden) for Swedish food items.

The subjects were instructed to maintain and follow their normal training regime. Heart rate (HR) monitors (Polar RS400®, Stockholm, Sweden) were used to assess exercise energy expenditure (EEE) based on individual prediction equations from measured HR and corresponding energy expenditure (EE) during an incremental maximal exercise test in the laboratory. The individual equation provided the basis for the calculation of EEE using HR measurement for each training session (20). Regression lines were calculated for the corresponding values of HR and EE during the incremental exercise test in the laboratory. Heart rates were highly correlated with O_2 consumption during increasing workloads (r = 0.94, 95% confidence interval 0.93-0.96). Energy availability was calculated by subtracting mean EEE from the mean energy intake, after calculating EEE by subtracting from total EEE the total EE during an equivalent time period without exercise. In order to calculate daily total EE, HR monitors (Polar RS400®) were used to assess EE during bicycle transportation, and actigraphy (ActiGraph GT3X®, Pensacola, Florida, USA) and the data analysis software ActiLife 5 (ActiGraph) were used for the assessment of non-exercise physical activity thermogenesis (NEAT). The subjects were instructed to wear an accelerometer on the wrist during sleep, and on the hip from getting up in the morning until bedtime, taking it off only during showering, swimming, bicycle transportation, and training.

Eating behaviour was assessed using the Eating Disorder Inventory (EDI-3), a questionnaire assessing behaviour and attitudes related to DE behaviour (21). Subjects were categorized as having DE behaviour when the subscale Drive for Thinness was ≥14 and/or Body Dissatisfaction ≥19 (21). The Eating Disorder Examination (EDE-16.0) (22) was used to determine whether subjects met the criteria for ED according to the DSM-IV-criteria for anorexia nervosa, bulimia nervosa and ED not otherwise specified (EDNOS). All interviews were performed by the same EDE-certified member of the research team.

Subjects using OC were requested to stop for a minimum of 6 weeks prior to examination in order to secure a sufficient wash-out period for exogenous oestrogen and progesterone. Subjects not recovering their menstrual bleeding within the 6 weeks were monthly contacted by the research team during a follow up period of minimum 3 months before gynaecological assessment. Menstruating athletes were examined in the early follicular phase, defined as the third to fifth day of menstruation. A pregnancy test was performed and menstrual function was examined by an experienced gynaecologist who performed a trans-vaginal ultrasound examination. The maximum number of ovarian follicles present in a single plane was counted and total volume of these was assessed. Sex hormone status (oestrogen, LH, follicle stimulating hormone (FSH) and androgens) and anamnestic assessment, including age of menarche, OC and previous MD were recorded. Subjects were then classified as *eumenorrhea* (menstrual cycles of 28 days ±7 days); *oligomenorrhea* (menstrual cycles >35 days); *FHA* (either primary; no menarche after 15 years of age, or secondary; absence of ≥3 consecutive menstrual cycles); *other MD* (anatomic defects, hyper-prolactinemia or other dysfunctional ovarian conditions); *polycystic ovary syndrome* (PCOS) (≥2 of the following criteria fulfilled: 1) enlarged ovaries

with a volume greater than 10 mL and/or ≥one ovary demonstrating ≥12 follicles in one plane, 2) oligomenorrhea/amenorrhea and 3) elevated androgen level or otherwise androgen stigmatized).

Body weight was measured in underwear on an electronic scale and height was measured using a fixed stadiometer. Dual-energy X-ray absorptiometry (DXA) (Hologic®, Model Discovery 2009, Hologic Inc. Waltham, MA) was used to determine fat-free, fat and bone mass. All measurements and scans were performed in the fasted state between 7:30 to 9:00 a.m. and were assessed by the same technician and performed on the same scanner. Calibration of the DXA was performed weekly using a phantom provided by the manufacturer. Subjects were classified as having normal bone mineral density (BMD) when *Z*-scores were >-1 in all measured sites, low BMD when *Z*-score was -1 to -1.9 in at least one site, and osteoporosis when *Z*-score was ≤-2 (1). Changes in blood pressure (BP) when moving from a supine to standing position were measured using a standardized tilt table and an electronic sphygmomanometer (Microlife BP A100, Widnau, Switzerland). After resting in a supine position for seven minutes, BP was measured three times before (the mean was used) and once directly after tilting the subjects to 70º standing position. Hypotension was defined as a supine systolic BP < 90 mmHg and/or a diastolic BP < 60 mmHg. Orthostatic hypotension was defined as a fall in systolic BP >20 mmHg and/or a fall in diastolic BP >10 mmHg when moving from supine to standing position (5;23).

Blood samples were drawn from an antecubital vein in a resting state after an overnight fast. Oestrogen, FSH, LH and androgen were analyzed using ADVIA Centaur Immunoassay Systems (Siemens Healthcare Diagnostics Products GmbH, Germany). The analytic sensitivity was 18.4-

15781 pmol/L for estrogen, 0.3-200 IU/L for FSH, 0.07-200 IU/L for LH and 0.1-28.0 nmol/L for androgen. The intra- inter assay precision coefficient of variability (CV) for estrogen was 2.6 % and 4.1%, for FSH; 1.2 % and 2.0 % and for LH; 2.9 % and 2.3 %. The CV for androgen was 4.7% at 7.5 nmol/L. T_3 was analysed using the ARCHITECT® system assay (Abbott Laboratories, Longford, Ireland) with an analytical sensitivity of ≤0.25-8 ng/mL and a total assay precision CV of <10%. Capillary blood glucose was analyzed using Biosen C Line (EKF Diagnostic, Germany) with a measurement range between 0.5-50 mmol/L and a CV of 1.5% at 12 mmol/L. Cortisol was analyzed using Roche Electro Chemiluminiscence Immunoassay (ECLI) (Roche Diagnostic, Bromma, Sweden). Analytical sensitivity for the cortisol assay was 0.5-1750 nmol/L with an assay precision CV of 1.3-2.1 %. Leptin was analyzed using Quantikine® ELISA (R&D Systems® Europe Ltd., Abingdon, UK). The lower analytic limit for leptin was 1.56 ng/mL; lower concentrations are associated with some uncertainty and therefore not automatically calculated. Seven subjects had results < 1.56 ng/mL, and in order to calculate their concentrations, the concentration was extrapolated using the formula for the standard curve: $\log_{10}(\text{ABS}) = A*\text{Log}_{10}(\text{concentration}) - B \rightarrow \text{concentration} = 10^{\log10(\text{ABS})+B))/A}$. The intra- and inter-assay precision CV for leptin was 3.0-3.3% and 3.5-5.4%, respectively.

Statistical methods

Distributional features of variables were measured on a continuous scale (height etc.) and variable scores were examined for central tendency. Skewed data were log transformed before further analysis. The intra-class correlation coefficient was used in order to calculate the difference between the test and test-retest scores. In this group of female endurance athletes,

oligomenorrhea/FHA existed both with and without the presence of low current EA (≤125 kJ/kg FFM). Due to this mismatch; subjects were divided into the 50% with the highest current EA and 50% with the lowest current EA in order to examine discriminant validity. The statistical procedure for questionnaire validation described by Black et al. was used (5), where discriminant validity was assessed by testing the mean item score for each of the six variables for significant difference (two-sample *t*-tests) between the group with lower EA versus higher EA, MD versus eumenorrhea and low BMD versus normal BMD. To measure concurrent validity and the degree of association among the total LEAF-Q score, LEAF-Q variables and *Triad* conditions, Pearson's correlation coefficient (r) was calculated. Furthermore, the contribution of LEAF-Q variables to the different *Triad* conditions was calculated using an ordinal logistic regression. Pearson's correlation coefficient (r) was also calculated to assess the degree of association between self-reported training hours per week and training hours registered during the week of data collection, as well as self-reported and measured body weight. For comparison of means between subjects categorized at risk for the *Triad* versus those categorized with low risk, Student's paired *t*-test was used. To test whether there was a difference between the two kinds of classifications, e.g., the number of subjects categorized at risk for the *Triad* versus those categorized with low risk having MD, Fisher's exact test was applied. Statistical significance was declared for $P<0.05$. Internal consistency between questionnaire variables was examined by Cronbach's alpha (24). The items found to significantly predict lower EA, MD and/or low BMD were retained and summed up to provide an overall LEAF-Q score. In order to test the validity of the LEAF scale, sensitivity and specificity were calculated.

Results

During the first part of development and testing reliability (n=37), internal consistency testing of the main variables resulted in an overall alpha of 0.86, suggesting a relatively high homogeneity of the LEAF-Q. The intra-class correlation coefficient was used in order to calculate the difference between the test and the test-retest score. Test-retest reliability was 0.79 after a two-week interval of retesting. The responders found all the questions relevant, easy to read and understand. The response options were in general adequate but an open response category to individual explanations was lacking in items regarding training, injuries, menstruation and OC and this was therefore addressed in the final version used for validation.

Energy availability were 28.7 ±11.0 kJ/kg FFM/day in the group with the lower current EA, and 49.1 ±6.9 kJ/kg FFM/day in the group with the higher current EA. Twenty-four subjects were diagnosed with FHA/oligomenorrhea, 3 with PCOS and 2 with other MD. Twenty-one had low BMD (*Z*-score ≤ -1). Eleven subjects were diagnosed with an ED and in addition one was classified as having DE. Seven subjects had hypotension in supine position, while 1 subject had orthostatic hypotension. One subject had hypothyroidism (T_3 <1.2 ng/mL), 32 had low leptin levels (<3.88 pg/mL), 16 had hypoglycaemia (<4 mmol/L) in the fasted and rested state and one had elevated cortisol (>800 nmol/L).

Self-reported training hours/week correlated with registered training hours/week (r=0.49, *P*<0.01) and self-reported body weight to measured body weight (r=0.97, *P*<0.001). The variable scores for gastrointestinal symptoms, injuries and MD showed significant differences between the groups—Gastrointestinal symptom: lower current EA versus higher current EA (*P*=0.023); Injury: low BMD versus high BMD (*P*=0.021); and menstrual function: MD vs. eumenorrhea

(P<0.001). These 3 variables were significantly associated with a lower current EA and/or MD and/or low BMD using a logistic regression (Table 1), and the mean variable score for the remaining variables correlated with the total LEAF-Q score (Table 2). The 3 variables had values of Cronbach's alpha ≥0.71 (Table 3); they were therefore retained and provided an overall LEAF-Q score.

Table 1 Variables associated with lower current energy availability, menstrual dysfunction and impaired bone health

Variable	Verifying variable	OR	95% CI
Gastrointestinal symptoms	Lower current EA	3.39	1.03–11.15*
Injuries	Low BMD	1.43	1.04–1.96*
Menstrual dysfunction	Menstrual dysfunction	1.65	1.23–2.20**

Abbreviation: OR: odds ratio >1 is a risk factor and indicates the contribution of the different LEAF-Q variables to lower current energy availability, menstrual dysfunction and low bone mineral density (BMD), respectively. CI: confidence interval. *P<0.05, **P<0.01

Table 2 Correlations between Total LEAF-Q and variable scores

Total LEAF scale score	Rho
Gastrointestinal symptoms	0.53**
Injuries	0.49**
Menstrual dysfunction	0.87**

Pearson's correlation coefficient, **P<0.01

Table 3 Cronbach's alpha for variables, and Total LEAF-Q test scale

Item	Alpha
Gastrointestinal symptoms	0.75
Injuries	0.79
Menstrual dysfunction	0.61
Test scale	0.71

The variable score producing the highest sensitivity and specificity for the corresponding *Triad* end point was used as the cut-off for each item score (≥ 2 for gastrointestinal symptoms, ≥ 2 for injuries and ≥ 4 for MD). The total LEAF-Q score ≥ 8.0 produced a sensitivity of 78% and a specificity of 90% for correctly classifying current EA and/or reproductive function and/or bone health. When excluding the athletes with PCOS and other MD than oligomenorrhea/FHA, the total LEAF-Q score produced a sensitivity of 83% and a specificity of 90%, with an overall validity of 73% for correctly classifying current EA and/or reproductive function and/or bone health (see Supplemental Digital Content 1; The LEAF-Q and 2; The LEAF-Q scoring key).

Summary statistics for all subjects (n=45) participating in the verification program and divided as subjects categorized at risk for the *Triad* (total LEAF-Q score ≥ 8) versus those categorized as having low risk (total LEAF-Q score < 8) are shown in Table 4. Subjects at risk for the *Triad* had lower body fat and there was a trend towards lower body weight and BMI compared to subjects categorized with low risk. Furthermore, subjects at risk had lower levels of leptin, T_3 as

well as fasting blood glucose compared to subjects with low risk, while there was no difference in cortisol levels.

Table 4 Descriptive details on all subjects and divided by total LEAF-Q score

	All (n=45)	LEAF-Q ≥ 8 (n=28)	LEAF-Q < 8 (n=17)	*P*-value
Age (yrs.)	26.6 ±5.4	26.0 ±5.6	27.5 ±5.1	0.35
Height (cm)	169.3 ±0.05	168.7 ±0.06	170.2 ±0.05	0.38
Weight (kg)	58.7 ±6.8	57.1 ±6.4	61.2 ±7.0	0.05
BMI (kg/m^2)	20.5 ±1.8	20.0 ±1.7	21.1 ±1.9	0.06
Body fat (%)	20.2 ±3.4	19.2 ±3.2	21.8 ±3.4	**0.01**
Whole body *Z*-score	0.1 [-0.7–0.4]	-0.1 [-0.7–0.5]	0.2 [-0.7–0.4]	0.91
Lumbar spine *Z*-score	-0.7 [-1.6 to -0.1]	-0.8 [-1.6 to -0.2]	-0.7 [-1.3–0.3]	0.39
EA (kJ/kg FFM/day)	161 ±58	156 ±55	169 ±64	0.49
Energy balance (%)	86 ±19	85 ±18	89 ±21	0.46
EDI-3				
DT-score	4.0 [1.0-9.0]	4.0 [1.5-9.0]	4.0 [0.0-9.5]	0.16
BD-score	5.5 [2.0-10.5]	5.0 [3.0-10.0]	6.0 [2.0-17.0]	0.29
Blood pressure (mmHg)				
Supine systolic	113.0 ±10.6	111.5 ±11.1	115.4 ±9.7	0.24
Supine diastolic	68.2 ±9.5	66.1 ±8.4	71.4 ±10.5	0.07
Standing systolic	111.7 ±10.5	109.2 ±9.9	115.4 ±10.5	0.06
Standing diastolic	72.3 ±9.1	70.4 ±8.9	75.1 ±9.0	0.10
T_3 (ng/mL)	1.61 ±0.27	1.54 ±0.22	1.74 ±0.29	**0.01**
Cortisol (nmol/L)	453 ±135	477 ±142	412 ±115	0.13
Leptin (pg/mL)	2.9 [1.6-4.1]	2.0 [1.4–3.2]	4.1 [1.4–5.3]	**<0.01**
(pg/mL/kg FM)	0.24 ±0.17	0.21 ±0.19	0.29 ±0.31	**0.02**
Glucose (mmol/L)	4.1 ±0.5	4.0 ±0.5	4.3 ±0.3	**0.02**
Exercise (hours/week)	11.3 ±4.3	11.7 ±4.8	10.6 ±3.2	0.39
VO_{2peak} (l/min)	3.2 ±0.4	3.2 ±0.4	3.2 ±0.4	0.78
VO_{2peak} (ml/kg/min)	55.7 [48.2–58.8]	56.7 [49.2–61.2]	52.4 [46.5–56.0]	0.13

Abbreviations and definitions: EA: energy availability, EDI-3: eating disorder inventory-3, DT-score: drive for thinness score, BD-score: body dissatisfaction score BMI: body mass index (kg/m^2), T_3: triiodothyronine, FM: fat mass, VO_{2peak}: maximal oxygen uptake, ¤ determined by DXA scan. Data are presented as mean ±SD for normal distributed data and as median and interquartile range [25–75] for skewed data. Student's paired *t*-test was used after skewed data were log transformed.

More subjects categorized as at risk for the *Triad* were diagnosed with MD and hypoglycaemia, while there were no significant differences in the number of subjects diagnosed with DE/ED, impaired bone health or hypotension between the groups (Table 5).

Table 5 Prevalence of Triad associated conditions in all subjects and divided by total LEAF-Q score

	All (N=45)	LEAF-Q ≥ 8 (n=28)	LEAF-Q < 8 (n=17)
BMD *Z*-score ≤-1¤	(N=17)	(n=13)	(n=4)
BMD *Z*-score ≤-2¤	(N=4)	(n=2)	(n=2)
Eating disorders	(N=11)	(n=7)	(n=4)
Disordered eating	(N=1)	(n=1)	(n=0)
Underweight (BMI <18.5)	(N=6)	(n=5)	(n=1)
Menstrual dysfunction	(N=29)	(n=23)	(n=6) **
Hypotension	(N=7)	(n=6)	(n=1)
Hypoglycemia	(N=16)	(n=13)	(n=3)*

Abbreviations and definitions: BMD: bone mineral density, BMI: body mass index, ¤ determined by DXA scan. *P<0.05, **P<0.01, tested with Fisher's exact test

Discussion

The intension of the LEAF-Q was to construct a brief questionnaire focusing only on self-reported physiological symptoms linked to persistent energy deficiency, with or without DE/ED, which can be routinely used in order to identify individuals at risk of the *Triad*. Overall, the LEAF-Q had an acceptable sensitivity and specificity as well as internal consistency, indicating that it has the potential to be a useful screening tool for the identification of female athletes at risk for the *Triad* and a relevant complement to existing validated DE screening instruments.

In contrast to pure psychometric scales, on which direct measuring of the underlying concept is impossible, the validity testing of the LEAF-Q included the verification of self-reported symptoms linked to the three endpoints of the *Triad* by objective confirmation. This created an opportunity to test the score of the LEAF-Q against measured current EA, MD and bone health, and calculated the sensitivity and specificity of the LEAF-Q. The sensitivity and specificity found indicate that the LEAF-Q has the ability to correctly characterize 8 out of 10 female athletes with lower current EA and/or oligomenorrhea/FHA and/or low BMD, but also to correctly classify 9 of those 10 athletes with higher current EA, eumenorrhea and normal BMD.

Disordered eating behaviour and ED are most often accompanied by energy deficiency and other end points of the *Triad* (25-27). The prevalence of DE/ED in this study was high (n=12, 27 %). Only 58% of these subjects had, however, lower current EA, while they all had additional MD and/or low BMD. The LEAF-Q was not constructed to discriminate between normal and pathological eating behaviour, but to detect female athletes at risk for the *Triad*. Findings of either one of the end points in the *Triad* when using the LEAF-Q should therefore implicate assessment for the other conditions, including a pathological eating behaviour.

Gastrointestinal problems are commonly reported in female endurance athletes (28) and in patients with DE/ED (1;5). In this group of female endurance athletes, the variable for gastrointestinal problems was verified by lower current EA. Persistent energy deficiency causes mucosal atrophy characterized by diminished intestinal function as well as morphological changes (29) linking lower current EA to gastrointestinal problems.

Thein-Nissenbaum and colleges showed an increased risk for muscular skeletal injuries in female athletes with restricted eating behavior and MD (30). These results confirm earlier findings of increased risk for muscular skeletal injuries in female athletes with restricted eating behaviour, MD as well as low BMD (11). In this group of female endurance athletes, the LEAF injury score correlated with impaired bone health and not with MD or current EA. A sub analysis showed that the group with impaired bone health exercised more compared to those with normal BMD (13.4 ±5.1 hours per week vs. 9.6 ±2.5 hours per week, $P<0.01$), supporting similar findings regarding impaired bone health in studies with female elite athletes from weight bearing endurance sports (26;31). It is furthermore, well established that prolonged, exhaustive endurance exercise is capable of inducing skeletal muscle damage and temporary impairment of muscle function (32), making the association between LEAF injury score and impaired bone health physiological plausible.

PCOS is reported as common among female elite athletes (13), and is not associated with hypothalamic inhibition because of energy deficiency. In this group of athletes, however, 3 out of 5 subjects diagnosed with PCOS or a MD other than oligomenorrhea/FHA had low BMD and 4 had lower current EA. Our findings emphasize the importance of differential diagnoses in order to secure proper intervention for MD in female athletes and indicate that assessment of coexisting *Triad* conditions could be relevant despite the type of MD.

Efforts were made to improve content validity thorough focusing on physiologic plausibility, examining the research literature to identify the most discriminating variables between subjects with long-term energy deficiency. The strength of this study is that standardized objective

methods have been used to assess DE/ED, orthostatic blood pressure and reproductive function as well as bone health. Gastrointestinal function, including stool frequency and consistency, as well as reproductive function and the reason for use of OC are personal issues, and questions in this regard may be considered offensive and as intimidating with regard to personal integrity. Therefore, the LEAF-Q and the results should be carefully administered and evaluated by medically qualified personal. The LEAF-Q has so far been validated only in this group of female endurance athletes. Since sports-specific issues may be relevant to assess, the LEAF-Q needs to be tested in other areas, such as aesthetic sports and sports that require weight categories for competition, but also in sports not focusing on weight, such as soccer and team handball.

The LEAF-Q is brief and could be considered for use to identify female athletes at risk of the *Triad* in order to promote early detection and treatment.

Acknowledgments, competing interests, funding

This study was supported by research grants from the Faculty of Science, University of Copenhagen, the World Village of Women Sports Foundation and Arla Foods Ingredients. AM, ÅT, SS, AS and JSB designed the study. AM analyzed data and was together with AS and JSB primarily responsible for writing the manuscript. AM and ÅT collected data. CR, JF and SS provided advice and consultation in their areas of expertise and contributed together with ÅT to the writing of the manuscript. We thank Katrine Haahr at Herlev Hospital for performing the gynaecological examinations; Anders Johansson at Lund University and Ulla Kjærulff-Hansen at Herlev Hospital for taking the blood samples; Fiona Koivola at Lund University for assisting during testing; Sara Sig Møller for assisting in data collection; Hanne Udengaard at Herlev

Hospital for logistic assistance; Helle-Marina Oxfeldt at Herlev Hospital for measuring blood pressure; Inge Birgit Rasmussen and Søren Andresen at the University of Copenhagen for blood analysis and Janni Borg Hoffmeyer for diet record analysis. We greatly appreciate the extraordinary cooperation of the athletes participating in this study, and the support of the Swedish and Danish National Sport Federations and Team Danmark. None of the authors have any conflict of interest in the results of the present study or any professional relationships with companies or manufacturers who might benefit from the results of the present study.

Reference List

(1) Nattiv A, et al. The female athlete triad. Medicine and Science in Sports and Exercise 2007 Oct;39(10):1867-82.

(2) Loucks AB, Thuma JR. Luteinizing hormone pulsatility is disrupted at a threshold of energy availability in regularly menstruating women. Journal of Clinical Endocrinology & Metabolism 2003 Jan;88(1):297-311.

(3) Nagel DL, Black DR, Leverenz LJ, Coster DC. Evaluation of a screening test for female college athletes with eating disorders and disordered eating. Journal of Athletic Training 2000 Oct;35(4):431-40.

(4) McNulty KY, Adams CH, Anderson JM, Affenito SG. Development and validation of a screening tool to identify eating disorders in female athletes. Journal of the American Dietetic Association 2001 Aug;101(8):886-92.

(5) Black DR, Larkin LJS, Coster DC, Leverenz LJ, Abood DA. Physiologic screening test for eating disorders/disordered eating among female collegiate athletes. Journal of Athletic Training 2003 Oct;38(4):286-97.

(6) Loucks AB. Energy availability, not body fatness, regulates reproductive function in women. Exerc Sport Sci Rev 2003 Jul;31(3):144-8.

(7) Ihle R, Loucks AB. Dose-response relationships between energy availability and bone turnover in young exercising women. Journal of Bone and Mineral Research 2004 Aug;19(8):1231-40.

(8) Cumming DC, Cumming CE. Estrogen replacement therapy and female athletes: Current issues. Sports Medicine 2001;31(15):1025-31.

(9) Sundgot-Borgen J, Torstveit MK. The female football player, disordered eating, menstrual function and bone health. British Journal of Sports Medicine 2007 Aug;41:I68-I72.

(10) Meczekalski B, Podfigurna-Stopa A, Genazzani AR. Hypoestrogenism in young women and its influence on bone mass density. Gynecological Endocrinology 2010 Sep;26(9):652-7.

(11) Rauh MJ, Nichols JF, Barrack MT. Relationships Among Injury and Disordered Eating, Menstrual Dysfunction, and Low Bone Mineral Density in High School Athletes: A Prospective Study. Journal of Athletic Training 2010 May;45(3):243-52.

(12) De Souza MJ, Williams NI. Physiological aspects and clinical sequelae of energy deficiency and hypoestrogenism in exercising women. Human Reproduction Update 2004 Sep;10(5):433-48.

(13) Hagmar M, Berglund B, Brismar K, Hirschberg AL. Hyperandrogenism May Explain Reproductive Dysfunction in Olympic Athletes. Medicine and Science in Sports and Exercise 2009 Jun;41(6):1241-8.

(14) Warren MP. Endocrine manifestations of eating disorders. J Clin Endocrinol Metab 2011 Feb;96(2):333-43.

(15) Rodriguez NR, DiMarco NM, Langley S, Denny S, Hager MH, Manore MM, et al. Nutrition and Athletic Performance. Medicine and Science in Sports and Exercise 2009 Mar;41(3):709-31.

(16) Bonci CM, Bonci LJ, Granger LR, Johnson CL, Malina RM, Milne LW, et al. National athletic trainers' association position statement: preventing, detecting, and managing disordered eating in athletes. J Athl Train 2008 Jan;43(1):80-108.

(17) Greydanus DE, Omar H, Pratt HD. The adolescent female athlete: current concepts and conundrums. Pediatr Clin North Am 2010 Jun;57(3):697-718.

(18) Torstveit MK, Sundgot-Borgen J. Eating disorders in male and female athletes. In: International Olympic Comittee, editor. Clinical Encyclopaedia of Sports Nutrition. 2013.

(19) Loucks AB. Low energy availability in the marathon and other endurance sports. Sports Medicine 2007;37(4-5):348-52.

(20) Tomten SE, Hostmark AT. Energy balance in weight stable athletes with and without menstrual disorders. Scandinavian Journal of Medicine & Science in Sports 2006 Apr;16(2):127-33.

(21) Garner DM. Eating Disorder Inventory-3. Professional Manual. Par Inc; 2004.

(22) Fairburn CG, Cooper Z, O'Connor M. Eating Disorder Examination (Edition 16.0D). Cognitive Behavior Therapy and Eating Disorders.New York: Guilford Press; 2008.

(23) Naschitz JE, Rosner I. Orthostatic hypotension: framework of the syndrome. Postgrad Med J 2007 Sep;83(983):568-74.

(24) Cortina JM. What is Coefficient Alpha? An Examination of Theory and Applications. Journal of Applied Physiology 1993;78(1):98-104.

(25) Torstveit MK, Rosenvinge JH, Sundgot-Borgen J. Prevalence of eating disorders and the predictive power of risk models in female elite athletes: a controlled study. Scandinavian Journal of Medicine & Science in Sports 2008 Feb;18(1):108-18.

(26) Barrack MT, Van Loan MD, Rauh MJ, Nichols JF. Physiologic and behavioral indicators of energy deficiency in female adolescent runners with elevated bone turnover. American Journal of Clinical Nutrition 2010 Sep;92(3):652-9.

(27) Gibbs JC, Williams NI, De Souza MJ. Prevalence of individual and combined components of the female athlete triad. Med Sci Sports Exerc 2013 May;45(5):985-96.

(28) Fallon K. Athletes with gastrointestinal disorders. In: Burke L, Deakin V, editors. Clinical Sports Nutrition. Third edition ed. McGraw Hill; 2006. p. 721-38.

(29) Shaw D, Gohil K, Basson MD. Intestinal mucosal atrophy and adaptation. World J Gastroenterol 2012 Nov 28;18(44):6357-75.

(30) Thein-Nissenbaum JM, Rauh MJ, Carr KE, Loud KJ, McGuine TA. Menstrual Irregularity and Musculoskeletal Injury in Female High School Athletes. Journal of Athletic Training 2012 Jan;47(1):74-82.

(31) Pollock N, Grogan C, Perry M, Pedlar C, Cooke K, Morrissey D, et al. Bone-Mineral Density and Other Features of the Female Athlete Triad in Elite Endurance Runners: A Longitudinal and Cross-Sectional Observational Study. International Journal of Sport Nutrition and Exercise Metabolism 2010 Oct;20(5):418-26.

(32) Grobler LA, Collins M, Lambert MI, Sinclair-Smith C, Derman W, St Clair GA, et al. Skeletal muscle pathology in endurance athletes with acquired training intolerance. Br J Sports Med 2004 Dec;38(6):697-703.

Supplemental Digital Content 1 The LEAF-Q

Supplement Digital Content 2 The LEAF-Q scoring key