Prediction Model for Stress Fracture in Young Female Recruits during Basic Training

DANIEL S. MORAN[1,2], ERAN ISRAELI[3], RACHEL K. EVANS[4], RAN YANOVICH[1,2],
NAAMA CONSTANTINI[5], NOGAH SHABSHIN[6], DRORIT MERKEL[2,7], ORIT LURIA[8],
TOMER ERLICH[1,2], ARIE LAOR[1,2], and AHARON FINESTONE[9]

[1]Heller Institute of Medical Research, Sheba Medical Center, Tel Hashomer, ISRAEL; [2]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, ISRAEL; [3]Division of Medicine, Department of Gastroenterology, Hebrew University-Hadassah Medical Center, Jerusalem, ISRAEL; [4]US Army Research Institute of Environmental Medicine, Natick, MA; [5]Department of Orthopaedic Surgery, The Hadassah-Hebrew University Medical Center, Ein Kerem, Jerusalem, ISRAEL; [6]Department of Diagnostic Imaging, Sheba Medical Center, Tel Hashomer, ISRAEL; [7]Department of Hematology, Sheba Medical Center, Tel Hashomer, ISRAEL; [8]Israel Surgeon General Headquarters, IAF, Tel Hashomer, ISRAEL; and [9]The Foot and Ankle Unit, Orthopaedic Department, Assaf Harofeh Medical Center, Zeriffin, ISRAEL

ABSTRACT

MORAN, D. S., E. ISRAELI, R. K. EVANS, R. YANOVICH, N. CONSTANTINI, N. SHABSHIN, D. MERKEL, O. LURIA, T. ERLICH, A. LAOR, and A. FINESTONE. Prediction Model for Stress Fracture in Young Female Recruits during Basic Training. *Med. Sci. Sports Exerc.*, Vol. 40, No. 11S, pp. S636–S644, 2008. **Purpose**: To develop a new prediction model for stress fractures (SF) in female recruits during basic training (BT) to identify risk factors and to try to prevent orthopedic injuries. **Methods**: Measurements and data collection were taken from three companies of gender-integrated recruited units before the BT program (a total of 227 females and 83 males). Measurements included anthropometric variables, blood samples for hematology profile and markers for bone metabolism, fitness tests, bone quality (peripheral quantitative computed tomography), nutritional and activity habits, psychological assessment, and medical evaluation. SF were diagnosed during BT by bone scintigraphy and/or magnetic resonance imaging. **Results**: All collected measurements were used to construct a new prediction model for the 27 and 192 female soldiers found with and without stress fracture, respectively. There were no SF in the male soldiers. The model successfully predicts 76.5% of the female soldiers with and without stress fractures (SF) as follows: $\text{PSF} = -13.98 + 0.079\ \text{Ht} - 0.014\ \text{Fe} + 0.464\ \text{BUR} - 0.105\ \text{BMI} + 0.035\ \text{Ferritin}$, where PSF is the SF prediction according to the log odds(SF); odds(SF) is the ratio between probability of SF existence and nonexistence; Ht is the height (cm); BUR is a subjective assessment of burnout on a scale of 1 to 7; Fe is the iron blood level ($\mu\text{g}\cdot\text{dL}^{-1}$); ferritin is the iron storage level ($\text{ng}\cdot\text{mL}^{-1}$); and BMI is the body mass index ($\text{kg}\cdot\text{m}^{-2}$). **Conclusion**: A young female recruited to an integrated light combat unit is at risk for stress fracture if she is tall, lean, feels "burnout," has iron deficiency, and is at the high end of the normal ferritin range. However, further evaluation is required in different populations, conditions, and training programs to evaluate these results. **Key Words:** BONE, IMAGING, RISK FACTOR, MILITARY, GENDER

Bone is a dynamic tissue, constantly undergoing breakdown and repair as it adapts to the loads to which it is exposed (6). To fulfill one of its primary functions, supporting the body for the purpose of movement, the bone must detect the mechanical signals of the loads it experiences and integrate them into changes in the bone architecture (34). Adaptations to mechanically induced stimuli are thus part of the normal functional response of healthy bone tissue.

Repeated application of loads below the fracture threshold over short periods of time, however, can contribute to stress fractures (SF) when the rate of extracellular matrix microcracking does not allow sufficient time for bone repair. Areas of bone that are subjected to repeated stress after the induction of targeted remodeling have been observed to undergo fatigue fracture (24), although it is unclear whether this remodeling is truly targeted or simply stochastic (9). Wasserman et al. (36) suggested that microcracking is more likely to occur in cortical bone, which is highly mineralized, as it is relatively brittle despite being resistant to strain. Tension loads, on the other hand, are typically borne by bone with lower mineral density, which tends to have a higher amount of remodeling activity than bone whose ambient state is compressive (3,33). Because more strain-resistant bone also tends to be more brittle, the magnitude of strain itself is not a good predictor of stress injury propensity (11,19).

The inability to predict stress injuries based on strain magnitude is unfortunate as SF is one of the most common and potentially debilitating overuse injuries seen in military recruits, causing morbidity ranging from pain to permanent

---

Address for correspondence: Daniel S. Moran, Heller Institute of Medical Research, Sheba Medical Center, Tel Hashomer 52621, Israel; E-mail: dmoran@sheba.health.gov.il.

0195-9131/08/4011S-S636/0
MEDICINE & SCIENCE IN SPORTS & EXERCISE®

DOI: 10.1249/MSS.0b013e3181893164

disability. This injury occurs much more frequently in women than in men, even when they are performing the same prescribed physical activities, and is a significant concern during recruit training in the Israeli Defense Forces (IDF). However, although clinicians are able to diagnose bone stress injury after it occurs with nearly 100% sensitivity using bone scintigraphy (17,20), a method for predicting the occurrence of bone overuse injuries before they occur remains elusive. This is despite a plethora of available data on military recruits that provide a wide range of indications as to each recruit's biological, physiological, behavioral, and psychological state over the course of training. Using some of the available recruit data to develop a predictive model would contribute to a better understanding of the mechanical environment that results during the basic training (BT) program and provide a key to effective prediction, intervention, and possibly treatment of SF.

The authors know of no phenomenological models of tibial SF that account for damage accumulation and remodeling, which are practical enough to be suitable on a patient specific basis. Several promising avenues of research have been published in recent years, including finite element-based approaches such as Garcia-Aznar et al. (15) and Negus and Impelluso (25). When coupled with research involving strategies such as statistical shape modeling (5,8,13,27), these have the potential to make patient-specific modeling a reality. But even then, they will require more computational resources than a simple weighted-factor statistical regression formula that can be used with readily available or easily obtained data from a group of potential military recruits.

The purpose of this study was to use a wide range of data collected from military recruits to develop a new prediction model for SF in female recruits during BT. This model would be developed from measurements taken from new recruits at the beginning of light combat BT and include parameters of physical fitness, anthropometry, hematology, inflammatory and bone serum markers, nutritional habits, psychological aspects, and bone quality. The newly constructed model should help with identifying individuals at risk for developing SF during BT and enable the prevention of orthopedic injuries through the adjustment of the training program for these individuals.

MATERIALS AND METHODS

Subjects

The research design was a prospective study of female army recruits who volunteered to serve in an integrated combat unit. Three different companies from a gender-integrated combat battalion participated in this study beginning on the first day of their recruitment. Collectively, there were 227 female soldiers, 18–19 yr old, from the recruitment periods August 2004, April 2005, and December 2005. Twenty-six percent of the female recruits in the combat unit had volunteered for a 1-yr national service program after graduating from high school and before induction in the army.

Inclusion criteria for all subjects included written and verbal approval for participating in the study and a medical examination before the start of the study by the study physician to determine the physiological condition of the subjects. Anthropometry criteria for the women in the combat unit included a minimum height of 1.50 m and a minimum of weight of 50 kg. Exclusion criteria for all subjects included medical consideration, subjects' objection for continuation, intermission of military service, or switching from combat duty to a position not requiring physical activity.

This study was reviewed and approved by the institutional review boards of the Committee for Research on Human Subjects, Israeli Defense Forces (IDF) Medical Corps, Tel Hashomer, Israel and the Human Use Review Committee, US Army Research Institute of Environmental Medicine, Natick, MA.

Protocol Design

Data from anthropometrics, nutrition, blood samples, and psychological factors were collected over a 1-yr training period at four time points: baseline (immediately on entry to basic training [BT]), after 2 months (the middle of BT), after 4 months (conclusion of BT), and after 16 months (1 yr after BT). The blood samples were drawn for the purpose of monitoring markers of bone metabolism. Fitness and bone quality were measured at the same time the other data were collected, excluding the 2-month time point. A medical evaluation was conducted at baseline and as necessary during the study to document the occurrence of SF and other overuse injuries. Only baseline data and measurements were used in the construction of the prediction model for SF.

Outcome Measures

Anthropometric variables. Height (cm) was measured using a stadiometer (accuracy ± 1 cm), and weight (kg) was determined with a metric scale (accuracy ± 100 g). Percent body fat was estimated by skinfold thickness measured using a four-site method referenced by Durnin and Rahaman (12), with measurements at the biceps, triceps, subscapula, and suprailiac sites, using Lange skinfold calipers (Beta Technology, Santa Cruz, CA). The same investigator performed all skinfold measurements on all the subjects.

Fitness variables

a. Maximum volume of oxygen consumption ($\dot{V}O_{2\max}$). This was measured to assess aerobic capacity using a continuous, uphill, stepwise, treadmill protocol. Warm-up exercises consisted of walking for 3 min at 3.1 mph ($5\ km \cdot h^{-1}$) on a level grade on the treadmill. The subject then began running at a 2% grade and at a speed determined to be easy to moderate based on the subject's heart rate during the warm-up walk. The subject wore a face mask connected by a flexible hose to a metabolic measurement system (SensorMedics Co., Yorba Linda, CA), which monitored oxygen uptake that was displayed and printed every 10 s. Every 2 min, the treadmill grade was increased by 2% with no change in the treadmill speed. The subject was considered to have reached maximal oxygen uptake if, 1 min after a speed increase, oxygen uptake had not increased by at least 2 $\text{mL}\cdot\text{min}^{-1}\cdot\text{kg}^{-1}$. The subjects reached maximum oxygen uptake within 10–15 min after starting the test.

b. Anaerobic capacity. This was measured by the 30-s Wingate Anaerobic Test on a cycle ergometer (1). This test is used to determine peak anaerobic power and anaerobic capacity. The testing device consists of a mechanically braked bicycle ergometer (Ergomedic 894 Ea; Monark, Varberg, Sweden). After a 10-min warm-up, the subject began pedaling as fast as possible without any resistance. Within 3 s, a fixed resistance was applied to the flywheel and the subject continued to pedal "all out" for 30 s. Electrical and mechanical counters continuously recorded flywheel revolutions in 5-s intervals. The period used for maximal effort was 30 s, where the major fuel source was anaerobic.

c. Lower extremity power and force (vertical jump). This was measured and assessed by the Leonardo Ground Reaction Force Platform (Orthometrix, Inc., White Plains, NY). The subject was positioned on the force platform (wearing sneakers) according to manufacturer's guidelines and was asked to use both legs to jump as high as possible. This exercise was repeated three times. The subject was then asked to perform three one-legged vertical jumps on each leg. Recorded parameters included power (W), velocity ($\text{m}\cdot\text{s}^{-1}$), peak velocity ($\text{m}\cdot\text{s}^{-1}$), displacement (m), and work (J).

d. Bar-Or basic fitness test. All volunteers completed the Bar-Or Fitness Test, a test to assess a soldier's basic level of fitness (1). This three-event test assessed the time to complete a 2-km run and the maximum number of continual push-ups and sit-ups able to be performed until exhaustion (stopping for 2 s). This test was administered by the cadre of the subjects' military unit as part of their military requirements, and results were provided to the research team members.

Bone quality

a) Peripheral quantitative computed tomography (pQCT) (Stratec/Medizintechnik XCT 2000, Pforzheim, Germany). This was used to measure bone and muscle characteristics of the tibia, according to previously established methods (31). Subjects were positioned on a chair with the nondominant leg extended through the scanning cylinder and were asked to maintain a convenient and stable position for the duration of the procedure (10–15 min). An initial scout scan was used to identify the distal end plate of the tibia. After this, scans of the tibia (single axial slices of 2.2-mm thickness, voxel size 0.5 mm, measure diameter 140 mm) were taken at a translation speed of 20 $\text{mm}\cdot\text{s}^{-1}$ at 4%, 38%, and 66% of the approximated segment length proximal to the distal endplate of the tibia. These sites have been determined to be the best sites to analyze trabecular bone, cortical bone, and muscle area, respectively, in the tibia (29).

Image processing and calculation of the various bone indices were performed using the manufacturer's software package (version 5.5D). For all distal 4% scans, a threshold algorithm (contour mode 1) was used to separate the tibia from the soft tissue background using <180 $\text{mg}\cdot\text{cm}^{-3}$ threshold. Once separated from the soft tissue background, total volumetric bone mineral density (vBMD) and total bone area were calculated. Trabecular bone of the tibia was analyzed after concentrically peeling off 45% of the bone pixels (peel mode 1). Trabecular vBMD and BMC were measured as the mean value within the remaining 55% of the total area.

Cortical bone properties were assessed at the 38% and 66% proximal sites using the default threshold value of 710 $\text{mg}\cdot\text{cm}^{-3}$. The following parameters were assessed at the proximal site: cortical vBMD ($\text{mg}\cdot\text{cm}^{-3}$), total area ($\text{mm}^2$; cross-sectional area of the bone after the soft tissue has been peeled off), cortical area (mm^2; the area that is assigned to be purely cortical), and cortical thickness (mm). The ability of a bone to resist bending or torsion was assessed using the axial and polar strength strain index (SSI), which provides an estimate of bone strength calculated using geometric and density parameters. The cortical bone SSI parameters were assessed at the 38% and 66% proximal sites using a threshold algorithm (contour mode 1) with a threshold value of at least 280 $\text{mg}\cdot\text{cm}^{-3}$.

Before the pQCT data collection, a cone phantom was measured to check the linearity of the results and to confirm the precision and quality assurance of the measurements. Soft tissue parameters were assessed at the 66% proximal site using an iterative contour detection (contour mode 3) with a user-defined threshold of at least 40 $\text{mg}\cdot\text{cm}^{-3}$ and using a 3×3 filter, first for voxels between -500 and 500 $\text{mg}\cdot\text{cm}^{-3}$ and then a 5×5 filter in the range between -500 and 300 $\text{mg}\cdot\text{cm}^{-3}$ (filter CO_2; the negative values are the outcome of the absorption values that calibrated for hydroxyapatite densities; fat is set to 0 and air is about -250 $\text{mg}\cdot\text{cm}^{-3}$). The following parameters were assessed at the 66% proximal site: total area (mm^2), cross-sectional area of the bone and soft tissue (mm^2), fat area (mm^2), and muscle area (mm^2).

b) Magnetic resonance imaging (MRI). The tibias bilaterally of all subjects were scanned on an open 0.5-T MR unit (SP; GE medical systems, Milwaukee, WI) using a flex coil, with the subjects in a supine position. The scans were obtained before and after the 4-month BT course. Subjects suspected of SF during the BT course were evaluated by bone scintigraphy according to the IDF Medical Corps protocol followed by another MRI. Imaging sequences were composed of coronal and axial short tau inversion recovery (STIR) and coronal gradient echo (GE) in and out of phase. The STIR sequences were used in two imaging planes because early stress-related changes in bones are seen as high signal in the bone marrow on fluid-sensitive sequences and because in relatively large fields of view the fat suppression in STIR is more homogenous than T2-weighted fat-suppressed images. GE in-phase sequence is a fast T1 sequence and can demonstrate a fracture line, the cortex, and the fine anatomy. Comparison between GE in phase and GE out of phase can provide information about red versus yellow marrow.

The MR images were evaluated by two trained musculoskeletal radiologists in consensus for periosteal reaction, bone marrow edema, and presence of a fracture line. Comparison of interval changes in the amount of red marrow was evaluated. Comparison between the right and the left tibias was obtained. Correlations were made between MRI images and bone scintigraphy.

Questionnaires

a. Activity assessment before recruitment. This was evaluated by a detailed physical activity questionnaire that included questions on the type and frequency of the activity and the age span during which the subject was involved in the activity. If subjects had not participated in sports activities the year before data collection, they were described as "inactive." If they had participated in sports activities during the year before data collection, they were described as "active." Furthermore, their activities were categorized as high impact (i.e., volleyball, jumping), odd impact (i.e., soccer, martial arts), high magnitude (i.e., weight training), low magnitude (i.e., walking, yoga), or nonimpact (i.e., swimming, bicycling). Also included was the average number of hours each activity was performed per week. Accordingly, these data were analyzed to describe the soldiers for three categories: inactive, moderately active, and very active.
b. Psychological assessment. This was evaluated by questionnaires, based on 0–3 or 1–7 scales, that were administered to collect data relating to stress/exhaustion, cohesion, burnout, motivation, self-confidence, and psychological parameters, which mediate between stress and its outcomes. Different psychological factors were measured to assess the soldier's overall function and included the effect of different levels of motivation and dropout rates, variations of the perceived pressure and their effects on psychological intermediates, and the overall effect on functioning.
c. Nutrition profile. Subjects were interviewed on site regarding their dietary intake using a Food Frequency Questionnaire developed specifically for the Israeli population. This is a common method used to assess individual long-term dietary intake of foods and nutrients. The questionnaire elicits a subjectively reported "usual frequency" of consuming an item from a list of foods (30). Interviews were conducted by professionals from the Ben-Gurion University in Beer-Sheba.
d. Health/injury history A health history assessment was conducted during the prerecruitment period. This questionnaire was based on a new recruit survey conducted in 2003 by the IDF's Military Combat Fitness Center.

Hematology profile

a. Blood drawing procedures. Recruits fasted for 10 h before a blood sample was collected while seated (between 0700 and 0800 h) according to the following procedure: approximately 30 mL of blood were drawn from an antecubital vein using sterile venipuncture techniques. Blood was collected in 5-mL silicone-coated tubes (BD Vacutainer SST II Advance; Becton, Dickinson & Company, Franklin Lakes, NJ), allowed to sit for 30 min at room temperature, and immediately centrifuged at 4°C at 2000*g* for 15 min. Serum samples were then separated and stored at −70°C immediately after collection and remained frozen until analysis.
b. Serum analysis. Assays were performed in duplicate with an average of both assays being used as the final measure. Samples from each subject were analyzed in the same assay to minimize the effects of assay variability. Assay results for albumin were used to adjust for plasma volume shifts, as appropriate.

Bone turnover markers. Bone alkaline phosphatase (BAP), which has been associated with osteoblastic activity and may be an indicator of mineralization, was assayed by enzyme-linked immunosorbent assay (ELISA; Octeia™ Octase® BAP Immunoenzymetric assay; IDS Ltd., England, UK), which was specific for the bone isoform. Interassay coefficient of variation (CV) was 6%. Reference value (mean ± SD) for healthy premenopausal women is 8.7 ± 2.9 $\mu g \cdot L^{-1}$ with a reference range of 3.7–20.9 $\mu g \cdot L^{-1}$. N-terminal propeptide of type I collagen (PINP), which is released before assembly of collagen fibers, was measured by the UNiQ radioimmunoassay from Orion Diagnostica (Espoo, Finland). Interassay CV was 9.5%. The reference range of values for healthy premenopausal women was 22.0–87.0 $\mu g \cdot L^{-1}$. Tartrate-resistant acid phosphatase (TRAP 5b), which may indicate the onset of bone resorption, was measured by way of ELISA using the

BoneTRAP® assay (IDS Ltd., England, UK). The interassay variation was 8%. Reference value (mean ± SD) for young premenopausal women is 2.59 ± 0.78 U·L^{-1}. The upper limit for normal women was 4.15 U·L^{-1} with a reference range of 1.3–4.8 U·L^{-1}. C-telopeptide cross-links of type I collagen, which are byproducts of collagen resorption (Serum Crosslaps), were measured by ELISA kits from Nordic Bioscience Diagnostics (Herlev, Denmark). Interassay CV was 5%. Reference value for healthy premenopausal women was 0.112 ± 0.738 ng·mL^{-1} with a reference range of 0.010–0.712 ng·mL^{-1}.

Nutrition markers. Albumin and calcium were both measured using a DXC600 Pro (Bechman Coulter, Fullerton, CA). Interassay CV were 1.4% and 1.7%, respectively. Reference value for healthy premenopausal women was 3.1–5.4 g·dL^{-1} for albumin and 8.9–10.4 mg·dL^{-1} for calcium. Vitamin D (25-OHD), or vitamin D, was measured by radioimmunoassay using a kit from DiaSorin (Stillwater, MN). Interassay CV was 98.6%. Reference values for healthy men and premenopausal women were from 8.9 to 46.7 ng·mL^{-1}. Parathyroid hormone (PTH) was measured by immunoassay with chemiluminescent detection on the Immulite 2000 (Diagnostics Products Corporation, Los Angeles, CA). Interassay CV was 4.7%. Reference values for healthy men and premenopausal women ranged from 12.0 to 72.0 pg·mL^{-1}.

Inflammatory markers. Tumor necrosis factor α (TNF-α), interleukin 1b (IL-1b), and IL-6 were measured by ELISA from Linco (Linco Research Inc., St. Charles, MO) on the Luminex Labmap 100 (Luminex Corp., Austin, TX). Interassay CV are 10.9%, 13.3%, and 12.7%, respectively. The sensitivity of these assays was 3.2 pg·mL^{-1}. The reference ranges for IL-1B, IL-6, and TNF-α are <3.2–52, <3.2–263, and <3.2–36.4 pg·mL^{-1}, respectively.

Medical evaluation/surveillance. Injury surveillance and SF diagnosis took place over the course of the entire training period. Data were inserted in a personal surveillance table and collected by the commanders of the clinics on the military base where the participants serve. Orthopedic monitoring was conducted every 2–3 wk by an orthopedic surgeon. The physician followed-up any orthopedic-related injuries and/or problems and made the appropriate decision regarding treatment based on clinical need. SF were diagnosed by bone scintigraphy, MRI, or both and were treated based on the IDF Medical Corps Protocol (23).

Statistical Analysis

We computed means and SD, minimum and maximum, for all continuous variables in the two groups of subjects with and without SF. To compare between these two groups, we used unpaired *t*-test.

The bivariate correlation between the SF existence (which is a dichotomous variable) and the continuous variables was computed by Pearson product–moment correlation, with the corresponding P value, which is computed after Fisher's z transformation.

For multivariate analysis, we used logistic regression. The dependent variable in our analysis is SF presence, which is a dichotomous variable. The predictor variables (independent) were a set of the measured variables derived from demographic, hematological, fitness, biomarkers, psychological, and bone density factors measured before the start of BT.

For a description of the accuracy of the SF prediction model and the true-positive rate (sensitivity) and the false-positive rate (1-specificity), we used the receiver operating characteristic curve (ROC) method (22). In this method, accuracy of a prediction model is measured by the area under the ROC curve, whereby an area of 1.0 represents a perfect prediction model, an area of 0.8–0.9 is good, an area of 0.7–0.8 is fair, and an area of 0.6–0.7 is considered as a poor prediction. In fact, the curve is constructed by computing the sensitivity and the specificity of the prediction model.

Backward stepwise elimination method was used to achieve a parsimonious model that adequately describes the data and does not include noncontributory factors at $P > 0.15$ level.

The computations were performed using SAS 9.1.3 software.

RESULTS

Data and measurement variables from 219 newly recruited female soldiers from three different companies were used in

TABLE 1. The 77 variables from eight different categories that were measured and assessed from the subjects and were analyzed for Pearson's correlation coefficients and significance value with stress fracture (SF) occurrence.

Category	Variable
Anthropometry	Height, weight, BMI, percent body fat
Psychology	General and cognitive burnout feeling, commitment, self-control, strain feeling, self-efficacy, 'no energy,' general and subscale Lazarus for problem solving, acceptance, growth, distance, and anxiety
Hematology	Ferritin, Fe, transferrin, Fe/transferrin, folate, CRP, B12, Hgb
Biomarkers in serum	PINP, TRAP, BAP, CTx, IL-1B, IL-6, PTH, TNF-α, vitamin D, Ca
Nutrition consumption	Calcium, iron, protein, fat, carbohydrate, total energy
Bone quality	Dual-energy x-ray absorptiometry measures: total body bone mineral density, percent body fat, tibia mineral density, radius mineral density. pQCT measures: 4% site—mass, total area, total density, trabecular density; 38% site—mass, total area, cortical density, cortical area, SSI; 66% site—bone muscle ratio, mass, total area, cortical density, cortical area, fracture load X, fracture load Y, SSI, muscle area.
Fitness	Total PT test score, 2-km run time, push-ups, sit-ups, $\dot{V}O_{2max}$, Wingate Anaerobic Test (peak power and average power), lower extremity power and force.
General	Past and present smoking, physical activity before recruiting, menstrual disturbances, contraceptive pill usage.

TABLE 2. The mean ± SD of the 20 variables that found with the higher Pearson's correlation coefficients and significance value from *t*-test with the stress fracture (SF) group and with the group with no SF (NSF).

	Variable	Correlation with SF (R)	SF Group (Mean ± SD)	NSF Group (Mean ± SD)	*t*-Test (*P* Value)
Anthropometry	Height	0.168	165.4 ± 5.4	162.1 ± 6.5	0.011
	BMI	−0.091	22.17 ± 3.15	23.10 ± 3.35	0.169
Psychology	Burnout feeling	0.147	3.87 ± 0.78	3.39 ± 1.08	0.025
	Commitment	−0.132	0.86 ± 0.28	0.98 ± 0.34	0.083
	Self-control	−0.061	1.13 ± 0.31	1.18 ± 0.27	0.429
	Strain feeling	0.005	4.24 ± 0.48	4.23 ± 0.55	0.950
Hematology	Ferritin	0.103	21.38 ± 16.03	17.31 ± 12.36	0.124
	Fe	−0.115	54.70 ± 24.52	70.40 ± 46.65	0.088
	CRP	0.122	3.63 ± 4.50	2.31 ± 3.32	0.071
	B12	0.049	396 ± 253	362 ± 219	0.462
Biomarkers in serum	TRAP	−0.019	3.03 ± 1.01	3.07 ± 0.71	0.789
	BAP	−0.060	16.96 ± 7.32	18.15 ± 6.94	0.395
	CTx	−0.065	0.80 ± 0.28	0.86 ± 0.36	0.360
	IL-1B	−0.006	4.27 ± 4.80	4.42 ± 5.80	0.899
	IL-6	0.064	57.43 ± 80.71	40.90 ± 87.71	0.366
	PTH	0.053	38.77 ± 12.49	36.55 ± 14.50	0.458
	TNF-α	−0.043	6.01 ± 2.72	6.52 ± 4.16	0.542
	Vitamin D	−0.017	28.45 ± 9.61	28.95 ± 9.89	0.813
	Calcium	−0.046	9.48 ± 0.40	9.52 ± 0.33	0.515
Nutrition	CA in diet	0.096	1000 ± 537	849 ± 508	0.169
Fitness	Push-ups	−0.030	36 ± 15	37 ± 12	0.669

These variables, when entered in their entirety to an SF prediction model, were able to correctly predict 90.8% of soldiers with and without SF.

this study from a total of 227 recruited soldiers that began BT. Twenty-seven soldiers (12.3%) were diagnosed with SF of different grades during the 4-month BT period. Thus, the prediction model for SF was constructed to predict the 27 and 192 soldiers with and without SF, respectively. Notably, not all the variables were collected from all the subjects due to logistic and technical reasons.

Model construction. The basic concept of constructing the SF prediction model was to calculate the probability of young recruited females developing SF under combat BT programs. Sensitivity is defined as the proportion of subjects with SF who tested positive, and specificity is defined as the proportion of subjects without SF who tested negative. The sensitivity and the specificity described how well the model discriminated between soldiers with and without SF. The variables that were collected and measured from the various categories were pooled to construct the new prediction models for SF. Logistic regression was used to predict a discrete outcome from the set of all the measured variables. In principal, the goal of logistic regression was to correctly predict the category of outcome for individual cases using the most parsimonious model. To accomplish this goal, statistical analysis for correlation and significance was performed between each of the 77 collected variables from eight different categories and the response variable, which is the group of subjects with SF (Table 1).

Next, we constructed the initial prediction SF model with the 20 potential predictor variables, using only baseline data (Table 2), that had the most significant correlations with SF, indicating meaningfulness in predicting SF. Subject status data were fit to this initial prediction model, constructed from the 20 variables, and found to correctly predict the presence or the absence of SF in 90.7% of soldiers. In Figure 1, we depict, by the receiver operating characteristic curve (ROC) method, the true-positive rate (sensitivity) against the false-positive rate (1-specificity) of SF for the different possible cut points of the prediction model. The area under the curve was found to be 0.907, which is considered as an excellent prediction model. This value is the percentage (90.7%) of randomly drawn pairs of subjects

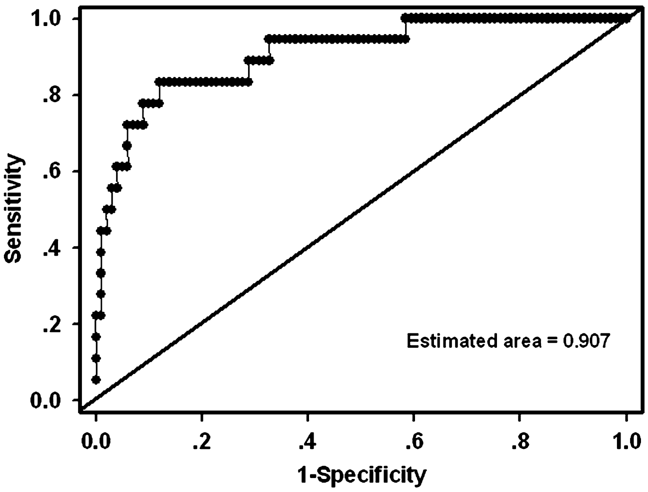

FIGURE 1—Accuracy of the prediction SF model constructed from 20 variables collected from 219 recruited females with and without SF (27 and 192, respectively) diagnosed during light combat BT. The accuracy of the model, which is measured by the area under the curve, was 90.7% as depicted by the receiver operating characteristic curve (ROC) method.

TABLE 3. Analysis of the five variables used in constructing the logistic prediction model for SF, including coefficients, chi-square value, and the corresponding *P* value.

Variable	Coefficient Estimate	Wald Chi-Square	*P* value
Intercept	−13.9838	5.8598	0.0155
Height	0.0789	5.7513	0.0165
Ferritin	0.0350	5.2725	0.0217
Burnout	0.4637	4.7062	0.0301
Fe	−0.0139	4.0074	0.0453
BMI	−0.1054	2.2109	0.1370

with and without SF and predicts correctly their status (with or without SF).

To reduce the number of variables required for the prediction SF model, we used a backward stepwise elimination analysis to find the most prevailing variables for SF. This method involved repeated iterations of model construction and elimination of the least contributory variable in each iteration. Thus, the least significant variable to each model was eliminated before construction of the subsequent model. The final model was completed when all remaining variables were found to contribute to the model significantly, when $P < 0.15$ (Table 3). This last model, which was constructed from five explanatory variables, predicted the presence or the absence of stress fractures (SF) in 76.5% of the female soldiers that participated in this study (Fig. 2).

The final model to predict stress fracture (PSF) was as follows:

$$\mathrm{PSF} = -13.98 + 0.079\ \mathrm{Ht} - 0.014\ \mathrm{Fe} + 0.464\ \mathrm{BUR} - 0.105\ \mathrm{BMI} + 0.035\ \mathrm{Ferritin},$$

where PSF is the SF prediction according to the log odds(SF); odds(SF) is the ratio between probability of SF existence and nonexistence; Ht is the height (cm); BUR is a subjective assessment of burnout on a scale of 1 to 7; Fe is the iron blood level ($\mu g \cdot dL^{-1}$); ferritin is the iron storage level ($ng \cdot mL^{-1}$); and BMI is the body mass index ($kg \cdot m^{-2}$).

To explain the specific contribution of each variable to the prediction model, we used the relative odds ratios for a risk factor, with its specific 95% confidence interval. In this case, the odds ratio is defined as the odds of a subject with SF being exposed to the risk factor divided by the odds of a subject without SF being exposed to the same risk factor. To quantify the relative influence of each of the explanatory variables to the prediction SF model, we computed the adjusted odds ratio (Table 4). For each variable, a change in one unit will be reflected in a corresponding change in the odds ratio for stress factor prediction. For example, a subject whose questionnaire responses registered a "burnout" score of four units is 1.590 times more likely to suffer an SF than a subject with a "burnout" score of three units. Thus, an increase in one unit, from three to four units, in "burnout" score will cause an increase of 1.590 times in the odds of SF. Accordingly, a newly recruited female soldier with lower BMI and Fe values, but with higher height and Ferritin values, will be more likely to develop SF than a matched individual with the opposite values.

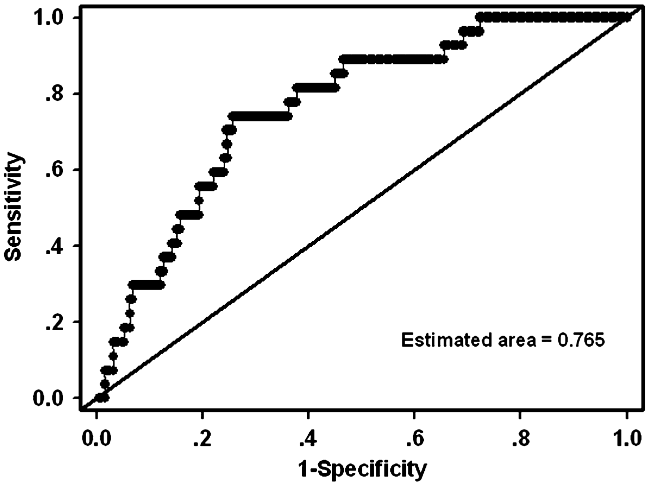

FIGURE 2—Accuracy of the prediction SF model constructed from five significant variables collected from 219 recruited females with and without SF (27 and 192, respectively) diagnosed during light combat BT. The accuracy of the model, which is measured by the area under the curve, was 76.5% as depicted by the receiver operating characteristic curve (ROC) method.

TABLE 4. The adjusted odds ratios with their 95% confidence intervals for the five variables that construct the prediction model for SF.

Effect	Odds Ratio Estimates		
	Adjusted Odds Ratio for SF	95% Confidence Limits	
Ferritin	1.036	1.005	1.067
BMI	0.900	0.783	1.034
Fe	0.986	0.973	1.000
Burnout	1.590	1.046	2.417
Height	1.082	1.015	1.154

When the adjusted odds ratio is >1, a higher value increases the odds for SF, and when the adjusted odds ratio is <1, a lower value increases the odds for SF.

DISCUSSION

Stress reaction and SF represent a spectrum of soft tissue and osseous injuries that result from abnormal repetitive stress applied to healthy bone, usually during prolonged periods of unaccustomed or strenuous activities (e.g., running, marching) (14,21). Repetitive vigorous stress creates a region of accelerated bone remodeling, which may progress to an SF if the stress continues. SF are particularly common in certain populations, principally athletes and military recruits, who are under constant and significant physical demands (7,10). Exercise-induced SF are common in the lower extremities, with 75% of stress reaction and SF occurring in the tibia (14). Because an individual with SF should temporarily cease physical activity, the identification of individuals at risk and early diagnosis are essential to ensuring that safety measures are taken to prevent orthopedic injuries. This is particularly true for new recruits entering combat BT.

Seventy-seven variables were collected and used for the SF prediction model. These variables, from eight different categories (Table 1), were statistically analyzed to determine the variables that most significantly contribute to SF occurrence. Before inclusion of the variables in the model, each variable was tested for bivariate correlation with SF existence. Notably, no correlation value was high ($R < 0.2$). However, multivariate logistic models provide much higher predictive power than single variable models because they reveal intercorrelations between variables and include independent contributions from more than one variable.

The prediction SF model constructed in this study includes five variables from three different categories. Height and BMI, which represent a female soldier's anthropometry, are the most dominant variables in predicting SF in a young (19 yr) Israeli female recruited soldier. However, these two variables have opposite effects in the prediction model. Height contributes positively and body mass negatively. As such, a young, tall, lean female entering BT is at a higher risk of developing SF. Height has also been found to be a significant risk factor in young male recruits (35). Thus, the physical activity and the exercise training executed in BT have a more pronounced impact on a tall, ectomorphic-type person. Nonetheless, soldiers of this age are still growing, so SF risk may also be related to the not fully formed state of the female soldier's skeleton.

Analysis of the hematology profile of the new soldier recruits who participated in this study revealed two parameters that were included in the prediction model. First, SF risk is inversely related to the level of iron (Fe) in the blood. This is consistent with iron deficiencies that have been found in Israeli soldiers during BT (26). Nonetheless, this study demonstrates that iron deficiencies existed in the soldiers before they were recruited and did not result from BT. Because iron deficiency can be easily treated, blood testing during the recruitment process should be considered because it could facilitate possible intervention and correction of these deficiencies. In contrast to iron, the ferritin level before recruiting was found to be positively correlated to SF risk. Ferritin represents the level of iron stored mainly in the liver but also in the bone marrow. Consequently, it is also an acute phase reactant that reflects inflammatory state. Therefore, a higher ferritin level can be assumed to be a possible indication of inflammation. This conclusion is supported by the significant positive correlation ($R = 0.2$, $P < 0.003$) found between ferritin and C-reactive protein (CRP), which is another marker for inflammation. Notably, the baseline ferritin values found in this work were at the high end of the normal ferritin range and decreased at the end of BT.

The final variable included in the prediction SF model was "burnout," as determined by the soldiers' responses to a psychology-discipline questionnaire. In this questionnaire, the subjects rated their "burnout" feeling on a scale between 1 and 7. The inclusion of this variable was unexpected because such significant feelings of "burnout" were not expected from new recruits in a voluntary combat unit. However, this variable contributed very significantly to the prediction model ($P = 0.0301$) with an odds ratio of 1.59 per unit of rated "burnout." Notably, iron deficiency was previously found in a few studies to be associated with depression, fatigue, and burnout feeling especially while exercising (4,32) and might also be the reason for the exhausted feeling found in this study in females with SF.

The high relevancy of the "burnout" variable in the model may be influenced by the background of the surveyed soldiers, as 26% of them served together in a national community program for 1 yr before their recruitment. Although most soldiers are recruited after graduation from high school, these soldiers served together for an additional year before BT in a stressful environment and may have experienced increased "burnout" feeling as a result.

The most consistent risk factor for SF injury reported in the literature is decreased aerobic fitness (18). Surprisingly, the only measure of fitness that factored significantly into the preliminary model depicted in Figure 1 was push-up performance, a measure of strength and endurance. Parameters of both strength and endurance have been found to be significant risk factors for SF in both US Army and Israeli soldiers (2,16), and a history of less than 7 months of lower-extremity weight training was significantly associated with SF in female Marine recruits (28). The relationship between muscle strength and endurance and risk for SF has yet to be fully elucidated.

The suggested prediction model constructed in this study was found to be accurate with high sensitivity and specificity and an area under the curve of 0.765 in the ROC curve analysis (Fig. 2). However, a trade-off was found between sensitivity and specificity, with any increase in sensitivity being accompanied by a decrease in specificity. Thus, for most cases with high sensitivity (0.85–1.00), the corresponding specificity was within a lower range of values (0.45–0.80). The practical interpretation of this distribution is that the model better predicts soldiers who will develop SF during BT than soldiers who will not. However, the suggested model should be looked at very carefully due to some limitations of this study. First, not all the methods and measurements were available for all the subjects from the three companies. Thus, there were missing values for some variables (e.g., dual-energy x-ray absorptiometry was used only in the first company whereas the pQCT and the Leonardo Ground Reaction Force Platform were used only in the second and the third companies). Next, the second company had a different training program than the first and third companies, and finally in the first company, there was a high rate of soldiers that served in a national community service program before recruitment.

In conclusion, the model that was refined in this study was found to be able to correctly predict the presence or the absence of SF in 76.5% of a sample population. A young female soldier is at greater risk of developing an SF during BT if she is tall, lean, feels "burnout," has iron deficiency, and is at the high end of the normal ferritin range. However, because this is an exploratory analysis-derived model, further evaluation is required for different populations and under different conditions and protocols (e.g., females that did not serve in national service before BT, non-Israeli female soldiers, different training regimes, and different climate conditions).

The opinions and assertions in this article are those of the authors and do not necessarily represent official interpretation, policy, or views of the US Department of Defense or the Israeli Defense Forces.

REFERENCES

1. Bar-Or O. The Wingate Anaerobic Test. An update on methodology, reliability and validity. *Sports Med*. 1987;4(6):381–94.
2. Beck TJ, Ruff CB, Shaffer RA, Betsinger K, Trone DW, Brodine SK. Stress fracture in military recruits: gender differences in muscle and bone susceptibility factors. *Bone*. 2000;27(3):437–44.
3. Bell KL, Loveridge N, Power J, Garrahan N, Meggitt BF, Reeve J. Regional differences in cortical porosity in the fractured femoral neck. *Bone*. 1999;24(1):57–64.
4. Benton D, Donohoe RT. The effects of nutrients on mood. *Public Health Nutr*. 1999;2(3A):403–9.
5. Boyd SK, Muller R. Smooth surface meshing for automated finite element model generation from 3D image data. *J Biomech*. 2006;39:1287–95.
6. Burger EH, Klein-Nulen J. Responses of bone cells to biomechanical forces in vitro. *Adv Dent Res*. 1999;13:93–8.
7. Coady CM, Micheli LJ. Stress fractures in the pediatric athlete. *Clin Sports Med*. 1997;16(2):225–38.
8. Cootes T, Taylor C, Cooper D, Graham J. Active shape models—their training and application. *Comput Vis Image Underst*. 1995; 61:38–59.
9. Costa Gomez TM, Barrett JG, Sample SJ, et al. Up-regulation of site-specific remodeling without accumulation of microcracking and loss of osteocytes. *Bone*. 2005;37(1):16–24.
10. Daffner RH. Anterior tibial striations. *Am J Roentgenol*. 1984;143(3):651–3.
11. Donahue SW, Sharkey NA. Strains in the metatarsals during the stance phase of gait: implications for stress fractures. *J Bone Joint Surg Am*. 1999;81(9):1236–44.
12. Durnin JV, Rahaman MM. The assessment of the amount of fat in the human body from measurements of skinfold thickness. *Br J Nutr*. 1967;21(3):681–9.
13. Fleute M, Lavallee S. Building a complete surface model from sparse data using statistical shape models. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI'98, First International Conference*: Cambridge, MA: Springer; 1998. p. 879–87. ISBN: 3-540-65136-5.
14. Gaeta M, Minutoli F, Scribano E, et al. CT and MR imaging findings in athletes with early tibial stress injuries: comparison with bone scintigraphy findings and emphasis on cortical abnormalities. *Radiology*. 2005;235(2):553–61.
15. Garcia-Aznar JM, Rueberg T, Doblare M. A bone remodeling model coupling microdamage growth and repair by 3D BMU-activity. *Biomechan Model Mechanobiol*. 2005;4:147–67.
16. Giladi M, Milgrom C, Simkin A, Danon Y. Stress fractures. Identifiable risk factors. *Am J Sports Med*. 1991;19(6):647–52.
17. Hod N, Ashkenazi I, Levi Y, et al. Characteristics of skeletal stress fractures in female military recruits of the Israel defense forces on bone scintigraphy. *Clin Nucl Med*. 2006;31(12):742–9.
18. Jones BH, Thacker SB, Gilchrist J, Kimsey CD Jr, Sosin DM. Prevention of lower extremity stress fractures in athletes and soldiers: a systematic review. *Epidemiol Rev*. 2002;24(2):228–47.
19. Lai YM, Qin L, Hung VW, Chan KM. Regional differences in cortical bone mineral density in the weight-bearing long bone shaft-a pQCT study. *Bone*. 2005;36(3):465–71.
20. Lassus J, Tulikoura I, Konttinen YT, Salo J, Santavira S. Bone stress injuries of the lower extremity: a review. *Acta Orthop Scand*. 2002;73(3):359–68.
21. Matheson GO, Clement DB, McKenzie DC, Taunton JE, Lloyd-Smith DR, MacIntyre JG. Stress fractures in athletes. A study of 320 cases. *Am J Sports Med*. 1987;15(1):46–58.
22. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med*. 1978;8:283–98.
23. Milgrom C, Finestone A, Shlamkovitch N, et al. Stress fracture treatment. *Orthopedics*. 1995;3:363–7.
24. Muir P, Johnson KA, Ruaux-Mason CP. In vivo matrix microdamage in a naturally occurring canine fatigue fracture. *Bone*. 1999;25(5):571–6.
25. Negus CH, Impelluso TJ. Continuum remodeling revisited: deformation rate driven functional adaptation using a hypoelastic constitutive law. *Biomech Model Mechanobiol*. 2007;6(4): 211–26.
26. Novack V, Finestone AS, Constantini N, Shpilberg S, Weitzman S, Merkel D. The prevalence of low hemoglobin values among new infantry recruits and nonlinear relationship between hemoglobin concentration and physical fitness. *Am J Hematol*. 2007;82(2): 128–33.
27. Rajamani K, Joshi SC, Styner M. Bone model morphing for enhanced visualization. In: *IEEE Symposium on Biomedical Imaging*. 2004. p. 1255–8. ISBN: 0-7803-8388-5.
28. Rauh MJ, Macera CA, Trone DW, Shaffer RA, Brodine SK. Epidemiology of stress fracture and lower-extremity overuse injury in female recruits. *Med Sci Sports Exerc*. 2006;38(9): 1571–7.
29. Rittweger J, Beller G, Ehrig J, et al. Bone-muscle strength indices for the human lower leg. *Bone*. 2000;27:319–26.
30. Shahar D, Shai I, Vardi H, Brener-Azard A, Fraser D. Development of a semi-quantitative Food Frequency Questionnaire (FFQ) to assess dietary intake of multiethnic populations. *Eur J Epidemiology*. 2003;18:855–61.
31. Sievanen H, Koskue V, Rauhio A, Kannus P, Heinonen A, Vuori I. Peripheral quantitative computed tomography in human long bones: evaluation of in vitro and in vivo precision. *J Bone Miner Res*. 1998;13(5):871–82.
32. Silverberg DS, Iaina A, Wexler D, Blum M. The pathological consequences of anaemia. *Clin Lab Haematol*. 2001;23(1):1–6.
33. Skedros JG, Bloebaum RD, Mason MW, Bramble DM. Analysis of a tension/compression skeletal system: possible strain-specific differences in the hierarchical organization of bone. *Anat Rec*. 1994;239(4):396–404.
34. Turner CH, Pavalko FM. Mechanotransduction and functional response of the skeleton to physical stress: the mechanisms and mechanics of bone adaptation. *J Orthop Sci*. 1998;3(6): 346–55.
35. Valimaki VV, Alfthan H, Lehmuskallio E, et al. Risk factors for clinical stress fractures in male military recruits: a prospective cohort study. *Bone*. 2005;37(2):267–73.
36. Wasserman N, Yerramshetty J, Akkus O. Microcracks colocalize within highly mineralized regions of cortical bone tissue. *Eur J Morphol*. 2005;42(1–2):43–51.