

1 **TITLE:**

2 Development of a prediction model for stress fracture during an intensive physical training

3 program: The Royal Marines Commandos

4

5 ABSTRACT

6 **Background:** Stress fractures (SF) are one of the more severe overuse injuries in military
7 training and therefore knowledge of potential risk factors is needed to assist in developing
8 mitigating strategies.

9 **Purpose:** To develop a prediction model for risk of SF in Royal Marines (RM) recruits
10 during an arduous military training program.

11 **Study Design:** Cohort study

12 **Methods:** 1,082 recruits (age range: 16-33 years) enrolled between September 2009 and July
13 2010, were prospectively followed through the 32-week RM training program. SF diagnosis
14 was confirmed from a positive X-Ray or Magnetic Resonance Imaging (MRI) scan. Potential
15 risk factors assessed at week-1 included recruit characteristics, anthropometric assessment,
16 dietary supplement use, lifestyle habits, fitness assessment, blood samples, 25(OH)D, bone
17 strength as measured by heel Broadband Ultrasound Attenuation (BUA), history of physical
18 activity, and previous and current food intake. A logistic least absolute shrinkage selection
19 operator (LASSO) regression with 10-fold cross-validation method was used to select
20 potential predictors among 47 candidate variables. Model performance was assessed using
21 measures of discrimination (c-index) and calibration. Bootstrapping was used for internal
22 validation of the developed model and to quantify optimism.

23 **Results:** A total of 86 (8%) volunteer recruits presented with at least one SF during training.
24 Twelve variables were identified as the most important risk factors of SF. Variables strongly
25 associated with SF were age, body weight, pre-training weight bearing (WB) exercise, pre-
26 training cycling and childhood intake of milk and milk products. The c-index for the
27 prediction model was 0.73 (optimism-corrected c-index 0.68), which represents the model
28 performance in future volunteers. Although 25(OH)D and VO_{2max} had only a borderline
29 statistical significant association with SF, the inclusion of these factors improved the

30 performance of the model.

31 **Conclusion:** These findings will assist in identifying recruits at greater risk of SF during
32 training, and support interventions to mitigate this injury risk. However, external validation
33 of the model is still required.

34 **Key words:** Stress Fracture, risk factors, prediction, Military, Royal Marines

35

36 **What is known about the subject:**

37 Many of the variables identified as important risk factors have been shown to predict the risk
38 of SF in earlier studies, such as age and body weight, VO2 max and 25(OH)D concentrations.

39

40 **What this study adds to existing knowledge:**

41 This study has produced the first risk prediction model for stress fracture during elite military
42 training, using a wide variety of risk factors and excellent follow up data. We used a
43 powerful penalized regression method (Lasso regression) to select important predictors. A
44 novel finding from this study was to identify new exercise and diet variables as predictors of
45 SF.

46

47 **Introduction**

48 The 32-week Royal Marines (RM) recruit training is one of the most arduous and longest
49 initial military training programs in the world. Recruits are at relatively high risk of
50 musculoskeletal overuse injuries (30%), including stress fracture (SF), which represents one
51 of the more severe overuse injuries in military training ⁴².

52 SF is a partial or complete fracture of bone and it occurs when bones are repetitively loaded
53 with vigorous weight-bearing exercise, over short time periods without sufficient time for
54 repair. Existing research in military and athletic populations have documented that the
55 incidence of SF ranges from 0.7% to 31% ^{14, 17, 24}. The most prevalent site of SF is generally
56 the tibia, followed by the metatarsal bones ¹⁸. Furthermore, it is the single most common
57 cause for lost training days, and represents a significant cost in terms of medical support and
58 rehabilitation time, as well as increasing the likelihood of recruits leaving training prior to
59 completion.

60 The aetiology of SF is multifactorial, and knowledge of potential risk factors is required to
61 assist in developing mitigating strategies. Previous prospective studies in military populations
62 have assessed SF risk factors in the Indian Army ⁸, US military ^{1, 27, 28, 32, 47, 48}, Israel military
63 ^{14, 15, 33}, and Finnish Army ^{31, 55}, all of which range between 8 and 16 weeks of training.
64 However, there are only a few data assessing risk factors during longer training programs,
65 such as RM training ^{9, 10, 36}. Previous systematic reviews have identified the following SF-risk
66 factors ^{18, 37, 57}: older-age, female gender, lifestyle habits (smoking and alcohol ingestion),
67 low bone mineral density (BMD), previous lower limb injury and poor nutrition. Because all
68 of these variables are of limited predictive value when considered in isolation, the
69 combination of these variables is needed for better predictive accuracy.

70 The problem of selecting a set of potential risk factors to include in regression modelling is
71 well known, but it is also among the most controversial and difficult tasks in epidemiologic

72 analysis ⁵⁶. Selecting a practical number of predictors to be included in the model is the
73 natural first step, and they are generally selected based upon subject knowledge from clinical
74 expertise and reviews of the literature ^{23, 50, 51}. Currently, there is a lack of research on the risk
75 factors for SF in an elite military setting; therefore a more statistically driven method of
76 predictor selection was utilised. For this study, given the sample size and incidence of
77 outcome, standard variable selection methods (forward selection and/or backward
78 elimination) could lead to only a limited number of predictors being considered to avoid
79 model overfitting. An alternative method that overcomes this limitation is lasso regression
80 (least absolute shrinkage and selection operator) and is the most widely used ^{38, 49}. Lasso is a
81 powerful penalized regression method used in predictor selection ³⁴.
82 The purpose of this study was to identify the most probable SF risk factors in RM recruits at
83 the start of military training, and construct a SF prediction model using advanced statistical
84 methods. The resulting better understanding of the interrelationship between SF risk factors
85 would assist in developing evidence-based preventive interventions and safety promotion
86 programs for mitigating SF in the military.

87

88

89 **Methods**

90 In this study, we followed the TRIPOD (transparent reporting of a multivariable prediction
91 model for individual prognosis or diagnosis) statement ⁶ to report the prediction model
92 including model development, model performance and model internal validation.

93 **Study Population**

94 The present study utilised data from Phase-II of *Surgeon General's Bone Health Project*
95 (SGBHP). SGBHP adopted a prospective, observational study design, to assess the
96 relationship between nutritional influences on bone health and SF occurrence during the 32-
97 week recruit training program at the Commando Training Centre Royal Marines (CTCRM),
98 [country deleted to maintain the integrity of the review process]. The RM is an all-male elite
99 fighting force; therefore, there were no females in the sample. Recruits who successfully
100 completed the physical and professional selection tests, and were deemed medically fit and
101 healthy to undertake RM training, were eligible to participate in the study.

102 A total of 1,113 recruits from twenty troops of RM recruits, commencing training between
103 September 2009 and July 2010, were invited to participate in this study. Written informed
104 consent was obtained from volunteer recruits ($n=1,090$; 98% response rate) (age range 16-
105 33); $n=23$ recruits declined to participate in the study. A further eight recruits were
106 discharged during the first week due to pre-existing medical conditions, leaving 1,082
107 participants.

108

109 **Ethical Considerations**

110 The study was approved by the [text deleted to maintain the integrity of the review process].

111

Primary Outcome – Stress Fracture

Recruits reporting to the Medical Centre with symptoms of a potential SF underwent examination and X-Ray or Magnetic Resonance Imaging (MRI) scanning to confirm SF diagnosis as part of routine care. SF diagnosis was based on a positive X-Ray or MRI scan. Depending on the fracture site, a negative initial X-Ray was followed-up by a further X-Ray or MRI to confirm diagnosis. All recruits with SF were removed from RM training and underwent rest and rehabilitation *in situ* under medical supervision.

Potential Risk Factors

Based on the information available in the study and on previous research in the scientific literature, 47 potential predictors from measurements collected, at the start training, were selected to be included in the modelling analyses in the areas of physical fitness, diet, lifestyle, education, season for commencing RM training, and measures of bone health.

Recruit information

Age, education and the season of the year at the start training were available. As the association of age on SF was non-linear, it was divided into three categories based on distributions from a previous study associated with discharge in US Marine Corps recruits⁴⁰: less than 19 years (range 16-18 years); 19 to 23 years; 24 years or older (range 24-32 years). Education was defined as secondary school versus further education/degree, and year season when the recruit started training was divided into the standard 4 seasons (autumn, winter, spring and summer).

Anthropometric assessments

Height and body weight were measured. Height was measured in centimetres (to the nearest 0.1 cm) in a standing position, with shoes/boots removed, on the stadiometer (Invicta, Leicester, England), with feet together. Body weight was measured in kilograms (Seca, Hamburg, Germany) in standard issue t-shirt and shorts.

Self-reported dietary supplement usage self-reported

Multivitamin and minerals, creatine, sports/energy drinks/ energy gels and protein bars/ powder/ shakes were included in an assessment of self-reported pre-RM training dietary supplement usage. Categorical variables of intake frequency for each dietary supplement were generated (never, sometimes and everyday).

Lifestyle habits

Smoking habit and alcohol intake were assessed. Recruits were classified as never smokers, ex-smokers, and current smokers. Alcohol intake was considered relative to intake in units per week and it was used as a continuous variable.

Broadband Ultrasound Attenuation Measurement (BUA)

Broadband Ultrasound Attenuation Measurement (BUA) was assessed on the dominant and non-dominant foot (dB.MHz-1) as an indicator of bone strength²⁶. It is considered to be a rapid, safe and relatively inexpensive technique for measuring skeletal status²⁰. This measure was taken across the *calcaneum* of a seated recruit. A continuous score was used for the analysis, where a greater BUA was indicative of higher bone mass and greater bone strength.

Royal Marine Fitness Assessment (RMFA)

The RMFA is composed of four parts, which include the Multistage Fitness Test (MSFT)³ to estimate maximum oxygen uptake ($\text{VO}_{2\text{max}}$)²⁹, a push-up test, a sit-up test and a pull-up test. The four fitness tests were undertaken in the gymnasium with recruits wearing shorts, t-shirt and training shoes. The recruits were required to do the maximum number of push-ups, sit-ups and pull-ups in 60 seconds. All measures were treated as continuous for the analysis.

Exercise pre-RM training preparation

Mode, duration, frequency and volume of exercise pre-RM training preparation were included. Mode of training was assessed by the amount of weight and non-weight bearing (WB) exercises from a list of three WB exercises (i.e. running, circuit training and weight training) and two non-WB exercises (i.e. cycling and swimming). Duration was assessed by the number of weeks of pre-RM training preparation, frequency by the number of training sessions per week and volume by the minutes per week training.

Previous lower limb injuries (dominant leg and non-dominant leg) were also self-reported by recruits.

Assessment of micronutrient and vitamin D status

A non-fasting blood sample was drawn by medical personnel, using serum separation vacutainers. Serum samples were provided for magnesium (as marker of micromineral status), zinc, selenium, copper (as markers of trace element status), and serum 25(OH)D concentration (as marker of vitamin D status). A threshold of 50 nmol L⁻¹ for 25(OH)D was used for the analysis¹⁰.

Physical activity and dietary intake measurements

A validated questionnaire, the Food Frequency Questionnaire (FFQ)^{11, 35}, which examines childhood, adolescence and current diet and physical activity levels was administered to recruits at the beginning of RM training.

Dietary intake focused on recruit eating choices just prior to commencing RM training (the last month), as well as during childhood and adolescence, as an assessment of habitual dietary patterns. Milk, milk products, vegetables and fruit were included. Intake of each group was determined in times/week (except vegetables and fruit were determined in portions/week).

Current activity levels were assessed by number of minutes walked per day, number of minutes cycled per day³⁹ and by the following two questions: “during your working time and during your non-working time, how often during a normal week were you physically active for at least 20 minutes during which time you became short of breath and sweat?” Recruits were classified into three groups: Once or less per week, 2-3 times per week, and more than 3 times per week. Physical activity throughout childhood and early adulthood (i.e. 0-12 years and 12-18 years) was assessed by asking recruits how often they were normally physical active for at least 20 minutes during which time they became short of breath and sweaty: Once or less per week, 2-6 times per week, more than 6 times per week.

Table 1 provides a detailed description of all covariates recorded.

204 **Table 1.** List of prognostic variables for stress fractures available for analysis

Variables	Additional information	Variables	Additional information
Age	1) 16-18 y; 2) 19 to 23 y; 3) 24-32y	RMFA	
Education	1) Secondary school; 2) Further education/Degree	VO _{2max}	ml.kg ⁻¹ min ⁻¹ (continuous)
Season at start Training	1) Autumn; 2) Winter; 3) Spring; 4) Summer	Push-up	Counts (continuous)
Anthropometric assessment		Sit-up tests	Counts (continuous)
Height	Meters ² (continuous)	Pull-up test	Counts (continuous)
Body weight	Kilograms (continuous)	Exercise pre-RM training	
Dietary supplement usage self-reported		Amount of weight bearing exercise	1) 0-2 WB exercise; 2) 3 WB exercise
Multivitamins with Minerals	1) Never; 2) Sometimes; 3) Everyday	Amount of non-weight bearing exercise	1) 0-1 non WB exercise; 2) 2 non WB exercise
Creatine	1) Never; 2) Sometimes; 3) Everyday	Duration	Numbers of weeks (continuous)
Sports/energy-drinks/energy gels	1) Never; 2) Sometimes; 3) Everyday	Frequency per week	Times.week (continuous)
Protein bars/powder/shakes	1) Never; 2) Sometimes; 3) Everyday	Weekly training volume	Minutes (continuous)
Lifestyle Habits		Previous injury	
Smoke	1) Never; 2) Ex-smoker; 3) Current	Lower limb injury in dominant leg	Yes or No
Alcohol	Units per week (continuous)	Lower limb injury in non-dominant leg	Yes or No
BUA			
Dominant foot	dB.MHz ⁻¹ (continuous)		
Non dominant foot	dB.MHz ⁻¹ (continuous)		
BUA, Broadband Ultrasound Attenuation Measurement; RMFA, Royal Marine Fitness Assessment, VO _{2max} , Maximum Oxygen Uptake; RM, Royal Marine			

205
206

207 **Table 1.** List of prognostic variables for stress fractures available for analysis (continued).

Variables	Additional information	Variables	Additional information
Blood Sample		Dietary intake FFQ	
Magnesium	$\mu\text{mol/L}^{-1}$ (continuous)	<i>Current</i>	
Zinc	$\mu\text{mol/L}^{-1}$ (continuous)	Milk intake	1) Low ($<285 \text{ ml.d}^{-1}$); 2) Moderate (426 ml.d^{-1}); 3) High ($\geq 852 \text{ ml.d}^{-1}$)
Selenium	$\mu\text{mol/L}^{-1}$ (continuous)	Milk products intake	1) Low ($<2 \text{ times.week}$); 2) Moderate ($2-5 \text{ times.week}$); 3) High ($\geq 5 \text{ times.week}$)
Copper	$\mu\text{mol/L}^{-1}$ (continuous)	Fruit intake	1) Low ($<8 \text{ portions.week}$); 2) Moderate ($8-13 \text{ portions.week}$); 3) High ($\geq 14 \text{ portions.week}$)
25(OH)D	Less than 50 nmol.L^{-1} vs. 50 nmol.L^{-1} or more	Vegetables	1) Low ($<14 \text{ portions.week}$); 2) Moderate ($14-22 \text{ portions.week}$); 3) High ($\geq 23 \text{ portions.week}$)
FFQ: Physical activity		<i>Adolescence and childhood</i>	
<i>Current</i>		Milk intake	1) Low ($<285 \text{ ml.d}^{-1}$); 2) Moderate (426 ml.d^{-1}); 3) High ($\geq 852 \text{ ml.d}^{-1}$)
Minutes per day of Walking	1) Less than 30 min per day; 2) 30 or more min per day	Milk products intake	1) Low ($<4 \text{ times.week}$); 2) Moderate ($4-5 \text{ times.week}$); 3) High ($\geq 6 \text{ times.week}$)
Minutes per day of Cycling	1) Less than 30 min per day; 2) 30 or more min per day	Fruit intake	1) Low ($<4 \text{ times.week}$); 2) Moderate ($4-5 \text{ times.week}$); 3) High ($\geq 6 \text{ times.week}$)
Physical activity during working time	1) Once or less a week; 2) 2-3 times a week; 3) More than 3 times a week	Vegetables	1) Low ($<4 \text{ times.week}$); 2) Moderate ($4-5 \text{ times.week}$); 3) High ($\geq 6 \text{ times.week}$)
Physical activity during non-working time	1) Once or less a week; 2) 2-3 times a week; 3) More than 3 times a week		
<i>Adolescence and childhood</i>			
Physical activity	1) Once or less a week; 2) 2-6 times a week; 3) More than 6 times a week		

FFQ= Food Frequency Questionnaire

Statistical Analyses

Descriptive statistics of all potential predictors, according to whether the recruit presented or did not present a SF during RM training, were examined using means (standard deviation) or medians (interquartile range) for quantitative measures, and frequency (percentage) for categorical variables.

Linearity assumption for continuous variables (using fractional polynomials or linear splines) was assessed, and the presence of interactions between age and the other variables was tested⁴³. To fill in variables with missing values, and because there was less than 15% missing data (see Online Resource 1 - Table S1), a stochastic simple imputation (SI) method was used. It was created as the first of a series of 10 multiple imputations (MI) using MICE (Multiple Imputation by Chained Equation)⁴⁵ (see Online Resource 1 - Table S2). All pre-specified predictors were included in the imputation model, together with the outcome.

The predicting model was achieved in two steps. First, Lasso shrinkage logistic regression method⁵⁴ was used to reduce the final model to the most important variables to predict SF. It shrinks the coefficient estimates toward zero, with the degree of shrinkage depending on an additional parameter, lambda (λ) (this study uses $\lambda = 1$). A single model adjusted for all potential variables was fitted with a 10-fold cross-validation and the minimum average mean-squared error (MSE) to extract the non-zero coefficients and therefore the significant predictors. This method focuses on the overall fit (best model fit) rather than statistical significance of individual predictors. As a consequence, predictors with a p-value > 0.05 could still be included in the final model. Second, odds ratios (OR) and 95% confidence of intervals (CI) comparing the proportion of recruits in a high-risk group at baseline with the proportion of recruits in a referent group, using classic logistic regression model were estimated for the principal risk factors selected in the previous step.

Internal validity

To check the internal validity of the model, 200 bootstrap samples with replacement was used to assess bias-corrected estimates of predictive ability²². The evaluation of the model

performance considered measures of discrimination and calibration^{5, 44}. Discrimination was assessed using the c-index (this value varies between 0 and 1, where 1 represents perfect discrimination)²¹. In logistic regression c-index is identical to the area under the receiver operating characteristic (ROC) curve. Calibration was assessed by calibration plots. All calculations were performed using Stata statistical software version 13.1 (StataCorp, College Station, Texas, USA) and R statistical software, version 3.2.3 (R Foundation for Statistical Computing, Vienna, Austria). Variable selection and internal validation of the model were performed using the ‘glmnet’¹⁹ and the ‘rms’¹⁶ packages, respectively.

Results

Recruit characteristics

During Phase-II of SGBHP, a total of eighty six recruits (8% of study cohort) suffered at least one SF during the 32-week training period, with the metatarsal as the most common injury site (44 recruits), followed by the tibia and fibula (34 recruits). The majority of SF occurred in the latter 15 weeks of RM training (~80%) (Figure 1). The highest frequency was in week-31 (17.3%), followed by week-17 (12.4%) and week-22 (11.1%).

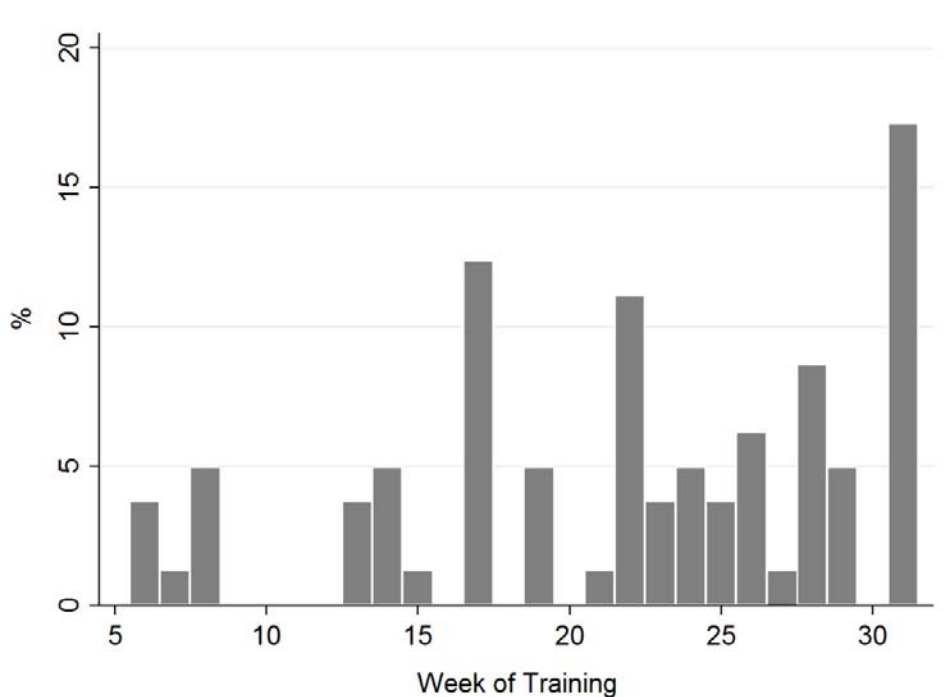


Figure 1. Distribution (%) of first stress fracture in 1,082 recruits during 34-week training

The proportion of missing data for each variable included in this study is shown in Online Resource 1 - Table S1. The missing data represented less than or equal to 10% for all variables.

Distributions of the all-potential predictors by SF status are presented in Online Resource 1 - Table S3. Significant differences ($p < 0.05$) between recruits who sustained a SF and those who did not were found for age, body weight, BUA of the dominant foot, volume of weight bearing exercise pre-RM training, minutes per day of cycling, and milk intake during childhood. Compared with the recruits without SF, recruits with at least one SF during training were older, had a lower body weight and lower units of BUA of the dominant foot. In addition, recruits who reported a high frequency of WB exercise pre-RM training preparation (3 WB exercise), cycling 30 minutes or more per day and low intake of milk during their childhood were more likely to incur a SF.

Selection of principal predictors of Stress Fracture

Twelve variables were selected by the Lasso selection approach. The model contained age, body weight, smoking habit, BUA of the dominant foot, VO_{2max} , 25(OH)D concentration, amount of WB exercise, minutes cycling per day, physical activity during childhood, vegetables intake during adolescence, and milk and milk products intake during childhood.

Predictors ORs and 95%CI are presented in Table 2. Higher SF risk was associated with older age group (range 24-32 years) (OR: 1.98, 95%CI: 1.07 to 3.55) compared with recruits between 19 to 23 years old. In addition, increased WB exercise pre-RM training (OR: 2.46, 95%CI: 1.51 to 4.00), high frequency of cycling per day pre-RM training preparation (OR: 1.71, 95%CI: 1.07 to 2.74) and high milk products intake during childhood (OR: 1.84, 95%CI: 1.03 to 3.30) were also associated with increased risk of SF.

In contrast, variables strongly associated with a decreased risk of SF were high body weight (OR: 0.96, 95%CI: 0.93 to 0.99); and high milk intake (OR: 0.45, 95%CI: 0.23 to 0.86)

281 compared with low intake, during childhood. Variables with a borderline statistical
282 significant association with SF included VO_{2max} and 25(OH)D. Recruits with poor aerobic
283 fitness at the start training and low concentrations of 25(OH)D (less than 50 nmol.L⁻¹) were
284 at an increased risk of SF during training.

285 Smoking habit, bone strength of the dominant foot, physical activity during childhood and
286 vegetables intake during adolescence contributed to the overall model performance although
287 they did not have a statistically significant association with SF.

288

Table 2. Estimation of OR and 95%CI for principal risk factors of stress fracture during Royal Marine recruit training

Predictor variables (reference category)	OR	95% CI	P-value
Age, years (19 to 23 year)			
<i><19 years</i>	1.66	0.97 to 2.85	0.066
<i>More than 23 year</i>	1.98	1.07 to 3.55	0.030
Body weight, kg	0.96	0.93 to 0.99	0.018
Smoke (Never)			
<i>Ex-smoker</i>	1.55	0.91 to 2.64	0.109
<i>Current</i>	0.73	0.32 to 1.64	0.447
BUA of the dominant foot, dB.MHz ⁻¹	0.99	0.98 to 1.00	0.150
VO _{2max} , ml.kg ⁻¹ min ⁻¹	0.93	0.86 to 1.01	0.074
Amount of weight bearing exercise (0-2 WB exercise)			
<i>3 WB exercise</i>	2.46	1.51 to 4.00	<0.001
25(OH)D, nmol.L ⁻¹ (<i>50 nmol.L⁻¹ or more</i>)			
<i>Less than 50 nmol.L⁻¹</i>	1.56	0.95 to 2.56	0.077
Minutes per day of Cycling (Less than 30 min per day)			
<i>30 or more min per day</i>	1.71	1.07 to 2.74	0.026
Physical activity when childhood (Once or less a week)			
<i>2-6 times a week</i>	1.27	0.54 to 2.96	0.584
<i>More than 6 times a week</i>	1.76	0.79 to 3.93	0.165
Adolescence			
Vegetables, times.week (Low (<4))			
<i>Moderate (4-5)</i>	0.67	0.36 to 1.24	0.200
<i>High (≥6)</i>	0.91	0.51 to 1.60	0.735
Childhood			
Milk intake, ml.d (Low (<285))			
<i>Moderate (426)</i>	0.95	0.57 to 1.61	0.862
<i>High (≥852)</i>	0.45	0.23 to 0.86	0.016
Milk products, times.week (Low (<4))			
<i>Moderate (4-5)</i>	1.19	0.67 to 2.11	0.550
<i>High (≥6)</i>	1.84	1.03 to 3.30	0.039
Model intercept	30.8		
C-index- fitted model (95% CI)	0.73 (0.67-0.78)		
Optimism	0.5		
Bias-corrected C-index	0.68		

Abbreviations: OR, odd ratio; 95% CI, 95% confidence intervals; BUA, Broadband Ultrasound Attenuation; RMFA, Royal Marine Fitness Assessment; RM, Royal Marine; WB, Weight bearing; FFQ, Food Frequency Questionnaire.

The performance of the model showed adequate calibration and discrimination with a c-index of 0.73 (95%CI: 0.67-0.78) (Figure 2A and 2B, respectively). Using bootstrap validation, the optimism-corrected c-index was 0.68, which indicated a moderate predictive model in future volunteers.

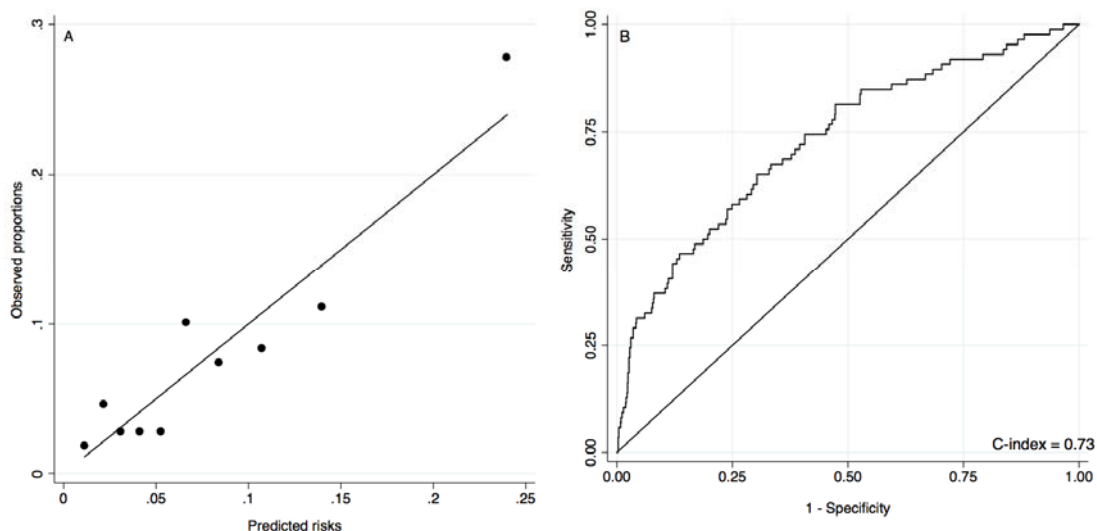


Figure 2. A) Calibration curve of the prediction model of the risk of stress fracture during 34-week Marines training. The solid line indicates perfect calibration. B) Receiver operating characteristic (ROC) curve plot to assess discrimination of the predictive model.

Discussion

This study identified pre-training predictors of developing SF during the 32 weeks of RM recruit training, by using advanced statistical methods.

Age, body weight, exercise pre-RM training and childhood milk intake were the strongest predictors of SF in the model. Although 25(OH)D, VO_{2max} , smoking habit, bone strength of the dominant foot, physical activity during childhood and vegetables intake during adolescence were weakly associated with SF ($p>0.05$), the inclusion of these factors improved the performance of the model, hence, the effects attributable to these factors were small but important to explain the outcome. The predictive model had a reasonable prediction capacity and validity to identify SF in RM recruits during the training period (see Table 2).

What do we already know?

The occurrence of SF during training in this population was 8% and it was consistent with rates reported for other military recruit populations in Europe⁵⁵. Many of the variables identified as important risk factors have been shown to predict the risk of SF in earlier studies^{24, 53}, supporting the plausibility of this model.

Age and body weight were significantly associated with SF. Older age (24-32 years) predicted SF in the present population, compared with recruits between 19-23 years old, which is in agreement with previous findings in other military and non-military populations^{4, 31}. Higher body weight was significantly associated with a decreased risk of SF, and this association has been well documented^{15, 36}

Within this study, VO_2 max and 25(OH)D concentrations were borderline significantly associated with SF during training. Recruits with lower aerobic fitness assessed in week-1 of training were more likely to have an increased risk of SF during the following 32 weeks of training. These finding were consistent with a previous report that used Cooper's test as a measure of physical fitness⁵⁵. However, a lack of association between VO_{2max} and SF incidence has been reported in a study of Israeli infantry recruits⁵². It should be noted that

recruits in the present study were volunteers joining one of the UK's elite Service Arms, where recruitment is partly based on having a high aerobic fitness at the start of training.

There is controversy with respect to 25(OH)D; previous studies found an association between low serum 25(OH)D levels and increased SF risk¹⁰, whereas others have found no evidence of an association in military personnel⁷. The present study found that low levels of 25(OH)D (less than 50 nmol.L⁻¹) were associated with greater risk of SF during training, compared with recruits with higher levels of 25(OH)D, and this result was consistent with previous studies¹⁰.

The present study aimed to produce a statistical model with the optimum predictive ability, and not to formally assess the association of 25(OH)D with SF *per se*. Thus, the scope of this present study, in developing an expansive prediction model, may explain the weak association for 25(OH)D.

What does this study add?

This study has produced the first risk prediction model for stress fracture during elite military training, using a wide variety of risk factors and excellent follow up data. This model had a reasonable predictive capacity and validity to identify occurrence of SF. After appropriate external validation, this model may be useful in helping to identify recruits as well as sports men at greater risk of SF, and hence contribute to the development of strategies to mitigate SF risk.

Although there are several well established risk factors for SF^{18, 37, 57}, a novel finding from this study is to identify new exercise and diet variables as predictors of SF, using an advanced variable selection method⁵⁴. Increased WB exercise pre-training was a significant risk factor of SF occurrence. This result does not agree with the majority of previous basic military studies that have reported that recruits who have physically active lifestyle in the past would be less likely to suffer SF when starting a vigorous exercise program^{8, 25, 47}. Possible explanations for this may be first, as most SF of foot and ankle are caused by repetitive vigorous WB activities such as running and marching, and they usually occur when

individuals change their activities (e.g. such as trying a new exercise, increasing the intensity of their workout, or changing the workout surface), well-conditioned individuals who have been preparing to join the RM could suffer injuries during the military training. Second, differences could be due to collecting pre-RM previous training data. High frequency of cycling per day pre-RM training preparation was associated with increased risk of SF. Cycling is an aerobic, non-weight bearing sport that has been associated to lower bone mineral density (BMD) ⁴⁶. Since BMD has been found to be a predictor of high risk of SF ⁹, high intensity of cycling pre-RM training may contribute to the development of SF due to its influence on bone development.

An interesting finding in the present was the association between SF and dietary intake during childhood. The models show that recruits who reported a high intake of milk had a smaller risk of SF than recruits with low intake of milk. A possible explanation for this is that in growing children, long-term avoidance of milk is associated with smaller stature and poorer BMD ⁴¹. As mentioned above, BMD is associated with SF, hence inadequate childhood calcium intake may impact on SF development during training due to its influence on bone health. A significant skeletal growth phase at childhood age may be particularly important to the prevention of SF.

In contrast, high intake of milk products during childhood was statistically associated with increased risk of SF compared with a low intake. Milk products (e.g. cheese, yoghurt, etc.) could be related to fat products, so a high fat dietary pattern could be associated with greater risk of SF, as inadequate nutritional intake may alter bone metabolism and predispose toward appearance of SF. Another possible explanation for this would be the lack of precision on reporting milk products.

To the authors' knowledge, this is the first study to report an association between past dietary intake and risk of SF during male military training.

Strengths and potential limitations

This study has a number of strengths including unique prospective data, a wide range of potential risk factors and low proportion of missing data. The study used a rigorous and robust variable selection method to reduce the number of SF potential risk factors. For prognostic studies, the lasso regression could select the most important variables much more efficiently than the standard variable selection methods, by omitting additional and redundant variables³⁴. Lasso controls multicollinearity and is also applicable in settings where the number of variables is higher than the sample size where traditional logistic regression would fail⁵⁴. Twelve factors were identified to be included into the final model, but eight factors were identified as the most important risk factors of SF occurrence during RM training. Established measures of prediction performance, including the overall model fit, discrimination, and calibration, suggested that the final model had a satisfactory performance. A further strength of this study was the use of the RM data set, where there had been very high recruitment and retention rates.

There are several potential limitations to this study. First, by using the lasso methods, which variables are "clinically important" cannot be defined because if two predictors were perfectly collinear, the lasso will pick one of them essentially at random. Second, there are currently no common methods to incorporate multiple imputation with lasso, therefore single imputation (SI) was used. SI may underestimate associations, and point estimates are potentially unstable, however the low level of missingness (<10%) in the data combined with the size of the dataset make this unlikely⁴⁹. Third, although the inclusion of recruits who did not complete training for non-injury reason (n=465) in logistic regression analysis may introduce bias, excluding them would reduce the statistical power and validity of the study. An analysis excluding these recruits was performed and no significant impact on the results and interpretation of this study was found. Fourth, the results of this study were restricted to male military personnel, which may not be generalizable to women or the normal population. Future research should focus on the relative contribution of general population and gender-

specific conditions. Fifth, residual bias may exist. The development of the model in the present study only took into account variables at the start training. Other possible factors during training could explain the high variability in outcome. Sixth, self-reported past physical activity and diet is subject to the weakness of recall bias. However, past measures have been found to be positively correlated with those recorded objectively at the same time period^{2, 12, 13, 30, 58}, and hence the FFQ, used in this study, represents a valid instrument of assessing past physical activity and food intake. Seventh, a SF may have remained undiagnosed for several weeks and even have been not reported in some recruits, so SF may have been underreported in this study. Finally, these results require validation through further prospective studies to improve the predictive capacity of the model. However, the results of the present study provide new important predictors of SF cases.

Conclusion

In conclusion, this model has provided an important contribution to the prediction of SF during RM training, identifying high-risk recruits for targeted injury prevention studies. SF risk during training may be modified through adjustments to selection. Information from this study could be used to determine recruits at risk of developing a SF. Further replication in additional data sets may lead to further enhancement of the current model for RM and other military training programs.

- 429 1. Almeida SA, Williams KM, Shaffer RA, Brodine SK. Epidemiological patterns of
430 musculoskeletal injuries and physical training. *Med Sci Sports Exerc.*
431 1999;31(8):1176-1182.
- 432 2. Besson H, Harwood CA, Ekelund U, et al. Validation of the historical adulthood
433 physical activity questionnaire (HAPAQ) against objective measurements of physical
434 activity. *Int J Behav Nutr Phys Act.* 2010;7:54.
- 435 3. Brewer J, Ramsbottom R, Williams C. Multistage Fitness Test: A Progressive
436 Shuttle-Run Test for the Prediction of Maximum Oxygen Uptake: Loughborough
437 University; National Coaching Foundation, Leeds, UK. 1988.
- 438 4. Buist I, Bredeweg SW, Bessem B, van Mechelen W, Lemmink KA, Diercks RL.
439 Incidence and risk factors of running-related injuries during preparation for a 4-mile
440 recreational running event. *Br J Sports Med.* 2010;44(8):598-604.
- 441 5. Collins GS, Mallett S, Altman DG. Predicting risk of osteoporotic and hip fracture in
442 the United Kingdom: prospective independent and external validation of
443 QFractureScores. *BMJ.* 2011;342:d3651.
- 444 6. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a
445 multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the
446 TRIPOD statement. *Ann Intern Med.* 2015;162(1):55-63.
- 447 7. Dao D, Sodhi S, Tabasinejad R, et al. Serum 25-Hydroxyvitamin D Levels and Stress
448 Fractures in Military Personnel: A Systematic Review and Meta-analysis. *Am J Sports*
449 *Med.* 2015;43(8):2064-2072.
- 450 8. Dash N, Kushwaha A. Stress fractures-a prospective study amongst recruits. *Med J*
451 *Armed Forces India.* 2012;68(2):118-122.
- 452 9. Davey T, Lanham-New SA, Shaw AM, et al. Fundamental differences in axial and
453 appendicular bone density in stress fractured and uninjured Royal Marine recruits--a
454 matched case-control study. *Bone.* 2015;73:120-126.
- 455 10. Davey T, Lanham-New SA, Shaw AM, et al. Low serum 25-hydroxyvitamin D is
456 associated with increased risk of stress fracture during Royal Marine recruit training.
457 *Osteoporos Int.* 2016;27(1):171-179.
- 458 11. Delaney S, Lanham-New SA, Leiper R, et al. Validation of a modified FFQ for
459 assessing food and nutrient intake in military personnel. *Proceedings of the Nutrition*
460 *Society.* 2010;69.
- 461 12. DuBose KD, Edwards S, Ainsworth BE, Reis JP, Slattery ML. Validation of a
462 historical physical activity questionnaire in middle-aged women. *J Phys Act Health.*
463 2007;4(3):343-355.
- 464 13. Dwyer JT, Gardner J, Halvorsen K, Krall EA, Cohen A, Valadian I. Memory of food
465 intake in the distant past. *Am J Epidemiol.* 1989;130(5):1033-1046.
- 466 14. Finestone A, Milgrom C. How stress fracture incidence was lowered in the Israeli
467 army: a 25-yr struggle. *Med Sci Sports Exerc.* 2008;40(11 Suppl):S623-629.
- 468 15. Finestone A, Milgrom C, Evans R, Yanovich R, Constantini N, Moran DS. Overuse
469 injuries in female infantry recruits during low-intensity basic training. *Med Sci Sports*
470 *Exerc.* 2008;40(11 Suppl):S630-635.
- 471 16. Frank E, Harrell J. rms: Regression Modeling Strategies. Available at:
472 <http://biostat.mc.vanderbilt.edu/rms>.
- 473 17. Fredericson M, Jennings F, Beaulieu C, Matheson GO. Stress fractures in athletes.
474 *Top Magn Reson Imaging.* 2006;17(5):309-325.
- 475 18. Friedl KE, Evans RK, Moran DS. Stress fracture and military medical readiness:
476 bridging basic and applied research. *Med Sci Sports Exerc.* 2008;40(11 Suppl):S609-
477 622.
- 478 19. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear
479 Models via Coordinate Descent. *J Stat Softw.* 2010;33(1):1-22.

- 480 20. Greenspan SL, Bouxsein ML, Melton ME, et al. Precision and discriminatory ability
481 of calcaneal bone assessment technologies. *J Bone Miner Res.* 1997;12(8):1303-1313.
- 482 21. Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models,*
483 *Logistic Regression, and Survival Analysis.* New York; 2001.
- 484 22. Harrell FE, Jr. *Regression modeling strategies with applications to linear models,*
485 *logistic regression, and survival analysis:* New York: Springer; 2001.
- 486 23. Harrell FE, Jr., Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling
487 strategies for improved prognostic prediction. *Stat Med.* 1984;3(2):143-152.
- 488 24. Jones BH, Thacker SB, Gilchrist J, Kimsey CD, Jr., Sosin DM. Prevention of lower
489 extremity stress fractures in athletes and soldiers: a systematic review. *Epidemiol Rev.*
490 2002;24(2):228-247.
- 491 25. Knapik JJ, Graham B, Cobbs J, Thompson D, Steelman R, Jones BH. A prospective
492 investigation of injury incidence and injury risk factors among Army recruits in
493 military police training. *BMC Musculoskelet Disord.* 2013;14:32.
- 494 26. Langton CM, Palmer SB, Porter RW. The measurement of broadband ultrasonic
495 attenuation in cancellous bone. *Eng Med.* 1984;13(2):89-91.
- 496 27. Lappe J, Davies K, Recker R, Heaney R. Quantitative ultrasound: use in screening for
497 susceptibility to stress fractures in female army recruits. *J Bone Miner Res.*
498 2005;20(4):571-578.
- 499 28. Lappe JM, Stegman MR, Recker RR. The impact of lifestyle factors on stress
500 fractures in female Army recruits. *Osteoporos Int.* 2001;12(1):35-42.
- 501 29. Leger LA, Lambert J. A maximal multistage 20-m shuttle run test to predict VO2
502 max. *Eur J Appl Physiol Occup Physiol.* 1982;49(1):1-12.
- 503 30. Maruti SS, Feskanich D, Colditz GA, et al. Adult recall of adolescent diet:
504 reproducibility and comparison with maternal reporting. *Am J Epidemiol.*
505 2005;161(1):89-97.
- 506 31. Mattila VM, Niva M, Kiuru M, Pihlajamaki H. Risk factors for bone stress injuries: a
507 follow-up study of 102,515 person-years. *Med Sci Sports Exerc.* 2007;39(7):1061-
508 1066.
- 509 32. Montain SJ, McGraw SM, Ely MR, Grier TL, Knapik JJ. A retrospective cohort study
510 on the influence of UV index and race/ethnicity on risk of stress and lower limb
511 fractures. *BMC Musculoskelet Disord.* 2013;14:135.
- 512 33. Moran DS, Israeli E, Evans RK, et al. Prediction model for stress fracture in young
513 female recruits during basic training. *Med Sci Sports Exerc.* 2008;40(11 Suppl):S636-
514 644.
- 515 34. Musoro JZ, Zwinderman AH, Puhan MA, ter Riet G, Geskus RB. Validation of
516 prediction models based on lasso regression with multiply imputed data. *BMC Med*
517 *Res Methodol.* 2014;14:116.
- 518 35. New SA, Bolton-Smith C, Grubb DA, Reid DM. Nutritional influences on bone
519 mineral density: a cross-sectional study in premenopausal women. *Am J Clin Nutr.*
520 1997;65(6):1831-1839.
- 521 36. Nunns M, House C, Rice H, et al. Four biomechanical and anthropometric measures
522 predict tibial stress fracture: a prospective study of 1065 Royal Marines. *Br J Sports*
523 *Med.* 2016.
- 524 37. Patel DS, Roth M, Kapil N. Stress fractures: diagnosis, treatment, and prevention. *Am*
525 *Fam Physician.* 2011;83(1):39-46.
- 526 38. Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk
527 prediction model when there are few events. *BMJ.* 2015;351:h3868.
- 528 39. Pucher J, Buehler R, Merom D, Bauman A. Walking and cycling in the United States,
529 2001-2009: evidence from the National Household Travel Surveys. *Am J Public*
530 *Health.* 2011;101 Suppl 1:S310-317.
- 531 40. Reis JP, Trone DW, Macera CA, Rauh MJ. Factors associated with discharge during
532 marine corps basic training. *Mil Med.* 2007;172(9):936-941.

- 533 41. Rizzoli R. Dairy products, yogurts, and bone health. *Am J Clin Nutr.* 2014;99(5
534 Suppl):1256S-1262S.
- 535 42. Ross RA, Allsopp A. Stress fractures in Royal Marines recruits. *Mil Med.*
536 2002;167(7):560-565.
- 537 43. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model
538 continuous risk variables in epidemiology. *Int J Epidemiol.* 1999;28(5):964-974.
- 539 44. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research:
540 Developing a prognostic model. *BMJ.* 2009;338:b604.
- 541 45. Royston P, White IR. Multiple Imputation by Chained Equations (MICE):
542 Implementation in Stata. *Journal of Statistical Software.* 2011;45(4).
- 543 46. Scofield KL, Hecht S. Bone health in endurance athletes: runners, cyclists, and
544 swimmers. *Curr Sports Med Rep.* 2012;11(6):328-334.
- 545 47. Shaffer RA, Brodine SK, Almeida SA, Williams KM, Ronaghy S. Use of simple
546 measures of physical activity to predict stress fractures in young men undergoing a
547 rigorous physical training program. *Am J Epidemiol.* 1999;149(3):236-242.
- 548 48. Shaffer RA, Rauh MJ, Brodine SK, Trone DW, Macera CA. Predictors of stress
549 fracture susceptibility in young female recruits. *Am J Sports Med.* 2006;34(1):108-
550 115.
- 551 49. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development,*
552 *Validation, and Updating;* 2008.
- 553 50. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modelling
554 with logistic regression analysis: a comparison of selection and estimation methods in
555 small data sets. *Stat Med.* 2000;19(8):1059-1079.
- 556 51. Steyerberg EW, Eijkemans MJ, Van Houwelingen JC, Lee KL, Habbema JD.
557 Prognostic models based on literature and individual patient data in logistic regression
558 analysis. *Stat Med.* 2000;19(2):141-160.
- 559 52. Swissa A, Milgrom C, Giladi M, et al. The effect of pretraining sports activity on the
560 incidence of stress fractures among military recruits. A prospective study. *Clin*
561 *Orthop Relat Res.* 1989(245):256-260.
- 562 53. Tenforde AS, Sayres LC, Sainani KL, Fredericson M. Evaluating the relationship of
563 calcium and vitamin D in the prevention of stress fracture injuries in the young
564 athlete: a review of the literature. *PM R.* 2010;2(10):945-949.
- 565 54. Tibshirani R. Regression shrinkage and selection via the Lasso. *Journal of the Royal*
566 *Statistical Society Series B.* 1996;58:267-288.
- 567 55. Valimaki VV, Alfthan H, Lehmuskallio E, et al. Risk factors for clinical stress
568 fractures in male military recruits: a prospective cohort study. *Bone.* 2005;37(2):267-
569 273.
- 570 56. Walter S, Tiemeier H. Variable selection: current practice in epidemiological studies.
571 *Eur J Epidemiol.* 2009;24(12):733-736.
- 572 57. Warden SJ, Burr DB, Brukner PD. Stress fractures: pathophysiology, epidemiology,
573 and risk factors. *Curr Osteoporos Rep.* 2006;4(3):103-109.
- 574 58. Wolk A, Bergstrom R, Hansson LE, Nyren O. Reliability of retrospective information
575 on diet 20 years ago and consistency of independent measurements of remote
576 adolescent diet. *Nutr Cancer.* 1997;29(3):234-241.
- 577

