Check for updates

Evaluating an Algorithm and Clinical Prediction Rule for Diagnosis of Bone Stress Injuries

Nathaniel S. Nye, MD,*† Carlton J. Covey, MD,‡ Mary Pawlak, MD,§ Cara Olsen, PhD,‖ Barry P. Boden, MD,¶ and Anthony I. Beutler, MD‖

Background: A novel algorithm and clinical prediction rule (CPR), with 18 variables, was created in 2014. The CPR generated a bone stress injury (BSI) score, which was used to determine the necessity of imaging in suspected BSI. To date, there are no validated algorithms for imaging selection in patients with suspected BSI.

Hypothesis: A simplified CPR will assist clinicians with diagnosis and decision making in patients with suspected BSI.

Study Design: Prospective cohort study.

Level of Evidence: Level 3.

Methods: A total of 778 military trainees with lower extremity pain were enrolled. All trainees were evaluated for 18 clinical variables suggesting BSI. Participants were monitored via electronic medical record review. Then, a prediction model was developed using logistic regression to identify clinical variables with the greatest predictive value and assigned appropriate weight. Test characteristics for various BSI score thresholds were calculated.

Results: Of the enrolled trainees, 204 had imaging-confirmed BSI in or distal to the femoral condyles. The optimized CPR selected 4 clinical variables (weighted score): bony tenderness (3), prior history of BSI (2), pes cavus (2), and increased walking/running volume (1). The optimized CPR with a score ≥3 yielded 97.5% sensitivity, 54.2% specificity, and 98.2% negative predictive value. An isolated measure, bony tenderness, demonstrated similar statistical performance.

Conclusion: The optimized CPR, which uses bony tenderness, prior history of BSI, pes cavus, and increased walking/running volume, is valid for detecting BSI in or distal to the femoral condyles. However, bony tenderness alone provides a simpler criterion with an equally strong negative predictive value for BSI decision making.

Clinical Relevance: For suspected BSI in or distal to the femoral condyles, imaging can be deferred when there is no bony tenderness. When bony tenderness is present in the setting of 1 or more proven risk factors and no clinical evidence of high-risk bone involvement, presumptive treatment for BSI and serial radiographs may be appropriate.

Keywords: stress fracture; algorithm; clinical prediction rule; MRI

Bone stress injury (BSI) is common among highly active individuals. BSI causes significant morbidity, time lost from exercise, and, if not diagnosed and treated properly, can lead to devastating fractures. BSI presents a fiscal and readiness challenge in all branches of the US military,[8] especially during initial-entry training. The cumulative incidence of lower extremity BSI during initial-entry military training ranges from 0.8% to 6.9% for men and 3.4% to 21.0% for women, with the tibia being the predominant site.[8,13] In US Air Force military basic trainees, the incidence of BSI increased by 56% from fiscal year 2012 to 2014.[17] This was associated with a high utilization rate of bone scans for trainees with lower extremity pain. The result was overdiagnosis (many asymptomatic “hot spots” in low-risk trainees diagnosed as BSI), often leading to further testing, time lost from training, increased patient anxiety, surveillance bias, and discharge from

From †Fort Belvoir Community Hospital, Fort Belvoir, Virginia, ‡Travis Family Medicine Residency, Travis Air Force Base, California, §559th Trainee Health Squadron, JBSA-Lackland, Texas, ‖Uniformed Services University of the Health Sciences, Bethesda, Maryland, and ¶The Orthopaedic Center, a Division of CAO, Rockville, Maryland
*Address correspondence to Nathaniel S. Nye, MD, Fort Belvoir Community Hospital, 9300 DeWitt Loop, Fort Belvoir, VA 22060 (email: nathaniel.s.nye.mil@mail.mil) (Twitter: @NathanielNye).
The authors report no potential conflicts of interest in the development and publication of this article.
This project was completed without grants; however, Department of Defense equipment and facilities were utilized in data collection, analysis, and writing.
The opinions herein are those of the authors. They do not represent official policy of the Department of the Air Force or the Department of Defense. Discussion or mention of any commercial products within this presentation does not create or imply any Federal/Department of Defense endorsement.
DOI: 10.1177/1941738120943540
© 2020 The Author(s)

training. This suggested a need for a standardized diagnostic algorithm for medical providers.

Until recently, there were only a few existing algorithms that made recommendations on imaging in patients suspected of having a BSI.[10,16,18-20] Though representing valuable work, none of these algorithms has previously been validated. Recent work by Wright et al[20] is helpful for the selection of proper imaging modalities, but offers no minimum criteria or guidance on patient characteristics for when advanced imaging is indicated or can be deferred; nor does it guide initial treatment steps. Our 2016 article[16] was the first, to our knowledge, to provide a clinical prediction rule with a threshold for when to obtain imaging and initiate treatment for a suspected BSI. Our proposed algorithm started with a clinical prediction rule (CPR) ("BSI score"). If scoring above a threshold of 4 points, including at least 1 positive physical examination finding (threshold determined by expert opinion), the algorithm proceeded with imaging and initial treatment steps.

CPRs or scores exist to aid point-of-care decision making for a diverse array of medical conditions, including venous thromboembolism (Wells rule),[7] cardiovascular disease risk (Framingham),[5] acute coronary syndrome (Thrombolysis in Myocardial Infarction trial),[3] hip fracture in osteoporosis,[11] ankle fracture (Ottawa Ankle Rules),[1] and many others. The number of CPRs has increased dramatically over the past decade; however, few are validated, which limits incorporation into clinical practice. We developed the BSI CPR[16] to help stratify patients presenting with lower extremity pain according to probability of a BSI.

The primary intent of the BSI CPR was to identify individuals with low BSI probability that would not benefit from imaging, thus sparing the radiation exposure as well as fiscal and time costs, while limiting patient morbidity due to delayed diagnosis of clinically significant BSI. Proper identification of high versus low BSI probability cohorts could improve early diagnosis and treatment, prevent progression to a displaced fracture, and shorten time to return to training.[6,9,10,12,15] Additionally, an effective CPR for BSI may reduce unnecessary imaging with associated costs.[20] Importantly, it may lend confidence in the decision to defer advanced imaging in patients with lower extremity pain when it is unlikely to adversely affect management of the injury.

The purpose of this study was to validate and optimize the algorithm and CPR. As per the previously published algorithm,[16] we hypothesized that a BSI score ≥4 (including ≥1 physical examination finding) would provide a safe and effective threshold for designation of low- and high-probability groups. We further hypothesized that an optimized CPR would provide better statistical performance and ease of use.

METHODS

To validate the BSI algorithm and CPR,[16] a prospective cohort study was designed. This study was approved by the institutional review board at Wilford Hall Ambulatory Surgical Center, Joint Base San Antonio–Lackland.

The Trainee Health Clinic (THC) provides primary medical care for all Basic Military Trainees, including those with musculoskeletal injuries. Trainees presenting to the THC with lower extremity pain between July 15, 2015, and July 15, 2016, were administered a simple questionnaire regarding BSI risk factors and symptoms by a medical technician during the clinic check-in process (Appendix 1, available in the online version of this article). During the appointment, the provider reviewed the questionnaire and, after obtaining additional history and performing a physical examination, assigned the patient a BSI score and proceeded with treatment according to the algorithm. A total of 778 trainees were enrolled. There were no exclusion criteria. Of note, all female trainees are administered a pregnancy test on arrival at the Basic Military Training, and any with a positive pregnancy test are promptly discharged home for civilian medical care.

The a priori BSI score consisted of 3 parts: a risk factor score, a symptom score, and a focused physical examination score.[16] A total BSI score of ≥4, including at least 1 positive physical examination finding, was selected a priori as the threshold to begin diagnostic evaluation and activity restriction. As noted by Nye et al,[16] this threshold was determined by consensus among faculty physicians in the fields of sports medicine, orthopaedic surgery, family medicine, and musculoskeletal radiology.

All participants were followed via electronic medical record review for at least 60 days after their initial BSI score was determined, or until the outcome of interest was confirmed. The outcome of interest was a recorded International Classification of Diseases, Ninth or Tenth Revision (ICD-9/10) diagnosis of BSI in conjunction with positive findings on imaging (radiography, magnetic resonance imaging, or scintigraphy), as read by the on-duty radiologist (unassociated with the study). To clarify, neither an ICD-9/10 code nor imaging findings alone were sufficient to be included as an outcome of interest. Participants were considered an incident case only once during the surveillance period, even if the individual had stress injuries in multiple locations. Some trainees presented to the THC multiple times with undifferentiated lower extremity pain, usually as scheduled follow-up appointments, and thus had multiple BSI scores. In these cases, the most recent BSI score was used for analysis. All medical providers at the THC were provided refresher training on the BSI algorithm[16] in June and November 2015.

A power analysis based on recent epidemiological data for this population was performed to determine our sample size. Approximately 10% of patients who present to the THC with lower extremity pain are eventually diagnosed with a BSI.[17] If all patients with lower extremity pain are divided into "low BSI probability" and "high BSI probability," and then assume 5% incidence of BSI in the "low BSI probability" group and 20% incidence in the "high BSI probability" group (yielding approximately 10% overall incidence), we would require 250 patients to have 80% power to detect this 15% risk difference at the $\alpha = 0.05$ level. Analyzing the clinical prediction rule by sex required 500 total participants (250 male and 250 female) to have adequate statistical power.

Table 1. Locations of bone stress injury recorded in the study cohort

Stress Injury Location	n	%
Knee (femoral condyles, tibial plateau)	50	24.5
Lower leg (tibial shaft, fibula)	99	48.5
Ankle (tibial plafond, talus, medial/lateral malleolus)	16	7.8
Foot (metatarsals, tarsals, excluding talus)	39	19.1
Total	204	100.0

Statistical Analysis

Once data collection was complete, the BSI CPR was analyzed in its a priori form to calculate the test characteristics (sensitivity, specificity, positive and negative predictive value [PPV, NPV], receiver operating characteristic curve with area under the curve [ROC with AUC]). All calculations were completed using SPSS software (Version 22; IBM Corp). The prediction model was then refined using logistic regression modeling. Variables identified in the previous study were assessed for bivariate association with stress fracture using chi-square tests. Variables with $P < 0.25$ were entered into a logistic regression model, and a backward stepwise selection algorithm was used to select variables that were independently associated with stress fracture ($P < 0.05$). Once the final model was determined, simplified weights were assigned based on log odds ratios to determine the final scoring system.

The decision rule was developed on a 60% random sample of the data (stratified by image-confirmed BSI status) and validated on the remaining 40% of the data. Area under the ROC curve was estimated in the validation sample, along with sensitivity and specificity of the decision rule across a range of possible cutoff values.

An alternative approach was then used to see whether the same model would be selected using a different algorithm. A logistic regression model was estimated with all predictors from the clinical tool in question, not just those that were significant in chi-square tests. Backward stepwise selection was used to select from the full set of variables. These include slow 1.5-mile run, female sex, missed periods, history of stress fracture, history of vitamin D deficiency, history of eating disorder, pes cavus (on examination), gradual onset of pain, exertional pain, dull aching pain, recent increase in walking/running, night pain, antalgic gait, focal bony tenderness, pain with hip internal rotation, pain with single-leg hop test, pain with tuning fork test, and pain with fulcrum test. Then, an ROC curve was fit to data in the validation sample, and a new model was developed. Sensitivity and specificity for different cutoff scores were calculated, along with AUC.

RESULTS

Of the 778 trainees who were enrolled, 273 (35%) were female and 505 (65%) were male. Overall, 218 patients (28.0%) had 1 or more BSI based on ICD-9/10 codes and imaging confirmation. The vast majority (204; 93.6%) of these BSIs were located in or distal to the femoral condyles. Only 14 patients (6.4%) of all those with confirmed BSIs had a BSI proximal to the femoral condyles. Given the low numbers of pelvic/hip/thigh BSI in our sample, these were excluded from all other analyses. Table 1 lists the number of patients with BSI at each location, after excluding pelvic/hip/thigh BSI.

The a priori prediction model (all 18 variables included with equal weight) with a cutoff score of ≥4 (including at least 1 positive physical examination finding) yielded a sensitivity of 93.1%, specificity of 31.1%, PPV of 44.1%, and NPV of 88.6%. In optimizing the model, 4 of these 18 variables were identified as independent predictors of BSI using logistic regression. A weighted sum of these 4 variables constituted the revised BSI score, with weights as described in Table 2. Bony tenderness on physical examination yielded the highest odds ratio (OR) for BSI (OR, 17.43), followed by prior history of stress fracture (OR, 6.57).

Sensitivity and specificity for diagnosis of BSI for the optimized model are described in Table 3, and score distributions for the a priori model and optimized model are shown in Figures 1 and 2, respectively. Based on these data, the optimized model provides the best statistical performance when using a BSI score ≥3 as a threshold to begin diagnostic evaluation, with the validation sample yielding the following accuracy estimates: sensitivity, 97.5%; specificity, 54.2%; PPV, 45.9%; and NPV, 98.2%. An ROC curve, fit to data in the validation sample, is shown in Figure 3. The AUC was 0.767 (95% CI, 0.714-0.820; $P < 0.001$).

As an isolated measure, bony tenderness demonstrated nearly identical statistical performance (97.5% sensitivity, 54.7% specificity, 46.2% PPV, and 98.2% NPV) as the optimized BSI score model with a cutoff score of 3. The vast majority of patients who scored ≥3 in the optimized model had bony tenderness (376/387; 97.2%).

Table 2. Statistical derivation of an optimized bone stress injury clinical prediction model[a]

				95% CI for Odds Ratio		
	Log Odds Ratio	*P*	Odds Ratio	Lower	Upper	Points
History of stress fracture	1.883	0.028	6.571	1.224	35.287	2
History of pes cavus	1.581	0.032	4.861	1.145	20.627	2
Increased walking/running volume	0.594	0.019	1.811	1.101	2.979	1
Focal bony tenderness	2.858	0.000	17.430	8.424	36.065	3

[a]Weighting points were assigned based on the log odds ratio. To force weights to be whole numbers, the log odds ratios were multiplied by 2 then rounded to the nearest whole number.

Table 3. Sensitivity, specificity, predictive values, and likelihood ratios for various cutoff scores under the optimized bone stress injury clinical prediction model

Positive if ≥	Sensitivity	Specificity	PPV	NPV	LR+	LR−
0	1.000	0.000	0.285	NA	1.000	NA
1	1.000	0.281	0.357	1.000	1.390	0.000
2	0.975	0.522	0.449	0.981	2.041	0.047
3	0.975	0.542	0.459	0.982	2.129	0.046
4	0.556	0.754	0.474	0.810	2.256	0.590
5	0.037	0.975	0.375	0.717	1.504	0.987
6	0.012	0.995	0.500	0.716	2.506	0.993
7	0.000	1.000	NA	0.715	NA	1.000
Presence of focal bony tenderness	0.975	0.547	0.462	0.982	2.152	0.047

LR+, positive likelihood ratio; LR−, negative likelihood ratio; NA, not applicable; NPV, negative predictive value; PPV, positive predictive value.

DISCUSSION

This study analyzed a proposed CPR for diagnosis of BSI (referred to as "a priori model"), then used data to develop a new, optimized CPR. The a priori model and optimized model were then compared against each other and against a simple test (bony tenderness on examination) using several measures of statistical performance. The primary finding was that all 3 models had high sensitivity and NPV, which is likely due to the high sensitivity and NPV of bony tenderness that is included in all models. The simplicity and high sensitivity of bony tenderness as a screening test for BSI support its central role in making this diagnosis. BSI should be suspected in patients with lower extremity pain and corresponding bony tenderness to palpation, particularly with an atraumatic/overuse history and when known risk factors are present (eg, prior history of BSI, pes cavus, underweight, rapid progression of physical activity/training). Such patients should be imaged and prescribed impact activity restrictions with or without crutches, according to the treatment algorithm as shown in Figure 4. Those without bony tenderness or with subthreshold BSI scores should be followed clinically and be treated for any suspected alternative diagnosis as indicated. A high NPV is an important marker for a good screening test for BSI, as the clinical goal is to avoid false negatives with the screen.

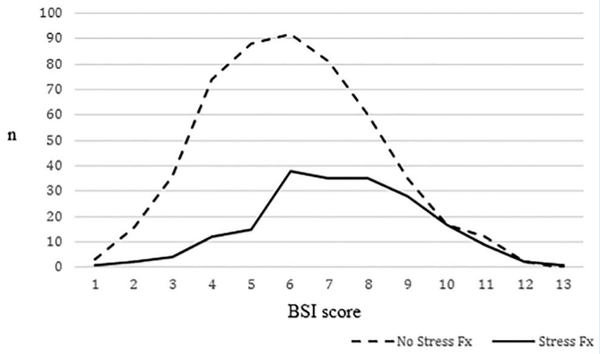

Figure 1. Bone stress injury (BSI) score distribution for the a priori model, with separate plots for those who did versus did not have an outcome of stress injury. Fx, fracture.

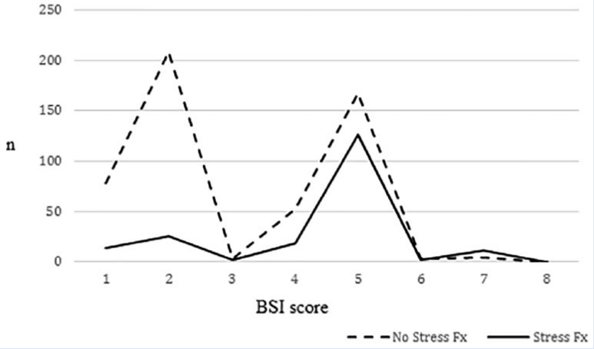

Figure 2. Bone stress injury (BSI) score distribution for the optimized model, with separate plots for those who did versus did not have an outcome of stress injury. Fx, fracture.

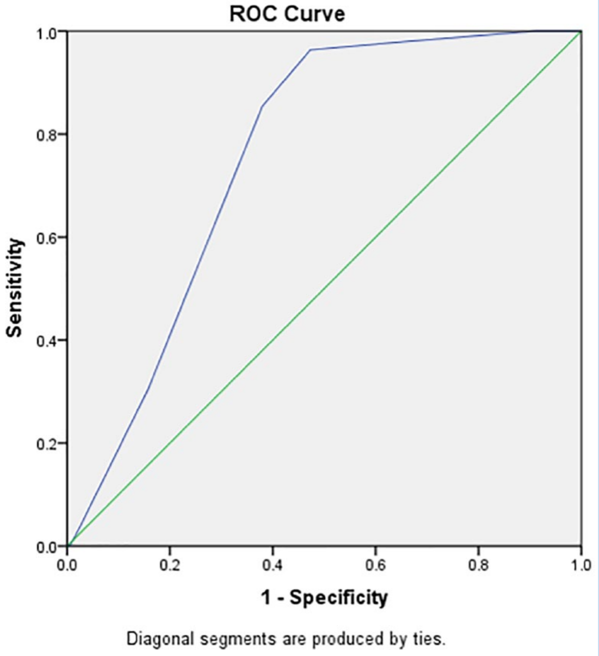

Figure 3. Receiver operating characteristic (ROC) curve for the optimized bone stress injury score model. Area under the curve = 0.767.

Because bony tenderness alone provides equivalent NPV, the described CPRs (a priori model and optimized model) did not add value for ruling out BSI in or distal to the femoral condyles. The a priori and optimized CPRs did provide increased PPV at higher scores, which may be helpful in certain clinical scenarios, particularly in populations with a high incidence of BSI and, therefore, higher pretest probability. When advanced imaging is ordered, a higher BSI score (eg, ≥5) may be useful in counseling the patient to comply with activity restrictions while awaiting imaging. Furthermore, the higher specificity and positive likelihood ratio (LR+) of higher BSI scores (see Table 3) may be useful in settings where magnetic resonance imaging is difficult to obtain or unavailable, such as deployed military clinics or humanitarian efforts.

It is not surprising that bony tenderness is a highly sensitive test for BSI in or distal to the femoral condyles, due to the superficial nature of most bones in this region. The most notable exception to this is the fibular diaphysis, which lies deep to the lower leg musculature. Another exception is the talar body and dome, which are largely obscured by the medial and lateral malleoli. BSI of the fibular diaphysis and talar body/dome are relatively uncommon in our population. Though the talus is generally considered a high-risk location for BSI,[2] a recent analysis of our population showed no increase in complications or time to return to running compared with low-risk injuries when a fracture line was not present.[4] When a patient presents with deep or vague pain of the calf (particularly lateral aspect) or ankle, it is important to maintain a high index of suspicion for BSI of the fibular diaphysis or talus, respectively.

Bony tenderness was the only physical examination finding that independently demonstrated predictive value for BSI based on these data. However, other physical examination maneuvers may also have an important role. In our experience, the single-leg hop test is valuable as a reproducible way to measure and compare the patient's progress at follow-up visits. An antalgic gait should raise concern for possible high-grade BSI. Our experience with the tuning fork test is that it depends greatly on how much pressure is used in applying the tuning fork to the bone (likely low interrater reliability) and adds little value beyond simply assessing for bony tenderness.

As with physical examination, only 3 items from history and risk factor assessment independently yielded predictive value in our data set. These included prior history of BSI, a recent increase in physical activity, and pes cavus (though assessed on examination, we consider this a risk factor). However, other factors such as being underweight (body mass index <18.5 kg/m^2),

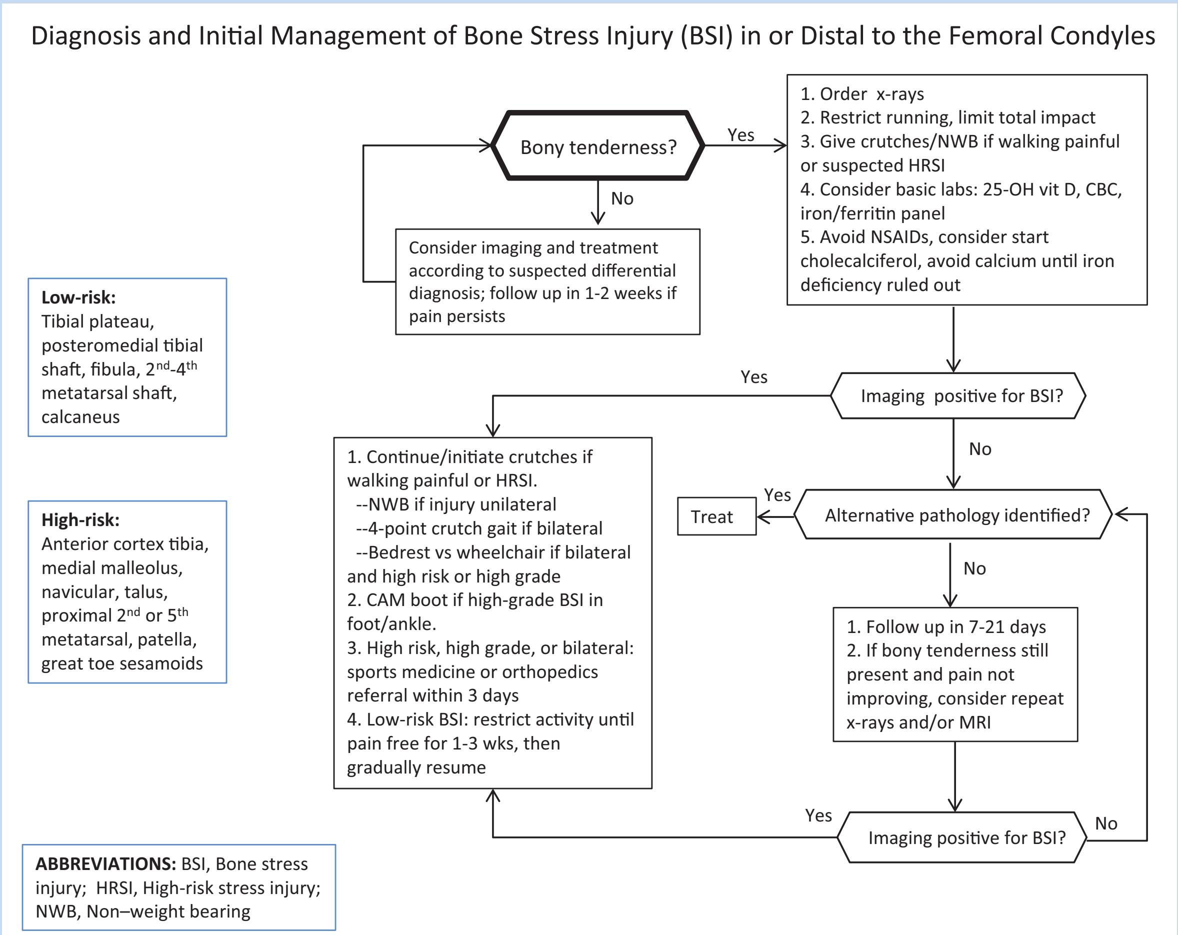

Figure 4. Updated algorithm for diagnosis and initial management of bone stress injury (BSI) in or distal to the femoral condyles. Low-risk: tibial plateau, posteromedial tibial shaft, fibula, second to fourth metatarsal shaft, calcaneus. High-risk: anterior cortex tibia, medial malleolus, navicular, talus, proximal second or fifth metatarsal, patella, great toe sesamoids. CAM, controlled ankle motion; HRSI, high-risk stress injury; MRI, magnetic resonance imaging; NSAIDs, nonsteroidal anti-inflammatory drugs; NWB, nonweightbearing.

smoking, amenorrhea, and female sex have been shown to be associated with increased risk of BSI in other studies, and should also be considered when caring for patients with lower extremity injuries.[8,13]

The BSI algorithm and CPR were initially designed for all BSI of the lower extremity (including high- and low-risk injuries); however, this study sample did not include sufficient numbers of BSI proximal to the knee to validate the CPR for these injuries. Furthermore, the very high–risk nature of certain BSIs in the hip region (ie, femoral neck, especially tension-sided[14]) does not lend well to lumping with lower risk injuries in a single CPR. This algorithm and CPR are valid, based on this analysis, for both high- and low-risk BSI in or distal to the femoral condyles. The algorithm provides a uniform threshold (presence of bony tenderness) (Figure 4) to initiate treatment and obtain radiographs, regardless of whether the injury is suspected in a low- or high-risk location. However, the components of initial treatment do depend on the nature of injury. Crutches and nonweightbearing are recommended for all suspected high-risk BSI until ruled out, whereas those with suspected low-risk BSI who ambulate without pain may be allowed to continue full weightbearing, but running and high-impact activities are nonetheless restricted. When radiographs are negative but a high-risk injury (eg, navicular, fifth metatarsal, talus) is still suspected, a shorter follow-up interval should be used (eg, 7 days rather than 14-21 days) (Figure 4). Decisions about advanced imaging can be made at the follow-up appointment based on clinical progress and continued level of concern for high-risk BSI.

The above CPRs and assessment tools do provide valid aids for clinical decision making; however, the findings must be considered in light of a few limitations. The studied population consisted mostly of young (average, 21 years old) novice runners, undergoing basic military training. Findings in this population may not be generalizable to nonmilitary populations.

BSI scores used in this study were completed by providers who, despite training provided by research staff, may be subject to their own biases. It must also be remembered that clinical prediction rules are designed to enhance but not replace clinical judgment. Finally, although bony tenderness is highly sensitive (97.5%), a small percentage of patients with BSI will not display bony tenderness, or it may be present only briefly (perhaps prior to or after the date of examination).

CONCLUSION

Although the a priori and optimized BSI CPRs are valid, a simple assessment of bony tenderness to palpation provides equally strong NPV (98.2%). The lack of bony tenderness on examination serves as a clinically useful aid for limiting unnecessary diagnostic testing in low-probability patients. In patients with bony tenderness, multiple risk factors, and no evidence of high-risk bone involvement (navicular, talus, fifth metatarsal, great toe sesamoids), a trial of presumptive treatment with serial radiographs instead of advanced imaging is appropriate. Because bony tenderness is difficult to assess for the fibular diaphysis and talar body/dome, these may be exceptions to this algorithm and CPR. This algorithm and clinical prediction rule should be studied in other populations for further validation and clinical impact analysis.

Clinical Recommendations

SORT: Strength of Recommendation Taxonomy
A: consistent, good-quality patient-oriented evidence
B: inconsistent or limited-quality patient-oriented evidence
C: consensus, disease-oriented evidence, usual practice, expert opinion, or case series

Clinical Recommendation	SORT Evidence Rating
Unless other indications exist, defer imaging in patients presenting with pain in or distal to the femoral condyles, who lack bony tenderness, as this indicates low probability of BSI.	B
Assess for prior history of BSI, pes cavus, recently increased physical activity, and bony tenderness in patients with pain in or distal to the femoral condyles. These factors carry independent predictive value for BSI.	B
When bony tenderness is present in the setting of one or more proven risk factors (eg, prior history of BSI, pes cavus) and there is no evidence of high-risk bone involvement, consider a trial of presumptive treatment with serial radiographs instead of advanced imaging.	C

REFERENCES

1. Beckenkamp PR, Lin CC, Macaskill P, Michaleff ZA, Maher CG, Moseley AM. Diagnostic accuracy of the Ottawa Ankle and Midfoot Rules: a systematic review with meta-analysis. *Br J Sports Med*. 2017;51:504-510.
2. Boden BP, Osbahr DC. High-risk stress fractures: evaluation and treatment. *J Am Acad Orthop Surg*. 2000;8:344-353.
3. Chesebro JH, Knatterud G, Roberts R, et al. Thrombolysis in Myocardial Infarction (TIMI) Trial, phase I: a comparison between intravenous tissue plasminogen activator and intravenous streptokinase. Clinical findings through hospital discharge. *Circulation*. 1987;76:142-154.
4. Curell AM, Nye NS, Webber BJ, Pawlak MT, Boden BP. Treatment and prognosis of high and low-risk Kaeding grade II bone stress injuries. *Transl J Am Coll Sports Med*. 2019;4:114-118.
5. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P; CHD Risk Prediction Group. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA*. 2001;286:180-187.
6. Fredericson M, Bergman AG, Hoffman KL, Dillingham MS. Tibial stress reaction in runners: correlation of clinical symptoms and scintigraphy with a new magnetic resonance imaging grading system. *Am J Sports Med*. 1995;23:472-481.
7. Geersing GJ, Zuithoff NP, Kearon C, et al. Exclusion of deep vein thrombosis using the Wells rule in clinically important subgroups: individual patient data meta-analysis. *BMJ*. 2014;348:G1340.
8. Jacobs JM, Cameron KL, Bojescul JA. Lower extremity stress fractures in the military. *Clin Sports Med*. 2014;33:591-613.
9. Kaeding CC, Miller T. The comprehensive description of stress fractures: a new classification system. *J Bone Joint Surg Am*. 2013;95:1214-1220.
10. Kaeding CC, Yu JR, Wright R, Amendola A, Spindler KP. Management and return to play of stress fractures. *Clin J Sport Med*. 2005;15:442-447.
11. Kanis JA, Borgstrom F, De Laet C, et al. Assessment of fracture risk. *Osteoporos Int*. 2005;16:581-589.
12. Kijowski R, Choi J, Shinki K, Del Rio AM, De Smet A. Validation of MRI classification system for tibial stress injuries. *AJR Am J Roentgenol*. 2012;198:878-884.
13. Knapik J, Montain SJ, McGraw S, Grier T, Ely M, Jones BH. Stress fracture risk factors in basic combat training. *Int J Sports Med*. 2012;33:940-946.
14. Kupferer KR, Bush DM, Cornell JE, et al. Femoral neck stress fracture in Air Force basic trainees. *Mil Med*. 2014;179:56-61.
15. Nattiv A, Kennedy G, Barrack MT, et al. Correlation of MRI grading of bone stress injuries with clinical risk factors and return to play: a 5-year prospective study in collegiate track and field athletes. *Am J Sports Med*. 2013;41:1930-1941.
16. Nye NS, Covey CJ, Sheldon L, et al. Improving diagnostic accuracy and efficiency of suspected bone stress injuries. *Sports Health*. 2016;8:278-283.
17. Nye NS, Pawlak MT, Webber BJ, Tchandja JN, Milner MR. Description and rate of musculoskeletal injuries in Air Force basic military trainees, 2012-2014. *J Athl Train*. 2016;51:858-865.
18. Roberts CL, Meyering CD, Zychowicz ME. Improving the management of tibia stress fractures: a collaborative, outpatient clinic-based quality improvement project. *Orthop Nurs*. 2014;33:75-83.
19. Scott SJ, Feltwell DN, Knapik JJ, et al. A multiple intervention strategy for reducing femoral neck stress injuries and other serious overuse injuries in U.S. Army basic combat training. *Mil Med*. 2012;177:1081-1089.
20. Wright AA, Hegedus EJ, Lenchik L, Kuhn KJ, Santiago L, Smoliga JM. Diagnostic accuracy of various imaging modalities for suspected lower extremity stress fractures: a systematic review with evidence-based recommendations for clinical practice. *Am J Sports Med*. 2016;44:255-263.

For article reuse guidelines, please visit SAGE's website at http://www.sagepub.com/journals-permissions.