**Title:** Derivation and sensitivity validation of a clinical decision tool to predict the need for spinal MRI in non-low risk patients suspected of pyogenic spinal infection

**Author List:** (Presented as given name and then family name for each author)

Steven Shroyer, MD[d]
StevenShroyerMD@gmail.com

Michael D. April, MD, DPhil, MSc[b,c]
Michael.D.April@post.harvard.edu

Sumeru Mehta, MD[d]
sumeru.mehta@gmail.com

Brit Long, MD[a,b]
brit.long@yahoo.com

Gregory Boys, MD[d]
greg.boys@gmail.com

William T. Davis, MD[a,b]
Wtdavis2@gmail.com

**Author Affiliations:**

a. Department of Emergency Medicine, San Antonio Military Medical Center, 3551 Roger Brooke Dr, Fort Sam Houston, TX 78234.
b. Department of Military and Emergency Medicine, Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814.
c. 2nd Brigade Combat Team, 4th Infantry Division, 10101 Warfighter Road, Fort Carson, CO 80913. Greater San Antonio Emergency Physicians, Methodist Hospital System, 7700 Floyd Curl Dr, San Antonio, TX 78229.
d. Department of Radiology, Methodist Hospital System, 7700 Floyd Curl Dr, San Antonio, TX 78229.

**Corresponding Author:** William T. Davis, MD, wtdavis2@gmail.com

Present address:

San Antonio Military Medical Center
Department of Emergency Medicine
3551 Roger Brooke Dr
Fort Sam Houston, TX 78234

Phone: 210-916-5512

**Conflict of Interest Disclosure:** WTD, MDA, SM, BL, GB, SS report no conflicts of interest.

**Keywords:** c-reactive protein, pyogenic spinal infection, spinal epidural abscess, vertebral osteomyelitis, back pain, diagnostic accuracy, sensitivity.

**Disclaimers:** The views expressed herein are those of the authors and do not reflect the official policy or position of Brooke Army Medical Center, the U.S. Army Medical Department, the U.S. Army Office of the Surgeon General, the 4th Infantry Division, the Fort Carson Post Command, the Department of the Army, the Department of the Air Force and Department of Defense or the U.S. Government. The views of the stated manufacturers are not necessarily the official
views of, or endorsed by, the U.S. Government, the Department of Defense, or the Department of the Air Force. No Federal endorsement of the stated manufacturers is intended. The voluntary, fully informed consent of the subjects used in this research was obtained as required by 32 CFR 219 and DODI 3216.02_AFI 40-402. This original contribution has not been published, it is not under consideration for publication elsewhere, its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

**Author Contributions:** WTD, SM, and SS were responsible for study concept and design. WTD and SS were responsible for acquisition of data. All authors were responsible for critical revision of the article for important intellectual content. WTD, MDA, and SS were responsible for analysis and interpretation of data. WTD was responsible for drafting the manuscript. WTD, MDA, and SS were responsible for statistical analysis. WTD had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. WTD takes responsibility for the paper as a whole.

Abstract

Objective/Background: Pyogenic spinal Infection (PSI) is not diagnosed at the initial presentation in most cases, and delay in diagnosis is associated with increased morbidity and medicolegal risk. We derived a sensitive clinical decision tool to estimate the risk of spinal infection and inform MR imaging decisions.

Methods: This study is a 2-part prospective observational cohort study. The six-year derivation phase collected 40 variables from 179 patients (134 uninfected, 40 infected, 5 excluded). We evaluated variables with univariate analysis, then logistic regression, to derive a four variable clinical prediction tool, SIRCH (Spine Infection Risk Calculation Heuristic). We calculated the statistical performance, MRI utilization and model fit in the derivation phase. We used the same protocol in collecting data for a second phase, enrolling only confirmed infections (49 infected, 4 excluded), validating sensitivity only.

Results: The SIRCH score's ROC (receiver operator curve) area, sensitivity, specificity, positive and negative predictive value for predicting PSI were respectively: 0.877, 100% (40/40), 56% (75/134), 40% (40/99), and 100% (75/75). The calibration plot demonstrated good model fit. The SIRCH score prompted 45 fewer MRI scan orders compared to CRP, and 54 fewer MRI scans than clinical suspicion, among the 174 patients. All three image prompts had equal sensitivities of 100% (40/40) in the derivation. The validation sensitivities for the SIRCH score and CRP were 92% (44/49) and 98% (48/49), respectively.

Conclusion: The SIRCH score provides a sensitive estimate of spinal infection risk and prompts less MR imaging than using CRP or clinical suspicion of spinal infection.

## 1. Introduction

**1.1.** Background

Pyogenic spinal Infection (PSI), which includes spinal epidural abscess, is an uncommon condition concealed among patients with a common chief complaint—back or neck pain.[1,2,3] This presenting complaint is also the 5$^{th}$ leading chief complaint among emergency department (ED) patients.[4] While diagnosing some cases of this infection are simplified by an obvious presentation of back pain and fever, or back pain and intravenous drug use (IVDU), most are not. [1,2,3,5] The challenge of detecting this uncommon signal from a great deal of background noise can result in diagnostic delay, which can lead to the progression of unrecognized sepsis, permanent neurologic deficit for patients, and increasing medical-legal risk for physicians. [5-10] Although magnetic resonance imaging (MRI) with gadolinium contrast is 96% sensitive and 93% specific for the infection, it is not an easily administered test. It requires 4 to 8 hours for test results [11], is uncomfortable, contributes to ED crowding, and is not 100% available at all facilities where back pain is evaluated. [6,10,12,13]

Currently, there are no prospective validated prediction tools for PSI, [14,15,16,20,] no agreement on CRP cut-off levels to indicate imaging [17], and no recommendations regarding MRI over-utilization. [13,18] Recent publications recommend imaging spine pain patients who have any of the following PSI risk features: historical risk factors, fever, history of fever, or progressive neurological deficit, [2,6,7,19,20,21] and to consider an alternate diagnosis if none of these are present. [2,6,19,21]

**1.2.** Goals of this investigation

We aimed to identify findings from the history, physical examination, and pain character, that when combined with a CRP level, can be used in a simple, intuitive four variable prediction score—SIRCH (Spinal Infection Risk Calculation Heuristic)—to provide a sensitive estimate of the likelihood of PSI and appropriately recommend MR imaging.

## 2. Methods

2.1 Design, setting, selection, population

Our study was a single-center observational prospective cohort study conducted in an academic community ED, with 18 emergency physicians (mean of 11 years' experience), treating 50,000+ adult patients annually in a city of 2.3 million people located in the southwestern united states. Further description of cohort characteristics and methods can be found in earlier publications. [22, 33] Our data was obtained from a convenience sample for use in developing a multivariable risk prediction tool that was conducted in two phases; in the first phase (Jan 2006-March 2010), we enrolled patients who were suspected of spinal infection that included both uninfected and PSI patients. From this phase we selected predictors and derived a risk prediction score. In the second phase (April 2010-August 2018), we enrolled only patients with infection to validate the sensitivity of our prediction tool.

2.2 Eligibility and data collection.

We considered back or neck pain patients for enrollment if they were ≥ 17 years old, had no diagnoses such as pyelonephritis or pneumonia to explain their pain prior to MRI order, and were suspected of having pyogenic spinal infection based on the presence of any of the following: historical risk factor [6] or first ED temperature ≥ 38˚C or a recently measured fever before arrival, progressive neurologic deficit, or clinician gestalt which included bounce-back for a spine-related visit, or unusually severe spine pain. We excluded patients who presented less than five days following a spinal surgical procedure [23,24]; if they had a fungal or tuberculous spinal infection; if the diagnosis could not be determined; or if the patient had an incomplete follow-up. All patients were independent visits. We educated the emergency department clinicians to cited risks for PSI at the beginning of the study period using illustrative cases. These were distributed by email and at monthly department meetings periodically throughout the study period. Once an MRI and or CRP was ordered for the purpose of evaluating spinal infection, the primary investigator (PI) was notified who evaluated all patients for enrollment and completed a standardized examination to obtain historical, physical examination findings and pain characteristics, and then recorded these and laboratory data on a data collection form prior to MRI results. Enrollment was based on the availability of the PI, which remained constant throughout the study period. We counseled discharged patients to return to the ED if they had any progression, new or concerning symptoms. The PI reviewed medical records to obtain CRP, imaging interpretations, blood culture results, operative findings, and culture results from surgery and needle aspiration samples.

The sensitivity validation phase (2010-2018) protocol remained the same for data collection and identifying patients for possible enrollment but included only infected patients. We compared data from the derivation and sensitivity validation phases to evaluate this protocol change. Our investigation followed the TRIPOD guidelines for risk prediction model development, and the study was approved by the hospital system's institutional review board. [25]

2.3 Outcome measures

Our primary outcome measures were the sensitivity and MRI utilization of our risk score for the presence of pyogenic spinal infection. PSI was defined as the presence of any of the following infections: spinal epidural abscess, discitis, vertebral osteomyelitis, paravertebral abscess, paraspinal abscess, or septic facet infection. [3,22] Isolated psoas muscle infection, without another spinal infection, was not considered a PSI but it's presence or absence was recorded. A PSI's presence could be reached by any of the following: 1. MRI evidence of spinal infection as read by a neuroradiologist, 2. surgical findings of spinal infection on the operative report, or 3. needle aspiration culture results consistent with a spinal infection. The pool of ten neuroradiologists interpreting images were blinded to the data collected except age, gender, and chief complaint. Due to the observational nature of our study, we did not obtain an MRI in all patients in the derivation phase. Clinical follow-up was used instead by telephone two to four weeks after their index visit, and for up to three years through available medical and imaging records to verify that no findings of pyogenic spinal infection had developed. This extended follow up was chosen due to the indolent course of some pyogenic spinal infections. We conducted an obituary search on those subjects not followed up.

For outcome measures, we calculated the area under the prediction tool's ROC, sensitivity, specificity, positive and negative predictive values, and likelihood ratios. We defined MRI utilization as the number of unnecessary MRI scan orders in uninfected patients prompted by the SIRCH score (false positives), CRP value, or by clinician suspicion. For the validation phase, we only calculated sensitivities.

C-reactive protein (CRP) was ordered according to individual physician preference. The details of our hospital's laboratory CRP autoanalyzer and MRI machines are located in Appendix 1.

## 3. Data Analysis

3.1 Statistical analysis

Two investigators double entered all data from the data collection sheets into Excel database (version 14; Microsoft, Redmond, WA) and then exported the data into both R version statistical software (Foundation for Statistical Computing, Vienna, Austria, 2018) and MedCalc Statistical Software version 19.5.3 (MedCalc Software Ltd, Ostend, Belgium; https://www.medcalc.org; 2020). In the derivation phase, we report all collected data for each of 179 cases, except 5 excluded patients. We assessed the distributions of infected and uninfected patient characteristics and their differences using the ~~Wilcoxon test~~ for continuous variables and Pearson's chi-squared test for categorical predictors. We selected sensible clinical candidate predictors and assessed all 174 cases with univariate and multivariate models. Missing CRP variables were multiply imputed using predicted mean matching (1000 imputations), and imputed models were combined using Rubin's Rules. We report all missing data in Appendix Table 3. Multivariate analysis was a backward logistic regression using candidate predictors with a biologically plausible association with PSI. Variables were entered with a difference of $P < 0.05$ and removed if $P > 0.2$, with pyogenic spinal infection considered the dependent variable. [28] We then simplified the derived, full model by rounding the logistic β coefficients and assigning these as points to each variable for an easily calculated bed-side risk score. We explored several models using their statistical performance, area under the ROC, and calibration plot to compare them (Appendix 5). [35] Each metric used bootstrapped 95% confidence intervals. We estimated the probability of PSI at various SIRCH criterion cut-offs (Table 4, Figure 3, Appendix 7). We estimated MRI overutilization by counting the number of MRIs that would be prompted in uninfected patients (false positives) using the SIRCH score, using CRP level at a standard laboratory cut-off (3.5 mg/L) and using the number prompted by clinician suspicion, and then compared these (Table 2).

We performed a sensitivity analysis comparing complete case models with imputed models (Appendix 6, Appendix 8).

Following the derivation phase, a second cohort of patients was studied to understand the sensitivity of the derived SIRCH score.

A P-value difference of < 0.05 was considered significant. We did not measure interoperator variability but used the most objective measures available or used variables with previously measured kappas. Since there is no existing prevalence data for PSIs in an at-risk population, we based our study size on obtaining at least ten outcome events for each chosen clinical predictor and on a sample size based on an estimated prevalence of 20%, 95% sensitivity, and a confidence interval of 8%, providing an estimated 143 subjects. A type-I error of 5% was used for all confidence intervals and hypothesis tests.

### 3.1 Predictor variables and derivation

Candidate variables were defined a priori (Appendix 1) and listed on our data collection sheets. We combined like variables to simplify our model as follows: 1. We collapsed the ten historical risk factor variables described by Davis [6] into one variable based on their common plausible pathophysiology—bacteremia seeding a spinal structure (exception, spinal surgical procedure); 2. We combined fever in the ED and recent history of measured fever into one variable, and defined fever as a temperature ≥ 38˚C (100.4˚F); and 3. We combined any abnormal new spinal neurological deficit into the variable, progressive neurological deficit. For our final model, we dichotomized the CRP at a cut-off of ≥ 50 mg/L (Youden's method) and entered the CRP and the above three variables into a logistic regression analysis. We also explored other models using CRP alone, clinical variables alone, and using CRP (at several cut-offs) in combination with clinical variables (Appendix 5). The final best full prediction model was the following: Probability of PSI = 1/(1+exp(Risk Score), Risk Score = - 5.16 + (2.88xCRP) + (1.60xRF) + (1.27xF or Hx of F) + (0.84xPND), AUC 0.886 (95% CI, 0.839-0.934), sensitivity 95% (95% CI, 85-100) and specificity 75% (95% CI, 60-85) (Appendix 5b and c). We simplified the full model to a bed-side model (SIRCH score), by rounded the logistic coefficients to the nearest integer and assigned those point values to each variable,[28] resulting in the following four-variable scoring system: historic risk factors-2, fever or history of fever-1, progressive neurological deficit-1 and CRP (≥ 50mg/L)-3, (Appendix 2). We performed a sensitivity analysis comparing our multiple imputation models to complete case models (Appendix 6 and 8). The final simplified SIRCH score model ranged from 0 to 7, and from its ROC we identified its most accurate cut-off (Youden's) distinguishing PSI from non-infection as SIRCH ≥ 3.

## 4. RESULTS

4.1 Baseline characteristics

Baseline characteristics are listed in Table 1, and further descriptive characteristics can be found in prior publications [22,33]. Of 179 patients enrolled in the derivation phase, we excluded three patients due to incomplete follow-up, another who died without an autopsy to confirm the absence of spinal infection, and another with a fungal spinal infection. We imputed missing CRPs in 13 of the 134 non-infections and 2 of the 40 infections. One hundred and thirteen of 174 patients had MR or alternate imaging, and 21 were followed clinically without imaging. Thirty-nine of 40 infected patients were imaged with MRI, and one infection was confirmed in the operating room (OR) without imaging.

Of 53 patients in the sensitivity validation phase, we excluded four who were adjudicated not to have a true pyogenic spinal infection. We imputed missing CRPs in six patients, leaving 49 infected patients (flow chart, Figure 1). We imaged 48 patients and confirmed one infection in the OR without imaging. We identified positive blood cultures in 47/82 and a microorganism in 77/89 infected patients. A total of 189 MRIs and 30 CT scans were obtained among the 232 studied subjects.

4.2 Model and performance

A derivation SIRCH score of ≥ 3 was 100% (40/40) sensitive for infection and would prompt unnecessary imaging in 59 of the total 174 (34%) patients evaluated (Table 2). The CRP, using a standard cut-off of 3.5 mg/L was also 100% (40/40) sensitive but would prompt nearly twice as many MRIs (104 of 174, 60%). The score would therefore reduce the number of MRIs ordered by 45 compared to CRP, and by 54 compared to clinician suspicion among the 174 patients evaluated. The SIRCH score's sensitivity declined in the validation phase to 92% (45/49). The AUC, Calibration intercept, and slope were, 0.877 (0.829 to 0.925), -5.23 (-7.14 to -3.77), and 0.938 (0.652 to 1.30), respectively.

The mean CRP among the 134 uninfected patients was 47.9 mg/L (95% CI, 36-59)—more than tenfold higher than the cut-off for our hospital system's standard laboratory of 3.5 mg/L. The mean CRP for the 40 infected patients was 140.2 (95% CI, 108-172)—nearly 50 fold higher. The mean CRP for the 49 patients in the sensitivity validation was similar to the derivation, 145.5 (95% CI, 117-174), and consistent with recent studies. [29, 30]

All 89 infected patients (except patient # 295V, Table 3) had at least one of the following eligibility criteria: historical risk factor, fever, progressive neurologic deficit, or bounce-back. Although severe pain prompted clinical suspicion of PSI and represented 43% (38/89) of PSI patients, other risk features were invariably present in these patients. Eighty-three percent (25/30) of PSI patients with neurological deficits had no fever to prompt consideration of infection among the 89 spinal infections, highlighting a potential diagnostic pitfall for emergency physicians.

The mean temperatures in the derivation (uninfected and infected), and in the validation phases (infected only) are indicated in Table 1. The proportion of infected patients with an ED fever was low in both phases but consistent with recent publications. [6] Several trends were noted in infected patients over the two phases of the study: 1. The proportion of patients having historical risk factors declined from 90% (36/40) to 76% (37/49); 2. The mean SIRCH score for PSIs declined from 5.53. The proportion with ED fever declined from 30% (12/40) to 10% (5/49); 4. The proportion of bounce-backs minimally increased from 63% (25/40) to 69%(34/49); and 5. There was a slight decline in the proportion having a progressive neurological deficit over the study period from 38% (15/40) to 31% (15/49). The only variable having a significant difference between the two phases was ED fever (p=0.030).

Table 1 indicates the spectrum of spinal infections and non-infection diagnoses. Table 3 indicates that less unnecessary MRIs would be ordered using the SIRCH score, compared to CRP or clinician suspicion (numbers do not account for seven patients having PSI, but required MR imaging for a neurological deficit). Figure 3 demonstrates that an increasing SIRCH score is associated with a higher probability of PSI. Table 3 lists the characteristics of infected patients who were missed or nearly missed using SIRCH. Eight of nine patients in this group were bounce-backs.

## 5. Limitations

This study's small number of evaluating physicians and single-center design may restrict the generalizability of our findings. Our sample was not consecutive and included patients considered to be intermediate or high risk for spinal infection as defined in prior publications. [19,34] The high PSI prevalence in our sample may subject our study to spectrum bias, which could result in overestimating the SIRCH score's sensitivity. Additionally, our enriched sample could overestimate the SIRCH score's MRI utilization benefits (fewer false positives) compared to lower prevalence populations. It is unknown if the same patient selection process in other settings would result in a similar prevalence.

Although blinding clinicians to the CRP results could have reduced potential work-up bias, this was inconsistent with the observational nature of our study. We estimate this bias was small based on the following: there is no widely accepted cut-off recommendation for CRP use in predicting PSIs; no diagnostic accuracy study validating its value in PSI [17]; the test is widely known to have poor specificity. This knowledge may have led to less CRP test orders in PSI patients as the study progressed (2 in the derivation and 6 in the validation). Despite this, there is potential for this bias to overstate the accuracy of our prediction score.

Since we did not identify all eligible patients who were not enrolled, we cannot exclude that we missed eligible patients who had an infection and did not return to our hospital system. However, all discharged patients were counseled to return if they developed worsening of symptoms, and previous publications indicate that most, if not all, infected patients have at least one of our enrollment criteria. This suggests that few patients would be missed using this screen. [1,2,6,19] There is also the potential of patients presenting with both a PSI and another infectious condition such as endocarditis, where the treatment of the primary infectious condition coincidentally treated the PSI, leading to an undiagnosed spinal infection. We estimate this occurred rarely. We also could have missed patients who had a PSI without back or neck pain. We estimate this occurred rarely.

Not all patients were evaluated using a single reference standard (MRI). However, two investigators reviewed all radiology reports and images, confirmed equivocal readings with culture and operating reports, we defined PSI precisely, and the 21 uninfected cases who had no MR imaging were followed clinically for a prolonged duration to verify no occurrence of infection. This likely represents the best available gold standard. Despite telephone follow-up, extended medical record follow-up, and obituary search, the potential for missed infections exists. Since the study spanned 14 years, it also may have been subjected to temporal bias due to increased MRI availability or improved clinician confidence in the selection and diagnosis of spinal infections over this lengthy period. Over time clinicians may have depended less on well-known high-risk features of PSI and more on acquired expertise, leading to the identification of more PSI patients in the second half of the study who had no fever, no historical risk factors, and who had more missing CRP orders.

Our derivation did not measure interobserver variability of key predictors such as severe spine pain and abnormal neurological examination. However, we mitigated this by using the most objective variables available and those with previously published measurements of interobserver variability (e.g., spine pain so severe the patient is unable to sit up, kappa 0.67 [26] and neurological deficit, kappa 0.69). [27] A small number of patients entered the study with back or neck pain and were later found to have posterior lower lobe pneumonia or pyelonephritis as the cause of their back pain. Had these conditions been recognized prior to spinal MRI order, the study would have resulted in greater CRP and SIRCH score specificities. Our study's small size required us to combine several variables into composite variables, possibly concealing the strength of crucial individual risk factors.

## **6. Discussion**

In 2004 Davis identified diagnostic failure occurred in 75% (47/63) of patients with spinal infection, and delay in treatment was associated with worse sequelae. [1] Our study demonstrates that clinical sensitivity remains weak, with 66% (59/89) of infected patients returning to the ED after an incorrect previous diagnosis. Additionally, Street's study suggests clinician specificity is also unsatisfactory, finding 33 to 1000 MRIs are ordered to diagnose one infection. [18] Authors have recommended various methods to improve the sensitivity, including the use of risk factors or red flags. [7,19,21,31] However, guideline inconsistencies for spinal infection and imprecise risk factor definitions [14,15,16,20] are likely responsible for incomplete adoption of any single recommendation for imaging decisions, and this may be associated with the high diagnostic failure rate cited by Bhise. [7] Our study supports several findings from Davis's seminal study. [6] Both studies are similar in size (89 PSIs versus 86 in Davis's), both have a low proportion of infections with fever (19% (17/89) versus 7.3%) and both studies focused on evaluating patients in whom there was a concern for PSI, and did not assess patients at very low-risk for infection, which is consistent with current guidelines that do not recommend imaging these patients. [2,6,19,21,34] However there are two critical differences: 1. The Davis screen, using risk factors only, was 82% (72/89) sensitive for PSI in our cases, compared to a SIRCH sensitivity of 96% (85/89), indicating SIRCH may be a more sensitive screen; and 2. The SIRCH algorithm considers progressive neurologic deficit a risk factor to be used in screening for PSI and recommends a contrast MRI for patients with a SIRCH score ≥ 3, whereas the Davis protocol considers the CRP unnecessary in the case of neurologic deficit. However, adding contrast to the MRI in this instance avoids the following pitfall: Most patients presenting with a PSI in our study did not have a fever, and likewise, 83% (25/30) who had a neurological deficit did not have a fever either. Clinicians not actively looking for PSI may not suspect infection in this group, and imaging an infected patient without contrast may lead to a missed PSI or an equivocal reading. This circumstance may prompt a neuroradiologist to recommend repeating the MRI scan, but with the addition of contrast, which adds another 4-8 hours [11] to the ED evaluation and to the patient's time in the ED.

MRI utilization is important since its lengthy turn-around-time contributes to ED crowding. In our derivation cohort, the SIRCH score would have reduced the number of unnecessary imaging for PSI by 54 compared to physician suspicion, and by 45 compared to CRP (using the standard laboratory cut-off, Table

2). While all three image prompts were 100% (40/40) sensitive for the infection, given the ubiquity of back pain, CRP testing in unselected patients would likely result in increased MRI overutilization. Various CRP cut-offs have been recommended in the literature. We selected a cut-off unique to our at-risk spine pain cohort to maximize its accuracy for this population, and clinicians using this cut-off should be aware of instances in which the CRP may be lower than our cut-off in PSI patients, especially those with cirrhotic liver disease or concurrent antibiotic use (Table 3). [23,24,33] In this study, the presence of other risk variables heightened suspicion of infection, which maintained our high sensitivity for these cases. Adding further uncertainty to CRP use is the fact that its units are reported in the literature in both mg/L and mg/dl, which differs by a factor of 10, leading readers to erroneous conclusions when units are not reported in a study. The Center for Disease Control recommends standardizing the reporting of CRP units to mg/L. [32]

"To date, there has been no risk prediction tool to assist ED physicians in assessing patients with low back pain." [14] The SIRCH score was 100% sensitive for PSI and prompted fewer MRIs than clinician suspicion or CRP use in our derivation, but was less sensitive in the validation (92%) compared to CRP (98%). Although our investigation characterizes the value of findings from the history, physical examination, and pain character used to estimate the presence of PSI and the need for MR imaging, we cannot recommend its clinical application until large-scale validation in other ED settings are completed. Promising additional candidate variables for future study and refinement of our model include bounce-back visits, arrival by EMS, bilateral paresthesias, bilateral radicular pain, and current antibiotic use.

References:

[1] Davis DP, Wold RM, Patel RJ, Tran AJ, Tokhi RN, Chan TC, et al. The clinical presentation and impact of diagnostic delays on emergency department patients with spinal epidural abscess. J Emerg Med. 2004;26:285-91.

[2] Alerhand S, Wood S, Long B, Koyfman A. The time-sensitive challenge of diagnosing spinal epidural abscess in the emergency department. Intern Emerg Med. 2017;12:1179–423 83. DOI: 10.1007/s11739-017-1718-5.

[3] Babic M, Simpfendorfer CS. Infections of the Spine. Infect Dis Clin North Am. 2017 Jun;31(2):279-297. Doi: 10.1016/j.idc.2017.01.003. Epub 2017 Mar 30.

[4] Pitts SR, Niska RW, Xu J, Burt CW. National hospital ambulatory medical care survey: 2006 emergency department summary. Natl Health Stat Rep 2008;7:1-38.

[5] Dubosh N, Edlow J, Goto T, Camargo C, Hasegawa. Missed serious neurologic conditions in emergency department patients discharged with nonspecific diagnoses of headache or back pain. Ann Emerg Med. 2019;74(4):549-61.

[6] Davis DP, Salazar A, Chan TC, Vilke GM. Prospective evaluation of a clinical decision guideline to diagnose spinal epidural abscess in patients who present to the emergency department with spine pain. J Neurosurg Spine. 2011;14(6):765-70. Doi:10.3171/2011.1.SPINE1091.

[7] Bhise V, Meyer A, Singh H, Wei L, Russo E, Al-Mutairi A, Murphy D. Errors in diagnosis of spinal epidural abscesses in the era of electronic health records. Am J Med. 2017 Aug;130(8):975-981. Doi: 10.1016/j.amjmed.2017.03.009. Epub 2017 Mar 31.

8] Shantharam G, DePasse JM, Eltarai AE, Durand WM, Palumbo MA, Daniels AH. Physician and patient factors associated with outcome of spinal epidural abscess related malpractice litigation. Orthopedic Reviews. 2018;10:7693.

[9] Babic M, Simpfendorfer CS, Berbari EF. Update on spinal epidural abscess. Curr Opin Infect Dis. 2019 Jun;32(3):265-271.

[10] Talbott JF, Shah VN, Uzelac A, NarvidJ, Dumont RA, Chin CT, Wilson DM. Imaging-based approach to extradural infection of the spine. Semin Ultrasound CT MR. 2018 Dec:29(6):570-587. Doi: 10.1053/j.sult.2018.09.003. Epub 2018 Sep 26.

13[11] Madhuripan N, Hicks R, Feldman D, Rathlev N, Salvador D, Artenstein A. A protocol-based approach to spinal epidural abscess imaging improves performance and facilitate early diagnosis. ACR. 2018;15(4):648-651.

[12] Ledbetter LN, Salzman KL, Shah LM. Imaging psoas sign in lumbar spinal infections: evaluation of diagnostic accuracy and comparison with established imaging characteristics. Am J Neuroradiol. 2016 Apr;37(4):736-41. Doi: 10.3174/ajnr.A4571. Epub 2015 Nov 19.

[13] El Sayed M, Witting MD. Low yield of ED magnetic resonance imaging for suspected epidural abscess. Am J Emerg Med 2011;29:978-82.

[14] Galliker G, Scherer DE, Trippolini MA, Rasmussen-Barr E, LoMartire R, Wertli MM. Low back pain in the emergency department: Prevalence of serious spinal pathologies and diagnostic accuracy of red flags – A systematic review. Am J Med. 2019 Jul 3. Doi: 10.1016/j.amjmed.2019.06.005. [Epub ahead of print]

[15] Verhagen A, Downie A, Nahid P, Maher C and Koes B. Red Flags presented in current low back pain guidelines: a review. Eur Spine J. 2016. 25:2788-2802. DOI 10.1007/sOO586-4684-0.

[16] Raison N, Alwan W, Abbot A, Farook M, Khaleel Arshad. The reliability of red flags in spinal cord compression. Arch Trauma Res. 2014;3(1):e17850.

[17] Harris JC, Caesar DH, Davison C, Phibs R, Than MP. How useful are the laboratory investigations in emergency department evaluation of possible osteomyelitis? Suspected osteomyelitis, Emerg Med Australasia 2011;23:317-30.

[18] Street KJ, White SG, Vandal AC. Clinical prevalence and population incidence of serious pathologies among patients undergoing magnetic resonance imaging for low back pain. Spine J. 2019 Sep 10. Pii: S1529-9430(19)30966-0. doi: 10.1016/j.spine.2019.09.002. [Epub ahead of print].

[19] Singleton J, Edlow J. Acute nontraumatic back pain. Risk stratification, and review of serious pathologies. Emerg Med Clin N Am. 2016;24:743-747. http://dx.org/10.1016/j.emc.2016.06.015.

[20] Premkumar A, Godfrey W, Gottschalk M, Boden S. Red flags for back pain are not always really red. J Bone Joint Surg Am. 2018;100(5):368-74.

[21] Patel ND, Broderick DF, Burns J, Deshmukh TK, Fries IB, Harvey HB, Holly L, Hunt CH, Jagadeesan BD, Kennedy TA, O'Toole JE, Perlmutter JS, Policeni B, Rosenow JM, Schroeder JW, Whitehead MT, Cornelius RS, Corey AS. ACR Appropriateness Criteria Low Back Pain. 2016 Sep;13(9):1069-78. Doi: 10.1016/j.jacr.2016.06.008. Epub 2016 Aug 3.

[22] Davis WT, April MD, Mehta S, Long B, Shroyer S. High risk clinical characteristics for pyogenic spinal infection in acute neck or back pain: Prospective cohort study. Am J Emerg Med. 2019 May 17. Pii: S0735-6757(19)30336-5. DOI: 10.1016/j.ajem.2019.05.025. [Epub ahead of print]

[23] Meyer B, Shaller K, Rohde V, Hassler W. The C-reactive protein for detection of early infections after lumbar microdiscectomy. Acta Neurochir 1995;136:145-50.

[24] Rosahl SK, Gharbahi A, Zink PM, Samii M. Monitoring blood parameters following anterior cervical fusion. J Neurosurg 2000;92:169-74. DOI:10.3171/spi.2000.92.2.0169

25] Collins G, Reitsma J, Altman D, Moons K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015 Jan 7;350:g7594. DOI: 10.1136/bmj.g7594. PMID: 25569120.

[26] Gorelick M, Atabaki S, Hoyle J, Dayan P, Holmes J, Kuppermann N, et al. Interobserver agreement in assessment of clinical variables in children with blunt head trauma. Acad Emerg Med. 2008:15(9);812-818.

[27] Viikari-Juntura E, Takala E, Riihimaki H, Malmivaara A, Martikainen R, Jappinen P. Standardized physical examination protocol for low back disorders: feasibility of use and validity of symptoms and signs. J Clin Epidemiol. 1998;51(3):245-255.

[28] Stiell IG, Wells GA: Methodologic standards for the development of clinical decision rules in emergency medicine. Ann Emerg Med. 1999;33:437-447.

[29] Jean M, Irisson J, Gras G, Bouchand F, Simo D, Duran C, Perronne C, Mulleman D, Bernard L, Dinh A. Diagnostic delay of pyogenic vertebral osteomyelitis and its associated factors, Scand J Rheumatol. 2017;46:64-68.

[30] Jeong DK, Lee HW, Kwon YM. Clinical value of procalcitonin patients with spinal infection. J Korean Neurosurg. 2015; 58(3):271-275. http://dx.doi.org/10.3340/jkns.2015.58.3.271

[31] Berbari E, Kanj S, Kowalski T, Darouiche R, Widmer A, Schmitt S, et. al. 2015 Infectious Diseases Society of America (IDSA) Clinical Practice Guidelines for the Diagnosis and Treatment of Native Vertebral Osteomyelitis in Adults. CID. 2015;61(6):859–63

[32] Myers G, Rifai N, Tracy R et al. CDC/AHA Workshop on Markers of Inflammation and Cardiovascular Disease: Application to Clinical and Public Health Practice Report from the Laboratory Science Discussion Group. Circulation. 2004; 110:e545-e549.

[33] Davis WT, April M, Mehta S, Long B, Boys G, Shroyer S. Sensitivity of C-reactive protein cut-off values for pyogenic spinal infection in the emergency department. CJEM. 2020 Jun 15;1-8. doi: 10.1017/cem.2020.402. Online ahead of print.

[34] Gurley K, Edlow J, Burstein J, Grossman S. Errors in decisionmaking in emergency medicine: The case of the landscaper's back and root cause analysis. Ann Emerg Med. 2020. 1-7. In press. doi.org/10.1016/j.annemergmed.2020.05.031.

[35] Steyerberg E, Vickers A, Cook N, Gerds T, Gonen M, Obuchowski N, Pecina M, Kattan M. Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology. 2010;21(1):129-138. Doi:10.1097/EDE.0b013e3181c30fb2.

**Table 1. Patient Characteristics in 223 patients suspected of pyogenic spinal infection**

| | | Derivation | | | | | Sensitivity Validation | |
|---|---|---|---|---|---|---|---|---|
| | Potential Predictor Variables | No Infection N=134 | % | PSI N=40 | % | P-value | PSI N=49 | % |
| 1 | Mean age (SD), years | 53.8 | 18.2 | 52.3 | 14.3 | 0.57 | 56.3 | 12.5 |
| 2 | Gender, male | 40 | 30% | 30 | 75% | <0.001 | 32 | 65% |
| 3 | Any historical risk factor | 84 | 63% | 36 | 90% | 0.001 | 37 | 76% |
| | IVDU | 0 | 0% | 3 | 7.5% | 0.001 | 7 | 14% |
| | Dialysis | 4 | 3.0% | 3 | 7.5% | 0.202 | 2 | 4.0% |
| | Prolonged indwelling IV (PICC, temporary dialysis catheter etc.) | 0 | 0% | 4 | 10% | <0.001 | 7 | 14% |
| | Hx consistent w/ bacteremia or SSTI within 2 wks of Sx onset | 3 | 2.4% | 15 | 38% | <0.001 | 13 | 27% |
| | Immunocompromise | 4 | 3% | 2 | 4.1% | 0.54 | 2 | 4.0% |
| | Diabetes | 39 | 29% | 17 | 43% | 0.11 | 19 | 39% |
| | Cirrhosis | 0 | 0% | 3 | 7.5% | 0.001 | 4 | 8.2% |
| | Spinal implant present (spinal pump, cord stimulator, etc.) | 7 | 5.2% | 0 | 0% | 0.14 | 2 | 4.1% |
| | Spinal fracture diagnosed ≥2 wks prior to presentation | 0 | | 0 | | | 0 | |
| | Spine procedure in past 3 months | 45 | 34% | 14 | 35% | 0.87 | 15 | 31% |
| 4 | ED fever or Hx or measured fever | 30 | 22% | 23 | 58% | <0.001 | 10 | 20% |
| | ED fever (≥38˚C in ED)† | 18 | 13% | 12 | 30% | 0.017 | 5 | 10% |
| | Hx of measured fever (≥38˚C) and no ED fever | 12 | 9% | 11 | 28% | 0.002 | 5 | 10% |
| 5 | New neuro deficit (spine related) | 28 | 21% | 15 | 38% | 0.033 | 15 | 31% |
| | New extremity weakness | 21 | 16% | 9 | 21% | 0.316 | 8 | 16% |
| | Overflow incontinence by Hx | 8 | 6.0% | 8 | 20% | 0.007 | 7 | 14% |
| | Extremity numbness | 14 | 10% | 6 | 15% | 0.430 | 4 | 8.2% |
| | Reflex abnormality | 5 | 3.7% | 5 | 13% | 0.037 | 4 | 8.2% |
| 6 | Bounce-back within 2 wks | NA | NA | 25 | 63% | | 34 | 69% |
| 7 | Pain Characteristics | | | | | | | |
| | Worst pain ever | 15 | 11% | 9 | 23% | 0.070 | 17 | 35% |
| | Intermittent radicular | 23 | 17% | 2 | 5.0% | 0.008 | 12 | 24% |
| | Constant severe radicular | 30 | 22% | 7 | 18% | 0.561 | 19 | 39% |
| | Intermittent or constant radicular | 51 | 38% | 9 | 23% | 0.070 | 27 | 56% |
| | Unable to sit up independently due to pain | 30 | 22% | 15 | 38% | 0.044 | 23 | 47% |
| | Unable to ambulate due to Pain | 31 | 23% | 16 | 40% | 0.036 | 6 | 12% |
| 8 | Temperature, mean, (SD); ˚C | 37.0 | 0.86 | 37.5 | 1.2 | 0.01 | 37.0 | 1.5 |
| 9 | Mean arterial pressure, (SD); mmHg | 98.8 | 15.7 | 92.8 | 23.3 | 0.161 | 96.2 | 15.1 |
| 10 | HR, mean, (SD); beats/minute | 90.1 | 19.5 | 94.7 | 17.9 | 0.121 | 93.4 | 16.6 |
| 11 | WBC, mean, (SD); cells/µL | 10.0 | 4.6 | 12.1 | 5.1 | 0.001 | 12.8 | 6.3 |
| 12 | CRP, mean, (SD); mg/L | 47.9 | 64.9 | 140.2 | 98.1 | <0.001 | 145.5 | 93.9 |

†ED fever=1st temperature taken in ED ≥ 38˚C (100.4˚F), Inc=incomplete data collected; ‡the total % patients imaged is greater than 100% since some patients had multiple imaging modalities; IVDU=intravenous drug use; IV= intravenous; PICC=peripheral inserted; central line; NA=not available; SSTI=skin and soft tissue infection; wks=weeks; Sx=symptoms, Hx=history; HR=heart rate; PMR= polymyalgia rheumatica

Figure 1. Patient Flow

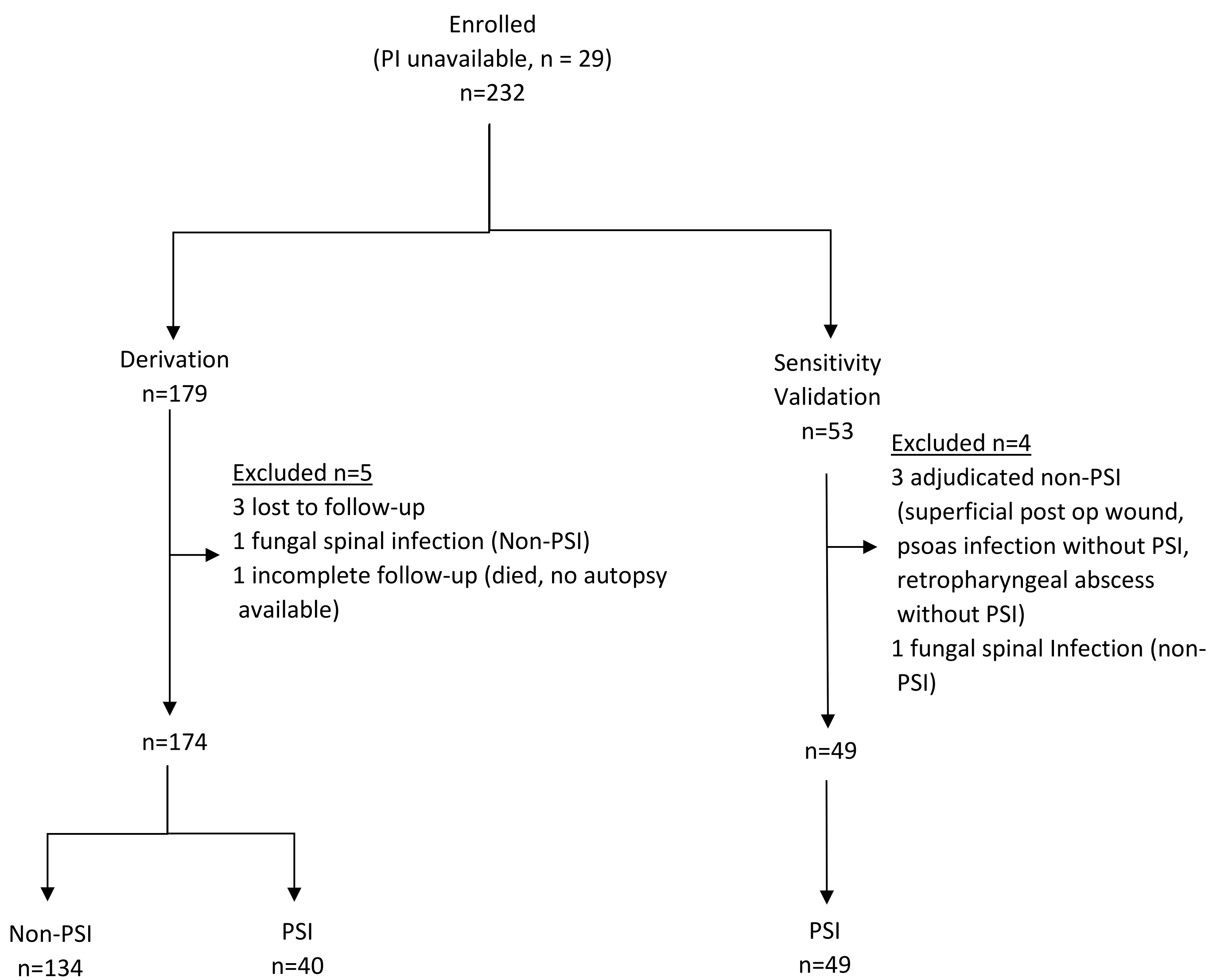
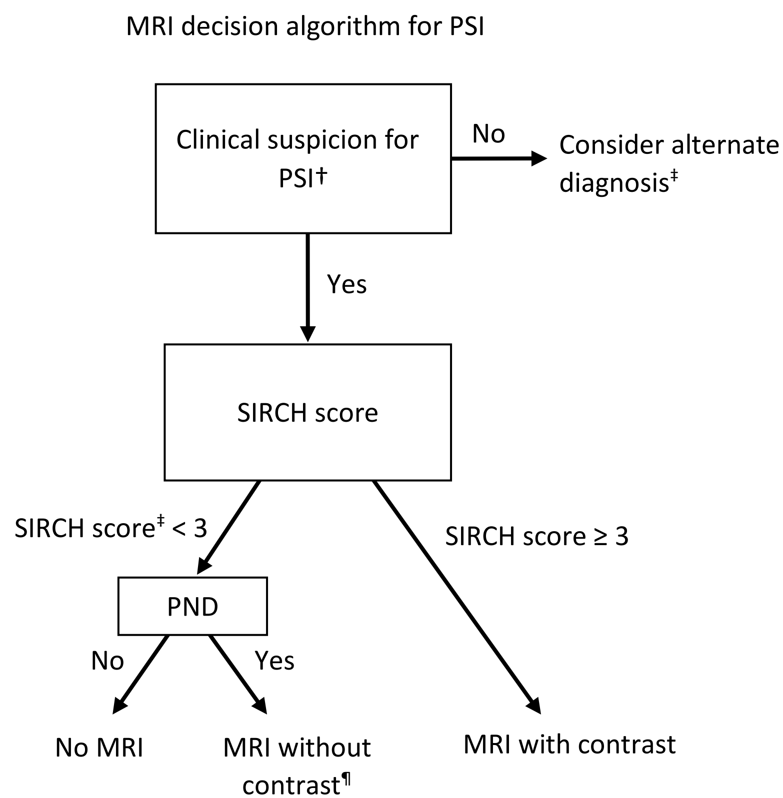


Figure 2. Recommended SIRCH Algorithm




† Clinical suspicion= enrollment criteria; ‡ Per published recommendations for patients at very low-risk for PSI[2,3,19,21]; § SIRCH Score = RF-2, F or HxF-2, PND-1, CRP≥50-2 (Range 0-7); ¶ Patients with a progressive neurologic deficit and a score <3 require MRI without contrast. PND=progressive neurological deficit, PSI=pyogenic spinal infection, MRI=magnetic resonance imaging

**Table 2. SIRCH score performance measures for PSI detection, MRI overutilization and validation sensitivity**

| Imaging prompt | Sensitivity | | | Specificity | | 95%CI | LR+ | 95% CI | LR- | 95% CI | PPV | 95% CI | Unnecessary MRI orders † | Sensitivity | | 95% CI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Derivation | | | | | | | | | | | | | Sensitivity Validation | | |
| **SIRCH score ≥ 3** | 40/40 | 100% | 91-100 | 75/134 | 56% | 47-65 | 2.3 | 1.9-2.7 | 0 | 0-5 | 40% | 31-51% | 59 | 45/49 | 92% | 80-97 |
| **CRP (cut-off of 3.5)‡** | 40/40 | 100% | 91-100 | 30/134 | 22% | 16-30 | 1.3 | 1.2-1.5 | 0 | NA | 28% | 21-36% | 104 | 48/49 | 98% | 89-100 |
| **Clinician suspicion§** | 40/40 | 100% | 91-100 | 0/134 | 0% | 0-0.03 | 1 | 0.91-1.0 | NA | NA | 23% | 17-30% | 113¶ | 49/49 | 100% | 93-100 |

†Numbers do not account for 7 patients who required MRI for neuro-deficits but had no PSI. ‡CRP > 3.5 is the standard laboratory cut-off for abnormal; §Clinical suspicion= enrollment criteria, ¶Of 134 non-PSI patients 21 had no imaging and were followed clinically. NA=Not Applicable.

**Table 3. Patient characteristics in SIRCH score failures and near-misses**

| Case No./Cohort | Historical risk factor | ED Fever | ED Fever or Hx of Fever | Progressive neuro-deficit | CRP (mg/L) | Bounce-back† | SIRCH Score | Infecting organism | Specific findings |
|---|---|---|---|---|---|---|---|---|---|
| 20D | 1 | 0 | 1 | 0 | 35.6 | 1 | 3 | GpD strep | >2 wk incr low back pain, Hepatitis C cirrhosis & diabetes; MRI: D, VO; +T9-10 disc aspirate |
| 67D | 1 | 0 | 1 | 0 | 32.3 | 1 | 3 | E coli | >3 wk incr low back pain, on fluoroquinolone for prostatitis, changed to DOX; MRI: L1-2 SEA, D, VO; +quinolone resistant E coli L1-2 disc aspirate |
| 133D | 1 | 0 | 1 | 0 | 33.0 | 1 | 3 | MSSA | 1 wk incr low back pain 4 wk after lumbar fusion surgery, MRI over-read the next day "concern for SEA", Pt called to return to the ED, CRP incr from 17.4 to 33.0 in 24 hours; Pt had recent MSSA PICC line infection |
| 254V | 0 | 0 | 0 | 0 | 266.0 | 1 | 3 | S. epi | >6 d incr low back pain, +chills & sweats, repeat temp 38.9˚(rectal); MRI: L5S1 D, SEA; +spinal canal culture |
| 262V | 0 | 0 | 0 | 0 | 146.0 | 1 | 3 | MSSA | >6 d incr low back & right leg pain in 19 year-old healthy male, too painful to walk, "worst pain ever"; MRI: L4-S1 SEA, PA, large iliacus abscess drained by interventional radiology, treated non-surgically |
| 265V | 1 | 0 | 0 | 1 | 10.6 | 1 | 3 | MRSA | 4 wk incr neck pain, radicular left arm pain & weak shoulder abduction, diabetes; MRI: left C3-4 SF, PSA; ESR 50, CRP level re-verified next day; +needle aspirate |
| 267V | 0 | 0 | 0 | 0 | 43.5 | 1 | 0 | GpB strep | 8 d incr back & bilateral leg pain "worst pain ever"; MRI: L3-4 SF, SEA, PA; ESR 40; +canal culture; follow up 2 years 5 mos later Pt diagnosed with end stage cirrhotic liver disease from NASH |
| 269V | 1 | 0 | 0 | 0 | 41.4 | 1 | 2 | MSSA | Incr chronic low back pain after temporary spinal cord stimulator placed 6 d prior, diabetes; repeat temp 38.8˚; MRI: T12-L1 SEA; +spinal canal culture |
| 275V | 1 | 0 | 0 | 1 | 29.7 | 1 | 3 | MSSA | 4 d incr thoracic back pain, leg paresthesias, leg weakness, hyper-reflexic; Pt on Abx for recent MSSA bacteremia; MRI: T5-6 SEA with cord compression, VO, D and PVA; +canal culture |
| 280V | 0 | 0 | 0 | 1 | 17.0 | 1 | 1 | MRSA | 14 d incr neck/thoracic spine pain, new stress urinary incontinence & use of a walker to ambulate prior 2 days; Pt recently treated with CRO & TMP for "UTI"; MRI: C2-4 SEA; +canal culture |
| 295V | 0 | 0 | 0 | 0 ‡ | 3.2 | 0 § | 0 | MSSA | Pt returned to the ER after syncope evaluation for incr pain & paresthesias to ulnar aspect of both arms. MRI ordered for possible central cord syndrome from neck injury. MRI:C5-6 VO, D, PVA; 2 of 2 blood cultures + MSSA; ESR 9; CRP re-sampled and verified, neg-IVDU status re-verified, treated non-surgically |

†Bounce back=patient not diagnosed on first visit; ‡MRI ordered based on suspicion of central cord syndrome but adjudicated not a PND; §Not considered a bounce back since patient's first visit unrelated to PSI; 1=present, 0=absent; SEA= spinal epidural abscess, VO= vertebral osteomyelitis, D= discitis, PVA= paravertebral abscess, PSA=paraspinous abscess, SF= septic facet, PA=psoas abscsess; Abx=antibiotics, CRO=ceftriaxone, DOX=doxycyline, SXT= trimethoprim-sulfamethoxazole, SSTI=skin and soft tissue infection, MSSA=methicillin sensitive Staph aureus, MRSA=methicillin resistant Staph aureus: S epi=Staph epidermidis; d=day, wk=week; incr=increasing; ESR=erythrocyte sedimentation rate; NASH=non alcoholic steatohepatitis; Pt=patient

**Table 4. SIRCH score statististical performance at varied criterion cut-offs**

| Score | Sensitivity | % | 95% CI | Specificity | % | 95% CI | PPV | Probability of Infection | 95% CI | NPV | Probability of No Infection | 95% CI | LR+ | 95%CI | LR- | 95%CI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **≥7** | 5/40 | 12.5 | 4.2-27 | 132/134 | 99 | 95-100 | 5/7 | 0.714 | 29-96 | 132/167 | 0.79 | 72-85 | 8.4 | 1.7-42 | 0.89 | 0.8-1.0 |
| **≥6** | 24/40 | 60 | 43-75 | 119/134 | 89 | 82-94 | 24/39 | 0.61 | 73-96 | 119/135 | 0.88 | 81-93 | 5.4 | 3.1-9.2 | 0.45 | 0.3-0.7 |
| **≥5** | 35/40 | 88 | 73-96 | 105/134 | 78 | 70-85 | 35/64 | 0.547 | 42-67 | 105/110 | 0.96 | 90-99 | 4.0 | 2.9-5.7 | 0.16 | 0.07-0.4 |
| **≥4** | 37/40 | 93 | 80-98 | 100/134 | 75 | 66-82 | 37/71 | 0.52 | 40-64 | 100/103 | 0.97 | 92-99 | 3.7 | 2.7-4.9 | 0.1 | 0.03-0.3 |
| **≥3** | 40/40 | 100 | 91-100 | 75/134 | 56 | 47-65 | 40/99 | 0.4 | 31-51 | 75/75 | 1 | 95-100 | 2.3 | 1.9-2.7 | 0 | |
| **≥2** | 40/40 | 100 | 91-100 | 41/134 | 31 | 23-39 | 40/133 | 0.3 | 22-39 | 41/41 | 1 | 91-100 | 1.4 | 1.3-1.6 | 0 | |
| **≥1** | 40/40 | 100 | 91-100 | 29/134 | 22 | 15-30 | 40/145 | 0.27 | 21-36 | 29/29 | 1 | 88-100 | 1.3 | 1.2-1.4 | 0 | |
| **≥0** | 40/40 | 100 | 91-100 | 0/134 | 0 | 0-0.03 | 40/174 | 0.23 | 17-30 | 0/0 | | | 1 | 1.0-1.0 | 0 | |

A prevalence less than this study's 23% should expect lower PPV estimates.

Figure 3. Probability of PSI at varied SIRCH score criterion cut-offs (error bars are 95% CI).