Value of Spinal Infection Treatment Evaluation Score, Pola Classification, and Brighton Spondylodiscitis Score from Decision to Surgery in Patients with Spondylodiscitis: A Receiver-Operating Characteristic Curve Analysis

Majid Rezvani1,*, Ali Ahmadvand2,*, Taravat Yazdanian3,*, Parisa Azimi4, Navid Askariardehjani2

1Department of Neurosurgery, Neuroscience Research Center, Al-Zahra Hospital, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
2Department of Neurosurgery, Isfahan University of Medical Sciences, Isfahan, Iran
3School of Medicine, Capital Medical University, Beijing, China
4Neuroscience Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Study Design: This was a retrospective study.

Purpose: This study aimed to assess the value of the Spinal Infection Treatment Evaluation (SITE) score, Brighton Spondylodiscitis Score (BSDS), and Pola classification to predict the need for surgical intervention in patients with spondylodiscitis.

Overview of Literature: Spondylodiscitis is a rare disease, and the prediction of its outcome is crucial in the decision-making process.

Methods: All case records were assessed to extract information on the American Spinal Injury Association (ASIA), Visual Analog Scale (VAS), and Japanese Orthopedic Association Back Pain Evaluation Questionnaire (JOABPEQ) scores before and after surgery. The SITE score, Pola classification, and BSDS were recorded. The receiver-operating characteristic (ROC) curve analysis and the area under the curve (AUC) were applied to estimate the predictive ability of the scoring systems. Patients’ satisfaction with surgery outcomes was evaluated using the VAS, ASIA, JOABPEQ, and Likert scale for quality-of-life evaluation.

Results: In all 148 patients, case records were reviewed. The mean±standard deviation age of the patients was 54.6±14.7 years. Of these, 112 patients underwent surgery. The AUC scores were 0.86, 0.81, and 0.73 for the SITE score, BSDS, and Pola classification, respectively. In the comparison of the AUC of ROC curves, SITE score vs. BSDS showed a significantly greater AUC, 0.13 (Z=2.1, p=0.037); SITE score vs. Pola classification, 0.05 (Z=0.82, p=0.412); and Pola classification vs. BSDS, 0.08 (Z=1.22, p=0.219). The optimal cutoff score was 8.5 (sensitivity, 80.6%; specificity, 81.2%) for the SITE score and 9.5 (sensitivity, 52.8%; specificity, 83.0%) for the BSDS in the decision to surgery. VAS back pain and JOABPEQ subscales showed a significant difference when compared with preoperative scores. According to ASIA grading, none of the patients experienced neurological deterioration. Overall, patients’ satisfaction was observed.

Conclusions: The findings suggest that the SITE score is a useful measure and helps clinicians make clinically sound decisions in patients with spondylodiscitis.

Keywords: Spondylodiscitis; Decision making; Spinal Infection Treatment Evaluation score; Brighton Spondylodiscitis Score; Pola classification
Introduction

Spondylodiscitis is a severe infectious spine disease that involves the intervertebral disk space and adjacent vertebrae and is an increasing healthcare problem [1]. Early diagnosis and targeted treatment are vital to reduce the risk of serious complications. However, the appropriate treatment is still debatable for these patients [2]. The main treatment support for spondylodiscitis tends to be conservative, such as bed rest, antibiotic therapy, and optimal spinal stabilization. Surgery is recommended when conservative treatment fails or when instabilities or complications occur to eliminate the infection focus and restore spinal stability [2]. However, the choice of treatment method is still controversial.

With an increasingly aging and multimorbid population, spondylodiscitis has become more common, and a clear understanding of which cases need surgery is essential [3]. However, surgical indications for patients with spondylodiscitis were determined according to published guidelines. For many patients, decision-making is challenging. The decision-making tools were developed and validated to examine the factors that affect the decision-making process [3-5]. Tools will never replace human experts; however, they can help in screening and be used by experts to double-check their decisions. Various tools such as Spinal Infection Treatment Evaluation (SITE) score [4], Brighton Spondylodiscitis Score (BSDS) [3], and Pola classification [5], are available for predicting the need for surgical intervention in patients with spondylodiscitis. The abovementioned tools were established to calculate patient-specific surgical risks based on variables associated with their disease. This question remains unanswered: which is the best classification for clinicians to use in surgical decision-making to double-check decisions? Hence, this study aimed to assess and compare the predictive value of the SITE score, BSDS, and Pola classification for surgery or conservative treatment in patients with spondylodiscitis. It will also discuss the clinical summary, surgical procedures, and outcomes.

Materials and Methods

This retrospective study included 148 patients who had undergone surgery or received conservative treatment for spondylodiscitis between September 2018 and January 2023 in a large teaching hospital in Isfahan, Iran. Data were collected through a review of patients’ records. The diagnosis of patients with spondylodiscitis was based on clinical findings, laboratory results, radiological assessment, and magnetic resonance imaging (MRI) [6]. All patients with spinal anomalies, aged <18 years, and had metastatic or primary tumors from the study were excluded.

All patients received conservative treatment including intravenous antibiotic therapy for at least 4 weeks, followed by a 2-week oral course until the normalization of the results of laboratory tests for infection [6]. The surgical indications for the cases included in the study were as follows: failure of conservative treatment, neurologic deficits, spinal instability, sepsis, and intraspinal empyema.

Cases were reviewed to determine whether the patients should undergo surgery or receive conservative treatment. Clinical information including age, sex, body mass index (kg/m²), level of infection, comorbidity, and complications were extracted. The start of the study was considered from the time of spondylodiscitis diagnosis.

1. Additional measures

(1) VAS and ASIA Impairment score
Clinical outcomes were evaluated by the Visual Analog Scale (VAS) [7] for back pain and the American Spinal Injury Association (ASIA) Impairment score [8] for neurological evaluation. The ASIA Impairment Scale for classifying spinal cord injury is as follows: A, complete injury; B, sensory incomplete; C, motor incomplete with a muscle grade <3; D, motor incomplete with a muscle grade >3; and E, normal.

(2) JOABPEQ score
The Japanese Orthopedic Association Back Pain Evaluation Questionnaire (JOABPEQ) score is a diseasespecific instrument for low back pain and contains 25 items tapping into five subscales: social function (four items), mental health (seven items), lumbar function (six items), walking ability (five items), and low back pain (four items). The scores for each subscale range from 0 to 100, with higher scores indicating better conditions [9]. The JOABPEQ subscale scores were calculated at baseline and the last follow-up surgery. In this study, low back pain, lumbar function, and walking ability were considered.

(3) SITE score
This novel scoring system was presented by Pluemer et
al. [4] for the evaluation of de novo spinal infection treatment. It consists of five categories, namely, neurology, location, radiology, pain, and host comorbidities. A detailed description of the SITE score and related points is provided in Table 1. The possible score on the SITE score ranges from 0 to 15 points, with higher scores indicating better health status. The SITE scores are classified as severe spinal infection (0–8), in which surgery is recommended; moderate spinal infection (9–12), in which medical treatment is recommended and surgical treatment is an option; mild spinal infection (13–15), in which medical treatment is recommended. The BSDS was calculated for each case, allowing for stratification into severe, moderate, and mild spinal infection groups.

(4) BSDS
The BSDS was established to recognize patients with spondylodiscitis who would likely fail conservative treatment and thus benefit from earlier surgical intervention. It consists of six categories, including distant site infection, medical comorbidity, immunocompromise, MRI characteristics, anatomical location, and neurological impairment. The possible BSDS ranges from 6 to 35 points, with higher scores indicating greater disability [3]. Scores are classified into low (6–14), moderate (15–20), or high (≥21) risk. The point system in the BSDS is shown in Table 2 [3]. The BSDS was calculated for each case, allowing for stratification into low-, moderate-, and high-risk groups.

(5) Pola classification
Pola et al. [5] proposed a classification of pyogenic spondylitis as a guide to surgical management [5]. This classification contains three main types (A, B, and C) according to the primary criteria of MRI: bone destruction of segmental instability, epidural abscess, and neurological impairment. Subclasses of type A consist of A1, A2, A3, and A4; those of type B include B1, B2, and B3; and those of type C consists of C1, C2, C3, and C4, with a detailed

Table 1. The Spinal Infection Treatment Evaluation score

<table>
<thead>
<tr>
<th>Variable</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurology</td>
<td></td>
</tr>
<tr>
<td>Acute plegia or bladder/bowel dysfunction</td>
<td>Surgery</td>
</tr>
<tr>
<td>Motor dysfunction</td>
<td>1</td>
</tr>
<tr>
<td>Sensory dysfunction</td>
<td>2</td>
</tr>
<tr>
<td>Neurologically intact</td>
<td>3</td>
</tr>
<tr>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>Junctional (occiput–C2, C7–T2, T11–L1, L5–S1)</td>
<td>1</td>
</tr>
<tr>
<td>Mobile (C3–6, L2–4)</td>
<td>2</td>
</tr>
<tr>
<td>Semirigid (T3–10)</td>
<td>3</td>
</tr>
<tr>
<td>Rigid (S2–5)</td>
<td>4</td>
</tr>
<tr>
<td>Radiology</td>
<td></td>
</tr>
<tr>
<td>Spinal canal stenosis with impingement of central neural elements with or without de novo deformity</td>
<td>1</td>
</tr>
<tr>
<td>Segmental angulation or translation with de novo deformity/foraminal stenosis or erosion of vertebral body on CT &gt;50% or PL involvement on both sides</td>
<td>2</td>
</tr>
<tr>
<td>Visible endplate erosion on CT or edema of vertebral body &gt;50% on MRI or intervertebral disc involvement on MRI or PL involvement on 1 side</td>
<td>3</td>
</tr>
<tr>
<td>None of these radiological findings</td>
<td>5</td>
</tr>
<tr>
<td>Pain</td>
<td></td>
</tr>
<tr>
<td>Standing axial pain or inability to ambulate</td>
<td>0</td>
</tr>
<tr>
<td>Other pain with the ability to ambulate</td>
<td>1</td>
</tr>
<tr>
<td>No pain</td>
<td>2</td>
</tr>
<tr>
<td>Host comorbidities</td>
<td></td>
</tr>
<tr>
<td>Intra-venous drug abuse or diabetes mellitus</td>
<td>0</td>
</tr>
<tr>
<td>Other comorbidities or no comorbidities</td>
<td>1</td>
</tr>
</tbody>
</table>

CT, computed tomography; MRI, magnetic resonance imaging; PL, posterior ligament.
In the absence of neurological deficits and significant instability, pyogenic spondylitis can be managed without surgical treatment (types A1–A4, B1–B2, and C1) [5].

(6) The 5-point Likert scale
Overall patient satisfaction was evaluated with a 5-point Likert scale. The questions included, “Are you satisfied with the results of the treatment?” Patients’ response on a Likert scale: 1, very satisfied; 2, somewhat satisfied; 3, neither satisfied nor dissatisfied; 4, somewhat dissatisfied; or 5, very dissatisfied.

2. Surgical technique
The surgical technique detail was the same as previously described [10]. Patients’ satisfaction with surgery outcomes was evaluated using the VAS, ASIA, JOABPEQ, and Likert scale for quality-of-life evaluation.

3. Statistical analysis
Descriptive statistics were used to explore the data. Differences in baseline characteristics were compared between patients who received surgery versus those who received conservative treatment. Student t-tests was used for continuous data and χ² tests for categorical data. Continuous variables (BSDS and SITE score) were categorized into classes by selecting the best cutoffs (receiver-operating characteristic analysis [ROC]). Discrimination was tested by using the ROC curves and evaluating areas under the curve (AUC). AUC values were compared as described by Hanley and McNeil [11] using MedCalc Software ver. 22.013 (MedCalc Software Ltd., Ostend, Belgium; accessed September 18, 2023). The optimal cutoff value was calculated using Youden’s index [sensitivity+(specificity−1)]. The AUC was interpreted follows: fail (0.50–0.60), poor (0.60–0.70), fair (0.70–0.80), good (0.80–0.90), and excellent (0.9–1.0). Values of ≤0.05 were considered significant. Statistical analysis of the data was performed using PASW SPSS ver. 18.0 (SPSS Inc., Chicago, IL, USA).

4. Ethics
The Ethics Committee of Isfahan University of Medical Sciences, Isfahan, Iran approved the study (ref number: IR.MUL.MED.REC.1400.567). The requirement for informed consent from individual patients was omitted because of the retrospective design of this study.

Results
Of the 148 patients, 112 underwent surgery, and the remaining 36 received conservative treatment. The characteristics of the patients with spondylodiscitis and their scores on the VAS, ASIA, JOABPEQ, BSDS, SITE, and Pola classification are shown in Table 3. No differences in sex or age were found between the surgery and non-surgery groups. The surgery group was followed up for at least 6 months postoperatively. The mean±standard deviation clinical follow-up was 18.7±2.1 months. As a clinical example, pre- and postoperative images of a 65-year-old patient are presented.
woman are shown in Fig. 1. The comorbidities identified in the surgery group included diabetes mellitus (DM; n=19), dialysis (n=10), brucellosis (n=2), kidney cancer (n=2), addiction (n=2), lupus (n=2), and prostate cancer (n=1), and that in the nonsurgery group was DM (n=3). For clinical assessment, VAS, JOABPEQ, and ASIA scores were compared before and after surgery. The results and patients’ satisfaction are shown in Table 4.

Regarding pain and functional improvement, the average change in the VAS score and JOABPEQ subscales scores were significantly different when compared with preoperative scores (all \( p < 0.0001 \)). According to the ASIA scale criteria, no patients had worse neurological findings postoperatively. Based on the Likert scale, the overall patient satisfaction was observed in 88 patients (78.6%) who underwent surgery.

In this study, 22 patients (19.6%) had postoperative complications, including deep vein thrombosis (n=10), wound infection (n=5), proximal junctional kyphosis, and distal junctional kyphosis (n=7).

The ROC curves of the three tools are shown in Fig. 2. The AUC values for the SITE score, BSDS, and Pola classification were 0.86 (95% confidence interval [CI], 0.78–0.94; standard error [SE], 0.041; \( p < 0.0001 \)), 0.73 (95% CI, 0.64–0.82; SE, 0.047; \( p < 0.0001 \)), and 0.81 (95% CI, 0.72–0.89; SE, 0.045; \( p < 0.0001 \)), respectively. These findings suggested that the SITE score and Pola classification had good sensitivity and specificity for discriminating between patients with spondylodiscitis when deciding on surgery. The BSDS had a lower AUC, which can be inter-
Fig. 2. Receiver-operating characteristic (ROC) of the Spinal Infection Treatment Evaluation (SITE), Brighton Spondylodiscitis Score (BSDS), and Pola scores with regard to decision to surgery. (A) The SITE score is inversely related to outcome and the BSDS and the Pola scores have a direct relation. (B) The BSDS and the Pola scores are inversely related to outcome and the SITE score has a direct relation. It is noted that a higher SITE score indicates better health status. However, a higher BSDS score indicates greater disability.

Fig. 1. Sagittal and axial computed tomography (CT) scan and sagittal and axial T2 weighted image (T2W) magnetic resonance imaging (MRI). Sagittal (A) and axial (B) CT scan and sagittal (C) and axial (D) T2W MRI showed preoperative findings of spondylodiscitis at T10–T11 level. (C) The operation was carried out with posterior instrumentation and allograft autograft combination bone chips interbody fusion at T10–T11 level. At 1 month and 5 months after surgery, sagittal CT scan (E) and sagittal T2W MRI (F), respectively, showed rigid fixation and fusion from T6 to L1.
pretend as fair sensitivity and specificity. In addition, the comparison of the AUC of the ROC curves revealed the following: SITE score versus BSDS, 0.13 (Z, 2.1; p=0.037) was statistically significant; SITE score versus Pola classification was slightly larger but not significant, 0.05 (Z, 0.82; p=0.412); Pola classification versus BSDS, 0.08 (Z, 1.22; p=0.219), suggesting similar AUC between the BSDS and Pola classification. Overall, these findings suggested that the prediction ability of the SITE score was better than that in another staging system for making surgical decisions. The optimal cutoff score was ≤8.5 (Youden index, 0.62; sensitivity, 80.6%; specificity, 81.2%) for the SITE score and ≥9.5 (Youden index, 0.36; sensitivity, 52.8%; specificity, 83.0%) for the BSDS when deciding for surgery.

**Discussion**

We reported the first case series of patients with spondylodiscitis to assess and compare the predictive value of the SITE score, BSDS, and Pola classification. Overall, our results showed that the SITE score is a good marker when making surgical decisions.

To the best of our knowledge, no guidelines exist for surgery versus conservative treatment relative to patients with spondylodiscitis, nor is there a threshold standard for operative intervention [3-5]. Our results show an acceptable accuracy performance for the SITE score in patients with spondylodictis. The discrimination between surgery and conservative treatment appeared to be effective for the SITE score when compared with the other scores. Pluemet al. [4] reported a cutoff value of ≤8 (sensitivity, 100.0%; specificity, 95.7%). This observation is in accordance with our study results that the optimal cutoff value was ≤8.5 (sensitivity, 80.6%; specificity, 81.2%) when the panel decided in favor of surgery. We included numerous cases, much higher than those included in the original study. The cutoff value of ≤8 may be used in clinical practice. Therefore, it may be an effective indicator for determining treatment strategies such as surgery or conservative treatments. However, to increase the reliability of the cutoff value, it is essential to apply it to other cohorts. Moreover, given that numerous factors may affect the accuracy of the cutoff value, clinicians should be cautious when applying our results to patients with spondylodiscitis.

Urrutia et al. [12] reported that neurological deficits and an abscess were associated with an increased rate of surgery. Researchers presented that distant site infection [13], DM [14], and an immune deficiency [15] are known risk factors for neurological deficit, which is a main indication for surgery in patients with spondylodiscitis. In this study, all patients with spondylodiscitis undergoing dialysis (n=10, 8.9%) had undergone surgery. Chronic kidney disease affects the central nervous system and the peripheral nervous system [16]. Uremic peripheral neuropathy is the most common neurological complication of patients undergoing dialysis, which manifests with pain, loss of sensation, and motor weakness [16]. Moreover, dialysis is an immunosuppressive disease that can affect bone quality and is associated with increased bone destruction [17]. Thus, dialysis in patients with spondylodiscitis is recommended to be considered as an independent factor in the SITE score and such cases be classified as DM.

Recently, Urrutia et al. [12] and Hunter et al. [18] conducted two validation studies for the BSDS; however, the external validations were not confirmed. Compared with
AUCs of 0.83 and 0.71 (95% CI, 0.50–0.88) in the original study and test populations [3], an AUC of 0.47 (95% CI, 0.22–0.71) was reported in their external validation [18], which is contradictory. However, in the present study, an AUC of 0.73 (95% CI, 0.64–0.82) was observed, which is in accordance with the original study [3]. The rates of surgical intervention in the moderate- and high-risk groups were 100% [3], 51% [18], 41.7% [12], and 68.2% [12] for the modified BSDS and 93.5% in the present study. In addition, in the present study, 67.6% of the patients in the low-risk group received surgical intervention. Thus, with the abovementioned considerations, it appears that BSDS classification and the modified BSDS are not precisely defined in the low-risk, moderate, and high-risk groups [12]. This could be the reason for the low AUC of the BSDS when compared with the SITE score and Pola classification in this study.

In the original article, patients (n=250) were classified based on the Pola classification: type A, 84 (33.6%); type B, 46 (18.4%); and type C, 120 (48.0%). The rates of surgical intervention in types A, B, and C were 20%, 73.7%, 93.0% in the present study and 11.9%, 63.0%, and 59.2% in the original article, respectively. The number of patients who received surgery were as follows: types A (n=10), B3 (n=29), C2 (n=15), C3 (n=30), and C4 (n=26) [5]. The rates of surgical intervention in the original article and the present study were 90.9% and 52.7% in types B3, C2, C3, and C4 and 0.0% and 44.6% in types B1, B2, and C1, respectively. A significant difference between the two studies was observed when deciding on surgery based on the types of spondylodiscitis. Therefore, it appears that the Pola classification is not precisely defined as a guide to surgical management based on types A, B, and C.

In the surgery group, improvement in VAS and JOAB-PEQ scores and patients' satisfaction demonstrate that an acceptable treatment algorithm ensures effective pain relief and good quality of life. Researchers found that based on ASIA grading, none of the patients had neurological deterioration after surgery [10,19], which is in line with our findings.

This study has some limitations. First, this study was performed in a single teaching hospital; thus, the generalizability of the findings to external populations remains uncertain. Second, this study is subject to potential bias because this study was not prospectively designed. Third, this study included patients with heterogeneous characteristics (e.g., age and sex), which might have resulted in a selection bias. Fourth, numerous factors may alter the SITE score, which were not considered. In addition, the SITE score is still primitive and needs to be re-evaluated and modified based on future research efforts. Fifth, the inconsistent cutoff value of the SITE score remains one of the limitations of validation in clinical practice. Thus, prospective studies are needed to define the cutoff value of the SITE score that will optimize score prediction in different diseases. Finally, when making clinical decisions, statistical or clinical methods may be used in making specific predictions. Given that these methods will not always yield the same results, clinicians must make a reasonable decision according to the use of their informal clinical judgment or a formal statistical process [20]. The SITE score will never substitute human expert decision-makers; however, it can assist in double-checking the routine decision-making process.

**Conclusions**

These results suggest that the SITE score determined at the time of spondylodiscitis diagnosis may be useful for predicting the need for surgical intervention of the affected patients. However, more studies with a larger sample are needed to evaluate these scoring systems as tools for improving decision-making.

**Conflict of Interest**

No potential conflict of interest relevant to this article was reported.

**Acknowledgments**

The authors thank the staff of the Neurosurgery Unit and the Neuroscience Research Center, Al-Zahra Hospital, Isfahan University of Medical Sciences, Isfahan, Iran.

**ORCID**

Majid Rezvani: https://orcid.org/0000-0002-9394-7150;
Ali Ahmadvand: https://orcid.org/0000-0003-2937-6514;
Taravat Yazdian: https://orcid.org/0009-0009-8978-2443;
Parisa Azimi: https://orcid.org/0000-0002-6934-8908;
Navid Askariardehjani: https://orcid.org/0000-0001-7117-4591
Author Contributions

Conceptualization: MR, AA, PA; data curation: MR, AA, NA; formal analysis: PA, TY; methodology: PA, AA, TY; project administration: AA, PA, MR; visualization: MR, PA, TY; writing--original draft: PA; TY; and final approval of the manuscript: all authors.

References