E-ISSN: 2147-673X

MEDITERRANEAN JOURNAL OF INFECTION, MICROBES AND ANTIMICROBIALS

Year: 2026
Volume: 15

www.mjima.org

CASE REPORT / OLGU SUNUMU

DOI: 10.4274/mjima.galenos.2025.25563.1
Mediterr J Infect Microb Antimicrob 2026;15:25563.1
Erişim: http://dx.doi.org/10.4274/mjima.galenos.2025.25563.1

Salmonella enteritidis-Associated Thoracic Spondylodiscitis with Paraspinal Abscess in an Immunocompetent Adult: First Case Report from Türkiye

İmmünkompetan Bir Erişkinde *Salmonella enteritidis* ile İlişkili Torasik Spondilodiskit ve Paraspinal Apse: Türkiye'den Bildirilen İlk Olgu

Rıdvan Dumlu*

İstanbul Medipol University Faculty of Medicine, Department of Infectious Diseases and Clinical Microbiology, İstanbul, Türkiye

Abstract

Salmonella enteritidis, a non-typhoidal Salmonella (NTS) serotype, rarely causes spondylodiscitis and typically affects immunocompromised individuals. We report the first documented case of *Salmonella enteritidis*-associated thoracic spondylodiscitis in an immunocompetent adult from Türkiye. A 32-year-old previously healthy male presented with chronic back pain. Magnetic resonance imaging revealed T7–T8 spondylodiscitis with a right-sided paraspinal abscess. Blood cultures were sterile; however, cultures from the drained abscess grew *Salmonella* spp., identified as *Salmonella enteritidis,* by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and serotyping. Histopathological examination showed abundant polymorphonuclear leukocytes. The patient received eight weeks of antibiotic therapy: two weeks of intravenous ceftriaxone followed by six weeks of oral ciprofloxacin, resulting in complete clinical and laboratory recovery. This case underscores the importance of considering NTS in the differential diagnosis of vertebral infections, particularly in regions endemic for tuberculosis and brucellosis.

Keywords: Immunocompetent, *Salmonella enteritidis,* spondylodiscitis, thoracic osteomyelitis, Türkiye

Öz

Tifo dışı Salmonella (NTS) serotiplerinden biri olan *Salmonella enteritidis*, nadiren spondilodiskite neden olmakta ve genellikle immünkompromize bireylerde görülmektedir. Bu raporda, Türkiye'den immünkompetan bir erişkinde *Salmonella enteritidis* ilişkili torasik spondilodiskit olgusu ilk kez sunulmaktadır. Otuz iki yaşındaki, önceden sağlıklı erkek hasta kronik sırt ağrısı ile başvurmuştur. Manyetik rezonans görüntülemede T7–T8 düzeyinde spondilodiskit ve sağ paraspinal alanda apse saptanmıştır. Kan kültürleri steril olmakla birlikte, drene edilen apse kültüründe *Salmonella* spp. üremiş, tür düzeyinde *Salmonella enteritidis* olarak tanımlanmıştır. Histopatolojik incelemede yoğun polimorfonükleer lökosit infiltrasyonu izlenmiştir. Hasta iki haftalık intravenöz seftriakson ve ardından altı haftalık oral siprofloksasin tedavisi olmak üzere toplam sekiz haftalık antibiyotik tedavisiyle tamamen klinik ve laboratuvar iyileşmesi göstermiştir. Bu olgu, özellikle tüberküloz ve brusellozun endemik olduğu bölgelerde, vertebral enfeksiyonların ayırıcı tanısında NTS etkenlerinin de göz önünde bulundurulması gerektiğini vurgulamaktadır.

Anahtar Kelimeler: İmmünkompetan, *Salmonella enteritidis*, spondilodiskit, torasik osteomyelit, Türkiye

Cite this article as: Dumlu R. *Salmonella enteritidis*-associated thoracic spondylodiscitis with paraspinal abscess in an immunocompetent adult: First case report from Türkiye. Mediterr J Infect Microb Antimicrob. 2026;15:1-3

Address for Correspondence/Yazışma Adresi: Rıdvan Dumlu MD, İstanbul Medipol University Faculty of Medicine, Department of Infectious Diseases and Clinical Microbiology, İstanbul, Türkiye
E-mail: r_dumlu@hotmail.com **ORCID ID:** orcid.org/0000-0001-8213-6064
Received/Geliş Tarihi: 21.07.2025 **Accepted/Kabul Tarihi:** 14.10.2025

Epub: 24.11.2025
Published: 06.01.2026

Introduction

Salmonella enteritidis, a non-typhoidal Salmonella (NTS) serotype, is primarily associated with foodborne gastroenteritis. Rarely, it can cause focal infections, including spondylodiscitis, via hematogenous dissemination. These infections typically occur in patients with predisposing conditions such as sickle cell disease, leukemia, immunosuppression, aortic aneurysms, or solid organ malignancies[1–4].

Several cases of vertebral osteomyelitis and epidural abscess due to *Salmonella enteritidis* have been reported, mainly from the United States, Europe, and Asia, and most involved immunocompromised hosts[1–9]. In Türkiye, only one case has been reported, involving a patient receiving immunosuppressive therapy for rheumatoid arthritis[10].

We present the first documented case of *Salmonella enteritidis*-associated thoracic spondylodiscitis with paraspinal abscess in an immunocompetent adult from Türkiye, contributing to the limited literature on this uncommon clinical presentation.

Case Report

A 32-year-old previously healthy male presented with a three-month history of progressive thoracic back pain. The pain was mechanical in character and was not associated with fever, weight loss, night sweats, or gastrointestinal symptoms. Physical examination revealed mild tenderness over the thoracic spine without deformity or scoliosis. Neurological examination was unremarkable, and other systemic findings were within normal limits.

Laboratory investigations showed a white blood cell count of 8,300/µL, erythrocyte sedimentation rate of 17 mm/h (reference < 20 mm/h), and C-reactive protein level of 8.7 mg/L (reference < 5 mg/L). Thoracic magnetic resonance imaging (MRI) demonstrated spondylodiscitis at the T7–T8 level with a right-sided paraspinal abscess. Imaging showed T1 hypointense and T2/short tau inversion recovery (STIR) hyperintense signal changes predominantly involving the T7 and T8 endplates, with increased T2/STIR signal intensity in the intervertebral disc space. Post-contrast images revealed marked enhancement of the endplates and disc. A right-sided paraspinal lesion measuring approximately 9 mm, consistent with an abscess extending anteriorly, was also observed. These findings were compatible with spondylodiscitis and paraspinal abscess.

Given the endemic nature of tuberculosis and brucellosis in Türkiye, these infections were initially considered; however, both Quantiferon-TB Gold and Brucella agglutination tests were negative. Blood cultures remained sterile.

The patient underwent surgical drainage, debridement, and posterior stabilization due to progressive symptoms and the presence of the paraspinal abscess. Cultures of intraoperatively obtained abscess material yielded bacterial growth. Colonies were identified as *Salmonella* spp. using matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS; Bruker Daltonics, USA), and serotyping according to the Kauffmann–White–Le Minor scheme (SSI Diagnostica, Denmark) confirmed the isolate as *Salmonella enteritidis*. Histopathological examination of the abscess wall revealed acute inflammatory infiltrates with abundant polymorphonuclear leukocytes. Antimicrobial susceptibility testing showed the isolate was susceptible to ampicillin, trimethoprim-sulfamethoxazole, chloramphenicol, ceftriaxone, and ciprofloxacin.

Further evaluation excluded hematologic disorders, immunosuppressive conditions, and vascular pathologies, including infected aortic aneurysm. Transthoracic echocardiography showed no evidence of endocarditis or cardiac involvement.

The patient received intravenous ceftriaxone for two weeks, followed by oral ciprofloxacin for six weeks, totaling eight weeks of targeted antimicrobial therapy according to the 2015 Infectious Diseases Society of America (IDSA) guidelines. This treatment resulted in full clinical, laboratory, and radiological resolution. Follow-up MRI demonstrated complete resolution, and no recurrence was observed during one year of follow-up.

Discussion

Spondylodiscitis is a severe spinal infection associated with significant morbidity, particularly when the thoracic spine is involved. Globally, the most common causative agents are *Staphylococcus aureus*, *Mycobacterium tuberculosis*, and *Brucella* species, with geographical and epidemiological factors influencing this distribution[1,11]. In Türkiye, both tuberculosis and brucellosis remain endemic, with Brucella seroprevalence estimated at approximately 4.5%[12]. Accordingly, these pathogens are typically prioritized in the initial differential diagnosis of spinal infections.

The 2015 IDSA Clinical Practice Guidelines for Native Vertebral Osteomyelitis recommend serologic testing for Brucella and interferon-gamma release assays for *Mycobacterium tuberculosis* in patients presenting with vertebral infections, particularly in endemic regions[1]. However, when these tests are negative or the clinical presentation is atypical, less common pathogens, including NTS, should be considered.

In our case, both blood cultures and standard serological tests for Brucella and tuberculosis were negative. The pathogen was identified only after surgical drainage of a paraspinal abscess, highlighting the limitations of relying solely on blood cultures and endemic pathogens for diagnosis. This emphasizes the importance of direct sampling from the infected site for microbiological confirmation.

Vertebral involvement due to *Salmonella enteritidis* is extremely rare, particularly in immunocompetent individuals. Most reported cases involve immunosuppressed patients or those with significant comorbidities such as hematologic malignancies, sickle cell disease, or vascular infections[2–9]. For example, Ikejiri et al.[8] described a case of *Salmonella enteritidis* vertebral osteomyelitis complicated by meningitis following influenza A infection, while Tomek et al.[5] reported a case associated with a mycotic aortic aneurysm in a patient with chronic lymphocytic leukemia. In contrast, our patient had no underlying disease or predisposing condition, making this presentation particularly unusual.

Diagnostic confirmation was achieved by MALDI-TOF MS and serotyping after specimen collection from the infection site. This approach aligns with the 2015 IDSA guidelines and the comprehensive review by Lew and Waldvogel, both emphasizing early MRI-based imaging and microbiological confirmation for pathogen-specific diagnosis[1,11].

The patient received two weeks of intravenous ceftriaxone followed by six weeks of oral ciprofloxacin, totaling eight weeks of targeted antimicrobial therapy in accordance with IDSA recommendations. This regimen led to full clinical, laboratory, and radiological recovery and is consistent with IDSA guidance suggesting 6–12 weeks of pathogen-specific therapy for vertebral osteomyelitis, depending on clinical response[1].

To our knowledge, this is the first reported case of *Salmonella enteritidis* spondylodiscitis in an immunocompetent adult from Türkiye, specifically involving the thoracic spine with a paraspinal abscess.

Clinical implications of this case include:

• In endemic regions, Brucella and *Mycobacterium tuberculosis* should be ruled out first; however, rare pathogens such as *Salmonella enteritidis* should be considered in atypical or culture-negative cases.

• Blood cultures may not always detect the causative pathogen; therefore, image-guided biopsy or surgical sampling is essential.

• Treatment should follow pathogen-specific guidelines with an adequate duration of antibiotics.

• Clinicians must recognize that invasive Salmonella infections can occur even in immunocompetent individuals.

This case broadens the clinical spectrum of *Salmonella enteritidis* infections and highlights the importance of guideline-based diagnostic and therapeutic strategies in spinal infections.

Ethics

Informed Consent: Written informed consent was obtained from the patient for publication of this case and accompanying clinical data.

Footnotes

Conflict of Interest: The author declare no conflict of interest.

Financial Disclosure: The author declared that this study received no financial support.

References

1. Berbari EF, Kanj SS, Kowalski TJ, Darouiche RO, Widmer AF, Schmitt SK, Hendershot EF, Holtom PD, Huddleston PM 3rd, Petermann GW, Osmon DR, Infectious Diseases Society of America. 2015 Infectious Diseases Society of America (IDSA) clinical practice guidelines for the diagnosis and treatment of native vertebral osteomyelitis in adults. Clin Infect Dis. 2015;61(6):e26–e46.
2. Pierrotti LC, Santos SdeS, Barone AA. Osteomyelitis by *Salmonella enteritidis* and sickle cell hemoglobinopathy. Rev Hosp Clin. 1996;51(3):96–8.
3. Schüler A, Schaumann D, Manns MP, Koch KM. Lumbale spondylodiszitis durch *Salmonella enteritidis*. Dtsch Med Wochenschr. 1994;119(41):1383–7.
4. Gupta SK, Pandit A, White DG, Evans PD. *Salmonella* osteomyelitis of the thoracic spine: an unusual presentation. Postgrad Med J. 2004;80(940):110–1.
5. Tomek M, Cheshire NJ, Rudarakanchana N, Samarasinghe D, Bicknell CD. *Salmonella* mycotic thoracoabdominal aortic aneurysm associated with chronic lymphocytic leukemia. Ann Vasc Surg. 2013;27(8):1186.e17–21.
6. Zaki M, Ariffin MH. Single-stage debridement and spinal instrumentation for *Salmonella* spondylodiscitis. *Cureus.* 2021;13(9):e18306.
7. Oki M, Ueda A, Tsuda A, Yanagi H, Ozawa H, Takagi A. *Salmonella enterica* vertebral osteomyelitis with epidural abscess and meningitis. Tokai J Exp Clin Med. 2016;41(3):169–71.
8. Ikejiri K, Suzuki K, Ito A, Yasuda K, Shindo A, Ishikura K, Imai H. Invasive *Salmonella enteritidis* infection complicated by meningitis and vertebral osteomyelitis shortly after influenza A infection in an immunocompetent young adult. J Infect Chemother. 2020;26(2):269–73.
9. Toyoshima H, Masuda N, Ishiguro C, Tanigawa M, Tanaka H, Nakanishi Y, Sakabe S. *Salmonella enterica* osteomyelitis with pulmonary involvement in an immunocompetent woman. IDCases. 2021;24:e01127.
10. Durmaz S, Doğan SA, Kandemir İ, Menkü A, Aygen B, Perçin D. A rare agent of spondylodiscitis in adult patient: *Salmonella enteritidis*. Dicle Medical Journal. 2012;39(1):139-41.
11. Lacasse M, Derolez S, Bonnet E, Amelot A, Bouyer B, Carlier R, Coiffier G, Cottier JP, Dinh A, Maldonado I, Paycha F, Ziza JM, Bemer P, Bernard L; Review group. 2022 SPILF—Clinical practice guidelines for the diagnosis and treatment of disco-vertebral infection in adults. Infect Dis Now. 2023;53(3):104647.
12. Aydın M, Aydın NN. Brusellozda osteoartiküler tutulumu olan hastaların değerlendirilmesi. ANKEM Derg. 2024;38(3):104-11.

RESEARCH ARTICLE / ARAŞTIRMA

DOI: 10.4274/mjima.galenos.2025.25536.2
Mediterr J Infect Microb Antimicrob 2026;15:25536.2
Erişim: http://dx.doi.org/10.4274/mjima.galenos.2025.25536.2

Pyogenic Spondylodiscitis: A 9-year Analysis

Piyojenik Spondilodiskit: 9 Yıllık Analiz

Merve Kaplan[1]*, Serpil Erol[2], Nurgül Ceran[2], Asuman İnan[2], Seniha Şenbayrak[2], Cansu Aşık[2]

[1]*Kırklareli Training and Research Hospital, Department of Infectious Diseases and Clinical Microbiology, Kırklareli, Türkiye*
[2]*University of Health Sciences, Haydarpaşa Numune Training and Research Hospital, Department of Infectious Diseases and Clinical Microbiology, İstanbul, Türkiye*

Abstract

Introduction: The management of pyogenic spondylodiscitis (PS) remains challenging due to the absence of clear, evidence-based guidelines. This study aimed to assess the clinical characteristics, diagnostic and follow-up challenges, and treatment outcomes of patients with PS.

Materials and Methods: The clinical, laboratory, and radiological data of all patients aged ≥18 years who were hospitalized with PS between January 2015 and June 2024 were retrospectively analyzed.

Results: Among 25 patients diagnosed with PS, 60% were male, with a mean age of 61 ± 10.6 years (range: 41–78). The most common symptoms were back or neck pain (88%), difficulty walking (24%), and fever (20%). The mean symptom duration was 2 months. The lumbosacral (60%), thoracic (44%), and cervical (12%) regions were the most frequently affected. A total of 84% of patients had at least one comorbidity, and 80% had a predisposing risk factor. Blood cultures were positive in 60% of patients. Among the 23 patients who underwent tissue and/or abscess culture, 43.7% and 30% yielded positive results, respectively. The most frequently isolated pathogen was methicillin-susceptible *Staphylococcus aureus* (MSSA) (48%). Among patients followed with contrast-enhanced magnetic resonance imaging (MRI), 41.2% demonstrated persistent contrast enhancement without significant change. The total treatment duration was 12 ± 4.1 weeks (range: 7–24). Treatment success was achieved in 86.3% of cases, while 3 (13.6%) patients experienced recurrence. In all recurrent cases, *Staphylococcus aureus* was the causative agent, and paraspinal abscess and bacteremia were present concomitantly. All recurrent cases had received at least 12 weeks of pathogen-targeted therapy.

Conclusion: Hospitalization and invasive procedures appear to be significant risk factors for PS. Obtaining blood and tissue/abscess cultures before initiating antimicrobial therapy enhances the likelihood of pathogen identification. Despite adequate treatment, MRI findings may persist without complete radiological resolution. Close monitoring is warranted for potential recurrence when *Staphylococcus aureus* is the causative pathogen, especially in the presence of abscess or bacteremia.

Keywords: Pyogenic spondylodiscitis, vertebrae, *Staphylococcus aureus*, recurrence

Öz

Giriş: Piyojenik spondilodiskitin (PS) tanısı, takibi ve tedavisi ile ilgili önerilerin sınırlı olması yönetimini güçleştirmektedir. Bu çalışmanın amacı PS olgularının klinik özelliklerinin, tanı ve takip sürecinde karşılaşılan zorlukların ve sonuçlarının değerlendirilmesidir.

Gereç ve Yöntem: Ocak 2015 ile Haziran 2024 tarihleri arasında PS tanısıyla hastaneye yatırılan ≥18 yaşındaki tüm hastaların klinik, laboratuvar ve radyolojik özellikleri retrospektif olarak analiz edildi.

Bulgular: PS tanısı alan 25 hastanın %60'ı erkekti ve yaş ortalaması 61±10,6 (41–78) idi. En sık görülen semptomlar sırt/boyun ağrısı (%88), yürüme güçlüğü (%24) ve ateşti (%20). Semptom süresi ortalama iki aydı. En sık tutulan bölgeler lumbosakral (%60), torakal (%44) ve servikaldi (%12). Hastaların %84'ünün en az bir ek hastalığı, %80'inin ise risk faktörü vardı. Kan kültürleri hastaların %60'ında pozitifti. Doku ve/veya apse kültürü alınan 23 hastanın sırasıyla %43,7'sinde ve %30'unda pozitif sonuç elde edildi. Tüm kültürlerde en sık izole edilen etken metisiline duyarlı *Staphylococcus aureus*'tu (%48). Kontrastlı manyetik rezonans görünteleme (MRI) ile izlenen hastaların %41,2'sinde belirgin iyileşme olmaksızın kontrast tutulumu devam etmekteydi. Toplam tedavi süresi ortalama 12±4,1 (7–24) haftaydı. Vakaların %86,3'ünde tedavi başarılıydı. Üç hastada (%13.6) nüks gelişti.

Cite this article as: Kaplan M, Erol S, Ceran N, İnan A, Şenbayrak S, Aşık C. Pyogenic spondylodiscitis: A 9-year analysis. Mediterr J Infect Microb Antimicrob. 2026;15:4–10

Address for Correspondence/Yazışma Adresi: Merve Kaplan MD, Kırklareli Training and Research Hospital, Department of Infectious Diseases and Clinical Microbiology, Kırklareli, Türkiye
E-mail: merveaykan1@gmail.com **ORCID ID:** orcid.org/0000-0001-5063-0922
Received/Geliş Tarihi: 16.07.2025 **Accepted/Kabul Tarihi:** 27.10.2025

Epub: 24.11.2025
Published: 13.01.2026

Öz

Nüks vakalarının tamamında etken *Staphylococcus aureus*'tu ve eşlik eden paraspinal apse ile bakteremi mevcuttu. Nüks vakalarının tamamı, etkene yönelik en az 12 haftalık tedavi almıştı.

Sonuç: Hastanede yatış ve invaziv girişimler, PS için önemli risk faktörleridir. Tedavi öncesi kan ve doku/apse kültürü alınması, etkenin belirlenme olasılığını artırır. Uygun tedaviye rağmen MRI bulguları tamamen düzelmeyebilir. *Staphylococcus aureus* kaynaklı, özellikle apse veya bakteremi ile seyreden vakalarda, nüks riski açısından yakın takip gereklidir.

Anahtar Kelimeler: Pyojenik spondilodiskit, vertebra, *Staphylococcus aureus*, nüks

Introduction

Spondylodiscitis is an infection of the spinal column caused by various pathogens, potentially involving the vertebral body, intervertebral discs, spinal canal, and paravertebral structures[1]. The etiology may be pyogenic (bacterial), granulomatous (tuberculous, brucellar, or fungal) or rarely parasitic; however, the majority of cases are bacterial in origin[1–3]. Pathogens may reach the vertebrae through hematogenous dissemination, direct inoculation (most commonly during spinal surgery), or contiguous spread from adjacent structures. A distant focus of infection is identified in about half of the cases, and infective endocarditis accompanies approximately 12% of them[3].

Spondylodiscitis accounts for 0.15%–7% of all osteomyelitis cases[1,2,4]. In Western countries, its annual incidence ranges from 0.4 to 2.4 per 100,000 population[5]. It is 1.5–3 times more common in males and occurs across all age groups, with a higher prevalence among individuals aged 50–70 years[1,2,4,6].

The incidence of spondylodiscitis has shown a concerning rise over the past two decades. This trend is attributed to increased life expectancy, the presence of comorbidities, greater use of invasive procedures and immunosuppressive therapies, expanding indications for spinal surgery, a growing vulnerable population, and advancements in diagnostic methods[2,6–8].

The clinical spectrum of the disease ranges from mild to rapidly progressive, potentially leading to severe morbidity and mortality, including vertebral collapse, permanent neurological deficits, and even death. Its nonspecific early symptoms often result in diagnostic delays of 30 to 90 days[6]. Early identification of the causative pathogen and prompt initiation of targeted antimicrobial therapy are therefore crucial[9]. However, diagnosis, follow-up, and treatment remain challenging due to the limited availability of evidence-based management guidelines. This study aimed to evaluate the clinical characteristics and challenges associated with microbiological and radiological assessments during diagnosis and follow-up, as well as the treatment outcomes of pyogenic spondylodiscitis (PS).

Materials and Methods

This study was designed as a single-center retrospective cohort study conducted at a tertiary care facility in Türkiye. Data were retrospectively collected from the hospital information system and patient records for all patients aged ≥ 18 years who were hospitalized with a diagnosis of PS between January 2015 and June 2024. Patients with tuberculous or brucellar spondylodiscitis were excluded.

The diagnosis of spondylodiscitis was based on typical clinical findings, characteristic changes on magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography–CT (PET–CT), and microbiological evidence of infection.

Demographic, clinical, microbiological, and radiological data of the patients were retrospectively reviewed. Microorganism identification and antimicrobial susceptibility testing (AST) were performed using fully automated systems (VITEK2 Compact, bioMérieux, France, and VITEK MS, bioMérieux, France). The AST results were interpreted according to the recommendations of the European Committee on AST.

The durations of intravenous and oral treatments were recorded. All patients received a total treatment duration of at least 6 weeks. Therapy was extended in patients with inadequate clinical or laboratory response or those with insufficient source control. Owing to the retrospective design of the study, the frequency of follow-up imaging was not standardized. Therefore, radiology reports obtained at the initiation and completion of therapy were compared for all patients.

Treatment success was defined as the absence of clinical or laboratory deterioration during a one-year follow-up. Patients who demonstrated clinical, laboratory, or radiological worsening after completion of therapy were classified as recurrent cases.

Statistical Analysis

Data were analyzed using IBM SPSS Statistics 26 (IBM SPSS, USA). The normality of parameter distribution was assessed using the Shapiro–Wilk test. Descriptive statistics were calculated for numerical data, including median, mean, and standard deviation values.

Data collection was conducted with the approval of the University of Health Sciences Türkiye, Haydarpaşa Numune Training and Research Hospital (approval number: 711/06/2024, dated: 05.06.2024).

Results

Between January 2015 and June 2024, 25 patients were diagnosed with PS. Of these, 15 (60%) were male and 10 (40%) were female, with a mean age of 61 ± 10.6 years (range: 41–78). At least one chronic disease was present in 21 (84%) patients. The most common comorbidities were diabetes mellitus in 14 (56%), hypertension in 12 (48%), hyperlipidemia in eight (32%), and chronic kidney failure in six (24%) patients (Table 1).

The most frequent presenting symptoms were back or neck pain in 22 (88%) patients, difficulty walking in six (24%), fever in five (20%), leg weakness in two (8%), incontinence in two (8%), weight loss in two (8%), night sweats in two (8%), and numbness in one (4%) at admission. The mean duration of symptoms before diagnosis was 2 months (range: 1 week–6 months). Motor deficit was detected in 11 (44%) patients, and a cardiac murmur in one (4%). No abnormal physical findings were observed in 13 (52%) patients. Pretreatment laboratory findings showed a mean white blood cell (WBC) count of 9,132 ± 3,986 cells/mm^3 (range: 1,780–19,690), C-reactive protein (CRP) level of 116 ± 92 mg/L (range: 3–317), and erythrocyte sedimentation rate (ESR) of 72 ± 22 mm/h (range: 26–104). By the end of treatment, the mean ESR had decreased to 29 ± 18.7 mm/h (range: 2–83) (Table 2). In recurrent cases, initial WBC counts were 7,700–15,300 and 19,690 cells/mm^3, the CRP levels were 10–190 and 29 mg/L, and the ESR values were 76–63 and 95 mm/h, respectively. At the end of treatment, WBC counts decreased to 4,890–8,170 and 8,120 cells/mm^3, the CRP levels to 0.4–4 mg/L, and the ESR values to 23–36 and 29 mm/h, respectively.

Discitis was detected in 20 (80%) patients, paravertebral soft tissue involvement in 21 (84%), and abscess formation in 12 (48%). The most frequently affected regions were the lumbosacral (60%), thoracic (44%), and cervical (12%) spine.

Table 1. Baseline characteristics of the study patients.

	n (%)		n (%)
Male/female	15 (60)/10 (40)	**Known risk factors**	20 (80)
Age (mean ± SD[§] (min.–max.)	61 ± 10.6 (41–78)	Hospitalization in the last 3 months	13 (52)
Chronic disease	21 (84)	ICU[†] admission in the last 3 months	4 (16)
Diabetes mellitus	14 (56)	Catheter infection	1 (4)
Hyperlipidemia	10 (40)	ERCP[‡]	2 (8)
Hypertension	8 (32)	Spinal procedure	3 (12)
Chronic renal failure	6 (24)	Other invasive procedures and operations	2 (8)
Chronic heart diseases	5 (20)	Hemodialysis	3 (12)
Pulmonary diseases	2 (8)		
Malignancy	1 (4)		

[†]ICU; [‡]ERCP; [§]SD; min.–max., minimum–maximum.

Table 2. Clinical characteristics and laboratory results.

	n (%)		Mean ± SD[||] (min.-max.)
Localization		**Laboratory results**	
Cervical	3 (12)	Pretreatment WBC[§] (cell/mm^3)	9,132 ± 3,986 (1,780–19,690)
Thoracal	11 (44)	Pretreatment CRP[‡] (mg/L)	116 ± 92 (3–317)
Lumbosacral	15 (60)	Pretreatment ESR[§] (mm/h)	72 ± 22 (26–104)
Paraspinal/psoas abscess	12 (40)	End treatment ESR (mm/h)	29 ± 18.7 (2–83)
Complications			
Spinal cord pressure	11 (44)		
Neurodeficit	2 (8)		
Height loss	10 (40)		
Endocarditis	2 (8)		

[†]WBC; [‡]CRP; [§]ESR; [||]SD.

As potential risk factors, 13 (52%) patients had a hospital admission within the past 3 months, and four (16%) had a history of intensive care unit stay. Three patients developed infection following vertebral surgery; two had undergone endoscopic retrograde cholangiopancreatography (ERCP), one had a catheter-related infection, and three were on hemodialysis. No risk factors were identified in 5 (20%) patients (Table 1).

Blood cultures were obtained before the initiation of antibiotic therapy in all patients, and 23 (92%) underwent tissue or abscess culture. Blood cultures were positive in 15 (60%) patients. Tissue cultures were obtained from 16 patients, with positive results in seven (43.7%). Abscess cultures showed bacterial growth in three of 10 (30%) patients. In seven (77.7%) of the tissue culture–negative cases and six (85.7%) of the abscess culture–negative cases, samples were collected after the initiation of empirical antibiotic therapy. No organism was isolated from any culture in 5 (20%) patients, who were treated empirically (Table 3).

The most frequently isolated pathogen was *Staphylococcus aureus*, identified in 12 (48%) cases with positive blood or tissue/abscess cultures. All isolates were methicillin-susceptible (MSSA); no methicillin-resistant *Staphylococcus aureus* (MRSA) was detected. Other identified pathogens included *Pseudomonas aeruginosa, Escherichia coli, Enterococcus faecium, Streptococcus pneumoniae, Streptococcus parasanguinis, Staphylococcus epidermidis, Bacteroides fragilis*, and *Enterobacter cloacae* (each in one case). Blood cultures were positive in 10 (83.3%) of the cases, where *Staphylococcus aureus* was the causative agent (Table 3).

Complications included spinal cord compression in 11 (44%) patients, vertebral height loss in 10 (40%), and endocarditis in two (8%).

The mean duration of intravenous antibiotic therapy was eight weeks (range: 4–24, standard deviation [SD]: 4.2). Oral sequential therapy lasted a mean of five weeks (range: 1–12, SD: 2.4), with a total treatment duration averaging 12 weeks (range: 7–24, SD: 4.1). The recurrent cases received total antibiotic therapy for 13, 12, and 24 weeks, respectively, with intravenous administration for the first 6, 8, and 24 weeks.

A total of 17 (68%) patients were followed with contrast-enhanced MRI, one (4%) with noncontrast MRI, two (8%) with CT, and one (4%) with PET–CT. Follow-up imaging data were unavailable for 4 (16%) patients. At the end of treatment, complete resolution of contrast enhancement was observed in 3 (17.6%) patients, partial regression in seven (41.2%), and no significant change in seven (41.2%). Among the three recurrent cases, regression on MRI was observed in two, while no change was detected in one at the end of treatment.

Of the 22 patients who completed treatment, three (13.6%) developed recurrence, and two (9%) required surgical intervention. Treatment was successful in 19 (86.3%) patients. Treatment outcomes could not be evaluated in 3 patients due to unavailable follow-up data. All recurrent cases were caused by *Staphylococcus aureus* and were accompanied by bacteremia and paraspinal abscess.

Discussion

The rising incidence of PS over the past 2 decades, along with the complexity of its management and the lack of standardized treatment guidelines, prompted this study to evaluate the clinical and etiological characteristics, diagnostic process, and prognosis of patients treated in our center.

A review of previous studies reported a mean patient age of approximately 60 years and a predominance of male (56%) patients with spondylodiscitis[5]. Our findings are consistent with these reports, with a mean age of 61 years and 60% of patients being male.

Table 3. Microbiological tests and results.

	n (%)		n (%)
Blood cultures obtained	25 (100)	**Organisms isolated from all cultures**	
Positive results	15 (60)	MSSA	12 (48)
Abscess and/or tissue culture obtained	23 (92)	*Staphylococcus epidermidis*	1 (4)
Abscess cultures obtained	10 (40)	*Enterococcus faecium*	1 (4)
Positive results	3 (30)	*Streptococcus pneumoniae*	1 (4)
Negative results	7 (70)	*Streptococcus parasanguinis*	1 (4)
Tissue cultures obtained	16 (64)	*Pseudomonas aeruginosa*	1 (4)
Positive results	7 (43.7)	*Escherichia coli*	1 (4)
Negative results	9 (57.3)	*Enterobacter cloacae*	1 (4)
		Bacteroides fragilis	1 (4)

The increasing prevalence of spondylodiscitis has been associated with longer life expectancy and the rise in comorbid conditions[3,10,11]. In support of this, 84% of our patients had at least one chronic disease. Furthermore, 52% of the cohort had been hospitalized in the preceding 3 months, indicating that hospitalization is a major risk factor. Catheter-related infections, hemodialysis, and prior spinal surgeries were common underlying causes, as expected. Notably, two cases of PS caused by gram-negative pathogens following ERCP highlight ERCP as a potential risk factor for gram-negative spondylodiscitis, a relationship described in only a few previous reports[12–14]. In 20% of our cases, no identifiable cause was detected, emphasizing that PS can occur even in patients without apparent risk factors.

Consistent with the literature, the lumbosacral region was the most frequently affected site (60%), followed by the thoracic (44%) and cervical (12%) regions[5]. Multilevel involvement was observed in 16% of all cases. Nearly half of the patients presented with concurrent abscess formation, consistent with the 19%–68% range reported in the literature[5,15,16]. Since abscesses can significantly influence both treatment approach and prognosis, screening for additional foci and the early detection of abscess formation are essential steps in clinical evaluation.

Identifying the causative pathogen is critical for ensuring targeted and effective treatment, especially for a disease such as spondylodiscitis, which requires prolonged therapy. Blood cultures yielded a high positivity rate of 60%, underscoring the value of this simple diagnostic tool and reaffirming the importance of obtaining blood cultures before initiating therapy[3]. Positive blood culture rates can reach up to 70% in patients without prior antibiotic exposure[1].

Tissue and abscess cultures are also valuable diagnostic tools. In our study, 44% of tissue cultures and 30% of abscess cultures yielded growth, which aligns with previously reported rates of 43%–78%[3]. The lower positivity rates observed in our study may be attributed to the exclusive use of percutaneous biopsies and the fact that most samples were obtained after empirical antibiotic therapy had begun. Tissue cultures obtained after antimicrobial initiation were negative in 77% of cases, while abscess cultures were negative in 85%. Antibiotic exposure significantly reduces culture positivity rates, and even without prior antibiotic use, biopsy cultures can remain negative in up to 39% of cases[17]. Moreover, CT-guided percutaneous biopsies may yield limited tissue, identifying the pathogen in only about half of cases[1]. Because the procedure is invasive, delays in sampling can also pose challenges in clinical practice. When the causative organism remains unidentified, determining optimal therapy becomes difficult, often necessitating broad-spectrum parenteral antibiotics and prolonged hospital stays. Therefore, obtaining appropriate cultures before initiating therapy is crucial. In nonemergent cases without neurological deficits or sepsis, treatment should be deferred until adequate samples are obtained[9].

The prevalence of *Staphylococcus aureus* in PS ranges from 20% to 84%, accounting for approximately half of all cases[18]. Consistent with the literature, *Staphylococcus aureus* was the most common causative pathogen in our study (48%). Methicillin susceptibility among *Staphylococcus aureus* isolates varies geographically and depends on patient risk factors and disease etiology[6]. Although most community-acquired strains are MSSA, the rising rate of MRSA raises concerns regarding its inclusion in empirical therapy[18]. A review of 14 studies in 2009 reported an MRSA prevalence of 2.6%[19], while more recent studies, particularly from high-income countries, have reported rates as high as 25%–30%[20–22]. In studies conducted in Türkiye, the prevalence ranges from 3.7% to 5.5%[4,23,24], and one study reported a rate of 12.5% in nosocomial cases[25]. Notably, all *Staphylococcus aureus* strains in our study were MSSA, despite a significant proportion of patients having prior hospitalizations or invasive procedures. This finding suggests that MRSA is not yet a major concern in PS cases in our region and that routine MRSA coverage in empirical therapy may not be necessary for most patients.

Spinal cord compression and vertebral height loss were the most common complications, while 8% of cases had concomitant endocarditis. This finding aligns with prior reports describing endocarditis in up to 12% of PS cases, highlighting the importance of screening for endocarditis in these patients[19].

MRI remains the gold standard for diagnosing spondylodiscitis, although its role in treatment follow-up remains debated[26]. In our study, 68% of patients underwent contrast-enhanced MRI. At the end of treatment, only 17% of patients demonstrated complete resolution of contrast enhancement, while 41% showed no significant change. As shown in previous studies, imaging abnormalities may persist despite clinical and laboratory improvement[26,27].

Treatment was successful in 86.3% of cases; however, 13% of patients experienced recurrence. All recurrent cases were caused by *Staphylococcus aureus* infections with concurrent paraspinal abscess and bacteremia. *Staphylococcus aureus* has been identified as an independent risk factor for treatment failure[6,15]. Bacteremia and paraspinal abscess have also been identified in some studies as predictors of recurrence and treatment failure[7,20,21]. One study reported that patients with bacteremia required longer treatment durations (> 8 weeks) [28], while another showed a significant decrease in recurrence among patients treated for more than eight weeks[22]. In our study, recurrent cases received at least 12 weeks of treatment, including a minimum of six weeks of intravenous therapy

targeting the causative agent. Despite the prolonged duration, recurrence still occurred. Furthermore, in these patients, WBC, CRP, and ESR values normalized by the end of treatment, and two patients demonstrated regression on contrast-enhanced MRI. However, neither laboratory nor imaging findings were reliable predictors of relapse. Therefore, close monitoring of patients with *Staphylococcus aureus* infections, particularly those with bacteremia and paraspinal abscess formation, is essential to minimize recurrence.

Study Limitations

This study has several limitations that should be considered when interpreting the findings. As a retrospective, single-center study with a relatively small sample size, the generalizability of the results is limited, and the statistical power may be insufficient to detect less common associations or outcomes. The analysis relied on hospital records, which may have lacked uniformity or completeness, particularly in documenting clinical symptoms and follow-up details. Moreover, follow-up imaging was not standardized in terms of timing or modality (MRI, CT, or PET–CT), introducing potential variability in the evaluation of radiological treatment response. Owing to the retrospective nature of the study, treatment decisions—including antibiotic duration and follow-up strategies—were not standardized and could have influenced treatment outcomes and recurrence rates.

Conclusion

In conclusion, recent hospitalization and invasive procedures remain significant risk factors for the development of PS. Obtaining blood, tissue, or abscess cultures before initiating antibiotic therapy substantially increases the likelihood of identifying the causative pathogen; at the very least, blood cultures should be obtained as a simple yet valuable diagnostic step before treatment. Despite appropriate therapy, MRI abnormalities may persist and should not be used in isolation to determine treatment success. *Staphylococcus aureus* continues to be the most frequently isolated pathogen. While methicillin resistance did not emerge as a concern in this cohort, all recurrent cases involved *Staphylococcus aureus* infections accompanied by abscess formation and bacteremia. Therefore, patients with *Staphylococcus aureus* spondylodiscitis, particularly those with concurrent bacteremia or paraspinal abscess, should be closely monitored for potential recurrence even after completion of therapy.

Ethics

Ethics Committee Approval: Data collection was conducted with the approval of the University of Health Sciences Türkiye, Haydarpaşa Numune Training and Research Hospital (approval number: 711/06/2024, dated: 05.06.2024).

Informed Consent: Not required.

Footnotes

Authorship Contributions

Surgical and Medical Practices: M.K., S.E., N.C., A.İ., S.Ş., C.A., Concept: M.K., S.E., N.C., A.İ., S.Ş., C.A., Design: M.K., S.E., Data Collection or Processing: M.K., C.A., Analysis or Interpretation: M.K., S.E., Literature Search: M.K., S.E., Writing: M.K., S.E., N.C., A.İ., S.Ş., C.A.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

The study was presented previously in 10th BUHASDER congress as an oral presentation.

References

1. Sobottke R, Seifert H, Fätkenheuer G, Schmidt M, Gossmann A, Eysel P. Current diagnosis and treatment of spondylodiscitis. Dtsch Arztebl Int. 2008;105(10):181–7.
2. Mavrogenis AF, Megaloikonomos PD, Igoumenou VG, Panagopoulos GN, Giannitsioti E, Papadopoulos A, Papagelopoulos PJ. Spondylodiscitis revisited. EFORT Open Rev. 2017;2(11):447–61.
3. Gouliouris T, Aliyu SH, Brown NM. Spondylodiscitis: update on diagnosis and management. J Antimicrob Chemother. 2010 Nov;65 Suppl 3:iii11–24.
4. Mete B, Kurt C, Yilmaz MH, Ertan G, Ozaras R, Mert A, Tabak F, Ozturk R. Vertebral osteomyelitis: eight years' experience of 100 cases. Rheumatol Int. 2012;32(11):3591–7.
5. Gentile L, Benazzo F, De Rosa F, Boriani S, Dallagiacoma G, Franceschetti G, Gaeta M, Cuzzocrea F. A systematic review: characteristics, complications and treatment of spondylodiscitis. Eur Rev Med Pharmacol Sci. 2019;23(2 Suppl):117–28.
6. Marchionni E, Marconi L, Ruinato D, Zamparini E, Gasbarrini A, Viale P. Spondylodiscitis: is really all well defined? Eur Rev Med Pharmacol Sci. 2019;23(2 Suppl):201–9.
7. Herren C, Jung N, Pishnamaz M, Breuninger M, Siewe J, Sobottke R. Spondylodiscitis: Diagnosis and Treatment Options. Dtsch Arztebl Int. 2017;114(51–52):875–82.
8. Turunc T, Demiroglu YZ, Uncu H, Colakoglu S, Arslan H. A comparative analysis of tuberculous, brucellar and pyogenic spontaneous spondylodiscitis patients. J Infect. 2007;55(2):158–63.
9. Berbari EF, Kanj SS, Kowalski TJ, Darouiche RO, Widmer AF, Schmitt SK, Hendershot EF, Holtom PD, Huddleston PM 3rd, Petermann GW, Osmon DR, Infectious Diseases Society of America. 2015 Infectious Diseases Society of America (IDSA) clinical practice guidelines for the diagnosis and treatment of native vertebral osteomyelitis in adults. Clin Infect Dis. 2015 Sep 15;61(6):e26–46.
10. Grammatico L, Baron S, Rusch E, Lepage B, Surer N, Desenclos JC, Besnier JM. Epidemiology of vertebral osteomyelitis (VO) in France: analysis of hospital-discharge data 2002–2003. Epidemiol Infect. 2008;136(5):653–60.
11. Conan Y, Laurent E, Belin Y, Lacasse M, Amelot A, Mulleman D, Rosset P, Bernard L, Grammatico-Guillon L. Large increase of vertebral osteomyelitis in France: a 2010–2019 cross-sectional study. Epidemiol Infect. 2021;149:e227.

12. Köklü S, Koçak E, Akbal E. Air cholangiography for severe hilar obstruction at ERCP. Gastrointest Endosc. 2011;73(6):1326.

13. Mastoraki A, Mastoraki S, Papanikolaou IS, Tsikala-Vafea M, Tsigou V, Lazaris A, Arkadopoulos N. Spondylodiscitis associated with major abdominal surgical intervention: Challenging diagnostic and therapeutic modalities. Indian J Surg Oncol. 2017;8(3):274–8.

14. Katsinelos P, Fasoulas K, Terzoudis S, Chatzimavroudis G, Zavos C, Kountouras J. Spondylodiscitis complicating cholangitis caused by stent occlusion. Gastrointest Endosc. 2011;73(6):1326–8.

15. Bernard L, Dinh A, Ghout I, Simo D, Zeller V, Issartel B, Le Moing V, Belmatoug N, Lesprit P, Bru JP, Therby A, Bouhour D, Dénes E, Debard A, Chirouze C, Fèvre K, Dupon M, Aegerter P, Mulleman D. Antibiotic treatment for 6 weeks versus 12 weeks in patients with pyogenic vertebral osteomyelitis: an open-label, non-inferiority, randomised, controlled trial. Lancet. 2015;385(9971):875–82.

16. Aagaard T, Roed C, Dragsted C, Skinhøj P. Microbiological and therapeutic challenges in infectious spondylodiscitis: a cohort study of 100 cases, 2006–2011. Scand J Infect Dis. 2013;45(6):417–24.

17. Michel SC, Pfirrmann CW, Boos N, Hodler J. CT-guided core biopsy of subchondral bone and intervertebral space in suspected spondylodiskitis. AJR Am J Roentgenol. 2006;186(4):977–80.

18. Fantoni M, Trecarichi EM, Rossi B, Mazzotta V, Di Giacomo G, Nasto LA, Di Meco E, Pola E. Epidemiological and clinical features of pyogenic spondylodiscitis. Eur Rev Med Pharmacol Sci. 2012;16 Suppl 2:2–7.

19. Mylona E, Samarkos M, Kakalou E, Fanourgiakis P, Skoutelis A. Pyogenic vertebral osteomyelitis: a systematic review of clinical characteristics. Semin Arthritis Rheum. 2009;39(1):10–7.

20. Patel AR, Alton TB, Bransford RJ, Lee MJ, Bellabarba CB, Chapman JR. Spinal epidural abscesses: risk factors, medical versus surgical management, a retrospective review of 128 cases. Spine J. 2014;14(2):326–30.

21. Li YD, Wong CB, Tsai TT, Lai PL, Niu CC, Chen LH, Fu TS. Appropriate duration of post-surgical intravenous antibiotic therapy for pyogenic spondylodiscitis. BMC Infect Dis. 2018;18(1):468.

22. Park KH, Cho OH, Lee JH, Park JS, Ryu KN, Park SY, Lee YM, Chong YP, Kim SH, Lee SO, Choi SH, Bae IG, Kim YS, Woo JH, Lee MS. Optimal duration of antibiotic therapy in patients with hematogenous vertebral osteomyelitis at low risk and high risk of recurrence. Clin Infect Dis. 2016;62(10):1262–9.

23. Okay G, Akkoyunlu Y, Bolukcu S, Durdu B, Hakyemez IN, Koc MM. Analysis of infectious spondylodiscitis: 7-years data. Pak J Med Sci. 2018;34(6):1445–51.

24. Kaya S, Kaya S, Kavak S, Comoglu S. A disease that is difficult to diagnose and treat: evaluation of 343 spondylodiscitis cases. J Int Med Res. 2021;49(11):3000605211060197.

25. Taşdemir C. Retrospective examination of patients with infective spondylodiscitis followed in Bursa Uludağ University Medical Faculty Hospital infective diseases clinic and outpatient clinic between 2000–2020. Medical Speciality Thesis. Bursa: Uludağ University Faculty of Medicine, Infectious Diseases and Clinical Microbiology; 2022.

26. Ahn KS, Kang CH, Hong SJ, Kim BH, Shim E. The correlation between follow-up MRI findings and laboratory results in pyogenic spondylodiscitis. BMC Musculoskelet Disord. 2020;21(1):428.

27. Zarrouk V, Feydy A, Sallès F, Dufour V, Guigui P, Redondo A, Fantin B. Imaging does not predict the clinical outcome of bacterial vertebral osteomyelitis. Rheumatology (Oxford). 2007;46(2):292–5.

28. Chae HJ, Kim J, Kim C. Clinical characteristics of spinal epidural abscess accompanied by bacteremia. J Korean Neurosurg Soc. 2021;64(1):88–99.

RESEARCH ARTICLE / ARAŞTIRMA

DOI: 10.4274/mjima.galenos.2025.25617.3
Mediterr J Infect Microb Antimicrob 2026;15:25617.3
Erişim: http://dx.doi.org/10.4274/mjima.galenos.2025.25617.3

Improvements in Surgical Site Infection Control in Albania: Impact of Infrastructure, Sterilization Standards, and Clinical Practices

Arnavutluk'ta Cerrahi Alan Enfeksiyonunun Kontrolünde İyileştirmeler: Altyapı, Sterilizasyon Standartları ve Klinik Uygulamaların Etkisi

Ardit Kacani[1], Krenar Lilaj[2], Vjollca Duro[3], Emigleo Sulku[4], Fjoralba Mulgeci[4], Sara Ferizaj[4], Irisi Sure[4], Lejdis Zeneli[4], Edmond Puca[5]*

[1]*University Hospital Center "Mother Theresa", Department of General and Digestive Surgery, Tirana, Albania*
[2]*University Hospital Center "Mother Theresa", Department of Anaesthesia and Reanimation, Tirana, Albania*
[3]*University Hospital Center "Mother Theresa", Laboratory Service, Tirana, Albania*
[4]*University of Medicine, Faculty of Medicine, Tirana, Albania*
[5]*University Hospital Center "Mother Theresa", Service of Infectious Diseases, Tirana, Albania*

Abstract

Introduction: Surgical site infections (SSIs) are a major cause of postoperative morbidity, prolonged hospitalization, and increased healthcare costs worldwide. In Albania, SSI rates have historically been high; however, recent infrastructural and organizational reforms in tertiary hospitals have led to measurable progress. This study evaluated the effects of infrastructure modernization, centralization of sterilization services, and the introduction of disposable materials and laparoscopic techniques on SSI incidence at the University Hospital Center "Mother Teresa" in Albania.

Materials and Methods: A prospective observational study was conducted in the Department of General and Digestive Surgery from October 2023 to October 2024. Data were collected on patient demographics (age, sex), surgical characteristics (upper or lower gastrointestinal; elective or emergency), comorbidities (hypertension, diabetes mellitus, malignancy), and SSI occurrence. Microbiological analyses included pathogen identification and antimicrobial susceptibility testing (AST) in accordance with the 2024 European Committee on AST guidelines.

Results: Among 1,179 patients (51.2% male; mean age, 57.8 years), 5.4% developed SSIs after abdominal surgery. Infection rates were significantly higher in lower gastrointestinal procedures (57.8%) than in upper gastrointestinal procedures (42.2%). Patients aged ≥ 50 years exhibited a greater risk of SSI ($p = 0.01$), as did those with comorbidities ($p = 0.0007$) and diabetes mellitus ($p = 0.0006$). Mean hospital stay was markedly longer among infected patients (4.4 vs. 2.0 days; $p < 0.0001$). *Escherichia coli* (39%) and *Enterococcus faecalis* (22%) were the most common isolates, demonstrating notable resistance to ciprofloxacin (34%) and trimethoprim–sulfamethoxazole (31%). Reductions in infection rates were closely linked to enhanced operating room ventilation, improved sterilization practices, and the use of single-use materials.

Conclusion: Albania has achieved substantial progress in SSI prevention through targeted infrastructural and procedural reforms. Nonetheless, persistent challenges—particularly antimicrobial resistance and the lack of a national SSI surveillance system—underscore the need for a coordinated, multidisciplinary strategy. Strengthening antimicrobial stewardship, standardizing perioperative protocols, and expanding the use of minimally invasive surgery are key priorities for sustaining improvements.

Keywords: Surgical site infection, Albania, abdominal surgery, hospital infrastructure, antimicrobial resistance, infection control

Cite this article as: Kacani A, Lilaj K, Duro V, Sulku E, Mulgeci F, Ferizaj S, Sure I, Zeneli L, Puca E. Improvements in surgical site infection control in Albania: Impact of infrastructure, sterilization standards, and clinical practices. Mediterr J Infect Microb Antimicrob.

Address for Correspondence/Yazışma Adresi: Edmond Puca MD, University Hospital Centre "Mother Theresa", Service of Infectious Diseases, Tirana, Albania
E-mail: edmond_puca@yahoo.com **ORCID ID:** orcid.org/0000-0002-0621-4865

Received/Geliş Tarihi: 16.09.2025 **Accepted/Kabul Tarihi:** 11.11.2025

Epub: 25.11.2025
Published: 26.01.2026

Öz

Giriş: Cerrahi alan enfeksiyonları (CAE), dünya çapında ameliyat sonrası morbidite, uzun süreli hastanede kalış ve artan sağlık hizmetleri maliyetlerinin başlıca nedenidir. Arnavutluk'ta CAE oranları tarihsel olarak yüksek olmuştur; ancak, üçüncü basamak hastanelerdeki son altyapısal ve organizasyonel reformlar ölçülebilir bir ilerleme sağlamıştır. Bu çalışmada, Arnavutluk'taki "Rahibe Teresa" Üniversite Hastanesi Merkezi'nde altyapı modernizasyonu, sterilizasyon hizmetlerinin merkezileştirilmesi ve tek kullanımlık malzemeler ile laparoskopik tekniklerin kullanılmaya başlanmasının CAE insidansı üzerindeki etkileri değerlendirilmiştir.

Gereç ve Yöntem: Ekim 2023 ile Ekim 2024 tarihleri arasında Genel ve Sindirim Sistemi Cerrahisi Anabilim Dalı'nda prospektif bir gözlemsel çalışma yürütüldü. Hastaların demografik özellikleri (yaş, cinsiyet), cerrahi özellikleri (üst veya alt gastrointestinal; elektif veya acil), eşlik eden hastalıklar (hipertansiyon, diabetes mellitus, malignite) ve CAE oluşumu hakkında veriler toplandı. Mikrobiyolojik analizler, 2024 Avrupa Antimikrobiyal Duyarlılık Testi (ADT) yönergelerine uygun olarak patojen tanımlama ve ADT içeriyordu.

Bulgular: Bin yüz yetmiş dokuz hastanın (%51,2'si erkek; ortalama yaş 57,8) %5,4'ünde abdominal cerrahi sonrası CAE gelişti. Enfeksiyon oranları alt gastrointestinal prosedürlerde (%57,8), üst gastrointestinal prosedürlere (%42,2) göre anlamlı derecede daha yüksekti. ≥ 50 yaş hastalarda CAE riski daha yüksekti ($p = 0,01$), eşlik eden hastalıkları ($p = 0,0007$) ve diyabeti ($p = 0,0006$) olanlarda da aynı risk mevcuttu. Enfekte hastalarda ortalama hastanede kalış süresi belirgin şekilde daha uzundu (4,4'e karşı 2,0 gün; $p < 0,0001$). *Escherichia coli* (%39) ve *Enterococcus faecalis* (%22) en sık görülen izolatlar olup, siprofloksasine (%34) ve trimetoprim-sülfametoksazol (%31) karşı belirgin direnç göstermekteydi. Enfeksiyon oranlarındaki azalmalar, gelişmiş ameliyathane ventilasyonu, iyileştirilmiş sterilizasyon uygulamaları ve tek kullanımlık malzemelerin kullanımıyla yakından bağlantılıydı.

Sonuç: Arnavutluk, hedefli altyapı ve prosedür reformları sayesinde CAE önlenmesinde önemli ilerleme kaydetmiştir. Bununla birlikte, özellikle antimikrobiyal direnç ve ulusal bir CAE gözetim sisteminin olmaması gibi devam eden zorluklar, koordineli ve çok disiplinli bir stratejiye olan ihtiyacın altını çizmektedir. Antimikrobiyal yönetimin güçlendirilmesi, perioperatif protokollerin standartlaştırılması ve minimal invaziv cerrahinin kullanımının yaygınlaştırılması, iyileştirmelerin sürdürülmesi için temel önceliklerdir.

Anahtar Kelimeler: Cerrahi alan enfeksiyonu, Arnavutluk, karın cerrahisi, hastane altyapısı, antimikrobiyal direnç, enfeksiyon kontrolü

Introduction

Healthcare-associated infections (HAIs) remain a persistent and critical challenge in modern surgical care. Also known as hospital-acquired or nosocomial infections, HAIs pose a substantial threat to patient safety and quality of care. Although largely preventable through evidence-based infection control strategies, they continue to occur with alarming frequency. HAIs are associated with extended hospital stays, adverse clinical outcomes, long-term disability, rising antimicrobial resistance (AMR), increased healthcare costs, and preventable deaths. Globally, they contribute significantly to both morbidity and mortality, representing a major burden for healthcare systems.

Among HAIs, surgical site infections (SSIs) account for approximately 16%–20% of all cases and exert profound effects on patient outcomes, including delayed wound healing, prolonged hospitalization, increased readmission rates, and elevated mortality[1–4]. The global prevalence of HAIs in acute care hospitals is estimated at 7.1%, but the burden is substantially higher in low- and middle-income countries (LMICs)[5,6]. In these settings, the financial consequences are particularly severe because patients frequently shoulder the majority of healthcare expenses. Developing an SSI can increase household expenditures by more than 10% of annual income, imposing a considerable economic strain[7].

In Albania, earlier reports documented SSI prevalence rates of up to 13.3%, underscoring long-standing infrastructural and procedural gaps in infection control[8,9]. Over the past decade, however, major reforms have been implemented at the University Hospital Center "Mother Teresa," including the renovation of surgical theaters, centralization of sterilization services, introduction of disposable surgical materials, and acquisition of laparoscopic systems. Recent prospective data indicate that these reforms have reduced the SSI incidence following abdominal surgery to 5.4%.

This study aims to evaluate the extent to which infrastructural modernization, improved sterilization standards, and updated clinical practices have contributed to this decline. Unlike previous investigations that primarily addressed clinical or epidemiological aspects, the present analysis highlights hospital modernization as a pivotal determinant in the prevention and control of SSIs in Albania.

Materials and Methods

Inclusion and Exclusion Criteria

All adult patients (≥ 14 years) who underwent abdominal surgery—either elective or emergency—at the Department of General and Digestive Surgery, University Hospital Center "Mother Teresa," Albania, between October 2023 and October 2024 were included. Abdominal procedures were classified as upper gastrointestinal, lower gastrointestinal, hepatobiliary, or hernia-related surgeries. Laparoscopic procedures were analyzed as a separate subgroup to assess their increasing utilization during the study period. Reoperations within the same hospitalization were included when associated with postoperative complications.

Patients admitted solely for diagnostic procedures, minor interventions (e.g., biopsies, wound revisions), or transferred postoperatively from other hospitals were excluded. Pediatric, thoracic, cardiac, trauma, and neurosurgical procedures were also excluded.

Study Design and Context

This research employed a hybrid design combining a narrative review of previously published national data with a prospective observational analysis. The prospective component included all abdominal surgery patients treated between October 2023 and October 2024 at the University Hospital Center "Mother Teresa."

Comparative analyses between pre- and post-reconstruction periods were based on institutional records obtained before (2015–2019) and after (2023–2024) modernization of operating theaters and central sterilization units. Published hospital reports and earlier prevalence surveys represented the pre-reconstruction baseline, whereas current data from the prospective survey characterized the post-reconstruction phase.

Descriptive statistics were used to compare SSI incidence, patient demographics, risk factors, and microbiological patterns across the two periods. The study emphasized systemic infrastructural and procedural improvements rather than individual patient-level matching.

The 2023–2024 prospective survey included 1,179 patients and assessed SSI incidence, associated risk factors, microbiological profiles, and AMR trends. Additionally, literature on SSI epidemiology in Europe and LMICs was reviewed to provide contextual background.

Data Extraction and Thematic Analysis

Data were analyzed under three domains:

• Infrastructure and organization: Upgrades in operating rooms, sterilization processes, use of disposable materials, and patient ward improvements.

• Clinical epidemiology: SSI incidence, patient demographics, comorbidities, and surgical categories.

• Microbiology and AMR: Pathogen distribution and antimicrobial susceptibility profiles.

Microbiological Sampling and Analysis

For patients clinically suspected of SSIs, wound specimens were collected aseptically. In superficial incisional infections, sterile cotton swabs were used after cleaning the wound with saline to eliminate surface contaminants. For deep or purulent infections, aspirates or tissue biopsies were collected to enhance diagnostic accuracy. All specimens were transported promptly to the Microbiology Laboratory of the University Hospital Center "Mother Teresa" and processed within 2 h of collection.

Microbial Identification and Antimicrobial Susceptibility Testing (AST)

Bacterial isolates were cultured on blood and MacConkey agar, followed by biochemical identification using conventional methods and the VITEK® 2 Compact system (bioMérieux, France). AST was performed according to the 2024 European Committee on AST (EUCAST) guidelines, using the automated VITEK® 2 platform. Breakpoints were interpreted per EUCAST standards. Resistance data were analyzed for each organism individually, and antibiotic names are reported as generic compounds.

Infrastructure and Organizational Upgrades

Comparison of conditions before and after departmental renovation revealed substantial improvements:

Before renovation:

• Operating theaters lacked adequate ventilation and temperature regulation.

• Sterilization was decentralized, performed independently by each department with variable standards.

• Reusable cloth drapes and gowns were employed, heightening infection risk.

• Laparoscopic instruments were unavailable.

• Wards were overcrowded (up to four patients per room) and lacked proper ventilation systems.

After renovation:

• Modern operating theaters with biofiltration and climate control systems were constructed.

• Sterilization was centralized under a specialized, trained team.

• Disposable gowns and drapes replaced reusable materials.

• Laparoscopic instruments became standard for minimally invasive surgery.

Table 1. Clinical and demographic characteristics of patients with SSIs.

Patients with SSIs	Number	Percentage
Female	27	42.19%
Male	37	57.81%
Lower gastrointestinal surgery	37	57.81%
Upper gastrointestinal surgery	27	42.19%
Comorbidities	45	70.31%
Arterial hypertension	41	64.06%
Diabetes mellitus	18	28.12%
Cancer	21	32.81%

SSIs, surgical site infections.

• Patient wards were renovated, featuring reduced bed density and air-conditioning systems.

Ethical Approval and Informed Consent

The study complied with the ethical principles of the Declaration of Helsinki (1964) and subsequent amendments. The research protocol was approved by the Ethics Committee of the University of Medicine, Tirana, and the National Agency for Scientific Research and Innovation (approval number: 20, dated: 02.11.2023). Written informed consent was obtained from all participants prior to enrollment. All data were anonymized to ensure confidentiality.

Statistical Analysis

Data analysis was performed using SPSS version 20.0 (IBM Corp., Armonk, NY, USA). Continuous variables were expressed as mean ± standard deviation, and categorical variables as frequencies (n) and percentages (%). Associations between categorical variables were evaluated using the chi-square test, while continuous variables were compared using the Student's t-test. Univariate logistic regression was applied to identify predictors of SSI development, with results reported as odds ratios (ORs) and 95% confidence intervals. Statistical significance was set at $p \leq 0.05$.

Results

A total of 1,179 patients were included, comprising 51.23% men and 48.77% women, with a mean age of 57.8 years (range: 19–92 years). SSIs occurred in 64 patients (5.4%), all of which were classified as superficial. Although infection rates were marginally higher in males, the difference was not statistically significant (OR 1.47, p = 0.12) (Table 1).

Table 2. Microorganisms isolated from SSIs.

Microorganisms isolated from the SSI	Number	Percentage
Escherichia coli	25	39.06%
Enterococcus faecalis	15	23.44%
Klebsiella pneumoniae	6	9.37%
Staphylococcus epidermidis	4	6.25%
Pseudomonas aeruginosa	3	4.69%
Staphylococcus hominis	3	4.69%
Enterobacter cloacae	2	3.13%
Staphylococcus aureus	1	1.56%
Staphylococcus capitis	1	1.56%
Staphylococcus caprae	1	1.56%
Streptococcus agalactiae	1	1.56%
Citrobacter	1	1.56%
Corynebacterium	1	1.56%

SSIs, surgical site infections.

Lower gastrointestinal surgeries carried a greater infection risk, accounting for 58% of SSIs, compared with 42% in upper gastrointestinal procedures. Laparoscopic operations represented 21% of all surgeries and showed a significantly lower SSI rate than open procedures (2.1% vs. 6.3%; $p < 0.05$), confirming the infection-preventive advantage of minimally invasive techniques.

Patients with SSIs were significantly older than those without infection (mean age, 66.2 vs. 57.0 years; $p < 0.0001$). Age ≥ 50 years increased the infection risk nearly threefold (OR: 2.69, p = 0.01). Comorbidities were also strong predictors (OR: 2.71, p = 0.0007), particularly type 2 diabetes mellitus (OR: 2.98, p = 0.0006). Older patients were significantly more likely to have comorbidities (OR 19.59, $p < 0.0001$) (Table 2).

Microbiological cultures confirmed that *Escherichia coli* and *Enterococcus faecalis* were the predominant pathogens, accounting for 39% and 23% of SSIs, respectively.

Resistance patterns are summarized in Table 3. Notably, over one-third of bacterial isolates exhibited resistance to ciprofloxacin, while resistance to trimethoprim–sulfamethoxazole, cefazolin, and levofloxacin exceeded 30%. In contrast, resistance to tetracycline, nitrofurantoin, penicillin, and fosfomycin remained low (< 5%). Patients who developed SSIs experienced hospital stays more than twice as long as those without infection, averaging 4.4 days versus 2.0 days ($p < 0.0001$). No deaths were directly attributable to SSIs during the study period. Overall postoperative mortality among patients undergoing abdominal surgery was 1.1% (13/1,179), and all deaths occurred in individuals with advanced malignancies or severe comorbidities rather than active SSIs. These findings indicate that improvements in infection control contributed not only to reduced morbidity but also to the maintenance of low mortality rates.

Discussion

SSIs are among the most frequent postoperative complications, contributing to increased patient morbidity, prolonged hospital

Table 3. AMR patterns among bacterial isolates from SSIs.

Antibiotic	Resistant isolates (%)
Ciprofloxacin	34%
Trimethoprim–sulfamethoxazole	31%
Cefazolin	30%
Levofloxacin	30%
Amoxicillin-clavulanate	18%
Gentamicin	12%
Tetracycline	3%
Nitrofurantoin	3%
Penicillin	3%
Fosfomycin	3%

AMR, antimicrobial resistance; SSIs, surgical site infections.

stays, and higher healthcare costs. Effective prevention and control of SSIs are widely recognized as key indicators of surgical quality and patient safety[10–12].

In LMICs, SSIs remain a major concern due to limited access to sterile equipment, overcrowded hospitals, and suboptimal infection control practices. The incidence in these settings is often two- to fivefold higher than in high-income countries, resulting in substantial preventable morbidity and mortality[7,13–15].

Prior to this study, national data on SSIs in Albania were scarce[8,16]. The observed decline in SSI rates from 13.3% to 5.4% strongly correlates with recent infrastructural and organizational improvements. These include renovation of operating theaters, centralization of sterilization services, adoption of single-use surgical materials, and the integration of laparoscopic techniques. International evidence supports that ventilated, filtered operating rooms, centralized sterilization units, and disposable materials significantly reduce infection rates. Minimally invasive surgery further lowers SSI risk by reducing wound exposure[17–21]. Our stratified analysis confirmed this protective effect, as laparoscopic procedures demonstrated substantially lower SSI rates, emphasizing the importance of expanding minimally invasive surgery programs in tertiary hospitals.

Despite these advancements, risk factors for SSI in Albania remain consistent with global trends, including advanced age, comorbidities, and diabetes mellitus. The predominance of *Escherichia coli* and *Enterococcus* species among isolates reflects the gastrointestinal origin of contamination[22–24]. Importantly, no SSI-related deaths were observed, and overall postoperative mortality remained low (1.1%), with all deaths occurring in patients with advanced malignancies or severe comorbidities rather than active infections. These findings underscore the clinical significance of modern infection control measures and laparoscopic techniques in reducing morbidity and maintaining low mortality in LMIC surgical settings.

A significant concern is AMR. Resistance rates exceeding 30% for commonly used antibiotics—including ciprofloxacin, trimethoprim–sulfamethoxazole, cefazolin, and levofloxacin—threaten the effectiveness of both prophylactic and therapeutic regimens. In contrast, resistance to tetracycline, nitrofurantoin, penicillin, and fosfomycin remained low (< 5%). This pattern mirrors reports from the European Centre for Disease Prevention and Control documenting rising AMR trends across Europe[25–27]. Accordingly, antimicrobial stewardship programs are as critical as infrastructural and procedural improvements in sustaining progress against SSIs.

Remaining challenges in Albania include:

• Absence of a national SSI surveillance system: Current data are limited to individual hospital reports.

• Unquantified economic burden: Although prolonged hospital stays suggest increased costs, comprehensive analyses are lacking.

• Need for continuous staff training and protocol standardization: Consistent implementation of perioperative infection control measures across hospitals is essential.

Study Limitations

This study has several limitations. First, it was conducted in a single tertiary hospital, which may restrict the generalizability of the findings to other healthcare settings in Albania. Second, only abdominal surgical procedures were evaluated, excluding thoracic, cardiac, traumatic, pediatric, and other surgical categories. Third, the comparison of pre- and post-reconstruction data relied partly on previously published institutional records rather than matched patient cohorts, which may introduce variability in reporting standards. Additionally, the study did not assess long-term postoperative outcomes beyond the hospitalization period, potentially underestimating late-onset SSIs. Finally, although antimicrobial resistance patterns were analyzed, the study did not evaluate the adequacy or adherence to antibiotic prophylaxis protocols, an important determinant of SSI prevention.

Conclusion

This study demonstrates that Albania has achieved substantial reductions in SSIs through hospital modernization, centralized sterilization services, and the adoption of disposable materials and laparoscopic techniques. The SSI incidence decreased from 13.3% to 5.4%, confirming that targeted infrastructural and organizational reforms can significantly enhance patient safety. Despite these advancements, AMR remains a major concern, underscoring the need for a national SSI surveillance system, standardized perioperative protocols, and strengthened antimicrobial stewardship programs. Sustained investment in hospital infrastructure, staff training, and infection prevention initiatives will be essential to maintain these gains and expand their impact across the national healthcare system.

Ethics Committee Approval: The study complied with the ethical principles of the Declaration of Helsinki (1964) and subsequent amendments. The research protocol was approved by the Ethics Committee of the University of Medicine, Tirana, and the National Agency for Scientific Research and Innovation (approval number: 20, dated: 02.11.2023).

Informed Consent: Written informed consent was obtained from all participants prior to enrollment. All data were anonymized to ensure confidentiality.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.K., Concept: A.K., K.L., Design: K.L., E.P., Data Collection or Processing: V.D., E.S., Analysis or Interpretation: A.K., K.L., V.D., E.S., F.M., S.F., I.S., L.Z., E.P., Literature Search: A.K., K.L., V.D., E.P., Writing: E.P.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Alamer A, Alharbi F, Aldhilan A, Almushayti Z, Alghofaily K, Elbehiry A, Abalkhail A. Healthcare-associated infections (HAIs): Challenges and measures taken by the radiology department to control infection transmission. Vaccines (Basel). 2022;10:2060.
2. Sandu AM, Chifiriuc MC, Vrancianu CO, Cristian R-E, Alistar CF, Constantin M, Paun M, Alistar A, Popa LG, Popa MI. Healthcare-associated infections: The role of microbial and environmental factors in infection control—a narrative review. Infect Dis Ther. 2025;14:933–71.
3. D'Alessandro D, Fara GM. Hospital environments and epidemiology of healthcare-associated infections. Indoor Air Quality in Healthcare Facilities. 2017;41–52.
4. Sartelli M, Marini CP, McNelis J, Coccolini F, Rizzo C, Labricciosa FM, Petrone P. Preventing and controlling healthcare-associated infections: The first principle of every antimicrobial stewardship program in hospital settings. Antibiotics. 2024;13:896.
5. Suetens C, Latour K, Kärki T, Ricchizzi E, Kinross P, Moro ML, Jans B, Hopkins S, Hansen S, Lyytikäinen O, Reilly J, Zaletel M, Plachouras D, Monnet DL. Prevalence of healthcare-associated infections, estimated incidence and composite antimicrobial resistance index in acute care hospitals and long-term care facilities: Results from two European point prevalence surveys, 2016 to 2017. Eurosurveillance. 2018;23:1800516.
6. Ripabelli G, Salzo A, Mariano A, Sammarco ML, Tamburro M. Healthcare-associated infections point prevalence survey and antimicrobials use in acute care hospitals (PPS 2016–2017) and long-term care facilities (HALT-3): A comprehensive report of the first experience in Molise Region, Central Italy, and targeted intervention strategies. J Infect Public Health. 2019;12:509–15.
7. Monahan M, Jowett S, Pinkney T, Brocklehurst P, Morton DG, Abdali Z, Roberts TE. Surgical site infection and costs in low- and middle-income countries: A systematic review of the economic burden. PLoS One. 2020;15:e0232960.
8. Faria S, Sodano L, Gjata A, Dauri M, Sabato AF, Bilaj A, Mertiraj O, Llazo E, Kodra Y, Schinaia N, Colarossi G, Pagano S, Petrosillo N. The first prevalence survey of nosocomial infections in the University Hospital Centre "Mother Teresa" of Tirana, Albania. J Hosp Infect. 2007;65:244–50.
9. Gillespie BM, Harbeck E, Rattray M, Liang R, Walker R, Latimer S, Thalib L, Erichsen Andersson A, Griffin B, Ware R, Fanning A, O'Connor M, Chaboyer W. Worldwide incidence of surgical site infections in general surgical patients: A systematic review and meta-analysis of 488,594 patients. Int J Surg. 2021;95:106136.
10. Pinchera B, Buonomo AR, Schiano Moriello N, Scotto R, Villari R, Gentile I. Update on the management of surgical site infections. Antibiotics (Basel). 2022;11:1608.
11. Nakhleh H, Samuel Fatokun B, Nakyanzi H, Mshaymesh S, Wellington J, Uwishema O. Surgical site infections in Sub-Saharan Africa: Epidemiology, risk factors, and prevention strategies. Ann Med Surg (Lond). 2025;87:3388–92.
12. Rezaei AR, Zienkiewicz D, Rezaei AR. Surgical Site Infections: A Comprehensive Review. J Trauma Inj. 2025;38:71.
13. Costabella F, Patel KB, Adepoju AV, Singh P, Attia Hussein Mahmoud H, Zafar A, Patel T, Watekar NA, Mallesh N, Fawad M, Hamza A, Thunga S, Mehmood S, Bhutta ZA. Healthcare cost and outcomes associated with surgical site infection and patient outcomes in low- and middle-income countries. Cureus. 2023;15:e42493.
14. Chaker SC, James AJ, Perdikis G, Nthumba P. Surgical care bundles for surgical site infection prevention in high-income and low-to-middle-income countries: A comparative review. Perioper Care Oper Room Manag. 2024;35:100406.
15. Mehtar S, Wanyoro A, Ogunsola F, Ameh EA, Nthumba P, Kilpatrick C, Revathi G, Antoniadou A, Giamarelou H, Apisarnthanarak A, Ling ML, Pittet D. Implementation of surgical site infection surveillance in low- and middle-income countries: A position statement for the international society for infectious diseases. Int J Infect Dis. 2020;100:123–31.
16. Gjerazi E, Gjata A, Kureta E. Risk factors of surgical site infections in a general surgery ward in Tirana. 2015.
17. Sadrizadeh S, Aganovic A, Bogdan A, Wang C, Afshari A, Hartmann A, Croitoru C, Khan A, Kriegel M, Lind M, Blomqvist P, Toprak S, Bivolarova MP, Nielsen PV. A systematic review of operating room ventilation. J Build Eng. 2021;40:102693.
18. Humphreys H. Infection prevention and control considerations regarding ventilation in acute hospitals. Infect Prev Pract. 2021;3:100180.
19. Pfeifer R, Hildebrand F, Halvachizadeh S. Operating room (OR) requirements. Eur J Trauma Emerg Surg. 2025;51:135.
20. Mullen AN, Wieser E. Improvement of operating room air quality and sustained reduction of surgical site infections in an orthopedic specialty hospital. Am J Infect Control. 2024;52:183–90.
21. Hailu S, Mulugeta H, Girma T, Asefa A, Regasa T. Evidence-based guidelines on infection prevention and control in operation theatres for anesthetists in a resource-limited setting: Systematic review/meta-analysis. Ann Med Surg (Lond). 2023;85:2858–64.
22. Santucci C, Carioli G, Bertuccio P, Malvezzi M, Pastorino U, Boffetta P, Negri E, Bosetti C, La Vecchia C. Progress in cancer mortality, incidence, and survival: A global overview. Eur J Cancer Prev. 2020;29:367–81.
23. Chang AY, Bolongaita S, Cao B, Castro MC, Karlsson O, Mao W, Norheim OF, Ogbuoji O, Jamison DT. Epidemiological and demographic trends and projections in global health from 1970 to 2050: A descriptive analysis from the third lancet commission on investing in health, global health 2050. Lancet. 2025;406:940–9.
24. McConn BR, Kraft AL, Durso LM, Ibekwe AM, Frye JG, Wells JE, Tobey EM, Ritchie S, Williams CF, Cook KL, Dungan RS, Ricke SC. An analysis of culture-based methods used for the detection and isolation of *Salmonella* spp., *Escherichia coli*, and *Enterococcus* spp. from surface water: A systematic review. Sci Total Environ. 2024;927:172190.
25. Ho CS, Wong CTH, Aung TT, Lakshminarayanan R, Mehta JS, Rauz S, McNally A, Kintses B, Peacock SJ, de la Fuente-Nunez C, Mathur P, Blaskovich M. Antimicrobial resistance: A concise update. Lancet Microbe. 2025;6.
26. Salam MA, Al-Amin MY, Salam MT, Pawar JS, Akhter N, Rabaan AA, Alqumber MAA. Antimicrobial resistance: A growing serious threat for global public health. Healthcare (Basel). 2023;11:1946.
27. Ahmed SK, Hussein S, Qurbani K, Ibrahim RH, Fareeq A, Mahmood KA, Mohamed MG. Antimicrobial resistance: Impacts, challenges, and future prospects. J Med Surg Public Health. 2024;2:100081.

RESEARCH ARTICLE / ARAŞTIRMA

DOI: 10.4274/mjima.galenos.2025.25555.4
Mediterr J Infect Microb Antimicrob 2026;15:25555.4
Erişim: http://dx.doi.org/10.4274/mjima.galenos.25555.4

A Laboratory Decision-Support System for Reflective Urine Culture Testing: Development of an Interpretable Artificial Intelligence Model

Reflektif İdrar Kültürü Testleri için Bir Laboratuvar Karar Destek Sistemi: Yorumlanabilir Bir Yapay Zekâ Modelinin Geliştirilmesi

Ferhat Demirci[1,2,*], Murat Akşit[2], İlkay Akbulut[3], Aylin Demirci[4]

[1]*Dokuz Eylül University, Institute of Health Sciences, Department of Neurosciences, İzmir, Türkiye*
[2]*University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, Department of Medical Biochemistry, İzmir, Türkiye*
[3]*University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, Department of Infectious Diseases and Clinical Microbiology, İzmir, Türkiye*
[4]*University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, Department of Family Medicine, İzmir, Türkiye*

Abstract

Introduction: Urinary tract infections are a common diagnostic challenge. Although urine culture remains the gold standard, it is time-consuming and often ordered reflexively. This study aimed to develop and validate an interpretable machine-learning–based Laboratory Decision-Support System (LDSS) to guide reflective urine culture prioritization using only structured laboratory data.

Materials and Methods: We analyzed a retrospective cohort of 51,923 adult patients. Seven machine learning algorithms were trained, with the Random Forest (RF) model demonstrating the highest accuracy. SHapley Additive exPlanations was employed to ensure model interpretability. A reduced RF model, using the top 10 predictive features, was used to construct three scoring systems: one emphasizing model fidelity, one optimizing diagnostic balance, and one maximizing sensitivity.

Results: The RF model demonstrated excellent performance (external receiver operating characteristic – area under the curve [ROC-AUC]: 0.956). The simplified 10-variable model maintained high accuracy (ROC-AUC: 0.947). Key predictors included bacterial count, leukocyte count, nitrite presence, and patient age. The scoring systems offered flexible options tailored to different diagnostic priorities, with the SAFE-Score achieving 95.3% sensitivity.

Conclusion: The developed LDSS supports rational antibiotic use by reducing unnecessary culture testing. Its explainable structure facilitates collaboration between laboratory professionals and clinicians, contributing to standardized reflective testing workflows and interdisciplinary decision-making and strengthens antimicrobial stewardship, while preserving the central role of urine culture in infection management.

Keywords: Urinary tract infections, machine learning, urine culture

Öz

Giriş: İdrar yolu enfeksiyonları sık karşılaşılan bir tanı sorunudur. Altın standart olan idrar kültürü, hem zaman alıcıdır hem de çoğu zaman gereksiz yere istenir. Bu çalışmada, yalnızca yapılandırılmış laboratuvar verilerini kullanarak reflektif idrar kültürü istemine rehberlik edecek, yorumlanabilir bir makine öğrenimi (ML) tabanlı Laboratuvar Karar Destek Sistemi (LKDS) geliştirilmesi ve doğrulanması amaçlandı.

Gereç ve Yöntem: Retrospektif olarak 51.923 erişkin hastaya ait veriler incelendi. Yedi ML algoritması eğitildi; en yüksek doğruluk Rastgele Orman (Random Forest, RF) modelinde elde edildi. Model şeffaflığı için SHapley Additive exPlanations kullanıldı. En iyi 10 özellikten oluşan sadeleştirilmiş

Cite this article as: Demirci F, Akşit M, Akbulut İ, Demirci A. A laboratory decision-support system for reflective urine culture testing: Development of an interpretable AI model. Mediterr J Infect Microb Antimicrob.

Address for Correspondence/Yazışma Adresi: Ferhat Demirci MD, Dokuz Eylül University, Institute of Health Sciences, Department of Neurosciences, İzmir, Türkiye
E-mail: drferhat5505@hotmail.com **ORCID ID:** orcid.org/0000-0002-5999-3399
Received/Geliş Tarihi: 18.07.2025 **Accepted/Kabul Tarihi:** 08.12.2025

Epub: 07.01.2026
Published: 03.02.2026

Öz

RF modeliyle üç farklı puanlama sistemi geliştirildi: Model doğruluğuna öncelik veren, tanısal dengeyi optimize eden ve hassasiyeti en üst düzeye çıkaran modeller.

Bulgular: RF modeli mükemmel performans gösterdi (harici testler – alıcı işletim karakteristiği eğrisi altında kalan alan [ROC-AUC]: 0,956). Basitleştirilmiş 10 değişkenli model yüksek doğruluğu korumuştur (ROC-AUC: 0,947). Temel öngörücüler arasında bakteri sayısı, lökositler, nitrit ve yaş yer almıştır. Skorlama sistemleri, farklı tanı hedeflerine göre uyarlanmış esnek seçenekler sunmuş ve SAFE-Skoru %95,3 hassasiyete ulaşmıştır.

Sonuç: Geliştirilen LKDS, gereksiz kültür sayısını azaltarak rasyonel antibiyotik kullanımını desteklemektedir. Açıklanabilir yapısı, laboratuvar profesyonelleriyle klinisyenler arasındaki iş birliğini kolaylaştırarak standartlaştırılmış reflektif test süreçlerine ve disiplinler arası karar vermeye katkı sağlar.

Anahtar Kelimeler: İdrar yolu enfeksiyonları, makine öğrenimi, idrar kültürü

Introduction

Urinary tract infections (UTIs) are among the most common infections in clinical practice, with an estimated global incidence exceeding 150 million cases annually[1]. They are associated with substantial healthcare costs, frequent antibiotic prescriptions, and increased diagnostic burden, particularly in outpatient and emergency settings[2,3]. Accurate diagnosis remains challenging due to nonspecific symptoms and reliance on time-consuming laboratory tests[4].

Urine culture is considered the gold standard for UTI diagnosis. However, its 24–48-hour turnaround often necessitates empiric antibiotic treatment before microbiological confirmation[5]. This practice contributes to antimicrobial resistance, now recognized by the World Health Organization as a global health threat[6]. Moreover, up to 60%–70% of urine cultures yield negative or clinically insignificant results, highlighting potential overuse of testing and therapy[7].

Rapid dipstick tests, detecting leukocyte esterase and nitrite, provide immediate screening but show variable performance across populations, with sensitivity and specificity ranging from 68% to 88% and 17% to 98%, respectively[8].

This diagnostic uncertainty has prompted efforts to improve laboratory decision-making, including the use of reflective testing. Reflective testing, increasingly recognized in modern laboratory medicine, involves laboratory physicians adding further analyses or interpretative comments after reviewing initial test results to enhance diagnostic reasoning[9]. In UTIs, this expert-led approach aids accurate interpretation and encourages more judicious use of microbiological testing. Laboratory physicians thus face the dual challenge of minimizing unnecessary culture requests while ensuring patients with a high likelihood of positive cultures are correctly identified.

In most laboratory information systems (LIS), detailed symptom information is not captured; only test orders and preliminary diagnoses, such as International Classification of Diseases (ICD) codes, are typically available. Consequently, the predictive modeling approach in this study relied solely on structured laboratory data. To address this, we developed a standardized, interpretable, and data-driven Laboratory Decision-Support System (LDSS) to optimize urine culture utilization using routine laboratory parameters. The LDSS is not intended to replace clinical diagnoses but to assist laboratory physicians in prioritizing reflex urine culture testing within laboratory workflows. Diagnostic responsibility remains entirely with the treating clinician, while the LDSS provides reproducible, standardized insights derived from LIS data.

Artificial intelligence (AI) and machine learning (ML) have gained increasing attention for developing predictive models in UTI diagnosis. Various algorithms—including Logistic Regression (LR), Random Forests (RFs), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and TabNet—have demonstrated robust performance using structured data such as urinalysis results, demographics, and clinical history[10–12]. Reported area under the receiver operating characteristic curve (AUROC) values commonly exceed 0.85, with some studies achieving 0.95 or higher in external validation cohorts[11,13].

Recent studies have highlighted the importance of model interpretability. By employing SHapley Additive exPlanations (SHAP), our LDSS not only ensures transparency but also facilitates clinical integration by illustrating the real-time contribution of each variable. Real-world implementations of ML-based LDSSs have shown reductions in unnecessary culture orders, accelerated treatment decisions, and improved antibiotic stewardship outcomes[12,14].

Despite these advances, challenges remain. Many predictive models are trained on single-center datasets and lack external validation, raising concerns about generalizability across institutions and diverse patient populations[13,15]. Additionally, variability in urinalysis platforms and clinical practice patterns may limit reproducibility and scalability.

Unlike existing tools, the proposed LDSS provides three distinct scoring systems tailored to different clinical priorities, ranging from high-sensitivity triage to specificity-focused decision-making. This flexibility promotes collaboration among biochemists, microbiologists, and clinicians while reducing diagnostic waste by minimizing unnecessary urine culture requests.

The aim of this study was to develop and externally validate a robust, interpretable ML-based LDSS to predict urine culture outcomes in patients with suspected UTIs. By standardizing reflective testing practices, the LDSS supports interdisciplinary decision-making, optimizes resource utilization, and ultimately contributes to rational antibiotic prescribing across healthcare settings.

Materials and Methods

Study Population/Subjects

This study was conducted at İzmir Tepecik Training and Research Hospital. Ethical approval was obtained from the University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital Non-Interventional Research Ethics Committee prior to study initiation (approval number: 2025/02-05, dated: 10.03.2025).

Eligible participants were adults aged ≥18 years who presented as inpatients or outpatients to the main hospital between January 1, 2014, and December 31, 2024, or to its affiliated hospital between January 1 and February 28, 2025. Inclusion criteria required patients to undergo their first urinalysis, complete blood count (CBC), and urine culture, ordered by a specialist physician based on clinical indication.

The study cohort included both culture-positive and culture-negative cases, capturing the full spectrum of patients for whom urine cultures were clinically indicated. Consequently, the dataset reflects real-world test-ordering practices rather than a biased subset of confirmed infections.

Patients were excluded if they had incomplete test results, missing sub-parameters, non-bacterial pathogens in their urine culture, delays exceeding one hour between urine sample collection and laboratory registration, delays exceeding 30 minutes for hemogram samples between phlebotomy and laboratory receipt, or a history of antibiotic treatment prior to testing.

CBC analyses were performed using UniCell DxH 800 analyzers (Beckman Coulter, Miami, FL, USA) from 2014 to 2020 and XN-2000 systems (Sysmex Corporation, Kobe, Japan) from 2020 onward. Urinalysis tests were conducted using fully automated analyzers across three periods: H-800 and FUS-200 systems (Dirui Industrial Co., Changchun, China) from 2014 to 2018; BT Uricell 1280–1600 (Bilimsel Products, İzmir, Türkiye) from 2018 to 2021; and U2610–U1600 (Zybio Corporation, Chongqing, China) from 2021 onward.

Midstream urine samples were collected in sterile containers simultaneously with urinalysis and processed according to standard microbiological procedures. Samples without detectable bacterial growth after 24 hours were incubated for an additional 24 hours; if no growth was observed, the result was reported as “no growth”.

Reagents and calibrators for urinalysis were obtained from authorized manufacturers and were certified and registered products. Quality control materials were sourced from Bio-Rad (California, USA). All results were reviewed and validated for accuracy and reliability by both a clinical biochemistry specialist and a clinical microbiology specialist.

Study Design

Patient identifiers were anonymized, and a dataset comprising age, sex, hemogram, urinalysis, and urine culture results from 55,385 patients (main hospital: 52,854; affiliated hospital: 2,531) was imported into Microsoft Excel 2021 (USA).

Symptom data were not included, as such information is not routinely recorded in LIS. In standard laboratory workflows, test orders are typically accompanied by preliminary diagnoses or ICD codes from the requesting physician, but detailed patient symptoms are not captured. Accordingly, the predictive model in this study was developed exclusively on structured laboratory data, aiming to forecast urine culture outcomes rather than to establish a clinical diagnosis of UTI.

After applying exclusion criteria, the final dataset included 49,720 patients, with an external validation cohort of 2,203 patients. The dataset was subsequently transferred to Python (version 3.13.1, USA) for ML analysis.

Following data cleaning, the main dataset was divided into training, internal test, and external test subsets using a 60:20:20 stratified sampling strategy based on the binary target variable, ensuring preservation of class distribution.

Patient flow throughout the study is depicted in Figure 1, in accordance with the Standards for Reporting Diagnostic Accuracy guidelines.

Data Preprocessing and Training of ML Algorithms

Patient data were initially exported from the LIS into Microsoft Excel. Hemogram values and flow cytometry parameters from urinalysis were used directly due to device standardization. Semi-quantitative dipstick results—reported by urinalysis analyzers as categorical values (e.g., “+,” “++,” “+/-,” “trace”)—were converted into numerical equivalents (e.g., “++” mapped to 2; “trace” standardized to 0.5) to ensure quantitative

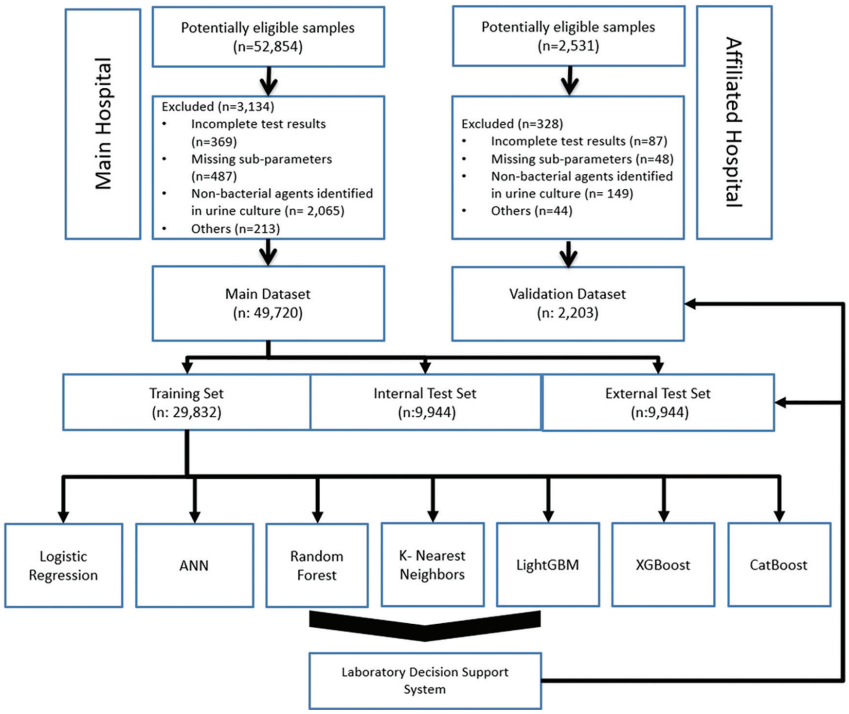

Figure 1. Standards for Reporting Diagnostic Accuracy flow diagram of study participants and urine culture testing.

consistency. Variables describing urine color and appearance were also recategorized by grouping similar classifications (e.g., light yellow to dark red; clear to very cloudy) to standardize the dataset.

Urine culture results were binarized as follows: samples with ≥10,000 colony-forming unit (CFU)/mL bacterial growth were defined as positive (label = 1), while samples with <10,000 CFU/mL, mixed flora, colonization, yeast, or no growth were classified as negative (label = 0).

The 10,000 CFU/mL threshold was selected based on recent evidence and the 2024 European Association of Urology guidelines, which acknowledge that lower colony counts ($\geq 10^3$–10^4 CFU/mL) may be clinically significant in symptomatic or catheterized patients[16]. Nelson et al.[17] demonstrated that these lower thresholds preserve diagnostic accuracy for symptomatic UTIs, supporting their use in reflective testing workflows. Additionally, Werneburg et al.[18] showed that urinalysis parameters reliably predict the absence of infection at this threshold, reinforcing its clinical validity. This definition also aligns with our institutional microbiology reporting standard for significant bacteriuria.

Yeast and colonization findings were labeled as negative (label = 0) based on established microbiological evidence and laboratory reporting standards. In urinary cultures, the presence of Candida species typically reflects colonization or contamination rather than true infection, even at colony counts exceeding 10^4–10^5 CFU/mL, unless accompanied by compatible clinical symptoms[19]. Classifying yeast as negative prevented false-positive propagation in the LDSS and improved the model's clinical specificity.

Similarly, cases labeled as "colonization"—including cultures with mixed flora or non-uropathogenic organisms—were considered negative. This approach aligns with standard microbiology practice, where such findings are reported as clinically non-significant. Although CLSI M100 (2025) does not define colony-count thresholds for colonization or candiduria, its terminology guided our categorization strategy. This interpretation reflects real-world laboratory workflows, ensuring that the LDSS mirrors standardized reporting logic and remains generalizable across institutions[20].

The cleaned dataset was transferred to Python for ML analysis. To enhance model robustness and address class imbalance, a stratified data partitioning scheme was applied, allocating 60% of samples to training and 20% each to internal and external testing. The dataset exhibited natural imbalance, with 22.4% culture-positive and 77.6% culture-negative samples. To mitigate majority-class bias, feature standardization and rebalancing strategies (class_weight='balanced') were applied uniformly across all classifiers.

As a preliminary check, a baseline LR model was trained and evaluated across all data splits. Receiver operating characteristic – area under the curve (ROC-AUC) scores (≈0.74, 0.73, 0.73 for training, internal, and external sets, respectively) and F1 scores (0.55, 0.54, 0.54) demonstrated consistent

generalization without evidence of overfitting or imbalance-driven inflation. The close alignment of these baseline metrics confirmed that stratified sampling preserved class proportions across all subsets (≈22.4% positive vs. 77.6% negative), ensuring reliable model development.

ML Model Selection and Development

The results confirmed that the methodological setup—including stratified sampling and proportional weighting—effectively mitigated class imbalance and provided a reliable foundation for model development. LR was used not as a primary model, but as a diagnostic tool to verify dataset integrity and the fairness of the training process[21].

Model development was performed in Python 3.13.1 using widely adopted libraries and workflows. Seven ML algorithms were evaluated for their suitability with the dataset and their potential effectiveness in predicting urine culture outcomes: RF, XGBoost, LightGBM, CatBoost, LR, Artificial Neural Network (ANN), and K-Nearest Neighbors (KNN).

Variables included in the analysis:

• Demographic: Age, sex

• Hemogram: White blood cell, neutrophil, lymphocyte, monocyte, eosinophil, basophil, hemoglobin (HGB)

• Urine Dipstick: Leukocyte esterase, nitrite, glucose, protein, pH, erythrocyte, bilirubin, urobilinogen, ketone

• Other Urinalysis: Urine color, urine density, appearance

• Flow Cytometry: Bacteria count, cylinder, yeast, urine leukocyte count

Data preprocessing, model training, evaluation, and visualization were conducted using open-source Python libraries:

• Data Processing and Analysis: pandas (v2.2.2), numpy (v2.0.2), optuna (v4.3.0)

• ML Model Development: scikit-learn (v1.6.1), XGBoost (v2.1.4), lightgbm (v4.5.0), catboost (v1.2.8), tensorflow (v2.10), keras (v2.10), torch (v2.6.0 + cu124)

• Model Evaluation and Visualization: matplotlib (v3.10), seaborn (v0.13.2), scipy.stats (v1.9), sklearn.metrics (v1.2), SHAP (v0.47)

Detailed hyperparameter optimization procedures, including search strategies and parameter configurations for each model, are provided in the Supplementary Table 1. Each model was retrained using the optimal hyperparameters identified during tuning. Final model evaluation was based on F1 and ROC-AUC scores derived from the internal test set.

Performance Evaluation

Performance evaluation was conducted using standard Python-based data science libraries. The modeling process was assessed comprehensively through internal cross-validation, hyperparameter tuning, and multiple performance metrics.

Classification Performance Metrics: Model discrimination and predictive capability were evaluated using:

• AUC-ROC

• Area under the precision-recall curve (AUC-PR)

• Sensitivity and Specificity

• Positive predictive value (PPV) and negative predictive value (NPV)

• Positive likelihood ratio (PLR) and negative likelihood ratio (NLR)

• F1 score

Model Interpretability Metrics: To enhance clinical transparency and foster trust in algorithmic decisions, interpretability was assessed using:

• Feature-Importance metrics

• SHAP graphs

This multidimensional evaluation approach balances predictive performance with explainability, providing a robust framework for forecasting urine culture outcomes based solely on laboratory and demographic data.

Development of the LDSS

The LDSS was built using the best-performing ML model identified during model selection. SHAP analysis was employed to select the ten most informative features, and a simplified model was retrained using only these variables. The reduced model maintained performance comparable to the full model, supporting its suitability for practical implementation.

Instead of the default probability threshold of 0.5, an optimized threshold based on Youden's J statistic was applied to improve sensitivity and minimize missed infections. Each selected feature was then converted into a binary indicator using individual cut-points derived from ROC analysis, enabling construction of a straightforward cumulative score.

Feature-importance values were normalized to derive clinically interpretable weights. Highly influential predictors received slightly higher weights, while moderately informative features were scaled conservatively to balance performance with interpretability. The final scoring system was recalibrated using internal data and externally evaluated, demonstrating preserved sensitivity and specificity. This streamlined, transparent design ensures that the LDSS is suitable for routine use within laboratory workflows.

Table 1. Baseline characteristics of the study population, including demographic, clinical, and laboratory variables.

Characteristics[a]	Unit	Main dataset (n = 49,720) mean ± SD	Training set (n = 29,832) mean ± SD	Internal test set (n = 9,944) mean ± SD	External test set (n = 9,944) mean ± SD	Validation set (n = 2,203) mean ± SD	p-value[b] (Main dataset vs. validation set)
Age		38.28 ± 26.85	38.07 ± 26.81	38.89 ± 26.99	38.29 ± 26.83	43.92 ± 28.53	<0.05
Male	Years	39.69 ± 28.20	39.33 ± 28.12	40.09 ± 28.39	40.37 ± 28.26	48.04 ± 28.38	<0.05
Female	Years	37.41 ± 25.96	37.29 ± 25.95	38.17 ± 26.07	37.03 ± 25.85	41.23 ± 28.33	<0.05
Gender							0.152
Male	n (%)	18,871 (38.0%)	11,358 (38.1%)	3,766 (37.9%)	3,747 (37.7%)	870 (39.5%)	
Female	n (%)	30,849 (62.0%)	18,474 (61.9%)	6,178 (62.1%)	6,197 (62.3%)	1,333 (60.5%)	
WBC	$\times 10^9$ cells/L	8.47 ± 4.63	8.5 ± 4.91	8.4 ± 3.86	8.45 ± 4.47	8.45 ± 3.48	0.795
Neutrophil	$\times 10^9$ cells/L	5.1 ± 3.4	5.11 ± 3.34	5.05 ± 3.13	5.11 ± 3.81	5.18 ± 3.14	0.244
Lymphocyte	$\times 10^9$ cells/L	2.45 ± 2.82	2.47 ± 3.27	2.42 ± 1.97	2.43 ± 1.99	2.36 ± 1.26	<0.05
Monocyte	$\times 10^9$ cells/L	0.68 ± 0.85	0.68 ± 1.01	0.68 ± 0.68	0.67 ± 0.37	0.67 ± 0.29	0.168
Eosinophil	$\times 10^9$ cells/L	0.2 ± 0.25	0.2 ± 0.25	0.2 ± 0.25	0.2 ± 0.24	0.19 ± 0.19	<0.05
Basophil	$\times 10^9$ cells/L	0.04 ± 0.06	0.04 ± 0.06	0.03 ± 0.05	0.04 ± 0.07	0.04 ± 0.03	1.000
HGB	g/dL	12.26 ± 1.91	12.26 ± 1.9	12.27 ± 1.92	12.27 ± 1.91	12.56 ± 1.98	<0.05
Bacteria count (urine)	/HPF	33.57 ± 124.45	33.55 ± 127.66	33.93 ± 120.89	33.24 ± 118.07	41.7 ± 157.49	<0.05
LYM (urine)	/HPF	53.46 ± 287.59	53.81 ± 288.38	52.32 ± 279.64	53.53 ± 293.02	64.28 ± 324.2	0.124
Yeast	/HPF	3.85 ± 133.83	5.04 ± 170.15	1.95 ± 36.1	2.2 ± 37.23	3.13 ± 55.43	0.587
Mucus	/HPF	11.32 ± 30.73	11.34 ± 30.69	10.97 ± 28.36	11.62 ± 33.03	22.14 ± 56.43	<0.05
Cylinder	/HPF	0.04 ± 0.22	0.04 ± 0.22	0.04 ± 0.23	0.05 ± 0.23	0 ± 0	<0.05
Density	-	1,016.98 ± 8.17	1,017.02 ± 8.14	1016.95 ± 8.23	1016.9 ± 8.22	1015.86 ± 7.19	<0.05
pH	-	5.9 ± 0.81	5.91 ± 0.82	5.89 ± 0.81	5.9 ± 0.81	6.05 ± 0.52	<0.05
Urine culture							<0.05
Positive	n	11,156 (22.4%)	6,694 (22.4%)	2,231 (22.4%)	2,231 (22.4%)	403 (18.3%)	1.000
Negative	n	38,564 (77.6%)	23,138 (77.6%)	7,713 (77.6%)	7,713 (77.6%)	1,800 (81.7%)	1.000

[a]Categorical variables were not included in this table. [b]Continuous variables were compared using Welch's t-test, and categorical variables were analyzed with Pearson's chi-square test. A p-value <0.05 was considered statistically significant.

Validation of the LDSS

An independent validation dataset, obtained from an affiliated hospital within the same healthcare network, was used to assess the generalizability and robustness of the LDSS through temporal validation. This temporally separated retrospective dataset was entirely independent of all model development phases, including training, feature selection, and score construction.

Performance of the reduced 10-variable RF model and the three derived scoring systems was evaluated within this separate clinical environment. Standard classification metrics were computed and compared with those from the original external test set, providing insight into the system's real-world applicability.

The validation strategy adheres to recommendations from the International Federation of Clinical Chemistry and Laboratory Medicine for evaluating diagnostic tools using independent datasets. This approach strengthens the clinical credibility of the LDSS by demonstrating reproducibility across diverse healthcare settings.

Statistical Analysis

Descriptive statistics are presented as means ± standard deviations (SDs) for continuous variables and as frequencies with percentages for categorical variables. Comparative analyses between the development and validation datasets were conducted using:

- Student's t-test for normally distributed continuous variables
- Welch's t-test for continuous variables with unequal variances or sample sizes
- Pearson's chi-square test for categorical variables
- Z-tests for proportions and McNemar's test for paired categorical outcomes, particularly for comparing model performance metrics across datasets

These statistical comparisons were used to evaluate diagnostic consistency and identify significant differences in classification outcomes, providing insight into the reproducibility and robustness of the LDSS across diverse clinical settings.

All p-values were two-sided, with statistical significance defined as $p < 0.05$. Analyses were conducted using Python 3.13 and its associated statistical packages.

Results

Dataset Description and Data Preprocessing

The analytical cohort comprised 51,923 patient encounters, including 49,720 records from the main institutional database and 2,203 from an affiliated tertiary center. The validation cohort was enriched with inpatients from high-acuity units, such as Palliative Care and Gynecologic Oncology, and was specifically used to assess the external validity of the LDSS.

The validation cohort demonstrated significantly higher age across all demographic strata (total: 43.92 vs. 38.28 years; males: 48.04 vs. 39.69; females: 41.23 vs. 37.41; all $p < 0.05$). Hematologic comparisons revealed statistically significant reductions in lymphocyte count (LYM) and eosinophil count, accompanied by a modest but significant increase in HGB levels ($p < 0.05$).

Among urinalysis variables, the validation group exhibited higher bacterial counts, increased mucus presence, and elevated pH levels, whereas urine specific gravity and cylinder counts were lower ($p < 0.05$ for all). No significant differences were observed in white blood cell (WBC), neutrophil, monocyte, or basophil counts, nor in leukocyte counts, yeast presence, or gender distribution (all $p > 0.05$). Although the proportion of urine culture-positive cases was numerically similar (22.4% vs. 18.3%), this difference reached statistical significance ($p < 0.05$), potentially reflecting distinct microbiologic or clinical characteristics in the validation population.

Overall, these findings indicate that while the two datasets are broadly comparable, the validation cohort exhibits distinct demographic and laboratory profiles, likely due to its inpatient composition. These differences should be considered when interpreting LDSS performance in more complex clinical settings. Detailed summary statistics and p-values for each variable are provided in Table 1.

Hyperparameter Tuning

Each ML model was trained and optimized to achieve optimal performance on our dataset. Final hyperparameter configurations, tailored to the structure of each algorithm, are summarized in the Supplementary Table 2.

Performance Metrics of ML Models

The performance of seven ML models was evaluated using both internal and external test datasets. Ensemble-based methods—RF, CatBoost, and XGBoost—consistently demonstrated high accuracy (≥0.929) and F1 scores (>0.83) across both datasets, highlighting their robustness for clinical prediction tasks.

On the external test set, RF outperformed all other models, achieving the highest ROC-AUC (0.956) and PR-AUC (0.907), indicating superior discrimination and precision-recall trade-off. CatBoost achieved the highest sensitivity (0.771) while maintaining balanced performance across other metrics.

KNN demonstrated exceptional specificity (0.988) and PPV (0.945) in the external set, making it particularly effective for ruling in cases. Conversely, LR, while computationally efficient, showed the lowest sensitivity and F1 scores, limiting its diagnostic utility.

Performance metrics from the external dataset closely mirrored those of the internal test set for all models, reinforcing their generalizability and stability. Comprehensive statistics for both datasets are provided in Table 2 and Figure 2.

Among all evaluated algorithms, RF exhibited the most consistent and highest overall performance, with an internal ROC-AUC of 0.952 (95% confidence interval [CI]: 0.948–0.956) and an external ROC-AUC of 0.956 [95% CI: 0.952–0.960], along with strong PR characteristics.

Given its superior accuracy, consistent generalizability, and interpretability, RF was selected as the core algorithm for integration into the LDSS. SHAP analysis was then performed on the final model to provide insight into the individual contribution of each feature to the predicted outcomes.

SHAP Analysis of the Optimal RF Model

Model interpretability was improved using SHAP, which quantifies the contribution of each feature to the predictions generated by the final RF model. As shown in Figure 3, the most influential features were

- Bacteria_Count (SHAP value: 0.061)
- Urine_Leu_Count (0.055)
- Nitrite (0.052)
- Age and Leukocyte Esterase (both 0.041)

Table 2. Classification performance metrics of the ML models, including accuracy, sensitivity, specificity, and AUC.

Model	Sensitivity	Specificity	PPV	NPV	Accuracy	F1 score	ROC-AUC	PR-AUC
Internal test set								
RF	0.758 (0.741–0.776)	0.985 (0.982–0.987)	0.934 (0.923–0.946)	0.934 (0.929–0.939)	0.934 (0.929–0.938)	0.838 (0.826–0.850)	0.952 (0.948–0.956)	0.897 (0.891–0.903)
XGBoost	0.768 (0.751–0.784)	0.976 (0.973–0.979)	0.902 (0.889–0.916)	0.936 (0.930–0.941)	0.929 (0.925–0.934)	0.830 (0.816–0.842)	0.930 (0.925–0.935)	0.861 (0.854–0.868)
LightGBM	0.681 (0.664–0.699)	0.972 (0.968–0.976)	0.876 (0.862–0.894)	0.913 (0.907–0.919)	0.907 (0.900–0.913)	0.766 (0.751–0.780)	0.916 (0.911–0.921)	0.825 (0.818–0.832)
CatBoost	0.764 (0.747–0.784)	0.980 (0.977–0.983)	0.918 (0.907–0.931)	0.935 (0.930–0.940)	0.932 (0.927–0.937)	0.834 (0.822–0.847)	0.930 (0.925–0.935)	0.861 (0.854–0.868)
LR	0.350 (0.330–0.370)	0.969 (0.965–0.973)	0.765 (0.738–0.791)	0.838 (0.830–0.846)	0.830 (0.823–0.837)	0.480 (0.459–0.501)	0.790 (0.782–0.798)	0.593 (0.583–0.603)
ANN (MLP)	0.561 (0.541–0.582)	0.943 (0.937–0.947)	0.738 (0.717–0.758)	0.881 (0.875–0.888)	0.857 (0.850–0.864)	0.637 (0.621–0.655)	0.844 (0.837–0.851)	0.698 (0.689–0.707)
KNN	0.723 (0.705–0.743)	0.984 (0.981–0.987)	0.929 (0.917–0.940)	0.925 (0.919–0.930)	0.925 (0.920–0.931)	0.813 (0.801–0.827)	0.947 (0.943–0.951)	0.903 (0.897–0.909)
RF (with top 10 variables)*	0.769 (0.761–0.777)	0.981 (0.979–0.984)	0.924 (0.919–0.930)	0.936 (0.931–0.941)	0.934 (0.929–0.939)	0.8397 (0.832–0.847)	0.947 (0.944–0.952)	0.890 (0.884–0.896)
External test set								
RF	0.76 (0.744–0.778)	0.987 (0.984–0.989)	0.943 (0.932–0.953)	0.935 (0.929–0.94)	0.936 (0.931–0.941)	0.842 (0.829–0.854)	0.956 (0.952–0.96)	0.907 (0.901–0.913)
XGBoost	0.767 (0.748–0.784)	0.980 (0.977–0.983)	0.917 (0.906–0.930)	0.936 (0.930–0.942)	0.932 (0.928–0.938)	0.836 (0.824–0.848)	0.932 (0.927–0.937)	0.877 (0.871–0.883)
LightGBM	0.686 (0.666–0.704)	0.976 (0.972–0.979)	0.892 (0.877–0.907)	0.915 (0.909–0.921)	0.911 (0.905–0.916)	0.776 (0.762–0.789)	0.919 (0.914–0.924)	0.840 (0.833–0.847)
CatBoost	0.771 (0.754–0.790)	0.982 (0.979–0.985)	0.924 (0.911–0.936)	0.936 (0.931–0.942)	0.934 (0.929–0.939)	0.840 (0.827–0.852)	0.929 (0.924–0.934)	0.875 (0.868–0.882)
LR	0.339 (0.321–0.358)	0.968 (0.964–0.972)	0.755 (0.725–0.781)	0.835 (0.828–0.842)	0.827 (0.819–0.834)	0.467 (0.445–0.487)	0.793 (0.785–0.801)	0.597 (0.587–0.607)
ANN (MLP)	0.565 (0.544–0.585)	0.937 (0.932–0.943)	0.722 (0.700–0.744)	0.881 (0.874–0.888)	0.854 (0.847–0.861)	0.634 (0.618–0.651)	0.846 (0.839–0.853)	0.707 (0.698–0.716)
KNN	0.719 (0.700–0.738)	0.988 (0.985–0.990)	0.945 (0.933–0.955)	0.924 (0.918–0.929)	0.927 (0.923–0.933)	0.817 (0.803–0.830)	0.947 (0.943–0.951)	0.905 (0.899–0.911)

*Reduced model including only the top 10 predictors selected by SHAP analysis: bacterial count in urine, urinary LYM, urinary nitrite test, patient age, leukocyte esterase activity in urine, HGB concentration, gender, LYM, urine density, and urinary erythrocyte count

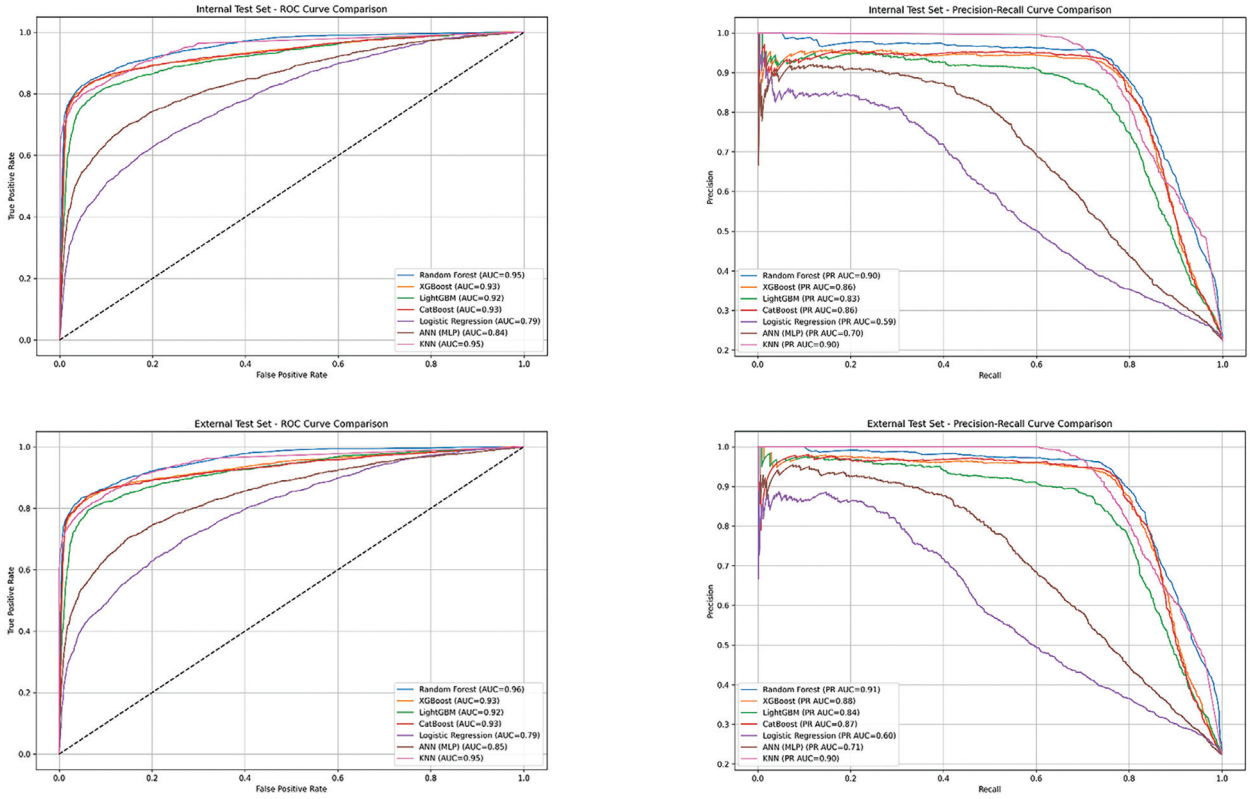

Figure 2. ROC and precision-recall PR curves illustrating the predictive performance of ML models.
PR, precision-recall.

These features correspond with well-established clinical markers of UTI, supporting the biological plausibility of the model.

Features with moderate importance included HGB, Gender, and LYM, with SHAP values ranging from 0.017 to 0.030. Features such as Bilirubin, Urobilinogen, and Ketone contributed minimally, each with SHAP values below 0.003.

Overall, the feature ranking confirms that the model primarily relies on clinically relevant variables, enhancing transparency and supporting its integration into laboratory decision-making.

Performance Metrics of the LDSS

A simplified RF model, built using the top 10 SHAP-derived features, maintained performance comparable to the full-feature model (ROC-AUC: 0.952 vs. 0.947; PR-AUC: 0.897 vs. 0.890), supporting its suitability for clinical implementation (Table 2). Based on these variables, three complementary scoring systems were developed to address distinct operational needs within laboratory workflows (Table 3):

• Model-Prioritized Score: Retains the behavior of the original ML model by assigning weights directly from normalized SHAP values. This version is ideal for institutions seeking high overall discrimination while remaining faithful to the underlying algorithm.

• Dual-Optimization Score: Adjusts feature weights to balance sensitivity and specificity, as reflected in stable metrics across both test datasets (Table 4, Figure 4). This score is intended for laboratories aiming to minimize both missed infections and unnecessary cultures.

• SAFE-Score: Optimized for high sensitivity and NPV, this score is suitable for safety-critical settings where missing true infections is unacceptable—such as high-acuity units, elderly populations, or immunocompromised patients. Its higher sensitivity comes at the expense of specificity, highlighting the trade-off between diagnostic conservatism and resource utilization.

Across all scoring systems, sensitivity remained consistent in external and independent validation cohorts, while specificity varied according to prioritization strategy (Table 4). Together, these tools provide laboratories with flexible options that can be tailored to local clinical priorities, test-ordering practices, and antimicrobial stewardship goals (Figure 4).

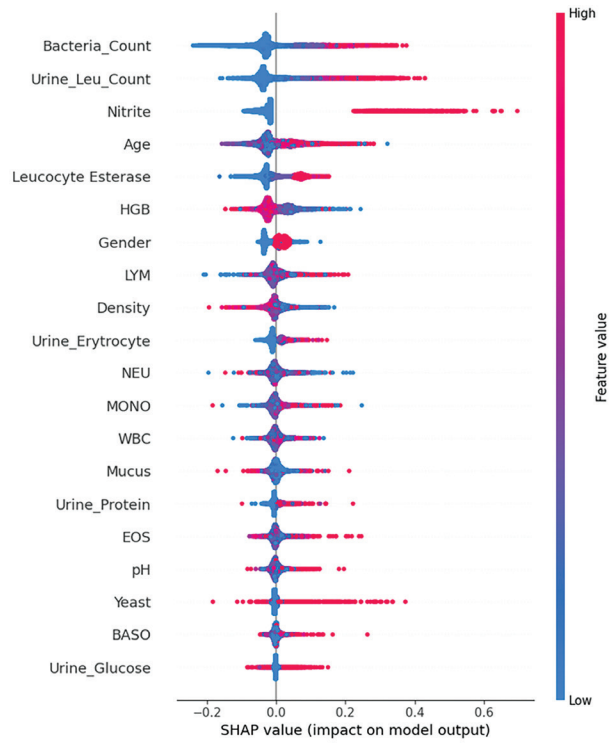

Figure 3. SHAP summary plot showing variable importance in the RF model.

Table 3. Confusion matrix–derived performance metrics of the ML models, including sensitivity, specificity, PPV, and NPV.

Feature	Threshold binarization	Normalized SHAP value	Model-Prioritized Score System[1]	Dual-Optimization Score[2]	SAFE-Score System (sensitive assessment for exclusion)[3]	Scientific justification
Bacteria count	>20, n	0.175	0.20	0.32	0.89	Major diagnostic marker for infection; emphasized clinically.
Urine LYM	>25, n	0.157	0.18	0.22	0.05	Strongly correlates with infection; slightly boosted for sensitivity.
Nitrite	= 1	0.147	0.17	0.15	0.77	Positive nitrite is a direct indicator of gram-negative bacterial activity.
Age	≥65 years	0.118	0.15	0.23	0.42	Increased risk in elderly population (>65 years).
Leucocyte esterase	>0	0.116	0.14	0.13	0.82	Biochemical indicator of leukocytes; moderate importance.
HGB	<12	0.085	0.10	0.12	0.71	Low HGB levels linked to increased infection susceptibility.
Gender	= 1 (Female)	0.062	0.08	0.04	0.06	Higher infection prevalence anatomically in females.
LYM	<1.5	0.051	0.06	0.1	0.65	Low lymphocyte count indicates immunosuppression risk.
Density	>1020	0.048	0.05	0.09	0.03	Higher urine density occasionally correlates with infection.
Urine erythrocyte	>0	0.047	0.05	0.1	0.77	Presence may suggest urinary tract pathology but less specific.

[1]The first system was developed using model-derived, data-driven thresholds and weighting. [2]The second system was designed to optimize both sensitivity and specificity, achieving balanced classification performance. [3]The third system prioritized minimizing false negatives, emphasizing maximum sensitivity and NPV.

Table 4. Performance metrics of the LDSS evaluated using both external test and validation datasets.

A. Results from the external test set.

Method	Sensitivity (%)	Specificity (%)	PPV	NPV	PLR	NLR	Accuracy	F1 score	ROC-AUC	PR-AUC
Model-Prioritized Score System[1]	55.94 (53.87–57.99)	85.83 (85.03–86.59)	53.31 (51.29–55.32)	87.07 (86.30–87.81)	3.95 (3.69–4.22)	0.51 (0.49–0.54)	79.1 (78.31–79.91)	54.59 (52.57–56.60)	70.88 (67.58–74.28)	54.62 (50.71–57.71)
Dual-Optimization Score System[2]	64.77 (62.76–66.72)	76.62% (75.67–77.55)	44.49 (42.79–46.20)	88.26 (87.47–89.01)	2.77 (2.47–3.07)	0.46 (0.44–0.49)	73.96 (73.09–74.82)	52.75 (51.03–54.46)	70.70 (68.70–72.70)	54.63 (52.72–56.72)
SAFE-Score System[3]	95.34 (94.38–96.14)	20.29% (19.41–21.20)	25.70 (24.77–26.66)	93.77 (92.51–94.83)	1.20 (1.03–1.136)	0.23 (0.21–0.25)	37.13 (36.18–38.08)	40.49 (39.44–41.55)	57.81 (57.80–57.83)	60.52 (55.55–65.58)

B. Results from the validation test set.

Method	Sensitivity	Specificity	PPV	NPV	PLR	NLR	Accuracy	F1 score	ROC-AUC	PR-AUC
Model-Prioritized Score System[4]	57.95 (53.00–62.78)	84.78 (83.04–86.41)	46.47 (43.09–49.89)	89.87 (88.74–90.85)	3.81 (3.32–4.37)	0.50 (0.44–0.56)	79.80 (78.06–81.46)	51.51 (47.30–55.37)	71.31 (68.60–73.94)	34.80 (31.05–38.61)
Dual-Optimization Score System[5]	66.50 (61.70–71.07)	76.48 (74.44–78.42)	39.19 (36.65–41.80)	90.92 (89.71–92.00)	2.83 (2.54–3.15)	0.44 (0.38–0.50)	74.63 (72.75–76.43)	49.36 (45.86–53.01)	71.40 (68.82–73.94)	32.35 (29.00–35.66)
SAFE-Score System[6]	94.87 (92.26–96.79)	24.30 (22.33–26.36)	22.22 (21.63–22.83)	95.40 (93.14–96.95)	1.25 (1.21–1.30)	0.21 (0.14–0.32)	37.40 (35.38–39.46)	36.08 (33.46–38.63)	59.58 (58.14–61.03)	22.03 (20.05–24.01)

[1]TP = 1,248; TN = 6,620; FP = 1,093; FN = 983. [2]TP = 1,445; TN = 5,910; FP = 1,803; FN = 786. [3]TP = 2,127; TN = 1,565; FP = 6,148; FN = 104.
[4]TP = 237; TN = 1,521; FP = 273; FN = 172. [5]TP = 272; TN = 1,372; FP = 422; FN = 137. [6]TP = 388; TN = 436; FP = 1,358; FN = 21.

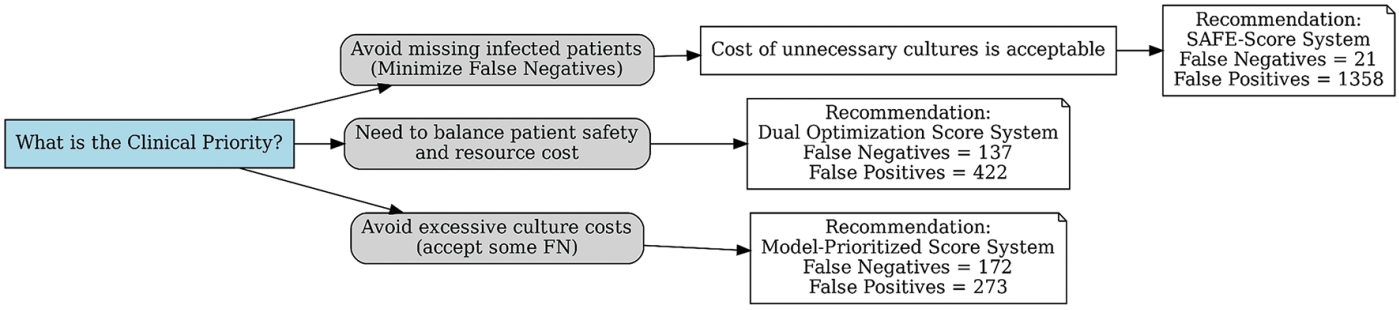

Figure 4. LDSS workflow illustrating selection criteria based on diagnostic accuracy and operational priorities.

Discussion

ML-based approaches offer substantial potential for the early diagnosis of UTIs. With the rising prevalence of antibiotic resistance, reducing unnecessary antibiotic use has become increasingly critical. Recent studies demonstrate that ML models improve diagnostic accuracy by integrating clinical symptoms, medical history, and urinary biomarkers, rather than relying solely on culture results[22].

Moreover, AI–driven decision-support systems can reduce diagnostic workload in hospitals, although their clinical validation remains limited[15]. Urinary biomarkers, such as nitrite and leukocyte esterase, exhibit high sensitivity for UTI diagnosis, yet their integration into ML models is essential to mitigate false-positive results[23]. AI-assisted methodologies are expected to be particularly beneficial for early detection of recurrent UTIs and multidrug-resistant pathogens, potentially improving patient outcomes and guiding more precise therapeutic interventions[23,24].

In this study, we evaluated the performance of multiple ML models in predicting urine culture outcomes and assessed their clinical applicability using explainable AI (XAI) techniques. Validation on a demographically and clinically distinct inpatient cohort further demonstrated the robustness and real-world adaptability of the LDSS. The incorporation of XAI enhanced interpretability, providing insight into the

decision-making process and supporting potential integration in complex healthcare settings.

The LDSS was developed using all physician-ordered urine culture requests, including both culture-positive and culture-negative cases. Consequently, the dataset reflects the complete real-world distribution of suspected UTIs encountered in laboratory practice, enabling the model to learn discriminative patterns for both infection and non-infection samples. Importantly, the LDSS functions solely as a laboratory-level decision-support tool rather than a diagnostic system. Its predictions are limited to variables available in the LIS and are intended to complement, not replace, physicians' diagnostic judgment.

Gender and Age-Related UTI Incidence

In our study, UTIs were significantly more common in female patients than in males. This finding aligns with existing literature and reinforces the well-established notion that women are more susceptible to UTIs due to urogenital anatomy, hormonal fluctuations, and lifestyle factors. Schmiemann et al.[1] reported that UTI incidence in women is four to five times higher than in men. Similarly, Hooton et al.[25] identified a higher risk in women attributable to a shorter urethra and variability in periurethral microbial flora. Additional risk factors include age, postmenopausal hormonal changes, and a history of recurrent infections.

Age also emerged as a critical determinant, with UTI incidence progressively increasing—particularly among women aged 65 years and older. While Foxman et al.[26] reported peak incidence in women aged 15–29, with a secondary rise in postmenopausal groups, and Møller et al.[11] linked estrogen depletion after age 50 to heightened susceptibility, our study identified older age (≥65 years) as an independent risk factor for positive urine culture in the LDSS model. This finding underscores the importance of incorporating age as a predictive variable and reflects the growing burden of UTIs in elderly populations.

Performance of ML Models

The predictive performance of the models developed in this study is consistent with, and in several cases surpasses, previously reported ML approaches for UTI prediction. Among the algorithms tested, ensemble-based models—particularly RF and CatBoost—demonstrated consistently high accuracy, balanced sensitivity and specificity, and favorable F1 scores. Compared to prior models reported by de Vries et al.[27] and Flores et al.[2], our RF model showed superior performance across multiple evaluation metrics. Likewise, our CatBoost implementation outperformed the model described by Mancini et al.[13], which exhibited lower AUC and F1 values in a comparable clinical context.

Tree-based gradient boosting methods, such as XGBoost and LightGBM, also performed robustly and yielded results similar to high-performing models developed by Choi et al.[5] and Lin et al.[28], indicating strong generalizability across diverse patient populations. In studies by Dhanda et al.[29] and Taylor et al.[30], RF and XGBoost models similarly demonstrated superior discriminatory capacity, achieving AUC-ROC values of 0.85 and 0.90, respectively.

The KNN model achieved precision metrics comparable to prior studies; however, its limited interpretability may constrain clinical adoption[7]. Conversely, LR, while highly interpretable, exhibited lower sensitivity and F1 scores—consistent with Ramgopal et al.[10], where the model tended to overpredict positive cases, reducing precision. ANN (MLP) models, though commonly employed in UTI prediction studies, demonstrated moderate performance in our dataset, slightly below previously reported benchmarks[2].

Overall, these results reinforce the value of ensemble ML methods in the context of a LDSS for UTI prediction. They offer high predictive accuracy and consistent performance across internal and external validation cohorts, supporting their applicability in real-world clinical settings.

Several studies have investigated ML–based urine culture prediction, varying in complexity and generalizability. Seheult et al.[31] developed a decision-tree algorithm across multiple institutions to identify urinalysis predictors of culture positivity, reporting ROC-AUC values of approximately 0.78–0.79; however, their study lacked external validation and interpretability assessment. By comparison, our model achieved higher discrimination during development (ROC-AUC = 0.94–0.96) under cross-validation. Following conversion into a simplified score-based LDSS, real-world performance remained consistent (ROC-AUC ≈ 0.70–0.72; F1 ≈ 0.50–0.55). This decline reflects the expected trade-off between model complexity and clinical interpretability, as the LDSS was designed for practical integration into LIS rather than maximizing algorithmic precision[31].

Sergounioti et al.[32] applied ensemble classifiers, including RF and XGBoost, to real-world laboratory data, achieving AUROC values of 0.79–0.82. However, their models combined clinical and laboratory parameters and lacked transparent feature-importance analysis. In contrast, our LDSS relied solely on structured laboratory data, achieved comparable discrimination (0.70–0.72), and preserved interpretability and reproducibility through rule-based score calibration via the Model-Prioritized and Dual-Optimization systems.

Sheele et al.[33] investigated bacteriuria prediction in an emergency-department cohort using mixed clinical–laboratory features, yielding AUC-ROC values of 0.86–0.93 depending

on the CFU/mL threshold. While their results were strong in a high-acuity population, our laboratory-only LDSS achieved comparable sensitivity (up to 95%) in routine diagnostic settings, highlighting its potential as a front-end decision-support tool for reflex culture testing.

Collectively, previous studies demonstrated the feasibility of ML-assisted urine culture prediction but often emphasized algorithmic performance over interpretability and clinical applicability. The present study addresses this gap by establishing a transparent, externally validated, and operational LDSS framework that maintains clinically acceptable performance while remaining fully interpretable and implementable within routine laboratory workflows.

Explainability and Feature Importance

SHAP-based feature-importance analysis in our study revealed a variable ranking that aligns with and extends existing literature. The most influential predictors were bacterial count, urine leukocyte count, nitrite, age, and leukocyte esterase. These findings are consistent with the meta-analysis by Devillé et al.[8], which reported that combining nitrite and leukocyte esterase yielded a sensitivity of 88% and specificity of 98% for UTI diagnosis. Similarly, Lachs et al.[34] demonstrated that integrating these parameters with clinical symptoms significantly improves diagnostic accuracy.

Notably, our model also identified HGB levels, sex, and LYMs as important features with relatively high SHAP values, suggesting sensitivity to broader systemic or demographic factors that may influence infection risk. This aligns with Zhao et al.[35], who reported age and sex among the top predictors in a SHAP-based post-urostomy UTI risk model, and Wang et al.[36], who found that systemic inflammatory markers and age were highly important in predicting post-surgical UTIs.

The predominance of microscopic urinalysis variables—particularly bacterial and leukocyte counts—over clinical or demographic features underscores the model's responsiveness to diagnostic biomarkers. This differentiates our approach from models such as Lee et al.[37], which focused on predicting antimicrobial resistance patterns but also leveraged SHAP analysis for interpretability.

Recent literature highlights the limitations of reflexive urine culture testing in the absence of clinical context. Munigala et al.[38] and others have shown that reflex algorithms triggered by markers like leukocyte esterase or nitrite may reduce test volume but compromise diagnostic precision when symptom data are unavailable. Fakih et al.[39] similarly argue that urinalysis alone is insufficient for accurate UTI diagnosis in asymptomatic patients, risking overdiagnosis and overtreatment.

Our study addresses the diagnostic gap through a reflective developed solely using structured laboratory data. Because symptom data are typically absent from LIS, the LDSS optimizes culture utilization within real-world laboratory constraints. Rather than functioning as an autonomous decision-maker or reflex trigger, the system serves as a reflective tool, providing SHAP-based analytical insights to support laboratory physicians' expert interpretation.

This reflective framework promotes standardized testing and interdisciplinary consultation. In equivocal cases, LDSS outputs can facilitate dialogue between laboratory and clinical teams, helping reconcile test reduction with diagnostic safety. Such an approach advances rational microbiological testing and provides a scalable model for clinician-laboratory collaboration[40].

The LDSS demonstrated robust predictive performance across internal and external datasets, supporting its seamless integration into routine laboratory workflows and reflective testing processes. The system is designed not to replace culture testing but to prioritize it based on evidence-driven probability, maintaining diagnostic stewardship.

To enhance accessibility for readers from diverse clinical and laboratory backgrounds, this study emphasizes the translational relevance of the LDSS over computational complexity. Its explainable design—supported by SHAP analysis and simplified scoring systems—enables non-technical users to interpret outputs transparently. While technical details were included to ensure methodological transparency and reproducibility, the interpretability of the system fosters trust, usability, and interdisciplinary communication between laboratory specialists and treating physicians. By promoting shared understanding of data-driven reasoning, the LDSS supports faster decision-making, improved test stewardship, and enhanced integration of laboratory insights into clinical workflows.

LDSS

Although symptom data were unavailable in the laboratory dataset, the LDSS was intentionally designed to function within the routine workflow of laboratory medicine, where test requests are frequently submitted without accompanying clinical narratives. By aligning the model with real-world laboratory constraints, the LDSS remains applicable and scalable across diverse clinical settings.

To improve interpretability and minimize unnecessary complexity, feature selection was applied to reduce the number of input variables. Prior studies have consistently demonstrated that parsimonious models are better suited for clinical implementation, as they are easier to interpret and maintain,

while preserving acceptable predictive performance[41,42]. Accordingly, subsequent model development was restricted to ten key parameters that did not result in a statistically or clinically meaningful decline in performance. This strategy ensured an optimal balance between model simplicity and predictive accuracy.

Several published studies have similarly developed LDSS frameworks based on urine culture data, including those reported by de Vries et al.[27], Dhanda et al.[29], Del Ben et al.[43], and Flores et al.[2] Among these, Del Ben et al.[43] employed a decision-tree-based approach, whereas the remaining studies selected RF as the primary algorithm. The LDSS developed by de Vries and colleagues demonstrated performance metrics comparable to those observed in the present study, with AUC-ROC values ranging from 0.70 to 0.80. Although their model achieved a higher PPV, its NPV was lower than that of our model, highlighting differences in clinical trade-offs between false-positive and false-negative predictions.

Notably, Dhanda et al.[29] and Flores et al.[2] implemented scoring systems that stratified patients into high- and low-risk groups, an approach that is conceptually aligned with the strategy adopted in the present study. Across key performance metrics, the predictive accuracy of their models was broadly comparable to that of our system.

What distinguishes our LDSS is the integration of three distinct predictive models within a unified decision-making framework. To our knowledge, this is the first study to report the implementation of such a multi-model structure for UTI prediction. This design enables clinicians and laboratory physicians to select among alternative strategies according to specific clinical priorities, such as maximizing case detection or minimizing unnecessary diagnostic testing.

Although the SAFE-Score achieved excellent sensitivity, its specificity was limited (approximately 20%), a trade-off that may raise concerns regarding potential overtesting. Importantly, the LDSS was intentionally designed to accommodate this limitation by offering three complementary scoring strategies, each reflecting a distinct clinical philosophy. These include prioritization of patient safety (SAFE-Score), balanced diagnostic performance (Dual Optimization), and strict adherence to model-derived predictions (Model-Prioritized). Rather than enforcing a one-size-fits-all solution, the LDSS functions as a flexible framework that facilitates consensus-based decision-making, allowing institutions to align model selection with local clinical expectations and operational priorities.

Crucially, the proposed system is not static. By continuously incorporating real-world data—particularly cases in which algorithmic recommendations are compared with expert laboratory physician judgments—the LDSS can be iteratively retrained and refined. As additional large-scale datasets are accumulated over time, improvements in specificity and overall diagnostic balance are anticipated, reflecting the inherent capacity of ML models to evolve with expanding data inputs. In this respect, the LDSS serves not only as an immediate decision-support tool but also as a scalable platform for continuous learning and performance optimization.

Within the Turkish healthcare context, reflective testing has not yet been systematically implemented. Nevertheless, the LDSS offers a structured and standardized framework that may facilitate its adoption, reduce inappropriate urine culture requests, and support antimicrobial stewardship initiatives. Moreover, the Ministry of Health of Türkiye has recently introduced a “Rational Laboratory Utilization” directive that explicitly promotes reflex and reflective testing practices [44]. This regulatory emphasis is expected to accelerate the integration of reflective testing into routine laboratory workflows, highlighting the timeliness and practical relevance of the proposed system.

Finally, the LDSS was designed for seamless integration into routine clinical practice through Microsoft Excel, a widely available and familiar platform in most healthcare settings. All three predictive models are embedded within a single interface and generate concurrent outputs, enabling direct comparison and transparent interpretation at the point of use.

Due to time constraints, the validation cohort was relatively small. Nevertheless, implementation of the LDSS within our hospital’s central laboratory is planned, where it will be deployed to support real-time microbiological decision-making. This implementation will allow prospective validation of the system within routine laboratory workflows, evaluation of its diagnostic impact, and quantification of downstream outcomes, including reductions in unnecessary urine cultures, shorter turnaround times, and improved antibiotic stewardship. In addition, future multicenter studies across diverse healthcare systems are planned, incorporating structured clinical variables such as symptomatology, comorbidities, and medication history to further enhance the model’s generalizability and clinical relevance.

Study Limitations

Although this study leveraged a large dataset and included external validation, several limitations should be acknowledged. First, all data were derived from a single healthcare network, which may limit generalizability to institutions with different patient populations, laboratory infrastructures, or clinical workflows. Second, the retrospective study design precluded assessment of the LDSS in real-time clinical decision-making; prospective implementation studies are therefore required to determine its effects on clinical practice and patient outcomes.

Third, the model relied exclusively on structured laboratory data and did not incorporate patient symptoms, comorbidities, medication history, or clinical notes—factors known to influence UTI risk assessment and antibiotic prescribing. In routine clinical care, integration of such information is primarily the responsibility of the treating physician, who orders diagnostic tests based on patient history, clinical presentation, and prevailing guidelines. In contrast, laboratory physicians are tasked with processing submitted specimens according to standardized pre-analytical and analytical protocols. Although pre-preanalytical factors, such as appropriate test selection, are important, these data are rarely available to LIS in a structured, analyzable format. Consequently, most LIS environments contain only coded test orders and limited demographic information, without access to patient symptomatology or detailed clinical context.

Within these real-world constraints, the LDSS was designed not as a replacement for clinical judgment but as a complementary, interpretable decision-support tool that standardizes reflective testing and promotes communication between laboratory and clinical teams. Accordingly, the system functions as a laboratory-based reflex testing prioritization tool rather than as a diagnostic or therapeutic decision-making platform.

Fourth, despite robust performance in both internal and external test sets, the relatively small independent validation cohort—enriched for high-acuity inpatients—may introduce spectrum bias and lead to overestimation of sensitivity in complex clinical populations. Fifth, although the conventional definition of significant bacteriuria is $\geq 10^5$ CFU/mL, this study adopted a $\geq 10^4$ CFU/mL threshold based on emerging clinical evidence and institutional practice. Future investigations should evaluate the effects of alternative thresholds on model calibration and performance across different clinical settings.

Sixth, scoring weights and feature thresholds were calibrated using a fixed probability cutoff and Youden's index derived from the present dataset. Optimal thresholds may vary across institutions and will require local adjustment to maintain the desired balance between sensitivity and specificity. Finally, while SHAP values were employed to enhance model interpretability, clinician acceptance, usability, and integration into routine workflows were not formally assessed. Future implementation studies are therefore essential to evaluate user engagement, potential alert fatigue, and cost-effectiveness prior to widespread clinical deployment.

Conclusion

We developed and preliminarily validated an interpretable, multi-model LDSS designed to improve the efficiency of urine culture utilization. By integrating ensemble ML approaches with SHAP-based interpretability, the system demonstrated strong discriminatory performance while offering flexible scoring strategies that prioritize sensitivity, specificity, or an optimized balance between the two. The LDSS has the potential to reduce unnecessary urine cultures, support antimicrobial stewardship efforts, and promote standardized, evidence-based laboratory decision-making.

Future work will focus on prospective, real-world implementation across diverse clinical settings. Planned enhancements include integration with electronic health record–derived clinical data, local calibration of decision thresholds, and systematic evaluation of clinical impact, user adoption, and cost-effectiveness. These steps are critical for translating this early-stage model into a scalable and clinically actionable decision-support tool.

Ethics Committee Approval: Ethical approval was obtained from the University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital Non-Interventional Research Ethics Committee prior to study initiation (approval number: 2025/02-05, dated: 10.03.2025).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: F.D., İ.A., Concept: F.D., İ.A., Design: F.D., M.A., İ.A., Data Collection or Processing: F.D., M.A., İ.A., Analysis or Interpretation: F.D., A.D., Literature Search: F.D., M.A., A.D., Writing: F.D., A.D.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Schmiemann G, Kniehl E, Gebhardt K, Matejczyk MM, Hummers-Pradier E. The diagnosis of urinary tract infection: a systematic review. Dtsch Arztebl Int. 2010;107(21):361–7.
2. Flores E, Martínez-Racaj L, Blasco Á, Diaz E, Esteban P, López-Garrigós M, Salinas M. A step forward in the diagnosis of urinary tract infections: from machine learning to clinical practice. Comput Struct Biotechnol J. 2024;24:533–41.
3. O'Brien M, Marijam A, Mitrani-Gold FS, Terry L, Taylor-Stokes G, Joshi AV. Unmet needs in uncomplicated urinary tract infection in the United States and Germany: a physician survey. BMC Infect Dis. 2023;23:281.
4. Ozkan IA, Koklu M, Sert IU. Diagnosis of urinary tract infection based on artificial intelligence methods. Comput Methods Programs Biomed. 2018;166:51–9.
5. Choi MH, Kim D, Park Y, Jeong SH. Development and validation of artificial intelligence models to predict urinary tract infections and secondary bloodstream infections in adult patients. J Infect Public Health. 2024;17:10–7.

6. World Health Organization. Global Antimicrobial Resistance and Use Surveillance System (GLASS) report 2022. Geneva: World Health Organization; 2022.
7. Dedeene L, Van Elslande J, Dewitte J, Martens G, De Laere E, De Jaeger P, De Smet D. An artificial intelligence-driven support tool for prediction of urine culture test results. Clin Chim Acta. 2024;562:119854.
8. Devillé WL, Yzermans JC, van Duijn NP, Bezemer PD, van der Windt DA, Bouter LM. The urine dipstick test useful to rule out infections: a meta-analysis of the accuracy. BMC Urol. 2004;4:4.
9. Verboeket-van de Venne WPHG, Aakre KM, Watine J, Oosterhuis WP. Reflective testing: adding value to laboratory testing. Clin Chem Lab Med. 2012;50:1249–52.
10. Ramgopal S, Horvat CM, Yanamala N, Alpern ER. Machine learning to predict serious bacterial infections in young febrile infants. Pediatrics. 2020;146(3):e20194096.
11. Møller JK, Sørensen M, Hardahl C. Prediction of risk of acquiring urinary tract infection during hospital stay based on machine-learning: a retrospective cohort study. PLoS One. 2021;16:e0248636.
12. Herter WE, Khuc J, Cinà G, Knottnerus BJ, Numans ME, Wiewel MA, Bonten TN, de Bruin DP, van Esch T, Chavannes NH, Verheij RA. Impact of a machine learning–based decision support system for urinary tract infections: prospective observational study in 36 primary care practices. JMIR Med Inform. 2022;10:e27795.
13. Mancini A, Vito L, Marcelli E, Piangerelli M, De Leone R, Pucciarelli S, Merelli E. Machine learning models predicting multidrug resistant urinary tract infections using "DsaaS." BMC Bioinformatics. 2020;21:347.
14. Shapiro Ben David S, Romano R, Rahamim-Cohen D, Azuri J, Greenfeld S, Gedassi B, Lerner U. AI driven decision support reduces antibiotic mismatches and inappropriate use in outpatient urinary tract infections. NPJ Digit Med. 2025;8:61.
15. Burton RJ, Albur M, Eberl M, Cuff SM. Using artificial intelligence to reduce diagnostic workload without compromising detection of urinary tract infections. BMC Med Inform Decis Mak. 2019;19:171.
16. Kranz J, Bartoletti R, Bruyère F, Cai T, Geerlings S, Köves B, Schubert S, Pilatz A, Veeratterapillay R, Wagenlehner FME, Bausch K, Devlies W, Horváth J, Leitner L, Mantica G, Mezei T, Smith EJ, Bonkat G. European Association of Urology Guidelines on urological infections: summary of the 2024 guidelines. Eur Urol. 2024;86(1):27–41.
17. Nelson Z, Aslan AT, Beahm NP, Blyth M, Cappiello M, Casaus D, Dominguez F, Egbert S, Hanretty A, Khadem T, Olney K, Abdul-Azim A, Aggrey G, Anderson DT, Barosa M, Bosco M, Chahine EB, Chowdhury S, Christensen A, de Lima Corvino D, Fitzpatrick M, Fleece M, Footer B, Fox E, Ghanem B, Hamilton F, Hayes J, Jegorovic B, Jent P, Jimenez-Juarez RN, Joseph A, Kang M, Kludjian G, Kurz S, Lee RA, Lee TC, Li T, Maraolo AE, Maximos M, McDonald EG, Mehta D, Moore WJ, Nguyen CT, Papan C, Ravindra A, Spellberg B, Taylor R, Thumann A, Tong SYC, Veve M, Wilson J, Yassin A, Zafonte V, Mena Lora AJ. Guidelines for the prevention, diagnosis, and management of urinary tract infections in pediatrics and adults: a WikiGuidelines Group consensus statement. JAMA Netw Open. 2024;7(11):e2444495. Erratum in: JAMA Netw Open. 2024;7(12):e2453497.
18. Werneburg GT, Lewis KC, Vasavada SP, Wood HM, Goldman HB, Shoskes DA, Li I, Rhoads DD. Urinalysis exhibits excellent predictive capacity for the absence of urinary tract infection. Urology. 2023;175:101–6.
19. Gharaghani M, Taghipour S, Halvaeezadeh M, Mahmoudabadi AZ. Candiduria: a review article with specific data from Iran. Turk J Urol. 2018;44(6):445–52.
20. Clinical and Laboratory Standards Institute (CLSI). Performance standards for antimicrobial susceptibility testing. 35th ed. CLSI document M100. Wayne (PA): CLSI; 2025.
21. Hosmer DW, Lemeshow S, Sturdivant RX. Applied logistic regression. 3rd ed. Hoboken (NJ): Wiley; 2013.
22. Yelin I, Snitser O, Novich G, Katz R, Tal O, Parizade M, Chodick G, Koren G, Shalev V, Kishony R. Personal clinical history predicts antibiotic resistance of urinary tract infections. Nat Med. 2019;25:1143–52.
23. Gadalla AAH, Friberg IM, Kift-Morgan A, Zhang J, Eberl M, Topley N, Weeks I, Cuff S, Wootton M, Gal M, Parekh G, Davis P, Gregory C, Hood K, Hughes K, Butler C, Francis NA. Identification of clinical and urine biomarkers for uncomplicated urinary tract infection using machine learning algorithms. Sci Rep. 2019;9:19694.
24. Werneburg GT, Rhoads DD, Milinovich A, McSweeney S, Knorr J, Mourany L, Zajichek A, Goldman HB, Haber GP, Vasavada SP. External validation of predictive models for antibiotic susceptibility of urine culture. BJU Int. 2025.
25. Hooton TM. Uncomplicated urinary tract infection. N Engl J Med. 2012;366:1028–37.
26. Foxman B. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. Am J Med. 2002;113:5–13.
27. de Vries S, ten Doesschate T, Totté JEE, Heutz JW, Loeffen YGT, Oosterheert JJ, Thierens D, Boel E. A semi-supervised decision support system to facilitate antibiotic stewardship for urinary tract infections. Comput Biol Med. 2022;146:105621.
28. Lin T-H, Chung H-Y, Jian M-J, Chang C-K, Lin H-H, Yu C-M, Perng C-L, Chang F-Y, Chen C-W, Chiu C-H, Shang H-S. Artificial intelligence-clinical decision support system for enhanced infectious disease management: accelerating ceftazidime-avibactam resistance detection in Klebsiella pneumoniae. J Infect Public Health. 2024;17:102541.
29. Dhanda G, Asham M, Shanks D, O'Malley N, Hake J, Satyan MT, Yedlinsky NT, Parente DJ. Adaptation and external validation of pathogenic urine culture prediction in primary care using machine learning. Ann Fam Med. 2023;21:11–8.
30. Taylor RA, Moore CL, Cheung K-H, Brandt C. Predicting urinary tract infections in the emergency department with machine learning. PLoS One. 2018;13:e0194085.
31. Seheult JN, Stram MN, Contis L, Pontzer RE, Hardy S, Wertz W, Baxter CM, Ondras M, Kip PL, Snyder GM, Pasculle AW. Development, evaluation, and multisite deployment of a machine learning decision tree algorithm to optimize urinalysis parameters for predicting urine culture positivity. J Clin Microbiol. 2023;61(6):e0029123.
32. Sergounioti A, Rigas D, Zoitopoulos V, Kalles D. From preliminary urinalysis to decision support: machine learning for UTI prediction in real-world laboratory data. J Pers Med. 2025;15(5):200.
33. Sheele JM, Campbell RL, Jones DD. Machine learning to predict bacteriuria in the emergency department. Sci Rep. 2025;15(1):31087.
34. Lachs MS, Nachamkin I, Edelstein PH, Goldman J, Feinstein AR, Schwartz JS. Spectrum bias in the evaluation of diagnostic tests: lessons from the rapid dipstick test for urinary tract infection. Ann Intern Med. 1992;117:135–40.
35. Zhao Q, Liu M, Gao K, Zhang B, Qi F, Xing T, Liu C, Gao J. Predicting 90-day risk of urinary tract infections following urostomy in bladder cancer patients using machine learning and explainability. Sci Rep. 2025;15:6807.
36. Wang H, Ding J, Wang S, Li L, Song J, Bai D. Enhancing predictive accuracy for urinary tract infections post-pediatric pyeloplasty with explainable AI: an ensemble TabNet approach. Sci Rep. 2025;15:2455.
37. Lee H-G, Seo Y, Kim JH, Han SB, Im JH, Jung CY, Durey A. Machine learning model for predicting ciprofloxacin resistance and presence of ESBL in patients with UTI in the ED. Sci Rep. 2023;13:3282.
38. Munigala S, Rojek R, Wood H, Yarbrough ML, Jackups RR, Burnham C-AD, Warren DK. Effect of changing urine testing orderables and clinician order sets on inpatient urine culture testing: analysis from a large academic medical center. Infect Control Hosp Epidemiol. 2019;40:281–86.
39. Fakih MG, Advani SD, Vaughn VM. Diagnosis of urinary tract infections: need for a reflective rather than reflexive approach. Infect Control Hosp Epidemiol. 2019;40:834–35.

40. Chambliss AB, Van TT. Revisiting approaches to and considerations for urinalysis and urine culture reflexive testing. Crit Rev Clin Lab Sci. 2022;59:112–24.

41. Chandrashekar G, Sahin F. A survey on feature selection methods. Comput Electr Eng. 2014;40:16-28.

42. Tonekaboni S, Joshi S, McCradden MD, Goldenberg A. What clinicians want: contextualizing explainable machine learning for clinical end use. Mach Learn Healthc. 2019:1–21.

43. Del Ben F, Da Col G, Cobârzan D, Turetta M, Rubin D, Buttazzi P, Antico A. A fully interpretable machine learning model for increasing the effectiveness of urine screening. Am J Clin Pathol. 2023;160:620–32.

44. General Directorate of Health Services (Türkiye Ministry of Health). Rational Laboratory Utilization Project (Akılcı Laboratuvar Kullanımı Projesi). Official Letter No. E.319. March 5, 2018. Ankara, Türkiye.

Supplementary Table 1. https://d2v96fxpocvxx.cloudfront.net/d363ec1e-9e5e-4591-a00a-d656bfcabb80/content-images/218f51a5-ad5b-4d21-939d-cba6f579c419.pdf