Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Predicting Acute Pyelonephritis With Clinical and Ultrasound Findings: Development of the Pyelo Prediction Score

Andrea Boccatonda

andrea.boccatonda@aosp.bo.it

IRCCS University Hospital of Bologna S Orsola-Malpighi Polyclinic: IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico di Sant'Orsola https://orcid.org/0000-0001-9303-7790

Alice Brighenti

Università di Bologna: Universita di Bologna

Valeria Tiraferri

University Hospital of Bologna Sant'Orsola-Malpighi Polyclinic: IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico di Sant'Orsola

Federica Di Natale

University Hospital of Bologna Sant'Orsola-Malpighi Polyclinic: IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico di Sant'Orsola

Daniela Agostinelli

University Hospital of Bologna Sant'Orsola-Malpighi Polyclinic: IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico di Sant'Orsola

Livia Masi

University Hospital of Bologna Sant'Orsola-Malpighi Polyclinic: IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico di Sant'Orsola

Nicola Venturoli

University Hospital of Bologna Sant'Orsola-Malpighi Polyclinic: IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico di Sant'Orsola

Elisa Gessaroli

University Hospital of Bologna Sant'Orsola-Malpighi Polyclinic: IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico di Sant'Orsola

Giorgia Comai

Università di Bologna: Universita di Bologna

Carla Serra

University Hospital of Bologna Sant'Orsola-Malpighi Polyclinic: IRCCS Azienda Ospedaliero-Universitaria di Bologna Policlinico di Sant'Orsola

Research Article

Keywords: pyelonephritis, CEUS, ultrasound, urinary infections, kidney

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11603687/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Background

Diagnostic uncertainty frequently complicates the assessment of suspected acute pyelonephritis (APN). We aimed to develop a bedside score combining clinical characteristics and B-mode ultrasound findings to support diagnostic stratification before contrast-enhanced ultrasound (CEUS).

Methods

This prospective, single-center observational study included consecutive episodes referred to a tertiary-care ultrasound service for suspected APN. All underwent B-mode ultrasound and CEUS. The reference outcome was the final discharge diagnosis, incorporating clinical, laboratory, microbiological, and imaging findings. A seven-predictor logistic regression model was obtained. Coefficients were converted into integer points. Discrimination and threshold-specific performance were evaluated, with exploratory analysis by renal transplant status.

Results

Among 329 episodes, 80 (24.3%) were classified as APN. The score included cortical hypo- and hyperechogenicity, globular renal morphology, flank pain, immunosuppression, renal calculi, and arterial hypertension, with a theoretical range of − 2 to 20 points. Area under the receiver-operating characteristic curve was 0.789. At ≥ 4 points, sensitivity was 51.3% and specificity 90.0%. At ≥ 1 point, sensitivity was 90.0% and negative predictive value 91.7%; eight APN episodes remained below threshold. At ≥ 7 points, specificity was 99.2% and positive predictive value 91.7%. Apparent discrimination was 0.813 in transplant recipients and 0.784 in non-transplant patients.

Conclusions

The Pyelo Prediction Score showed promising discrimination in a specialist referral cohort, but low scores did not reliably exclude APN. External validation is required before using the score to guide contrast-imaging decisions.

1.0 Introduction

Early diagnosis of acute pyelonephritis (APN) remains challenging due to its heterogeneous clinical presentation and the frequent overlap of symptoms with other causes of fever, abdominal pain, or systemic inflammation. Classical manifestations such as flank pain, fever, and dysuria may be absent or attenuated, particularly in elderly patients, immunocompromised individuals, or those with multiple comorbidities, leading to diagnostic uncertainty and potential delays in treatment (1–4).

Routine laboratory tests, including inflammatory markers and urinalysis, are helpful but lack sufficient specificity to reliably distinguish APN from other infectious or non-infectious conditions [3,5]. Recent European Association of Urology (EAU) guidelines recommend ultrasound as first-line imaging in suspected acute pyelonephritis, reserving contrast-enhanced computed tomography (CT) for patients with persistent fever or clinical deterioration (5). According to the American College of Radiology Appropriateness Criteria, imaging is usually not recommended in the initial evaluation of uncomplicated APN, while contrast-enhanced CT is reserved for complicated or high-risk patients (3). However, these recommendations do not provide structured criteria to guide early imaging decisions in patients with intermediate or uncertain risk profiles.

Conventional B-mode ultrasound has limited sensitivity for detecting renal parenchymal involvement, especially in early or focal forms of infection, and its diagnostic yield is highly operator-dependent (6). A 2024 systematic review and meta-analysis reported wide variability in ultrasound utilization for APN and showed that only a limited proportion of abnormal findings resulted in changes in clinical management, concluding that identification of clinical variables to guide appropriate use of ultrasound remains an unmet need (2).

Contrast-enhanced ultrasound (CEUS) emerged as a highly accurate imaging technique for the detection of renal parenchymal perfusion defects typical of APN, with diagnostic performance comparable to contrast-enhanced CT and without exposure to ionizing radiation or nephrotoxic contrast agents (3, 7). Nevertheless, CEUS is not always immediately available in emergency or acute care settings, and its indiscriminate use may lead to inefficient allocation of resources. In this context, there is a clear clinical need for a simple, bedside tool capable of stratifying patients according to their probability of APN, guiding further diagnostic work-up and optimizing the use of advanced imaging techniques. An integrated approach combining clinical features, routine laboratory indices, and basic ultrasound findings may enhance early diagnostic accuracy while maintaining feasibility in real-world practice.

The aim of this study was to develop and internally validate a pragmatic prediction score that integrates clinical data, laboratory parameters, and B-mode ultrasound signs to estimate the likelihood of APN. The proposed score is intended to support clinicians in identifying patients who may safely avoid contrast-enhanced imaging, those who are highly likely to have APN, and an intermediate-risk group in whom CEUS may provide the greatest incremental diagnostic value.

2.0 Materials and Methods

2.1 Study Design and Setting

This prospective, single-center observational study was conducted at the Diagnostic and Therapeutic Interventional Ultrasound Unit of IRCCS Azienda Ospedaliero-Universitaria di Bologna (Policlinico Sant’Orsola-Malpighi, Bologna, Italy). The study aimed to develop a prediction score integrating clinical,

laboratory, and B-mode ultrasound findings to estimate the probability of APN in patients referred for suspected disease.

2.2 Study Population

Consecutive patients referred to the ultrasound service for suspected APN were prospectively enrolled. Clinical suspicion was based on fever, flank pain, urinary symptoms, systemic inflammatory signs, or a combination of these findings. Patients were eligible for analysis if they underwent both conventional B-mode ultrasound and CEUS. Pregnancy and hemodynamic instability preventing completion of the ultrasound examination were exclusion criteria. The availability of clinical, laboratory, and imaging variables was assessed before statistical analysis, and the handling of missing data is described in the Statistical Analysis section.

2.3 Clinical and Laboratory Data Collection

Demographic characteristics, presenting symptoms, and relevant comorbidities were prospectively recorded. Clinical variables included fever, flank pain, urinary symptoms, and a history of APN. Comorbidities included diabetes mellitus, arterial hypertension, chronic kidney disease, heart failure, hypothyroidism, and immunosuppression. Immunosuppression was defined as ongoing immunosuppressive therapy, chronic corticosteroid use, active malignancy, or solid organ transplantation. Renal transplantation and the presence of urinary devices, including ureteral stents and nephrostomy tubes, were also recorded.

Laboratory variables included C-reactive protein (CRP), white blood cell count, serum creatinine, and procalcitonin when available. Urinalysis findings included leukocyte esterase and nitrite positivity. Urine and blood culture results, together with the microorganisms identified, were recorded when available.

2.4 Ultrasound and CEUS Protocol

All patients underwent conventional B-mode ultrasound followed by CEUS. Examinations were performed within a specialist hospital ultrasound service using high-end ultrasound equipment equipped with a multifrequency convex transducer, optimized for abdominal and renal imaging. All examinations were performed by five experienced operators, each with more than 5 years of experience in CEUS, particularly in abdominal and renal applications. Each examination was performed by a single operator.

B-mode ultrasound systematically evaluated predefined renal features, including cortical echogenicity abnormalities (hypoechoic or hyperechoic cortex), globular renal enlargement, presence of perirenal fluid collections, hydronephrosis, and renal stones.

CEUS examinations were performed in accordance with the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) guidelines for renal applications (8). A second-generation ultrasound contrast agent (sulphur hexafluoride microbubbles; SonoVue®, Bracco, Italy) was administered intravenously as a bolus injection of 2.4 mL, followed by a 5–10 mL saline flush, as recommended by current guidelines.

Low mechanical index real-time imaging was used to continuously assess renal perfusion. Image acquisition included the cortical phase, early parenchymal phase, and late parenchymal phase, allowing evaluation of both focal and diffuse perfusion abnormalities. The presence of focal or diffuse cortical or corticomedullary hypoperfusion areas was considered diagnostic for acute pyelonephritis. Well-defined non-enhancing areas persisting throughout all vascular phases were considered suggestive of renal abscess formation.

2.5 Reference Standard and Outcome Definition

The reference standard for APN was the final clinical diagnosis established at hospital discharge, based on the clinical course, laboratory findings and their evolution, microbiological results, and available imaging findings, including CEUS. The outcome was classified as the presence or absence of APN according to this diagnosis. CEUS examinations were performed and interpreted by ultrasound operators who were not involved in patient management or in establishing the discharge diagnosis. Conversely, the treating clinicians did not acquire or interpret the CEUS examinations. However, this separation of responsibilities did not constitute blinded outcome assessment, because CEUS findings were available to the clinicians and contributed to the final diagnosis. Consequently, the reference standard was not independent of CEUS, and incorporation bias cannot be excluded.

2.6 Statistical Analysis

Continuous variables were summarized as mean ± standard deviation or median with interquartile range, as appropriate, while categorical variables were reported as absolute numbers and percentages. Comparisons between patients with and without confirmed APN were performed using the Student's t test or Mann–Whitney U test for continuous variables, and the χ^2 test or Fisher's exact test for categorical variables.

An initial univariate analysis was conducted to identify variables associated with APN. Variables demonstrating statistical significance or deemed clinically relevant were entered into a multivariable logistic regression model to identify independent predictors of APN. Multicollinearity was assessed prior to model inclusion.

Model refinement was performed to obtain a parsimonious final model, balancing predictive performance and clinical applicability. Regression coefficients from the final multivariable model were transformed into a weighted points-based prediction score. Coefficients were scaled and rounded to ensure simplicity and usability at the bedside.

Internal validation was performed using bootstrapping techniques to evaluate model stability and reduce optimism bias. Model discrimination was assessed using receiver-operating characteristic (ROC) curve analysis, with calculation of the area under the curve (AUC) and corresponding 95% confidence intervals.

The optimal cut-off value for the prediction score was identified using the Youden index. Diagnostic performance metrics, including sensitivity, specificity, positive and negative predictive values, and

likelihood ratios, were calculated. Predefined score thresholds were explored to stratify patients into low-, intermediate-, and high-risk categories.

All statistical analyses were performed using SPSS software. A two-sided p value < 0.05 was considered statistically significant.

2.7 Ethical Considerations

The study was approved by the local Institutional Review Board (CE AVEC: 50/2024/Oss/AOUBo, approval number 0024778). Written informed consent was obtained from all participants. The study was conducted in accordance with the Declaration of Helsinki. Data collection and storage complied with the General Data Protection Regulation (GDPR).

3.0 Results

3.1 Study Population and Baseline Characteristics

The dataset comprised 329 observations, of which 80 (24.3%) were classified as APN and 249 (75.7%) as non-APN. The mean age was 58.0 ± 21.5 years, and 155 participants (47.1%) were male (Table 1). Patients classified as APN were significantly younger than those without APN (48.1 ± 21.0 vs 61.2 ± 20.6 years; p < 0.001). Women accounted for 62.5% of the APN group and 49.8% of the non-APN group, although the difference in sex distribution did not reach statistical significance using Fisher's exact test (p = 0.054). Renal transplantation was recorded in 19 patients with APN (23.8%) and 47 without APN (18.9%; p = 0.340). Ureteral stents were present in nine patients with APN (11.3%) and 32 without APN (12.9%; p = 0.846). Patients were referred from the emergency department and a range of medical, surgical, and critical care settings.

3.2 Clinical Features, Comorbidities, and Laboratory Findings

Fever was common in both the APN and non-APN groups (91.3% vs 84.3%; p = 0.140). A history of APN was reported in 16.3% and 8.8%, respectively (p = 0.093). Flank pain was more frequent in patients classified as APN (p < 0.001), whereas urinary symptoms did not differ significantly between groups (48.8% vs 37.8%; p = 0.090). Arterial hypertension was less prevalent in the APN group than in the non-APN group (25.0% vs 50.2%; p < 0.001). Heart failure (12.5% vs 22.1%; p = 0.075) and chronic kidney disease (25.0% vs 35.7%; p = 0.078) were also numerically less frequent in patients with APN. No statistically significant differences were observed for diabetes mellitus or hypothyroidism. Median CRP was higher in the APN group than in the non-APN group (13.38 mg/dL [IQR, 4.88–21.04] vs 7.99 mg/dL [IQR, 2.05–15.20]; p = 0.001). Conversely, serum creatinine was lower in patients with APN (0.91 mg/dL [IQR, 0.75–1.50] vs 1.25 mg/dL [IQR, 0.85–2.04]; p = 0.023). White blood cell count did not differ significantly between groups (10.33 × 10^9/L [IQR, 7.08–13.68] vs 9.38 × 10^9/L [IQR, 6.50–13.66]; p = 0.570), nor did recorded procalcitonin values (p = 0.409). Leukocyte esterase positivity was more

frequent in the APN group (58.8% vs 35.3%; p < 0.001), as was nitrite positivity (28.8% vs 12.4%; p = 0.001). Escherichia coli was the most frequently recorded microorganism in patients with APN, followed by Klebsiella pneumoniae. E. coli was recorded more frequently in the APN group than in the non-APN group (47.5% vs 29.7%; p = 0.004), whereas the difference for K. pneumoniae was not statistically significant (25.0% vs 17.3%; p = 0.142) (Table 1 Suppl.).

Table 1

Baseline demographic, clinical, and laboratory characteristics of the study population stratified according to the final diagnosis of acute pyelonephritis (APN). Continuous variables are presented as mean ± standard deviation (SD) or median (interquartile range [IQR]), as appropriate, whereas categorical variables are expressed as absolute numbers and percentages. Between-group comparisons were performed using Student's t-test or Mann–Whitney U test for continuous variables and chi-square or Fisher's exact test for categorical variables, as appropriate. Abbreviations: APN, acute pyelonephritis; IQR, interquartile range; SD, standard deviation; WBC, white blood cell count.

Variable	APN Absent (n = 249)	APN Present (n = 80)	P value
Demographic characteristics			
Age, mean ± SD (years)	61.2 ± 20.6	48.1 ± 21.0	< 0.001
Sex, n (%)			0.054
Male	125 (50.2)	30 (37.5)	—
Female	124 (49.8)	50 (62.5)	—
Renal transplant, n (%)	47 (18.9)	19 (23.8)	0.340
Admitting department, n (%)			
Emergency department	71 (28.5)	30 (37.5)	0.163
Nephrology	52 (20.8)	11 (13.7)	0.192
Internal medicine	47 (18.8)	15 (18.7)	1.000
Urology	25 (10.0)	7 (8.7)	0.831
Transplant unit	21 (8.4)	5 (6.2)	0.639
Other departments	33 (13.2)	12 (15.0)	0.710
Clinical features, n (%)			
Urinary symptoms	94 (37.8)	39 (48.8)	0.090
Fever	210 (84.3)	73 (91.3)	0.140
Previous history of APN	22 (8.8)	13 (16.3)	0.093
Comorbidities, n (%)			
Heart failure	55 (22.1)	10 (12.5)	0.075
Diabetes mellitus	55 (22.1)	13 (16.3)	0.341
Hypothyroidism	20 (8.0)	4 (5.0)	0.464
Arterial hypertension	125 (50.2)	20 (25.0)	< 0.001
Chronic kidney disease	89 (35.7)	20 (25.0)	0.078
Urinary devices, n (%)			

Variable	APN Absent (n = 249)	APN Present (n = 80)	P value
Nephrostomy	9 (3.6)	1 (1.3)	0.461
Ureteral stent	32 (12.9)	9 (11.3)	0.846
Laboratory parameters, median (IQR)			
C-reactive protein (mg/dL)	7.99 (2.05–15.20)	13.38 (4.88–21.04)	0.001
Procalcitonin (ng/mL)	0.30 (0.00–2.30)	0.30 (0.00–3.80)	0.409
WBC (×10⁹/L)	9.38 (6.50–13.66)	10.33 (7.08–13.68)	0.570
Serum creatinine (mg/dL)	1.25 (0.85–2.04)	0.91 (0.75–1.50)	0.023

3.3 B-mode Ultrasound Findings

B-mode ultrasound findings are summarized in Table 2. Hydroureteronephrosis, urinary retention, bladder wall thickening, and increased or reduced renal size did not differ significantly between the APN and non-APN groups. Renal calculi were numerically more frequent in patients with APN (10.0% vs 5.2%), although this difference was not statistically significant ($p = 0.185$). A perirenal fluid stripe was more frequent in the APN group (16.3% vs 4.8%; $p = 0.002$), whereas perirenal collections showed no significant difference (6.3% vs 5.6%; $p = 0.788$). Hypoechoic renal cortex (20.0% vs 2.0%; $p < 0.001$), hyperechoic renal cortex (17.5% vs 6.0%; $p = 0.003$), and globular renal morphology (15.0% vs 1.6%; $p < 0.001$) were significantly associated with APN in univariable comparisons. Ureteral wall thickening (15.0% vs 3.6%; $p < 0.001$) and bladder sediment (13.8% vs 5.6%; $p = 0.027$) were also more frequent in patients with APN.

Table 2
B-mode Ultrasound Findings According to Acute Pyelonephritis Status Values are reported as n (%) or, for findings with incomplete assessment, n/N (%), where N denotes the number of evaluable observations. Findings coded as “not assessed” were treated as missing and excluded from the corresponding comparison. P values were calculated using the two-sided Fisher’s exact test. Abbreviation: APN, acute pyelonephritis.

Variable	APN Present (n = 80)	APN Absent (n = 249)	*P* value
Bladder findings			
Urinary retention	3 (3.8)	10 (4.0)	1.000
Bladder wall thickening	2 (2.5)	15 (6.0)	0.382
Bladder sediment	11 (13.8)	14 (5.6)	**0.027**
Renal and ureteral findings			
Renal calculi	8 (10.0)	13 (5.2)	0.185
Ureteral wall thickening	12 (15.0)	9 (3.6)	**< 0.001**
Hydroureteronephrosis	10 (12.5)	25 (10.0)	0.535
Perirenal findings			
Perirenal collection	5 (6.3)	14 (5.6)	0.788
Perirenal fluid stripe	13 (16.3)	12 (4.8)	**0.002**
Renal parenchymal and morphologic findings			
Hypoechoic renal cortex	16 (20.0)	5 (2.0)	**< 0.001**
Hyperechoic renal cortex	14/80 (17.5)	15/249 (6.0)	**0.003**
Globular renal morphology	12 (15.0)	4 (1.6)	**< 0.001**
Reduced renal size	6 (7.5)	28 (11.2)	0.404
Increased renal size	15 (18.8)	29 (11.6)	0.130

3.4 Multivariable Analysis and Independent Predictors

A multivariable logistic regression analysis was performed on 329 episodes, including 80 with a final discharge diagnosis of APN (Table 2 Suppl. and Fig. 1 Suppl). Nine clinical and B-mode ultrasound variables were entered simultaneously into the model. Six variables were associated with higher odds of APN after adjustment for the other included predictors. Hypoechoic renal cortex showed the strongest association (adjusted odds ratio [aOR], 10.35; 95% CI, 3.06–35.01; p < 0.001), followed by globular renal

morphology (aOR, 6.49; 95% CI, 1.54–27.46; p = 0.011). Flank pain (aOR, 3.77; 95% CI, 1.91–7.44; p < 0.001), hyperechoic renal cortex (aOR, 3.62; 95% CI, 1.38–9.51; p = 0.009), immunosuppression (aOR, 3.36; 95% CI, 1.67–6.79; p < 0.001), and renal calculi (aOR, 3.32; 95% CI, 1.18–9.32; p = 0.023) were also associated with higher odds of APN. Arterial hypertension was associated with lower odds of APN (aOR, 0.41; 95% CI, 0.21–0.80; p = 0.009). Neither a perirenal fluid stripe (aOR, 2.02; 95% CI, 0.75–5.42; p = 0.162) nor hydroureteronephrosis (aOR, 1.10; 95% CI, 0.40–3.01; p = 0.848) showed a statistically significant association after adjustment.

3.5 Performance of the Pyelo Prediction Score and Risk Stratification

A model was subsequently fitted by retaining the seven predictors with statistically significant adjusted associations in the initial model (Table 3 Suppl.). All seven remained statistically significant after refitting. Hypoechoic renal cortex showed the strongest adjusted association with APN (adjusted odds ratio [aOR], 9.73; 95% CI, 3.52–26.92; p < 0.001), followed by globular renal morphology (aOR, 7.74; 95% CI, 2.42–24.76; p < 0.001). Flank pain, immunosuppression, hyperechoic renal cortex, and renal calculi were also associated with higher odds of APN. Arterial hypertension was associated with lower odds of APN (aOR, 0.39; 95% CI, 0.20–0.75; p = 0.005) (Table 3).

The Pyelo Prediction Score was derived from the model by dividing each regression coefficient by 0.5 and rounding to the nearest integer. Positive points were assigned to predictors associated with higher odds of APN, whereas arterial hypertension contributed − 2 points. The resulting seven-item score had a theoretical range of − 2 to 20 points.

The logistic regression model had an area under the receiver-operating characteristic curve (AUC) of 0.789 (Fig. 1). A threshold of ≥ 4 points maximized the Youden index, providing a sensitivity of 51.3%, specificity of 90.0%, positive predictive value of 62.1%, and negative predictive value of 85.2%.

Lower thresholds increased sensitivity at the expense of specificity. At ≥ 1 point, sensitivity was 90.0% and specificity was 35.3%, with a negative predictive value of 91.7%. At ≥ 7 points, specificity was 99.2% and the positive predictive value was 91.7%, although sensitivity decreased to 27.5%. Selected threshold-specific performance estimates are presented in Table 10. These thresholds illustrate alternative sensitivity–specificity trade-offs and do not establish validated clinical risk categories.

Table 3
Components and Point Allocation of the Pyelo Prediction Score- Points are assigned when the corresponding finding is present and summed to obtain the total score. Absent findings contribute zero points. Integer weights were obtained by dividing the reduced-model regression coefficients by 0.5 and rounding to the nearest integer.

Variable	**Points**
Hypoechoic renal cortex	+ 5
Flank pain	+ 3
Immunosuppression	+ 3
Hyperechoic renal cortex	+ 3
Globular renal morphology	+ 4
Renal calculi	+ 2
Arterial hypertension	−2
score range	**−2 to 20**

Diagnostic performance at selected score thresholds is reported in Table 4. A threshold of ≥ 1 point yielded a sensitivity of 90.0%, specificity of 35.3%, and negative predictive value of 91.7%, with a negative likelihood ratio of 0.28. Nevertheless, eight of the 96 episodes below this threshold had a final diagnosis of APN, indicating that a low score alone was insufficient to exclude disease. The threshold maximizing the Youden index (≥ 4 points) provided a sensitivity of 51.3% and specificity of 90.0%. A higher threshold of ≥ 7 points achieved a specificity of 99.2% and positive predictive value of 91.7%, with a positive likelihood ratio of 34.24, although sensitivity was limited to 27.5%. These findings illustrate the trade-off between sensitivity and specificity across score thresholds. The potential role of the score in prioritizing CEUS requires external validation and prospective evaluation.

Table 4

Apparent Diagnostic Performance at Selected Pyelo Prediction Score Thresholds. Each row defines a separate binary classification, with scores at or above the specified threshold considered positive. Performance was evaluated against the final discharge diagnosis of APN in 329 episodes, including 80 with APN. Estimates represent apparent performance in the development cohort. Thresholds are exploratory and do not constitute validated clinical risk categories or recommendations for CEUS use. Abbreviations: APN, acute pyelonephritis; LR+, positive likelihood ratio; LR−, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value.

Positive threshold	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	LR+	LR−
≥ 1	90.0	35.3	30.9	91.7	1.39	0.28
≥ 2	83.8	55.8	37.9	91.4	1.90	0.29
≥ 4	51.3	90.0	62.1	85.2	5.10	0.54
≥ 7	27.5	99.2	91.7	81.0	34.24	0.73

3.6 Exploratory Subgroup Analysis by Renal Transplant Status

The performance of the unchanged seven-item Pyelo Prediction Score was explored according to renal transplant status. Among 66 episodes involving renal transplant recipients, 19 (28.8%) had a final diagnosis of APN, compared with 61 of 263 episodes (23.2%) in patients without renal transplantation. The derived AUC was 0.813 (95% CI, 0.677–0.928) in transplant recipients and 0.784 (95% CI, 0.715–0.849) in non-transplant patients (Fig. 2). At the ≥ 4-point threshold, sensitivity and specificity were 68.4% and 83.0% in transplant recipients, compared with 45.9% and 91.6% in non-transplant patients. The ≥ 1-point threshold classified all transplant episodes as positive and therefore provided no specificity in this subgroup.

4.0 Discussion

This study developed a seven-item bedside score combining clinical characteristics and B-mode ultrasound findings to estimate the probability of a final diagnosis of APN in patients referred for suspected disease. The Pyelo Prediction Score showed an AUC of 0.789. These finding suggests potential for diagnostic stratification, although it does not demonstrate that score-guided imaging improves clinical outcomes or safely reduces CEUS use. The score comprises flank pain, immunosuppression, arterial hypertension, cortical hypo- and hyperechogenicity, globular renal morphology, and renal calculi; it does not require laboratory measurements. Hypoechoic renal cortex showed the strongest adjusted association with APN. Cortical echogenicity abnormalities and globular renal morphology are compatible with inflammatory parenchymal changes and renal swelling, although the present study did not directly investigate their underlying mechanisms. Flank pain and immunosuppression provided complementary clinical information, while renal calculi remained associated with APN after adjustment despite a non-significant univariable comparison. These associations should be interpreted within the selected referral population rather than as evidence of

causality. Similarly, the inverse association with arterial hypertension may reflect differences in age, comorbidity burden, or referral patterns. Threshold-specific performance illustrates the trade-off between sensitivity and specificity. A threshold of ≥ 1 point yielded 90.0% sensitivity and a negative predictive value of 91.7%; nevertheless, eight of the 96 episodes below this threshold had APN. Thus, a low score cannot independently exclude disease or establish that CEUS may safely be deferred. The threshold maximizing the Youden index (≥ 4 points) provided 51.3% sensitivity and 90.0% specificity, whereas ≥ 7 points achieved 99.2% specificity and a positive predictive value of 91.7%, but identified only 27.5% of APN episodes. These thresholds represent exploratory operating points rather than validated clinical risk categories. Decisions about further imaging should therefore continue to incorporate the clinical course, suspected complications, and individual patient characteristics.

Exploratory subgroup analysis showed AUCs of 0.813 in renal transplant recipients and 0.784 in non-transplant patients. However, only 19 APN episodes occurred in the transplant subgroup, resulting in substantial uncertainty. Threshold performance also differed: the ≥ 1-point threshold classified every transplant episode as positive and consequently provided no specificity. This finding highlights a potential limitation of low thresholds in populations with a high prevalence of immunosuppression. The subgroup results support further evaluation of transportability but do not establish superior performance in transplant recipients or justify subgroup-specific thresholds. The potential role of the score should be considered alongside the established imaging capabilities of CEUS. Published evidence describes CEUS as a useful technique for identifying renal perfusion abnormalities, characterizing inflammatory lesions and abscesses, and monitoring their evolution without ionizing radiation or nephrotoxic contrast agents [1]. Our study addresses a complementary question: whether information available before CEUS can help stratify diagnostic probability. It did not quantify the incremental diagnostic benefit of CEUS across score strata or test a strategy of withholding contrast imaging. Furthermore, CEUS findings contributed to the final discharge diagnosis, so the reference standard was not independent of imaging and incorporation bias cannot be excluded (1). Several prior studies highlighted the diagnostic capabilities of CEUS in the setting of APN. For example, an early report described characteristic CEUS enhancement patterns that differentiate focal or wedge-shaped hypoperfused areas consistent with APN from abscess formations and other complications, and showed diagnostic accuracy comparable to CT (9). A 2022 pediatric study reported good sensitivity and specificity of CEUS against both DMSA scintigraphy and contrast CT for the diagnosis of APN, supporting the generalizability of microbubble-based imaging in detecting active parenchymal infection through real-time perfusion assessment (10).

CEUS is also described as clinically valuable for follow-up and complication assessment (11). Prior reports note that CEUS can identify evolving abscesses and can be used to track resolution of inflammatory perfusion defects during therapy (12). Although guidelines such as those from EFSUMB primarily recommend CEUS for evaluation of renal abscess in complicated APN, they acknowledge that CEUS provides high-resolution perfusion information that can improve lesion characterization beyond B-mode and Doppler alone (12). A recent narrative review also underlines the expanded non-hepatic applications of CEUS, including infection and inflammation, where assessment of microcirculation offers diagnostic insights (13).

Our study evaluates a potential approach to diagnostic stratification before CEUS using a seven-item score that combines clinical characteristics and B-mode ultrasound findings, without requiring laboratory measurements. The score provides a quantitative estimate of diagnostic likelihood, but the thresholds explored in this cohort should be considered candidate operating points rather than validated SNOUT or SPIN pathways. Whether score-guided assessment can reduce unnecessary contrast imaging without increasing missed diagnoses remains to be established. The cohort included episodes referred from multiple hospital settings and patients with and without renal transplantation, providing insight into performance within a heterogeneous specialist-service population. Nevertheless, all included episodes underwent CEUS, and the findings may not generalize to unselected patients with suspected APN. The score could potentially complement existing imaging pathways, but this study did not test changes to guideline-based management or establish criteria for withholding CEUS, CT, or MRI. Multicenter external validation, followed by prospective evaluation of clinical impact, is required to determine whether its use improves diagnostic efficiency, patient outcomes, and resource utilization.

5.0 Limitations of the study

Several limitations should be acknowledged. This single-center study included a selected population referred to a tertiary-care ultrasound service, and eligibility for analysis required CEUS assessment. Its findings may therefore not generalize to unselected patients with suspected APN or settings with different referral patterns and ultrasound expertise. The reference standard was the final discharge diagnosis, which incorporated imaging findings, including CEUS; consequently, incorporation bias cannot be excluded. The cohort comprised 329 episodes with 80 APN diagnoses, limiting the precision of some estimates and increasing the risk of overfitting. Predictor selection was data-driven, and bootstrap assessment was conditional on the retained seven-predictor set rather than repeating the entire selection procedure. The reported optimism correction therefore did not capture all uncertainty associated with model development. Calibration and clinical net benefit were not evaluated, and conversion of regression coefficients into integer points introduced additional approximation. The inverse association with hypertension may also be affected by residual confounding, particularly because age was not included in the reduced model. Examinations were performed by five experienced operators, but inter-operator reproducibility was not assessed. Subgroup findings in renal transplant recipients were exploratory and based on only 19 APN episodes. Although three adolescents were included, this small number does not support conclusions regarding performance in pediatric populations. Finally, score thresholds were explored within the development cohort and do not constitute validated clinical risk categories. The study did not establish that low scores safely exclude APN or that score-guided decisions reduce CEUS use without adversely affecting outcomes. External validation, assessment of calibration and clinical utility, and prospective impact studies are required before the score can be recommended for routine imaging decisions.

6.0 Conclusions and Future directions

The seven-item Pyelo Prediction Score combines clinical characteristics and B-mode ultrasound findings to stratify the likelihood of a final diagnosis of APN in patients referred for suspected disease. The score showed an AUC of 0.789. Exploratory thresholds demonstrated different sensitivity–specificity trade-offs, but low scores did not reliably exclude APN. The score may offer a practical basis for further investigation of pre-CEUS diagnostic stratification; however, external validation, calibration assessment, and prospective impact studies are required before it can guide decisions to perform or defer contrast-enhanced imaging.

Declarations

Ethics approval:

The study protocol was reviewed by the Area Vasta Emilia Centro (AVEC) Ethics Committee, Bologna, Italy (Protocol No. 408/2026/Oss/AOUBo).

Conflicts of Interest:

The authors declare no conflicts of interest.

Funding:

None

Acknowledgment:

None

References

1. Boccatonda A, Stupia R, Serra C (2024) Ultrasound, contrast-enhanced ultrasound and pyelonephritis: A narrative review. World J Nephrol 13(3):98300
2. Yu J, Sri-Ganeshan M, Smit V, Mitra B (2024) Ultrasound for acute pyelonephritis: a systematic review and meta-analysis. Intern Med J 54(7):1106–1118
3. Smith AD, Nikolaidis P, Khatri G, Chong ST, De Leon AD, Ganeshan D et al (2022) ACR Appropriateness Criteria® Acute Pyelonephritis: 2022 Update. J Am Coll Radiol 19(11s):S224–s39
4. Jung HJ, Choi MH, Pai KS, Kim HG (2020) Diagnostic performance of contrast-enhanced ultrasound for acute pyelonephritis in children. Sci Rep 10(1):10715
5. Kranz J, Bartoletti R, Bruyère F, Cai T, Geerlings S, Köves B et al (2024) European Association of Urology Guidelines on Urological Infections: Summary of the 2024 Guidelines. Eur Urol 86(1):27–41

6. Rinaldo C, Grimaldi D, Di Serafino M, Iacobellis F, Verde F, Caruso M et al (2023) An update on pyelonephritis: role of contrast enhancement ultrasound (CEUS). J Ultrasound 26(2):333–342
7. Boccatonda A, Brighenti A, Tiraferri V, Loiacono R, Serra C (2025) Bedside CEUS: A feasible option to consider. Eur J Intern Med 136:146–147
8. Sidhu PS, Cantisani V, Dietrich CF, Gilja OH, Saftoiu A, Bartels E et al (2018) The EFSUMB Guidelines and Recommendations for the Clinical Practice of Contrast-Enhanced Ultrasound (CEUS) in Non-Hepatic Applications: Update 2017 (Short Version). Ultraschall Med 39(2):154–180
9. Fontanilla T, Minaya J, Cortés C, Hernando CG, Arangüena RP, Arriaga J et al (2012) Acute complicated pyelonephritis: contrast-enhanced ultrasound. Abdom Imaging 37(4):639–646
10. Jung HJ, Choi MH, Pai KS, Kim HG (2020) Diagnostic performance of contrast-enhanced ultrasound for acute pyelonephritis in children. Sci Rep 10(1):10715
11. Aggarwal A, Goswami S, Das CJ (2022) Contrast-enhanced ultrasound of the kidneys: principles and potential applications. Abdom Radiol 47(4):1369–1384
12. Boccatonda A, Venerato S, D'Ardes D, Cocco G, Schiavone C, Vicari S (2023) Contrast-Enhanced Ultrasound Follow-Up for Acute Pyelonephritis Patients. Healthc (Basel) 11:21
13. Eusebi L, Masino F, Bertolotto M, Montatore M, Sortino G, Pitoni L et al (2001) Contrast-enhanced ultrasound in the evaluation and management of solid renal lesions based on EFSUMB guidelines. J Med Ultrason 2025;52(3):293–303

Figures

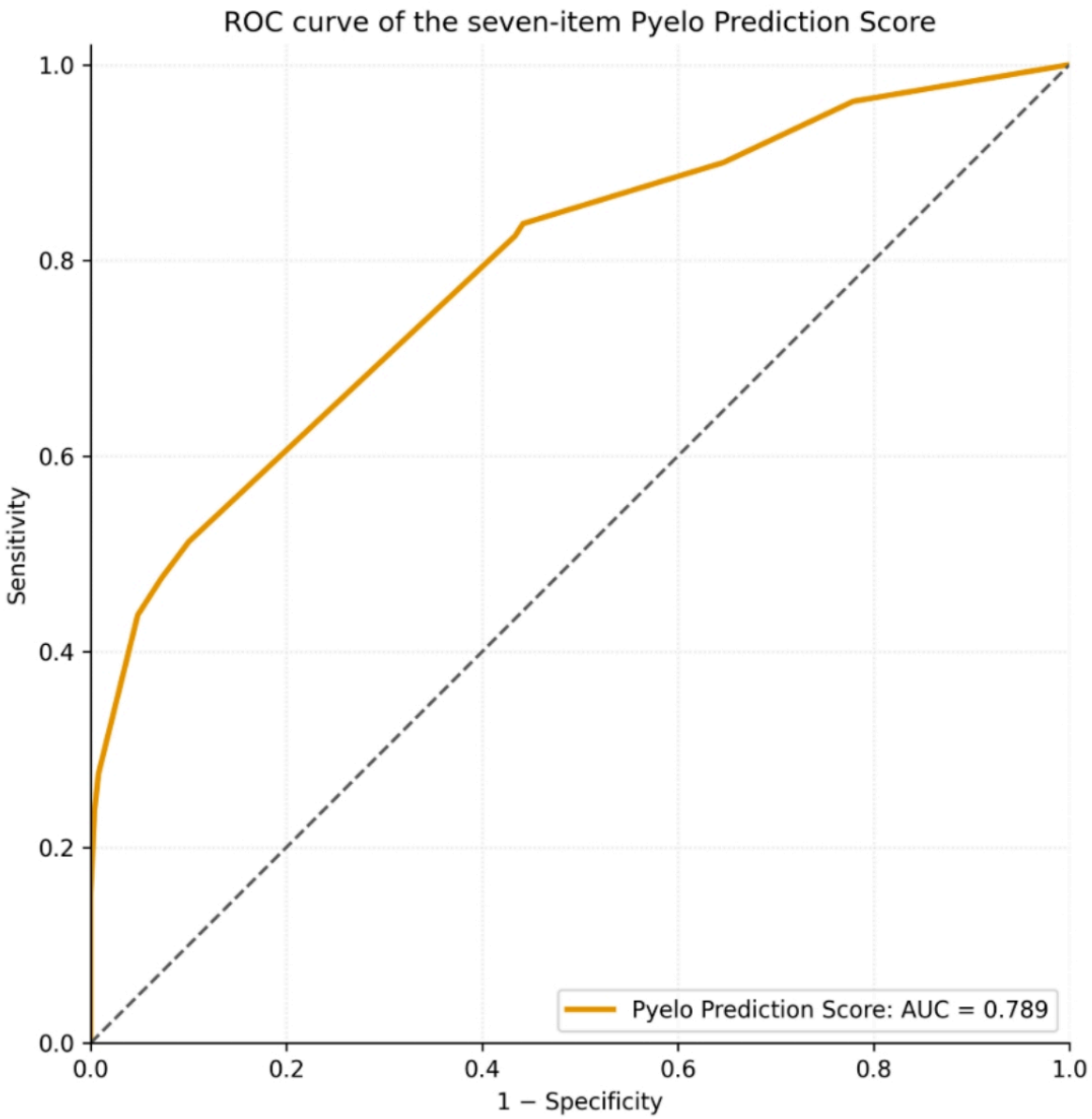


Figure 1

Receiver-operating characteristic curve of the seven-item Pyelo Prediction Score against the final discharge diagnosis of acute pyelonephritis in 329 episodes. The apparent AUC was 0.789. The dashed diagonal line represents chance discrimination.

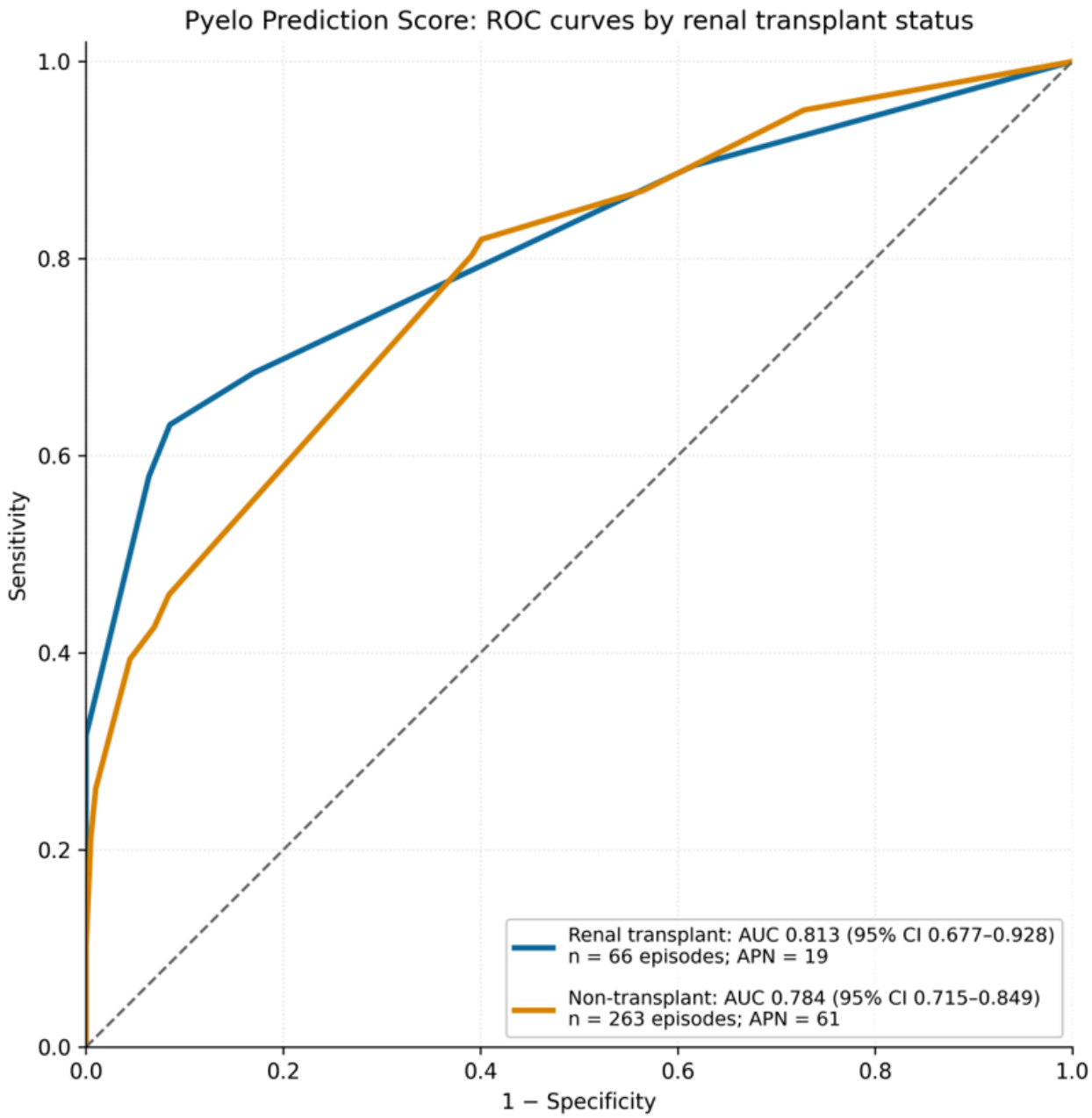


Figure 2

Receiver-operating characteristic curves of the seven-item Pyelo Prediction Score according to renal transplant status. Apparent AUCs were 0.813 (95% CI, 0.677–0.928) in renal transplant recipients and 0.784 (95% CI, 0.715–0.849) in non-transplant patients. Confidence intervals were estimated using 2,000 patient-level bootstrap resamples with the score held fixed. The dashed diagonal line represents chance discrimination.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- PyeloSupplfile.docx