

CAN THE ADDITION OF A URINE TEST STRIP FACILITATE THE DIAGNOSIS OF AN ACUTE APPENDICITIS? A DIAGNOSTIC ACCURACY STUDY

MOŽE LI DODATAK TEST TRAKICE ZA URIN OLAKŠATI DIJAGNOZU AKUTNE UPALE CRVULJKA? ISTRAŽIVANJE DIJAGNOSTIČKE TOČNOSTI

*Vanja Radišić Biljak^{1,2}, Ana Nikler¹, Dinko Bekić¹, Branko Bakula¹

<https://doi.org/10.64266/amu.2.3.1>

Abstract

Background: In our previous case-control retrospective pilot study we concluded that acute appendicitis may be successfully ruled in with high accuracy based on elevated WBCs and negative urine test strip in combination with presented clinical symptoms. We aimed to validate and assess the diagnostic accuracy of a newly proposed scoring system incorporating urinalysis.

Participants and methods: From June 2018 to May 2020, a total of 176 adult patients with suspected acute appendicitis were prospectively enrolled in a diagnostic accuracy study.

Results: The logistic regression on the entire group of patients revealed the significance of the addition of urine test strip results to standard laboratory variables (WBC and neutrophils%) only in the group of male patients.

Conclusion: Future studies should focus on validating the newly proposed scoring system incorporating urinalysis in larger, multicenter cohorts. The ultimate goal remains the development of accurate, practical tools that support timely decision-making while minimizing unnecessary imaging and surgery.

Keywords: acute appendicitis; Alvarado score; diagnostic accuracy study; scoring system; urine test strip

Sažetak

Uvod: U našem prethodnom retrospektivnom pilot-istraživanju s dizajnom slučaj-kontrola zaključili smo da se akutna upala crvuljka može s visokim stupnjem pouzdanosti potvrditi na temelju povišenog broja leukocita te negativne test trake mokraće u kombinaciji s prisutnim kliničkim simptomima. Cilj ovog istraživanja bio je provjeriti i procijeniti dijagnostičku točnost novopredloženog bodovnog sustava koji uključuje analizu mokraće.

Ispitanici i metode: Od lipnja 2018. do svibnja 2020. prospektivno je obuhvaćeno ukupno 176 odraslih bolesnika sa sumnjom na akutnu upalu crvuljka. Postupak je proveden kao ispitivanje dijagnostičke točnosti.

Rezultati: Logistička regresija na cijeloj skupini bolesnika pokazala je da dodavanje nalaza test trake mokraće standardnim laboratorijskim pokazateljima (broj leukocita i udio neutrofila) doprinosi dijagnostičkoj vrijednosti samo u skupini muških bolesnika.

1 University hospital Sveti Duh, Zagreb, Croatia

2 University of Zagreb, Faculty of Kinesiology, Zagreb, Croatia

* Corresponding author

Vanja Radišić Biljak
Department of Medical Laboratory Diagnostics
University Hospital Sveti Duh
Ulica Sveti Duh 64
10000 Zagreb
e-mail: vanja.radisic@gmail.com

Vanja Radišić Biljak
ID 0000-0002-3385-0533

Ana Nikler
ID 0000-0001-6624-863X

Dinko Bekić
ID 0000-0002-2941-4472

Branko Bakula
ID 0000-0001-6981-7877

Zaključak: Buduća istraživanja trebala bi biti usmjerena na potvrđivanje vrijednosti novopredloženog bodovnog sustava koji uključuje analizu mokraće u većim, multicentričnim istraživanjima. Cilj ostaje razvoj pouzdanih i praktičnih alata koji podupiru pravodobno donošenje kliničkih odluka uz smanjivanje potrebe za nepotrebnim podvrgavanjima slikovnim tehnikama i kirurškim zahvatima.

Ključne riječi: akutna upala crvuljka; Alvaradov bodovni sustav; bodovni sustav; dijagnostička točnost; test traka mokraće



Published under the Creative Commons Attribution 4.0 International License
<https://creativecommons.org/licenses/by/4.0>

Introduction

Suspected acute appendicitis is one of the most common indications for emergency surgery in abdominal surgery (1). Given that the procedure itself is technically relatively simple and considering the fact that neglected appendicitis is accompanied by a high morbidity rate, surgeons have a low threshold for operation if there is any suspicion of appendicitis, which results in a relatively high rate of acceptable negative appendectomies (up to about 15%) (2). Imaging methods in terms of ultrasound or MSCT can be of significant use in the diagnosis of acute appendicitis, but with certain limitations. In the adult population, unlike children, ultrasound has very limited value. On the other hand, MSCT has a high sensitivity in the diagnosis of acute appendicitis, but due to the unavailability and cost of the examination, as well as the significant harmful radiation, this examination could not be established as a routine examination in cases of suspected acute appendicitis (3). For all of the above, decision-making in cases of suspected acute appendicitis still relies predominantly on the surgeon's clinical assessment, which is based on data from the physical examination, medical history, and laboratory tests (4). Given the rather wide differential diagnosis of abdominal pain, numerous scoring systems have been proposed with the aim of facilitating the correct decision-making in the diagnosis of acute appendicitis. Unfortunately, none of them has proven to be ideal (5,6).

One of the most popular is certainly an Alvarado scoring system (7), designed in 1986, that combines certain clinical and laboratory parameters (Table 1). In our previously published case-control retrospective pilot study, we concluded that acute appendicitis in emergency settings may be successfully ruled in with high accuracy based on elevated WBCs and negative urine test strip in combination with presented clinical symptoms (8). We also observed that CRP did not contribute to the overall diagnostic accuracy, in contrast to previously published studies. Based on our preliminary findings we have prepared a prospective diagnostic accuracy study protocol with two distinct objectives:

(1) To evaluate the diagnostic accuracy of the Alvarado score within our patient population and determine its usefulness as a decision-support tool for surgeons when considering operative management. The results of this objective were published in *Acta Chirurgica Belgica Journal* in 2024 (9).

(2) To validate and assess the diagnostic accuracy of a newly proposed scoring system incorporating urinalysis in addition to other clinical and laboratory parameters (Table 2).

This study was built upon a previously published case-control pilot study, aiming to validate its preliminary findings.

Table 1. Alvarado score for diagnosis of acute appendicitis (7)

	Alvarado Score	Point*
Symptoms	Migratory right iliac fossa pain	1
	Anorexia	1
	Nausea/vomiting	1
Signs	Tenderness in right iliac fossa	2
	Rebound tenderness	1
	Elevated temperature	1
Laboratory findings	Leukocytosis (WBC >10x10 ⁹ /L)	2
	Shift to the left of neutrophils (>70%)	1
Total		10

* For each of the present symptoms, signs or laboratory findings, an adequate number of points is assigned. A score of 5 or 6 is compatible with the diagnosis of acute appendicitis. A score of 7 or 8 indicates a probable appendicitis. A score of 9 and 10 indicates a very probable appendicitis.

Table 2. A newly proposed score for diagnosis of acute appendicitis, including urinalysis

	Alvarado Score	Point*
Symptoms	Migratory right iliac fossa pain	1
	Anorexia	1
	Nausea/vomiting	1
Signs	Tenderness in right iliac fossa	2
	Rebound tenderness	1
	Elevated temperature	1
Laboratory findings	Leukocytosis (WBC >10x10 ⁹ /L)	2
	Shift to the left of neutrophils (>70%)	1
	Urine test strip negative	1
Total		11

* For each of the present symptoms, signs or laboratory findings, an adequate number of points is assigned. A score of 8 or 9 indicates a probable appendicitis.

Materials and Methods

Subjects

From June 2018 to May 2020, a total of 176 adult patients with suspected acute appendicitis were prospectively enrolled in a diagnostic accuracy study.

During the study period, all patients presenting to the Emergency Surgical Department of University Hospital Sveti Duh with right lower quadrant abdominal pain—and without a history of appendectomy—were evaluated by a single surgeon (author Branko Bakula, BB). Those deemed to have suspected appendicitis were included. The study received approval from the Hospital Ethics Committee, and written informed consent was obtained from all participants. The current paper addresses the second objective of the prospective study.

Methods

All patients underwent standard diagnostic evaluation, including a physical examination and collection of all variables required for the calculation of the Alvarado score (total white blood cell count (WBC), percentage of neutrophils). Additionally, the C-reactive protein concentrations (CRP) (immunoturbidimetry method on Beckman Coulter AU680 analyzer (Beckman Coulter, Brea, California, US) and urine test strip results (iChem Velocity Urine Chemistry Strips for in vitro use with the automated iChemVELOCITY System (Beckman Coulter, Brea, California, USA)) were also recorded. Positive findings of blood (≥ 0.3 mg/L) and/or leukocytes (≥ 25 WBC/ μ L) on the test strip were considered as a „positive“ result.

However, neither the Alvarado score nor the newly proposed scoring system incorporating urinalysis were not calculated during the initial assessment as surgical decision-making was based exclusively on clinical judgment and entirely independent of any scoring systems.

Urinalysis, traditionally used to exclude urinary tract pathology, may also play a complementary role in appendicitis assessment when integrated into structured diagnostic models.

Statistical analysis

Normality of data distribution was assessed using the Kolmogorov–Smirnov test. Patients were classified into groups with either confirmed or excluded acute appendicitis. Differences between the two groups were evaluated with the Mann–Whitney test for continuous variables and Fisher’s exact test for categorical variables. Diagnostic accuracy metrics—sensitivity, specificity, positive predictive value, negative predictive value, and positive and negative likelihood ratios—were derived from 2×2 contingency tables. Logistic regression analysis was performed to evaluate if the CRP concentrations and urine test strip results were significant predictors of acute appendicitis, among other laboratory parameters already included in the Alvarado score. Statistical analyses were conducted using MedCalc Statistical Software version 16.2.0 (MedCalc Software bvba, Ostend, Belgium). A p-value of <0.05 was considered statistically significant.

Results

Comparison of basic demographic characteristics, as well as clinical and laboratory parameters between groups of patients with and without appendicitis was described in details in the previously published article that addressed the first objective of the prospective study (9). The comparison of additional laboratory parameters (urine test strip analysis and CRP) between groups of patients with and without appendicitis is shown in Table 3. CRP was significantly higher in acute appendicitis groups (68.0 vs 6.0 mg/L, $P < 0.0001$). The comparison of proportions of

Table 3. Comparison of additional laboratory parameters between groups of patients with and without appendicitis

	APPENDICITIS NEGATIVE (N=126)	APPENDICITIS POSITIVE (N=50)	P*
CRP (mg/L) (median/IQR)	6.0 (1.4 – 28.7)	68.0 (29.9 – 131.0)	<0.0001
Urine test strip positive (N/proportion)	66 (0.52)	21 (0.42)	0.2813

*P<0.05 is considered statistically significant

positive findings of the urinalysis between groups revealed no statistically significant difference (P=0.2813).

Table 4.

The logistic regression on the entire group of patients revealed significance of all four included laboratory variables. However, when dividing patients according to gender, positive urine test strip remained statistically significant only in the group of male patients (Table 4). CRP, although significant in the whole group of selected patients, lost its significant in subgroups of patients stratified according to gender.

Table 5.

Considering the fact that urine test strip addition revealed significant in male subgroup of patients, a diagnostic accuracy study was performed for classical Alvarado

scoring system and for a newly proposed scoring system including urinalysis. The comparison of diagnostic accuracy metrics is shown in Table 5.

Discussion

In this prospective study, we evaluated whether the addition of urinalysis could improve diagnostic accuracy for acute appendicitis beyond the parameters included in the traditional Alvarado score. Although the presence of blood and/or leukocytes on urine test strip analysis did not differ significantly between groups of patients with confirmed/rejected diagnosis of acute appendicitis, the logistic regression revealed the urine test strip analysis as a statistically significant negative variable in predicting acute appendicitis. While microscopic hematuria or pyuria can theoretically occur due to irritation of the ureter or bladder by an inflamed appendix (10), these findings

Table 4. The logistic regression in identifying the clinical relevance of the expanded panel of laboratory variables that will contribute to the diagnosis of acute appendicitis

Significant variable	Coefficient	P	Odds ratio	95% CI	Percent of cases correctly classified
All patients (N=176)					
WBC, 10 ⁹ *L	0.20230	0.0017	1.2242	1.0792 – 1.3888	86.36 %
Neutrophils, %	0.10733	0.0007	1.1133	1.0461 – 1.1848	
CRP, mg/L	0.00911	0.0109	1.0092	1.0021 – 1.0163	
Urine test strip positive	-0.95009	0.0502	0.3867	0.1494 – 1.0009	
Female patients (N=103)					
WBC, 10 ⁹ *L	0.22363	0.0061	1.2506	1.0658 – 1.4674	86.41 %
Neutrophils, %	0.00849	0.0370	1.0887	1.0051 – 1.1791	
Male patients (N=73)					
WBC, 10 ⁹ *L	0.02332	0.0043	1.0236	1.0073 – 1.0401	93.15 %
Neutrophils, %	0.26302	0.0006	1.3008	1.1197 – 1.5113	
Urine test strip positive	-4.03151	0.0029	0.0177	0.0012 – 0.2521	

WBC – white blood cell count

CRP – C-reactive protein

The three-step logistic regression was performed to identify the optimal combination of the expanded panel of laboratory variables in predicting the dichotomous outcome: the presence/absence of acute appendicitis.

Step 1: All laboratory variables included in the original Alvarado score (WBC + neutrophils%) plus the expanded panel (CRP and urine test strip) were tested in the whole group of enrolled patients which identified the statistically significant ones.

Step 2: All laboratory variables included in the original Alvarado score (WBC + neutrophils%) plus the expanded panel (CRP and urine test strip) were tested only in the group of female patients which identified the statistically significant ones.

Step 3: All laboratory variables included in the original Alvarado score (WBC + neutrophils%) plus the expanded panel (CRP and urine test strip) were tested only in the group of male patients which identified the statistically significant ones.

Table 5. Comparison of diagnostic accuracy metrics for Alvarado score and newly proposed urinalysis scoring system, in a subgroup of male patients

N=73	Criterion	Se (%) (95% CI)	Sp (%) (95% CI)	LR+ (95% CI)	LR- (95% CI)	PPV (95% CI)	NPV (95% CI)
Alvarado score	≥7	84.6 (65.1 – 95.6)	87.2 (74.3 – 95.2)	6.63 (3.08 – 14.25)	0.18 (0.07 – 0.44)	78.6 (59.1 – 91.7)	91.1 (78.8 – 97.5)
Urinalysis score	≥8	76.9 (56.4 – 91.0)	91.5 (79.6 – 97.6)	9.04 (3.46 – 23.62)	0.25 (0.12 – 0.51)	83.3 (62.6 – 95.3)	87.8 (75.2 – 95.4)

Criterion - This value corresponds with the number of positive findings in the scoring systems presented in Tables 1 and 2

Se – Sensitivity; Sp – Specificity; LR+ - Positive Likelihood Ratio; LR- - Negative Likelihood Ratio; PPV – Positive Predictive Value; NPV – Negative Predictive Value

are nonspecific and commonly seen in urinary tract infections or other causes of abdominal pain (11). As such, urinalysis has traditionally been used more as a tool to exclude urinary tract pathology (12) rather than to confirm appendicitis, which could explain the observed negative correlation. Interestingly, our gender-stratified logistic regression analysis showed that a negative urine dipstick result remained a significant predictor only among male patients. The inclusion of urinalysis increased specificity in diagnosis of acute appendicitis, however the sensitivity was being reduced. This finding may reflect gender-specific differences in the prevalence of alternative diagnoses (13).

Women more frequently present with gynecological or urinary conditions that can produce positive urine dipstick results, potentially obscuring any association with appendicitis (14). In contrast, men have fewer common alternative causes of abnormal urinalysis, which may make positive findings more specific in this subgroup. Although this observation is clinically relevant, it should be interpreted with caution, as it may be influenced by sample size and requires further validation in larger cohorts. As recommended by Kollias et al., further research should evaluate the influences that can predict postoperative outcomes following appendectomies between sexes and how to prevent/reduce their occurrence as it has been shown that women had a higher rate of negative appendectomies (13, 15).

Consistent with existing literature, CRP was significantly elevated in patients with appendicitis, reinforcing its role as a useful inflammatory marker in the diagnostic workup (16). However, as we already showed in our preliminary pilot study (8), the logistic regression did not identify CRP concentration as a significant contributor to the acute appendicitis diagnosis in our study, when stratified into subgroups according to gender. Perhaps the underlying cause includes the same information that WBC and CRP offer, and thus one variable becomes redundant. It is worth mentioning that CRP spanned up to remarkably high 416 mg/L in the group of patients without the diagnosis of

acute appendicitis, while some acute appendicitis patients had normal CRP values. The observed is in line with the study performed by Atema et al. Who concluded that no WBC count or CRP level can safely and sufficiently confirm or exclude the suspected diagnosis of acute appendicitis in patients who present with abdominal pain of 5 days or less in duration (17). Hallan et al. cautions that CRP is a test of „medium accuracy“ with significant variability across studies (18). Critically, Shakhathreh et al. Concluded that while CRP is helpful, it does not replace clinical surgical expertise (19). The test is most valuable as a complementary diagnostic tool, not a standalone diagnostic method.

Incorporating urinalysis into a modified scoring system may support more precise clinical decision-making and help reduce unnecessary imaging and surgery, although further validation in larger multicenter cohorts is required.

The strengths of our study include its prospective design, the inclusion of consecutive patients evaluated by the same surgeon, and the independent assessment of clinical decision-making without influence from scoring systems. These features reduce selection bias and increase internal validity. However, several limitations should be acknowledged. First, this was a single-center study, which may limit generalizability. Second, the number of patients within gender subgroups may not have been sufficient to fully assess interaction effects. Third, urine dipstick analysis, while rapid and inexpensive, lacks the precision of microscopic urinalysis or urine culture, which might have provided additional insight.

Future studies should focus on validating the newly proposed scoring system incorporating urinalysis in larger, multicenter cohorts. The ultimate goal remains the development of accurate, practical tools that support timely decision-making while minimizing unnecessary imaging and surgery.

References

1. Guan L, Liu Z, Pan G, Zhang B, Wu Y, Gan T, Ouyang G. The global, regional, and national burden of appendicitis in 204 countries and territories, 1990-2019: a systematic analysis from the Global Burden of Disease Study 2019. *BMC Gastroenterol.* 2023;23(1):44. doi: 10.1186/s12876-023-02678-7.
2. Noureldin K, Hatim Ali AA, Issa M, Shah H, Ayantunde B, Ayantunde A. Negative Appendicectomy Rate: Incidence and Predictors. *Cureus.* 2022;14(1):e21489. doi: 10.7759/cureus.21489.
3. Comune, R., Tamburrini, S., Durante, A. et al. Ultrasonography (US) examination of acute appendicitis (AA): diagnosis of complicated and uncomplicated forms and when US is not enough. *J Med Imaging Intervent Radiol.* 2024;11,14. doi: 10.1007/s44326-024-00002-5.
4. Bergeron E. Clinical judgment remains of great value in the diagnosis of acute appendicitis. *Can J Surg.* 2006;49(2):96-100.
5. Ma T, Zhang Q, Zhao H, Zhang P. Diagnostic value of the MZXBTCH scoring system for acute complex appendicitis. *Sci Rep.* 2025;15(1):1366. doi: 10.1038/s41598-025-85791-9.
6. Alvarez-Lozada LA, Fernandez-Reyes BA, Arrambide-Garza FJ, García-Leal M, Alvarez-Villalobos NA, Martínez-Garza JH et al. Clinical scores for acute appendicitis in adults: A systematic review and meta-analysis of diagnostic accuracy studies. *Am J Surg.* 2025;240:116123. doi: 10.1016/j.amjsurg.2024.116123.
7. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med.* 1986;15(5):557-64. doi: 10.1016/s0196-0644(86)80993-3.
8. Soldo I, Radisic Biljak V, Bakula B, Bakula M, Simundic AM. The diagnostic accuracy of clinical and laboratory parameters in the diagnosis of acute appendicitis in the adult emergency department population - a case control pilot study. *Biochem Med.* 2018;28(3):030712. doi: 10.11613/BM.2018.030712.
9. Bakula B, Bogut A, Karačić A, Bakula M, Marić A, Radišić Biljak V. A prospective diagnostic accuracy study of the Alvarado score in a Croatian hospital – is it time for a final conclusion? *Acta Chir Belg.* 2025;125(4):197-201. doi: 10.1080/00015458.2025.2512279.
10. Avci MA, Akgün C, Uyanik MS. The relationship between abnormal urinalysis findings and appendicitis location. *Int J Colorectal Dis.* 2023;38(1):232. doi: 10.1007/s00384-023-04527-1.
11. Nagurney JT, Brown DF, Chang Y, Sane S, Wang AC, Weiner JB. Use of diagnostic testing in the emergency department for patients presenting with non-traumatic abdominal pain. *J Emerg Med.* 2003;25(4):363-71. doi: 10.1016/s0736-4679(03)00237-3.
12. Oh P, Lewis KC, Shoskes DA, Vasavada S, Goldman HB, Wood HM et al. Urinalysis is predictive for absence of urinary tract infection in men with and without catheters. *Neurourol Urodyn.* 2024;43(8):1850-1858. doi: 10.1002/nau.25549.
13. Kollias TF, Gallagher CP, Albaashiki A, Burle VS, Slouha E. Sex Differences in Appendicitis: A Systematic Review. *Cureus.* 2024;16(5):e60055. doi: 10.7759/cureus.60055.
14. Frazee BW, Enriquez K, Ng V, Alter H. Abnormal urinalysis results are common, regardless of specimen collection technique, in women without urinary tract infections. *J Emerg Med.* 2015;48(6):706-11. doi: 10.1016/j.jemermed.2015.02.020.
15. Flum DR, Koepsell T. The clinical and economic correlates of misdiagnosed appendicitis: nationwide analysis. *Arch Surg.* 2002;137(7):799-804. doi: 10.1001/archsurg.137.7.799.
16. Yu CW, Juan LI, Wu MH, Shen CJ, Wu JY, Lee CC. Systematic review and meta-analysis of the diagnostic accuracy of procalcitonin, C-reactive protein and white blood cell count for suspected acute appendicitis. *Br J Surg.* 2013;100(3):322-9. doi: 10.1002/bjs.9008.
17. Atema JJ, Gans SL, Beenen LF, Toorenvliet BR, Laurell H, Stoker J et al. Accuracy of White Blood Cell Count and C-reactive Protein Levels Related to Duration of Symptoms in Patients Suspected of Acute Appendicitis. *Acad Emerg Med.* 2015;22(9):1015-24. doi: 10.1111/acem.12746.
18. Hallan S, Asberg A. The accuracy of C-reactive protein in diagnosing acute appendicitis--a meta-analysis. *Scand J Clin Lab Invest.* 1997;57(5):373-80. doi: 10.3109/00365519709084584.
19. Shakhathreh HS. The accuracy of C-reactive protein in the diagnosis of acute appendicitis compared with that of clinical diagnosis. *Med Arh.* 2000;54(2):109-10.