The Medical Journal of Tikrit University (2026) 32 (1): 210 -219
DOI: http://dx.doi.org/10.25130/mjotu.32.1.22

IRAQI Academic Scientific Journals

ISSN: 1813-1638

The Medical Journal of Tikrit University

Available online at:www.mjtu.tu.edu.iq

MJTU

The Medical Journal of Tikrit University

Diagnostic Accuracy of Point-of-Care Ultrasound versus CT for Acute Appendicitis in the Emergency Department

Firas Abdullah Al-Baghdadi[1], Riyadh Adel Jaed Abdulazeez[2], Sarah Ghaleb Shati[3]

[1,2] *M.B.Ch.B., F.I.B.M.S. (DIAGNOSTIC RADIOLOGY) Ministry of Higher Education and Scientific Research, Department of Surgery, College of Medicine, University of Thi-Qar, Thi-Qar, Iraq*

[3] *CABMS (DIAGNOSTIC RADIOLOGY) Ministry of Higher Education and Scientific Research, Department of Surgery, College of Medicine, University of Thi-Qar, Thi-Qar, Iraq*

Keywords: *acute appendicitis; point-of-care ultrasound; computed tomography; diagnostic accuracy; emergency department; sensitivity and specificity.*

ARTICLE INFO

Article history:

Received 01/02/2026
Accepted 08/05/2026
Available online 30/06/ 2026

Citation

Corresponding author E mail: Dr.firas@utq.edu.iq

ABSTRACT

Background: Acute appendicitis is the most common surgical cause of acute abdominal pain. Contrast-enhanced computed tomography (CT) is highly accurate but exposes predominantly young patients to ionizing radiation and may delay disposition. Point-of-care ultrasound (POCUS), performed by the treating emergency or radiology physician at the bedside.

Objective: To determine the diagnostic accuracy of POCUS compared with contrast-enhanced CT, and to assess the effect of operator experience and body mass index (BMI) on performance.

Patients and methods: A prospective single-center diagnostic-accuracy cohort study was conducted at Al-Hussein teaching hospital in Nasiriyah, from 1 June 2023 through 1 December 2025. Adults aged 18 years or older with acute right-lower-quadrant pain underwent a structured index POCUS examination followed by reference-standard contrast-enhanced CT. POCUS operators were blinded to CT results, and CT readers were blinded to POCUS findings. The primary outcomes were sensitivity, specificity, positive and negative likelihood ratios, and the area under the receiver operating characteristic curve (AUC).

Results: Of 364 patients assessed, 298 formed the analytic cohort; 121 (40.6%) had CT-confirmed acute appendicitis. POCUS achieved a sensitivity of 84.3% (95% confidence interval [CI] 76.6–90.3%) and specificity of 89.3% (95% CI 83.7–93.4%), with positive likelihood ratio 7.87 and negative likelihood ratio 0.18 and AUC 0.86 (95% CI 0.81–0.91).

Conclusions: POCUS achieved good diagnostic accuracy for acute appendicitis and may serve as an effective first-line triage test, particularly when performed by credentialed operators in non-obese patients, potentially reducing CT utilization in selected adults.

INTRODUCTION

Acute appendicitis is the most common surgical cause of the acute abdomen, with a lifetime incidence of approximately 7–8% and a peak in the second and third decades of life [1],[2]. Prompt and accurate diagnosis reduces the risk of perforation and the rate of negative appendectomy, both of which carry measurable morbidity and cost [3]. Clinical evaluation supplemented by structured scoring systems such as the Alvarado score and the Adult Appendicitis Score has only moderate stand-alone accuracy, and cross-sectional imaging has become central to the contemporary diagnostic pathway [4],[5].

Contrast-enhanced computed tomography (CT) is the reference-standard imaging test for adult appendicitis, with pooled sensitivity and specificity each exceeding 90% in contemporary meta-analyses [6]. However, CT exposes a predominantly young population to ionizing radiation, requires intravenous iodinated contrast with attendant although low risk of adverse reaction and contrast-associated nephropathy, may not be immediately available in resource-constrained settings, and can prolong emergency department (ED) length of stay [7],[8]. These considerations have driven interest in ultrasound as a first-line imaging modality, and specifically in point-of-care ultrasound (POCUS) — focused, bedside ultrasonography performed and interpreted in real time by the treating physician rather than in a separate radiology department workflow [9].

The reported diagnostic accuracy of POCUS for appendicitis varies widely. A 2017 meta-analysis of physician-performed POCUS reported pooled sensitivity of 91% and specificity of 97%, with lower performance (sensitivity 80%, specificity 92%) in the subgroup of studies performed by emergency physicians [10]. A 2023 systematic review of ED ultrasound reported pooled sensitivity 0.81 and specificity 0.87 [11], and a 2024 meta-analysis of POCUS across ED applications reported appendicitis sensitivity 65% and specificity 89% [12]. This heterogeneity reflects differences in operator training, patient body habitus, disease prevalence, reference-standard definition, and the inherent operator dependence of ultrasound [13]. Two specific gaps motivate the present study. First, prospective diagnostic-accuracy cohorts with rigorous blinding and a robust composite reference standard remain comparatively few, and contemporary regional data are sparse. Second, the modifying effect of operator experience and patient BMI on POCUS performance is incompletely quantified in single-cohort designs that hold the reference standard and verification pathway constant [9],[10].

This study had three objectives: to determine the diagnostic accuracy (sensitivity, specificity, likelihood ratios, and AUC) of structured ED POCUS for acute appendicitis in adults, using contrast-enhanced CT supplemented by surgical, histopathological, or 30-day clinical follow-up as a composite reference standard; to quantify the effect of operator experience (credentialed sonographers versus trainees) on POCUS performance; and to characterize the effect of BMI on diagnostic accuracy. The contribution is a prospectively collected, STARD-compliant, blinding-rigorous single-center cohort that holds the reference standard constant while explicitly stratifying by the two principal known modifiers of ultrasound performance.

PATIENTS AND METHODS

STUDY DESIGN AND SETTING

A prospective single-center diagnostic-accuracy cohort study was conducted in the emergency department of Al-Hussein hospital, a teaching hospital with an annual ED census of approximately 70,000 visits, from 1 June 2023 through 1 December 2025. The study was reported in accordance with the Standards for Reporting of Diagnostic Accuracy Studies (STARD) 2015 statement [14]. The protocol was approved by the ethical committee of Thi-Qar college of medicine. Written informed consent was obtained from every participant before the index test.

PARTICIPANTS

Consecutive adults aged 18 years or older presenting to the ED with acute right-lower-quadrant or right-iliac-fossa pain of less than 7 days' duration, in whom the treating clinician judged cross-sectional imaging to be indicated, were screened for eligibility. Exclusion criteria were: pregnancy (confirmed by point-of-care human chorionic gonadotropin testing in women of childbearing potential); prior appendectomy; hemodynamic instability requiring immediate resuscitation or surgery without imaging; declared inability or refusal to provide consent; and absence of reference-standard CT. To minimize verification and spectrum bias, all eligible patients underwent both the index test and the reference standard regardless of POCUS result, and enrollment was consecutive during periods when a credentialed study sonographer or trained study investigator was available. Participant flow is summarized in figure 1.

INDEX TEST: POINT-OF-CARE ULTRASOUND

POCUS was performed at the bedside immediately after clinical assessment, before CT, using a GE Voluson E6 ultrasound system with a high-frequency linear transducer (7.5–12 MHz) and a curvilinear transducer (3.5–5 MHz) for deeper structures or larger body habitus. Operators followed a standardized graded-compression protocol over the point of maximal tenderness and the right iliac fossa. A positive POCUS examination for acute appendicitis was pre-defined as visualization of a non-compressible, blind-ending tubular structure with a maximal outer diameter greater than 6 mm, with or without an appendicolith, periappendiceal fluid, increased echogenicity of surrounding fat, or hyperemia on color Doppler. Examinations were classified prospectively as positive, negative, or non-diagnostic (appendix not visualized and no secondary signs). Operators recorded their interpretation on a structured form before the CT result was available and were blinded to the reference standard. Operators were categorized a priori as credentialed sonographers (radiologists or well trained emergency physicians) or trainees (residents).

REFERENCE STANDARD

All participants underwent contrast-enhanced CT of the abdomen and pelvis with intravenous iodinated contrast medium (iohexol or equivalent non-ionic agent) on a Toshiba (Aquilion 64 slices) scanner using a standardized abdominopelvic protocol. CT examinations were independently interpreted by a board-certified radiologist with more than 10 years of abdominal imaging experience, blinded to the POCUS interpretation. CT criteria for acute appendicitis were an appendiceal diameter greater than 6 mm with wall thickening, periappendiceal fat stranding, appendicolith, or abscess. The composite reference standard for the final diagnosis was: (i) for CT-positive patients,

surgical and histopathological confirmation where appendectomy was performed, or documented clinical course where managed non-operatively; and (ii) for CT-negative patients, an alternative discharge diagnosis with structured 30-day clinical follow-up to capture any subsequent appendicitis. This composite avoids incorporation bias by ensuring that the index test did not contribute to the reference standard.

SAMPLE SIZE

Sample size was calculated for the estimation of sensitivity. Anticipating a sensitivity of 0.85, a disease prevalence of approximately 40% among imaged patients, a desired 95% CI half-width of 0.07, and allowing for 15% non-evaluable or excluded cases, a minimum of 290 evaluable patients was required. The achieved analytic cohort of 298 satisfied this requirement.

STATISTICAL ANALYSIS

Sensitivity, specificity, positive predictive value, negative predictive value, and positive and negative likelihood ratios were calculated with 95% CIs by the Wilson method. The AUC was estimated by non-parametric methods with 95% CI by the DeLong method, and AUCs between operator-experience strata were compared using DeLong's test [15]. Non-diagnostic POCUS examinations were handled in two pre-specified ways: as a primary intention-to-diagnose analysis in which non-diagnostic studies were classified as negative (the conservative clinical interpretation), and as a secondary per-protocol analysis restricted to diagnostic examinations. Subgroup analyses by operator experience and by BMI category (less than 25, 25–29.9, and 30 kg/m² or greater) were pre-specified. Inter-observer agreement for POCUS interpretation on a random 15% subsample independently re-reviewed from recorded cine loops was assessed by the Cohen κ statistic. Two-sided p-values below 0.05 were considered statistically significant. Analyses were performed using IBM SPSS Statistics version 27.0 (IBM Corp., Armonk, NY) and R version 4.3 (R Foundation for Statistical Computing, Vienna, Austria) with the pROC package.

RESULTS

PARTICIPANTS AND DISEASE PREVALENCE

During the enrollment period, 364 patients were assessed for eligibility. After exclusion of 52 patients (18 declined consent, 11 pregnant, 9 with prior appendectomy, 7 hemodynamically unstable, and 7 without reference-standard CT) and 14 with non-diagnostic index or indeterminate reference findings, 298 patients formed the analytic cohort (see Figure 1). The mean age was 33.6 ± 13.2 years; 162 (54.4%) were male. Acute appendicitis was confirmed in 121 patients (disease prevalence 40.6%), of whom 112 underwent appendectomy with histopathological confirmation and 9 were managed non-operatively with documented clinical resolution or delayed appendectomy. Of the 177 CT-negative patients, 174 remained appendicitis-free at 30-day follow-up; 3 re-presented within 30 days with appendicitis and were classified as index false-negatives. Baseline characteristics are summarized in Table 1.

PRIMARY DIAGNOSTIC ACCURACY

In the primary intention-to-diagnose analysis (n = 298), POCUS correctly identified 102 of 121 patients with appendicitis and correctly excluded 158 of 177 without, yielding a sensitivity of 84.3% (95% CI 76.6–90.3%), specificity of 89.3% (95% CI 83.7–93.4%), positive predictive value of 84.3% (95% CI 76.6–90.3%), and

negative predictive value of 89.3% (95% CI 83.7–93.4%) (**see Table 2**). The positive likelihood ratio was 7.87 and the negative likelihood ratio was 0.18. The AUC was 0.86 (95% CI 0.81–0.91) (**see Figure 2**). In the secondary per-protocol analysis restricted to diagnostic examinations, sensitivity increased to 88.6% and specificity to 92.1%, consistent with the expected effect of excluding non-diagnostic studies.

EFFECT OF OPERATOR EXPERIENCE

Diagnostic performance differed significantly by operator experience (**see Figure 3 and Table 2**). Credentialed sonographers (n = 174 examinations) achieved a sensitivity of 90.2% (95% CI 81.2–95.8%) and specificity of 92.7% (95% CI 86.6–96.6%), with an AUC of 0.90 (95% CI 0.85–0.95). Trainees (n = 124 examinations) achieved a sensitivity of 74.4% (95% CI 60.4–85.4%) and specificity of 83.3% (95% CI 72.1–91.4%), with an AUC of 0.80 (95% CI 0.73–0.87). The difference in AUC between credentialed sonographers and trainees was statistically significant (DeLong p = 0.012). Inter-observer agreement for POCUS interpretation on the re-reviewed subsample was substantial (Cohen κ = 0.78, 95% CI 0.66–0.90).

EFFECT OF BODY MASS INDEX

Diagnostic accuracy declined with increasing BMI (see Figure 3 and Table 3). Among patients with BMI below 25 kg/m² (n = 121), sensitivity was 91.2% and specificity 94.1%. Among patients with BMI 25–29.9 kg/m² (n = 109), sensitivity was 83.3% and specificity 87.5%. Among patients with BMI 30 kg/m² or greater (n = 68), sensitivity fell to 66.7% and specificity to 78.6%, and the proportion of non-diagnostic examinations was highest in this group (13.2% versus 2.5% in the lowest BMI category). The interaction between BMI category and diagnostic accuracy was statistically significant (p = 0.008).

DISCUSSION

In this prospective STARD-compliant single-center cohort of 298 adults with suspected appendicitis, structured ED POCUS achieved a sensitivity of 84.3% and specificity of 89.3% against a composite reference standard, with an AUC of 0.86. Performance was significantly higher for credentialed sonographers (AUC 0.90) than trainees (AUC 0.80), and declined progressively with increasing BMI. These findings position POCUS as a useful first-line triage test in the diagnostic pathway for adult appendicitis, while quantifying the two principal modifiers — operator experience and body habitus — that determine where it can safely substitute for CT.

The overall accuracy observed here sits within the range reported by contemporary syntheses. The 2017 meta-analysis of physician-performed POCUS reported pooled sensitivity 91% and specificity 97%, with lower performance (sensitivity 80%, specificity 92%) in emergency-physician-performed studies — a subgroup that closely matches the present mixed-operator cohort [10]. The 2023 systematic review of ED ultrasound reported pooled sensitivity 0.81 and specificity 0.87, almost identical to the present estimates [11], and the 2024 multi-application ED POCUS meta-analysis reported appendicitis sensitivity 65% and specificity 89%, the lower sensitivity reflecting heterogeneous operator training [12]. The convergence of the present estimates with these independent syntheses supports internal validity.

Three findings deserve emphasis. First, the magnitude of the operator-experience effect

(AUC 0.90 versus 0.80, DeLong p = 0.012) confirms that POCUS accuracy for appendicitis is not a fixed property of the modality but a function of training and case volume [13]; this argues for formal credentialing and minimum-volume requirements before POCUS is used to defer CT. Second, the steep decline in sensitivity with BMI (91% below 25 kg/m² versus 67% at 30 kg/m² or greater), accompanied by a rising non-diagnostic rate, indicates that body habitus should explicitly gate the decision to rely on a negative POCUS in obese patients a negative POCUS should not defer CT [14]. Third, the negative likelihood ratio of 0.18 overall (0.11 for credentialed operators) indicates that a negative POCUS by an experienced operator in a non-obese patient meaningfully lowers post-test probability, [15] whereas the positive likelihood ratio of 7.87 supports a rule-in role that can expedite surgical referral without CT in selected patients [16].

For clinical practice, three implications follow. First, a tiered imaging pathway is supported in which experienced-operator POCUS is the first-line test in non-obese adults with suspected appendicitis, with CT reserved for non-diagnostic or discordant cases [17]; this has the potential to reduce CT utilization and associated radiation in a predominantly young population [18]. Second, POCUS performed by trainees should be interpreted with explicit acknowledgment of lower accuracy, and a negative trainee POCUS should not by itself defer CT [19] . Third, institutional POCUS credentialing pathways with minimum case volumes and periodic quality assurance are justified by the observed experience effect [20]. These recommendations are consistent with, and add quantitative regional precision to, current emergency-imaging guidance [21] .

The strengths of this study include its prospective design, consecutive enrollment, mandatory reference standard in all participants (minimizing verification and spectrum bias), pre-specified blinding of index and reference interpreters, a composite reference standard incorporating histopathology and structured follow-up (minimizing incorporation bias), and pre-specified subgroup analyses by the two principal known modifiers of ultrasound performance. Reporting followed the STARD 2015 framework.

LIMITATIONS

Several limitations apply. First, the single-center design limits external generalizability; diagnostic accuracy is influenced by local disease prevalence, operator training pathways, and equipment, and the estimates require external validation before transfer to other settings. Second, although all participants underwent the reference standard, the composite reference differed between CT-positive patients (surgical/histopathological) and CT-negative patients (30-day clinical follow-up), introducing differential verification that, while pragmatic and ethically necessary, may bias accuracy estimates. Third, POCUS is inherently operator-dependent, and although operators were categorized prospectively, residual variation within the credentialed and trainee strata cannot be excluded [22] . Fourth, the primary intention-to-diagnose analysis classified non-diagnostic POCUS examinations as negative; while this reflects conservative clinical use, it lowers apparent sensitivity [23] , and the per-protocol estimates should be interpreted as the upper bound of achievable performance. Fifth, the study did not formally compare time-to-diagnosis, ED

length of stay, cost, or radiation-dose-averted between POCUS-first and CT-first pathways; these patient-centered and system outcomes are the necessary subject of a subsequent management study [24]. Sixth, the cohort excluded pregnant patients, in whom ultrasound has a distinct and arguably greater role, so the findings do not extend to that population. Seventh, CT itself is an imperfect reference standard with a small false-negative rate, which may marginally bias accuracy estimates. Finally, prospective external validation, and ideally a randomized comparison of a POCUS-first versus CT-first diagnostic pathway with patient-centered endpoints, remain the necessary next steps.

REFERENCES

1. Ferris M, Quan S, Kaplan BS, Molodecky N, Ball CG, Chernoff GW, et al. The global incidence of appendicitis: a systematic review of population-based studies. *Ann Surg*. 2017;266(2):237–241. doi: 10.1097/SLA.0000000000002188
2. Bhangu A, Søreide K, Di Saverio S, Assarsson JH, Drake FT. Acute appendicitis: modern understanding of pathogenesis, diagnosis, and management. *Lancet*. 2015;386(10000):1278–1287. doi: 10.1016/S0140-6736(15)00275-5
3. Flum DR. Acute appendicitis—appendectomy or the antibiotics first strategy. *N Engl J Med*. 2015;372(20):1937–1943. doi: 10.1056/NEJMcp1215006
4. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med*. 1986;15(5):557–564. doi: 10.1016/S0196-0644(86)80993-3
5. Sammalkorpi HE, Mentula P, Leppäniemi A. A new adult appendicitis score improves diagnostic accuracy of acute appendicitis: a prospective study. *BMC Gastroenterol*. 2014;14:114. doi: 10.1186/1471-230X-14-114
6. Rud B, Vejborg TS, Rappeport ED, Reitsma JB, Wille-Jørgensen P. Computed tomography for diagnosis of acute appendicitis in adults. *Cochrane Database Syst Rev*. 2019;2019(11):CD009977. doi: 10.1002/14651858.CD009977.pub2
7. Smith-Bindman R, Lipson J, Marcus R, Kim KP, Mahesh M, Gould R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. *Arch Intern Med*. 2009;169(22):2078–2086. doi: 10.1001/archinternmed.2009.427
8. Pickhardt PJ, Lawrence EM, Pooler BD, Bruce RJ. Diagnostic performance of multidetector computed tomography for suspected acute appendicitis. *Ann Intern Med*. 2011;154(12):789–796. doi: 10.7326/0003-4819-154-12-201106210-00006
9. Moore CL, Copel JA. Point-of-care ultrasonography. *N Engl J Med*. 2011;364(8):749–757. doi: 10.1056/NEJMra0909487
10. Lee SH, Yun SJ. Diagnostic performance of emergency physician-performed point-of-care ultrasonography for acute appendicitis: a meta-analysis. *Am J Emerg Med*. 2019;37(4):696–705. doi: 10.1016/j.ajem.2018.07.025
11. Cho SU, Oh SK. Accuracy of ultrasound for the diagnosis of acute appendicitis in the emergency department: a systematic review. *Medicine (Baltimore)*. 2023;102(13):e33397. doi: 10.1097/MD.0000000000033397

12. Sharif S, Vlahaki D, Skitch S, Truong J, Atkinson P, Pageau P, et al. Utilization of point-of-care ultrasound as an imaging modality in the emergency department: a systematic review and meta-analysis. *Cureus*. 2024;16(1):e52371. doi: 10.7759/cureus.52371
13. Fu J, Zhou X, Chen L, Lu S. Abdominal ultrasound and its diagnostic accuracy in diagnosing acute appendicitis: a meta-analysis. *Front Surg*. 2021;8:707160. doi: 10.3389/fsurg.2021.707160
14. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527. doi: 10.1136/bmj.h5527
15. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–845. doi: 10.2307/2531595
16. Di Saverio S, Podda M, De Simone B, Ceresoli M, Augustin G, Gori A, et al. Diagnosis and treatment of acute appendicitis: 2020 update of the WSES Jerusalem guidelines. *World J Emerg Surg*. 2020;15(1):27. doi: 10.1186/s13017-020-00306-3
17. Mostbeck G, Adam EJ, Nielsen MB, Claudon M, Clevert D, Nicolau C, et al. How to diagnose acute appendicitis: ultrasound first. *Insights Imaging*. 2016;7(2):255–263. doi: 10.1007/s13244-016-0469-6
18. Matthew Fields J, Davis J, Alsup C, Bates A, Au A, Adhikari S, et al. Accuracy of point-of-care ultrasonography for diagnosing acute appendicitis: a systematic review and meta-analysis. *Acad Emerg Med*. 2017;24(9):1124–1136. doi: 10.1111/acem.13212
19. Benabbas R, Hanna M, Shah J, Sinert R. Diagnostic accuracy of history, physical examination, laboratory tests, and point-of-care ultrasound for pediatric acute appendicitis in the emergency department: a systematic review and meta-analysis. *Acad Emerg Med*. 2017;24(5):523–551. doi: 10.1111/acem.13181
20. Lee JH, Jeong YK, Park KB, Park JK, Jeong AK, Hwang JC. Operator-dependent techniques for graded compression sonography to detect the appendix and diagnose acute appendicitis. *AJR Am J Roentgenol*. 2005;184(1):91–97. doi: 10.2214/ajr.184.1.01840091
21. Giljaca V, Nadarevic T, Poropat G, Nadarevic VS, Stimac D. Diagnostic accuracy of abdominal ultrasound for diagnosis of acute appendicitis: systematic review and meta-analysis. *World J Surg*. 2017;41(3):693–700. doi: 10.1007/s00268-016-3792-7
22. Atkinson P, Bowra J, Lambert M, Lamprecht H, Noble V, Jarman B. International Federation for Emergency Medicine point of care ultrasound curriculum. *CJEM*. 2015;17(2):161–170. doi: 10.1017/cem.2015.8
23. Park JH, Salminen P, Tannaphai P, Cheng SC. Diagnostic imaging for acute appendicitis: a contemporary review. *Korean J Radiol*. 2022;23(11):1067–1078. doi: 10.3348/kjr.2022.0316
24. Reddan T, Corness J, Mengersen K, Harden F. Ultrasound of paediatric appendicitis and its secondary sonographic signs: providing a more meaningful finding. *J Med Radiat Sci*. 2016;63(1):59–66. doi: 10.1002/jmrs.154

TABLES

Table 1. Baseline characteristics of the analytic cohort (n = 298).

Characteristic	Value
Age, mean ± SD (years)	33.6 ± 13.2
Male sex, n (%)	162 (54.4%)
Symptom duration, median (IQR) (hours)	28 (16–46)
BMI, mean ± SD (kg/m²)	26.4 ± 4.8
BMI < 25 kg/m², n (%)	121 (40.6%)
BMI 25–29.9 kg/m², n (%)	109 (36.6%)
BMI ≥ 30 kg/m², n (%)	68 (22.8%)
Alvarado score, median (IQR)	6 (4–8)
White blood cell count, median (IQR) (× 10^9/L)	12.6 (9.4–15.8)
CT-confirmed acute appendicitis, n (%)	121 (40.6%)
POCUS by credentialed sonographer, n (%)	174 (58.4%)
POCUS by trainee, n (%)	124 (41.6%)
Appendectomy performed, n (%)	116 (38.9%)
Negative appendectomy, n (%)	4/116 (3.4%)

Table2 . Diagnostic accuracy of POCUS overall and by operator experience

Measure	All operators (n = 298)	Credentialed (n = 174)	Trainees (n = 124)
Sensitivity, % (95% CI)	84.3 (76.6–90.3)	90.2 (81.2–95.8)	74.4 (60.4–85.4)
Specificity, % (95% CI)	89.3 (83.7–93.4)	92.7 (86.6–96.6)	83.3 (72.1–91.4)
PPV, % (95% CI)	84.3 (76.6–90.3)	89.3 (80.1–95.3)	76.5 (62.5–87.2)
NPV, % (95% CI)	89.3 (83.7–93.4)	93.4 (87.4–97.1)	81.8 (70.4–90.2)
Positive likelihood ratio	7.87	12.36	4.46
Negative likelihood ratio	0.18	0.11	0.31
AUC (95% CI)	0.86 (0.81–0.91)	0.90 (0.85–0.95)	0.80 (0.73–0.87)

Table3 . Diagnostic accuracy of POCUS by body mass index category

BMI category	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Non-diagnostic, %
< 25 kg/m² (n = 121)	91.2 (81.8–96.7)	94.1 (86.8–98.1)	2.5
25–29.9 kg/m² (n = 109)	83.3 (70.4–92.4)	87.5 (77.6–94.1)	6.4
≥ 30 kg/m² (n = 68)	66.7 (46.7–82.9)	78.6 (63.2–89.7)	13.2

FIGURES

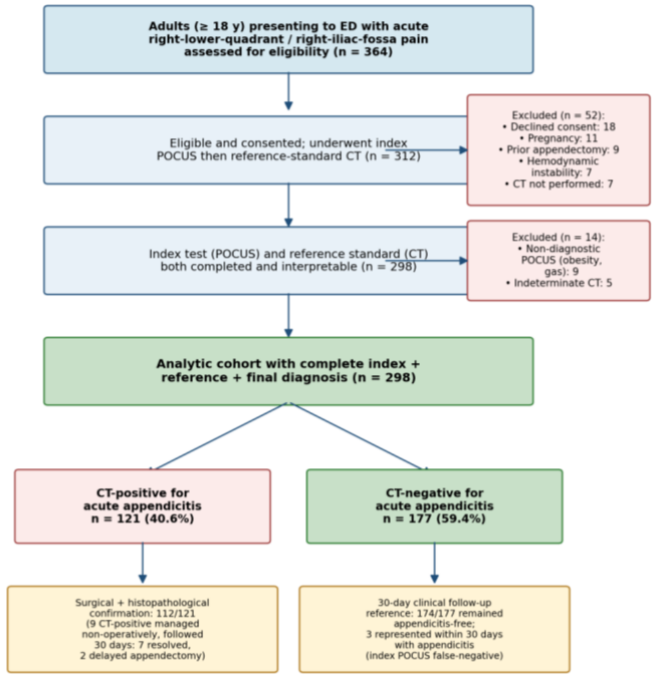

Figure 1. STARD participant flow

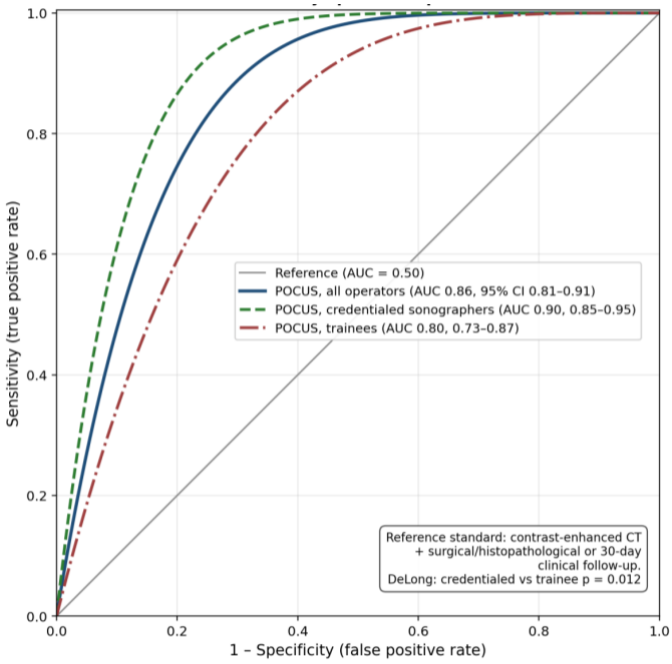

Figure 2. ROC curves for POCUS, stratified by operator experience

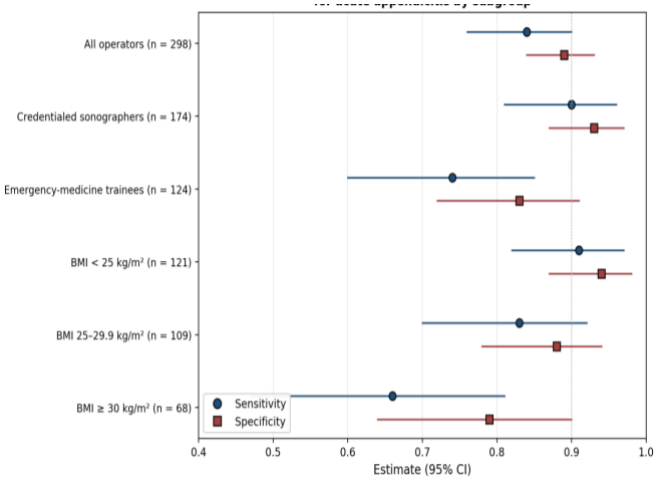

Figure 3. POCUS sensitivity and specificity by subgroup