Journal of Health and Rehabilitation Research JHRR

Vol. 6 · Issue 7 · July 2026
Open Access | DOI: 10.61919/jhrr.v6i7.2037

Review Article

Predictors of Complicated Acute Appendicitis in Adults: A Narrative Review with Emphasis on Saudi Arabia and Comparable Settings

Kashif Ali[1], **Bushra Rahmatullah**[2]

[1] UMLUJ General Hospital, Tabuk, Saudi Arabia
[2] Department of Gynecology and Obstetrics, UMLUJ General Hospital, Tabuk, Saudi Arabia

✉ **Correspondence:** Kashif Ali, kashifali135@gmail.com

Received: 08 April 2026	Revised: 12 June 2026	Accepted: 18 June 2026	Published: 30 July 2026

HOW TO CITE

Ali, K., & Bushra Rahmatullah. (2026). Predictors of Complicated Acute Appendicitis in Adults: A Narrative Review with Emphasis on Saudi Arabia and Comparable Settings. Journal of Health and Rehabilitation Research, 6(7), 1–15. https://doi.org/10.61919/jhrr.v6i7.2037

ARTICLE INFORMATION

Author Contributions	Concept: KA; Design: KA, BR; Data Collection: BR; Analysis: KA, BR; Drafting: KA; Critical Review: BR
Ethical Approval	NA
Informed Consent	NA
Conflict of Interest	The authors declare no conflict of interest.
Funding	No external funding.
Data Availability	Available from the corresponding author on reasonable request.
Acknowledgments	N/A.
AI Declaration	No generative AI was used for conceptualization, analysis, or content generation. AI use was limited to language refinement, and the authors remain fully responsible for the accuracy, originality, validity, and integrity of the manuscript's content.

ABSTRACT

Background: Differentiating complicated from uncomplicated acute appendicitis before definitive treatment is increasingly important because disease severity influences operative urgency, antimicrobial therapy, suitability for nonoperative management, and postoperative outcomes. Numerous clinical, biochemical, inflammatory, and computed tomography (CT) predictors have been proposed, but their performance varies across populations, and evidence from Saudi Arabia remains fragmented across individual predictors. Objective: To narratively synthesize contemporary evidence on predictors of complicated acute appendicitis in adults, with particular emphasis on Saudi Arabian studies and comparison with clinically relevant international settings. Methods: Evidence from Saudi Arabia and comparable populations was synthesized thematically, supplemented by recent literature on clinical predictors, inflammatory and biochemical biomarkers, CT characteristics, and multivariable prediction models. Greater interpretive weight was given to adjusted associations and studies reporting discrimination, calibration, or internal/external validation. Results: Prolonged symptom duration, peritoneal signs, higher appendicitis severity scores, C-reactive protein and other inflammatory abnormalities, hyperbilirubinemia, hyponatremia, appendicolith, increasing appendiceal diameter, periappendiceal fluid, bowel inflammatory changes, and other advanced CT findings were recurrently associated with complicated disease. Saudi studies specifically support associations involving symptom duration, Alvarado score, neutrophil-to-lymphocyte ratio, systemic inflammatory response, bilirubin, and CT abnormalities. Recent integrated prediction models combining clinical, laboratory, and imaging variables generally achieved greater discrimination than isolated biomarkers, with reported AUCs commonly approaching 0.80–0.86. However, most Saudi evidence remains retrospective, single-center, and without geographical external validation. Conclusion: Complicated appendicitis is best approached through multidomain risk assessment rather than a single biomarker. A prospectively developed, multicenter, externally validated Saudi prediction model integrating clinical, laboratory, and structured CT variables represents the principal remaining research priority. Keywords: Acute Appendicitis; Complicated Appendicitis; Appendiceal Perforation; Risk Assessment; Biomarkers; Computed Tomography; Saudi Arabia

INTRODUCTION

Acute appendicitis remains one of the most frequent abdominal surgical emergencies worldwide and continues to account for substantial emergency-department utilization, hospital admission, diagnostic imaging, antimicrobial use, and urgent operative intervention. Contemporary management increasingly recognizes acute appendicitis as a heterogeneous disease rather than a uniform pathological entity. Of particular clinical importance is the distinction between uncomplicated appendicitis and complicated disease characterized by gangrene, perforation,

periappendiceal abscess, phlegmon, localized or generalized peritonitis, or other advanced inflammatory changes. Updated international guidance emphasizes this distinction because disease severity directly influences urgency of treatment, appropriateness of nonoperative management, antimicrobial strategy, operative planning, duration of postoperative therapy, follow-up requirements, and the probability of adverse outcomes (1-4).

The increasing acceptance of selective nonoperative management for uncomplicated appendicitis has made accurate preoperative severity classification even more clinically consequential. Antibiotic-first strategies can be considered in appropriately selected patients with uncomplicated disease, whereas advanced appendiceal inflammation, free perforation, generalized contamination, or particular anatomical features may require substantially different management. Even within complicated appendicitis, localized abscess or phlegmon may be managed differently from gangrene or free perforation. Consequently, the question facing clinicians is no longer merely whether a patient has appendicitis, but also whether the disease can reliably be categorized before definitive treatment (1,2,4,5). This transition from diagnosis toward severity stratification has created a need for practical prediction approaches that are accurate, reproducible, rapidly available, and applicable across different healthcare environments.

Despite this need, the preoperative distinction between uncomplicated and complicated appendicitis remains imperfect. Clinical examination identifies inflammatory severity but lacks sufficient specificity when used alone. Routine biomarkers such as leukocyte count, neutrophil count, C-reactive protein (CRP), neutrophil-to-lymphocyte ratio (NLR), bilirubin, albumin, sodium, glucose, D-dimer, and other systemic inflammatory indices have each been associated with complicated disease, but reported thresholds and predictive performance differ considerably among studies. CT can identify appendicolith, increased appendiceal diameter, periappendiceal fluid, extraluminal gas, wall defects, abscess formation, bowel thickening, and other manifestations of advanced inflammation, yet individual radiological findings also show variable sensitivity and specificity. Recent validation work suggests that combinations of clinical, biochemical, and CT variables generally provide better discrimination than isolated predictors (6-9).

This variability reflects a deeper methodological problem within the literature. “Complicated appendicitis” is not defined uniformly. Some investigations classify only perforated disease as complicated, whereas others include gangrene, periappendiceal abscess, phlegmon, inflammatory mass, localized peritonitis, generalized peritonitis, or combinations of these features. Reference standards also vary between operative assessment, histopathology, radiological classification, or composite definitions. Differences in patient age, duration of symptoms, referral patterns, use of CT, thresholds for surgery, availability of interventional radiology, prior antibiotic exposure, and institutional practice may therefore materially affect apparent predictor performance (6-9). A biomarker or scoring threshold derived in one population cannot automatically be assumed to retain the same discrimination or calibration in another.

Recent predictive studies illustrate both the progress and the unresolved limitations of this field. Kaewlai et al. found that duration of symptoms, appendicolith, periappendiceal fat stranding, periappendiceal fluid, and extraluminal air contributed to discrimination of complicated appendicitis, with the better-performing clinical-CT scoring systems achieving areas under the receiver operating characteristic curve (AUCs) approaching 0.83 (6). Liang et al. subsequently developed a laboratory-CT nomogram incorporating white blood cell count, lymphocyte count, D-dimer, glucose, albumin, maximum appendiceal diameter, and appendiceal fecalith; the model achieved an AUC of 0.806 in development and 0.799 in an external dataset (7). More recently, an internally validated adult cohort identified age, guarding, rebound tenderness, CRP, appendiceal diameter, periappendiceal free fluid, and appendicolith as independent predictors, with an optimism-corrected AUC of approximately 0.84 (8). A separate model that intentionally avoided sectional imaging achieved an AUC of 0.861 using accessible clinical and laboratory parameters and produced a high negative predictive value in its lowest-risk category (9). Collectively, these studies show that useful prediction is feasible, but they also demonstrate that predictor composition differs substantially between models.

The Saudi Arabian evidence base is particularly relevant because adult appendicitis is encountered across institutions serving geographically, demographically, and operationally diverse populations. Saudi studies have investigated individual or limited groups of predictors rather than converging on a single validated adult severity model. Alfehaid et al. demonstrated independent associations between elevated total and direct bilirubin and

complicated appendicitis in Qassim (10). Al Amri et al. reported a strong association between NLR and complicated disease in Aseer (11). Alghamdi et al. identified symptom duration, Alvarado score, and CT abnormalities as predictors of complicated appendicitis in another Saudi cohort (12). A large Saudi database analysis linked prolonged prehospital symptom duration with perforation (13), while SIRS has also been evaluated as a marker of complicated disease in a Saudi clinical setting (14).

These studies are clinically informative but collectively expose an important national evidence gap. First, much of the Saudi evidence is retrospective and single-center. Second, different studies assess different predictor domains, making it difficult to determine whether bilirubin, NLR, symptom duration, clinical scores, systemic inflammatory responses, or CT abnormalities retain independent predictive value when considered together. Third, locally proposed associations have generally not undergone geographical external validation across Saudi regions. Fourth, the definitions of complicated appendicitis are not sufficiently standardized to permit straightforward comparison of effect sizes. Fifth, most Saudi reports focus on association or diagnostic performance of individual predictors rather than complete prediction-model development with calibration, internal validation, decision-curve analysis, and subsequent external validation. Finally, recent international evidence has shifted toward integrated multidomain risk prediction, whereas Saudi evidence remains comparatively fragmented.

This gap has become more important as management increasingly depends on reliable preoperative categorization. A highly significant association between a laboratory value and complicated disease does not necessarily mean that the marker adds clinically useful information beyond history, examination, CRP, or CT. Similarly, a model with strong apparent discrimination may perform less effectively when transported to a population with different disease prevalence, imaging practices, referral pathways, or case definitions. What is therefore lacking is not simply another statistically significant predictor, but a clinically transportable framework that integrates readily obtainable predictors, defines complicated disease consistently, quantifies discrimination and calibration, and undergoes validation in populations different from that in which it was developed.

Accordingly, this narrative review synthesizes evidence concerning predictors of complicated acute appendicitis in adults, with particular emphasis on Saudi Arabia and comparison with clinically relevant international settings. The review examines five interconnected predictor domains: delayed presentation and symptom duration; bedside clinical findings and appendicitis severity scores; systemic inflammatory and biochemical biomarkers; CT-based anatomical features; and multivariable prediction models. Particular attention is given to the consistency and independence of reported associations, differences between isolated biomarkers and integrated models, applicability of international models to Saudi practice, and the methodological priorities required for development of a robust Saudi-specific prediction strategy.

Literature Identification and Narrative Synthesis Approach

A narrative synthesis approach was adopted because the evidence relevant to complicated appendicitis was heterogeneous in design, disease definition, predictor selection, analytical methodology, reference standard, and reporting completeness. The core evidence set comprised studies from Saudi Arabia and clinically comparable settings addressing preoperative predictors of gangrenous, perforated, abscess-forming, phlegmonous, or otherwise complicated appendicitis. The original evidence base was supplemented with targeted retrieval of recent literature, prioritizing studies published from 2022 onward and emphasizing adult cohorts, prediction-model studies, validation studies, contemporary reviews, and clinical guidance relevant to differentiation of complicated from uncomplicated appendicitis.

The synthesis was structured conceptually rather than statistically. Particular emphasis was placed on multivariable-adjusted associations, diagnostic discrimination, calibration and validation when reported, while unadjusted associations were interpreted more cautiously. Because several studies used non-identical definitions of complicated appendicitis, prevalence values, thresholds, odds ratios, and model-performance estimates were not treated as directly exchangeable. Evidence was instead organized around recurring clinical and pathophysiological domains and interpreted according to the outcome definition and methodological context of each study.

Burden and Definition of Complicated Appendicitis

The frequency of complicated appendicitis varies substantially between published cohorts. Part of this variation likely represents true differences in population characteristics and presentation patterns, but much of it also reflects differences in case definition and study selection. In the Saudi cohort reported by Alfehaid et al., 33 of 158 patients, corresponding to 20.9%, had complicated appendicitis defined by perforation and/or periappendicular abscess (10). Al Amri et al. reported complications in approximately 31% of 103 patients evaluated at Aseer Central Hospital (11). In a considerably larger Saudi dataset, Shirah et al. identified 363 perforated appendicitis cases among 2,573 surgically treated patients, corresponding to approximately 14.1%, although that investigation focused specifically on perforation rather than a broader complicated-disease composite (13).

Comparable international cohorts demonstrate similarly wide variation. Khan et al. reported complicated appendicitis in 88 of 442 patients, approximately 20%, in a Pakistani cohort (15), whereas Bayissa et al. reported a proportion of approximately one-third among patients treated in Eastern Ethiopia (16). An Iraqi prospective study identified complicated disease in approximately 23% of adults (17). More recent studies similarly show substantial variation depending on definitions and selection criteria. Liang et al. reported gangrenous or perforated appendicitis in approximately one-quarter of their development cohort and approximately one-fifth of the validation cohort (7), whereas a recent Vietnamese adult cohort selected for surgery classified 40.3% of 496 patients as having complicated disease (8).

These differences demonstrate why raw prevalence cannot be interpreted as a simple measure of healthcare quality or geographical risk. A study defining complicated disease exclusively by perforation will produce a different prevalence from a study combining gangrene, perforation, phlegmon, abscess, and generalized peritonitis. Selection of only surgically managed patients may also enrich a cohort for more advanced disease compared with a population that includes patients treated nonoperatively. Increasing use of CT further changes preoperative case classification. Standardization of the outcome definition is therefore a fundamental prerequisite for development and comparison of predictive models.

Delay in Presentation and Duration of Symptoms

Duration of symptoms before definitive assessment is one of the most consistently recurring clinical predictors in the available evidence. The biological rationale is intuitive: persistent obstruction, intraluminal distension, progressive mural inflammation, impaired perfusion, bacterial invasion, tissue necrosis, and eventual perforation may occur as disease advances. Nevertheless, modern appendicitis research also cautions against assuming that every episode of uncomplicated appendicitis inevitably progresses linearly to perforation, and the relationship between time and progression appears more complex than traditional surgical teaching suggested (1,5).

In a Saudi cohort, Alghamdi et al. identified a clinically meaningful association between symptom duration and complicated appendicitis. Complicated disease occurred in 23.3% of patients whose symptoms persisted for more than 24 hours compared with 9.4% among patients presenting within 24 hours, with a reported P value of 0.004. After adjustment, symptom duration exceeding 24 hours remained associated with complicated disease, with an odds ratio of 2.83 (12). This finding is important because duration of symptoms can be established rapidly and without additional diagnostic cost.

The large Saudi analysis by Shirah et al. also supports the importance of prehospital delay. Among patients with perforation, mean symptom duration was approximately 68.8±12.4 hours. The investigators interpreted delayed presentation rather than in-hospital operative delay as the major temporal factor associated with perforation (13). Although this study predates several more recent prediction-model investigations, its large sample provides valuable Saudi-specific evidence that the period before hospital presentation may represent an important component of advanced disease.

Findings from Pakistan are concordant. Khan et al. reported a mean symptom duration of 2.0±1.2 days in complicated appendicitis compared with 1.5±0.8 days among uncomplicated cases. In multivariable analysis, each additional day of symptoms was associated with higher odds of complicated appendicitis, with an odds ratio of 1.57 and 95% confidence interval of 1.25-1.97 (15). In Ethiopia, duration of the presenting complaint and previous attendance at facilities without definitive surgical capacity were associated with complicated disease, suggesting that progression may be affected not only by patient delay but also by referral-system characteristics (16).

Recent international prediction studies continue to identify duration as relevant. The NoCtApp model incorporated symptom duration into a multivariable framework alongside age, ASA status, WBC, CRP, and ultrasonographic free fluid, although bilirubin lost independent predictive value in the adjusted model (9). Kaewlai et al. likewise found symptom duration greater than 12 hours independently associated with complicated appendicitis alongside several CT findings (6). A 2023 nomogram study reported abdominal pain duration as an independent predictor together with age, fever, and bilirubin, supporting the continued importance of temporal information when integrated with other markers.

The association is nevertheless not universal. Symptom duration was not significant in the Iraqi cohort (17), and some multivariable studies have excluded time-based variables after adjustment for more proximal inflammatory or anatomical predictors. Such inconsistency may result from inaccurate recall of symptom onset, different definitions of delay, early antibiotic exposure, variable disease biology, and correlations between duration and biomarkers of inflammation. Duration should therefore be interpreted as a risk component rather than a deterministic clock to perforation.

Clinical Findings and Appendicitis Severity Scores

Bedside clinical findings remain essential because they are available before advanced testing and can convey information about both probability and severity of appendicitis. Guarding, rebound tenderness, fever, and signs of localized or generalized peritoneal irritation may indicate extension of inflammation beyond the appendiceal lumen. However, their interpretation is influenced by patient age, analgesia, anatomical position of the appendix, obesity, comorbidity, and examiner variability.

In Saudi Arabia, Alghamdi et al. found that patients with complicated appendicitis had a median Alvarado score of 6 compared with 5 among uncomplicated cases, with a significant difference between groups. Each incremental increase in Alvarado score was independently associated with complicated disease, with a reported odds ratio of 1.339 (12). The finding suggests that a diagnostic score developed primarily to support appendicitis recognition may also contain clinically useful information regarding severity.

The Iraqi study similarly reported associations between complicated appendicitis and fever, rebound tenderness, and modified Alvarado score (17). More recent evidence strengthens the relevance of individual physical signs. In the 2026 adult cohort reported by Le et al., guarding was independently associated with complicated appendicitis with an adjusted odds ratio of 3.755, while rebound tenderness remained independently associated with an adjusted odds ratio of 2.248 (8). These effects persisted in a model that also included age, CRP, appendiceal diameter, periappendiceal free fluid, and appendicolith.

Table 1. Empirical Evidence on Predictors of Complicated Acute Appendicitis

Study	Setting / Sample	Main Predictor(s)	Key Empirical Finding
Alfehaid et al., 2023	Saudi Arabia; n=158	Total and direct bilirubin	Total bilirubin ≥15 μmol/L: aOR 3.79 (95% CI 1.67–8.48); direct bilirubin ≥5 μmol/L: aOR 4.74 (95% CI 2.07–10.86)
Al Amri et al., 2023	Saudi Arabia; n=103	NLR	NLR significantly associated with complicated appendicitis, *P*<0.00001
Alghamdi et al., 2025	Saudi Arabia; n=256	Symptom duration, Alvarado score, CT findings	Symptoms >24 h: OR 2.83; cecal/ileal thickening: OR 11.05
Khan et al., 2019	Pakistan; n=442	Symptom duration, appendicolith	Duration: OR 1.57 per additional day; appendicolith: OR 2.40 (95% CI 1.40–4.07)
Kaewlai et al., 2023	Adult CT cohort; n=325	Symptoms >12 h, appendicolith, fat stranding, fluid, extraluminal air	Best clinical-CT scores achieved AUCs up to approximately 0.83
Strohäker et al., 2023	Germany; n=1,132	Age, ASA status, symptoms, WBC, CRP, free fluid	NoCtApp model AUC 0.861; low-risk NPV 95.8%
Liang et al., 2024	Multicenter adults; n=796	WBC, lymphocytes, D-dimer, glucose, albumin, appendiceal diameter, fecalith	Development AUC 0.806; external-validation AUC 0.799
Le et al., 2026	Adult surgical cohort; n=496	Age, guarding, rebound tenderness, CRP, diameter, free fluid, appendicolith	Apparent AUC 0.864; optimism-corrected AUC 0.840

Abbreviations: aOR, adjusted odds ratio; OR, odds ratio; NLR, neutrophil-to-lymphocyte ratio; CRP, C-reactive protein; CT, computed tomography; AUC, area under the receiver operating characteristic curve; NPV, negative predictive value; ASA, American Society of Anesthesiologists.

Clinical scores therefore appear most useful as components of broader probability assessment rather than definitive severity classifiers. Existing scoring systems were generally designed to estimate the likelihood of appendicitis, not specifically to distinguish complicated from uncomplicated disease. Recent work has accordingly shifted toward dedicated severity models or modified scoring systems combining clinical and CT findings.

Validation of several such systems has produced AUCs ranging from approximately 0.69 to 0.83, illustrating clinically useful but imperfect discrimination (6).

Systemic Inflammation and Hematological Predictors

Progression toward gangrene, perforation, abscess, and peritonitis is associated with increasing systemic inflammatory activity. Consequently, considerable research has focused on inexpensive hematological markers obtainable from routine blood testing.

Al Amri et al. examined NLR among 103 Saudi patients and identified a highly significant relationship between NLR category and complicated appendicitis (11). The study supports the biological plausibility and clinical accessibility of NLR, but its utility as an independent diagnostic threshold is less certain because a fully developed multivariable prediction model, internally validated cutoff, and comprehensive discrimination analysis were not reported.

Saudi evidence relating systemic inflammation to complicated disease is also provided by Bukhari and Mirza, who evaluated SIRS among patients with acute appendicitis. SIRS was significantly associated with operative perforation accompanied by collection, while its relationship with gangrenous disease was less definitive (14). This pattern is compatible with the concept that systemic physiological disturbance becomes more pronounced with increasing inflammatory and infectious burden.

International studies broadly support the association between inflammatory biomarkers and severity but also demonstrate considerable heterogeneity in their independent predictive contribution. Uludağ and Erginöz reported that CRP and NLR were higher and lymphocyte counts lower among perforated appendicitis cases, whereas WBC alone did not significantly differentiate groups (18). A large study evaluating a modified systemic inflammation score found that a higher score, based partly on albumin and lymphocyte-to-monocyte ratio, was associated with complicated appendicitis, although specificity remained limited (19).

Recent multivariable models provide important perspective. Liang et al. retained WBC and lymphocyte count in their final laboratory-CT nomogram but did not retain every inflammatory variable evaluated (7). The model developed by Le et al. retained CRP rather than NLR, with CRP showing an independent relationship with complicated disease after adjustment for examination and CT findings (8). The NoCtApp model similarly incorporated WBC and CRP into its final non-CT framework (9). Kobayashi et al. identified CRP, maximal appendiceal diameter, and appendiceal fecalith as independent predictors of gangrenous or perforated appendicitis (20), while Suzuki et al. later incorporated a CRP threshold together with objective CT parameters in a gangrenous appendicitis scoring model (21).

These studies indicate that inflammatory markers are useful but correlated. NLR, neutrophil percentage, WBC, lymphocyte count, CRP, albumin, and composite inflammatory scores may reflect overlapping aspects of the same systemic response. Strong univariable significance does not guarantee independent predictive contribution once other biomarkers and imaging findings enter the model. This distinction is particularly relevant when considering development of a Saudi model: selection should be based on incremental predictive value, reproducibility, clinical feasibility, and model parsimony rather than the statistical significance of individual markers considered separately.

Bilirubin as a Biochemical Marker

Hyperbilirubinemia has been repeatedly investigated as a marker of complicated appendicitis. Proposed mechanisms include inflammatory cytokine effects, bacterial translocation, altered hepatocellular function, endotoxemia, and cholestatic changes accompanying more advanced intra-abdominal infection. Although these mechanisms provide biological plausibility, the diagnostic performance of bilirubin varies across populations.

The strongest Saudi-specific evidence is provided by Alfehaid et al. Among 158 patients, total bilirubin of at least 15 µmol/L was independently associated with complicated appendicitis, with an adjusted odds ratio of 3.79 and 95% confidence interval of 1.67-8.48. Direct bilirubin of at least 5 µmol/L produced an adjusted odds ratio of 4.74 with 95% confidence interval of 2.07-10.86 (10). Total bilirubin yielded sensitivity of 57.6%, specificity of 73.6%, positive predictive value of 36.5%, negative predictive value of 86.8%, and an AUC of approximately 0.69. Direct bilirubin achieved sensitivity of 54.6%, specificity of 80.0%, positive predictive value of 41.9%, negative predictive value of 86.9%, and an AUC of approximately 0.72 (10).

These results are important because they demonstrate independent associations in a Saudi population while simultaneously showing why bilirubin should not be treated as a stand-alone diagnostic test. Moderate sensitivity means that a substantial proportion of complicated cases would not be detected by bilirubin alone. The relatively high negative predictive values are potentially useful, but predictive values are prevalence-dependent and therefore require reassessment in other clinical settings.

Recent international evidence is consistent with a role for bilirubin as an adjunct. Shuaib et al. reported that both higher bilirubin and lower sodium levels were associated with complicated appendicitis and found that their combination improved discrimination compared with either variable considered alone (22). Barut and Ceylan identified age, delayed emergency presentation, elevated CRP, and elevated total bilirubin as independent risk factors and developed a scoring approach with an AUC of approximately 0.88 (23). Basukala et al. reported significant associations between total and conjugated bilirubin and perforation and concluded that combined biochemical markers provided greater diagnostic utility than isolated values (24).

Bilirubin is nevertheless not universally retained in multivariable models. In the NoCtApp development analysis, bilirubin was significant in univariable testing but did not remain independently predictive after adjustment for other clinical and laboratory parameters (9). This discrepancy illustrates the broader principle that bilirubin may act as a marker of inflammatory severity without always providing sufficient incremental information once CRP, duration, clinical status, and imaging variables are considered. The Saudi bilirubin findings therefore warrant replication within a comprehensive multidomain model rather than simple reproduction of another bilirubin cutoff study.

Sodium, Albumin, Glucose, D-Dimer, and Emerging Routine Biomarkers

Beyond traditional inflammatory measures, several routinely available laboratory parameters have emerged as candidate markers of complicated disease. Hyponatremia has attracted particular attention. Symeonidis et al. reported a significant association between preoperative hyponatremia and complicated appendicitis in adults (25). Other studies have similarly demonstrated associations between reduced sodium and disease severity, although the sensitivity of isolated sodium thresholds is often modest.

The potential pathophysiological explanation includes cytokine-mediated non-osmotic antidiuretic hormone secretion during severe inflammation. However, sodium levels may also be influenced by vomiting, hydration status, renal function, intravenous fluid administration, medications, and comorbid disease. Sodium should therefore be viewed as an accessible adjunctive marker rather than a disease-specific test.

Albumin, glucose, and D-dimer have entered more recent prediction models. Liang et al. incorporated serum glucose, albumin, and D-dimer together with WBC, lymphocytes, appendiceal diameter, and fecalith into their externally validated model (7). Albumin may reflect inflammatory severity and physiological reserve, whereas D-dimer may capture coagulation activation accompanying advanced inflammation. Their inclusion demonstrates the expanding conceptual shift from single-organ markers toward multidimensional signatures of systemic disease.

Such variables are especially attractive for settings seeking scalable risk models because many are already obtained during routine emergency evaluation. Before implementation, however, models require evaluation for missingness, timing of blood sampling, confounding by chronic disease, and reproducibility across laboratories.

Imaging Predictors and Structural CT Changes

CT provides direct visualization of appendiceal anatomy and surrounding inflammatory changes and consequently plays a central role in severity assessment. Imaging variables often demonstrate stronger associations with complicated disease than individual laboratory markers because they capture structural consequences of progression rather than indirect systemic responses.

In the Saudi cohort reported by Alghamdi et al., cecal and ileal thickening represented the strongest reported multivariable predictor, with an odds ratio of approximately 11.05 for complicated appendicitis (12). Although this large estimate requires cautious interpretation because of sample size and the available reporting of uncertainty, it highlights the potential significance of inflammatory extension beyond the appendix itself.

Appendicolith or fecalith is among the most consistently identified anatomical risk features. Khan et al. found appendicoliths in 42.0% of complicated cases compared with 23.7% of uncomplicated appendicitis, with an adjusted odds ratio of 2.40 (95% CI 1.40-4.07) (15). Kaewlai et al. also identified appendicolith as an independent predictor alongside symptom duration, periappendiceal fat stranding, fluid, and extraluminal air (6). Liang et al. retained fecalith in their final prediction model (7), and Le et al. reported an adjusted odds ratio of 2.084 for appendicolith after controlling for clinical, inflammatory, and other CT factors (8).

Maximum appendiceal diameter is another recurring structural predictor. Liang et al. included maximum outer diameter in their nomogram (7), Kobayashi et al. identified it together with CRP and fecalith (20), and Suzuki et al. used appendiceal diameter of at least 12 mm as one component of a gangrenous appendicitis score (21). Le et al. similarly found increasing appendiceal diameter independently associated with complicated disease, with an adjusted odds ratio of 1.123 per millimeter (8).

Other CT manifestations provide additional information. Akıncı et al. reported higher frequencies of appendicolith, free fluid, wall defect, abscess, free air, and retroperitoneal involvement in perforated appendicitis and identified several CT variables as predictive of perforation (26). Kaewlai et al. demonstrated the importance of periappendiceal fat stranding, fluid, and extraluminal air (6). Suzuki et al. identified cecal mucosal edema and quantitative intraluminal CT attenuation together with diameter and CRP as predictors of gangrenous disease (21).

Overall, the imaging literature suggests that no single CT sign fully defines complicated appendicitis. Instead, anatomical severity appears to be represented by combinations of luminal obstruction, appendiceal enlargement, wall disruption, periappendiceal inflammation, fluid or abscess formation, extraluminal gas, and extension to surrounding structures.

Multivariable Prediction and Integrated Risk Assessment

The most important evolution in recent evidence is the movement away from isolated biomarkers toward integrated prediction models. This reflects both statistical and clinical reality: complicated appendicitis is the product of multiple interacting dimensions, including disease duration, host response, anatomical obstruction, local tissue injury, and systemic inflammation.

Liang et al. provide one of the clearest examples. Their seven-variable nomogram combined WBC, lymphocyte count, D-dimer, serum glucose, albumin, maximum appendiceal diameter, and appendiceal fecalith. The model achieved an AUC of 0.806 (95% CI 0.763-0.849) in the development cohort, with sensitivity of 82.1% and specificity of 66.9%. External validation produced an AUC of 0.799 (95% CI 0.741-0.856), with sensitivity of 70.5% and specificity of 75.3%. Calibration was acceptable in both datasets, and decision-curve analysis suggested potential clinical utility (7).

The externally tested performance is important because models often appear stronger in derivation samples than in new populations. The modest reduction in performance observed by Liang et al. indicates a degree of transportability while also showing why external validation matters.

The 2026 study by Le et al. demonstrated another integrated approach. Independent predictors included older age, guarding, rebound tenderness, higher CRP, greater appendiceal diameter, periappendiceal free fluid, and appendicolith. The apparent AUC was 0.864 and the optimism-corrected AUC approximately 0.840 following internal validation (8). This predictor composition is clinically attractive because it combines information from history/examination, laboratory testing, and radiology.

An alternative strategy is to construct models that do not rely on CT when cross-sectional imaging is unavailable or undesirable. Strohäker et al. developed a model based on accessible clinical and laboratory parameters from more than 1,100 patients. Age, ASA status, symptom duration, WBC, CRP, and ultrasonographic free fluid contributed to the final model, which achieved an AUC of 0.861. The simplified NoCtApp score had a negative predictive value of 95.8% among patients in its lowest-risk category (9). This model is relevant to lower-resource or high-volume settings in which universal CT may be impractical.

Kobayashi et al. developed a more parsimonious model incorporating CRP, appendiceal diameter, and fecalith, with an AUC of 0.792 and good reported calibration (20). Suzuki et al. later proposed a four-component score

incorporating CT attenuation, appendiceal diameter, cecal mucosal edema, and CRP, with very high observed probability of gangrenous appendicitis in the highest-risk category (21).

The collective evidence indicates that prediction accuracy can be clinically meaningful, but there is no universally accepted model. Predictor sets differ, definitions differ, imaging protocols differ, and external validation remains limited. These limitations are directly relevant to the Saudi context because importing a model without local validation risks miscalibration even when discrimination appears acceptable.

DISCUSSION

The present narrative synthesis demonstrates that complicated appendicitis is associated with a broad constellation of clinical, inflammatory, biochemical, and radiological features, but the strength and independence of individual predictors vary considerably among populations. The most consistent general finding is therefore not the superiority of any single marker; rather, it is the value of integrated risk assessment. Duration of symptoms, peritoneal findings, CRP and other inflammatory measures, bilirubin, selected routine biochemical parameters, appendicolith, appendiceal enlargement, periappendiceal fluid, and advanced CT changes repeatedly appear across the literature, but contemporary prediction studies increasingly achieve their best performance by combining these variables rather than using them independently (6-9,20,21).

This interpretation is consistent with the current clinical landscape of acute appendicitis. Recent guidelines emphasize that management decisions differ according to whether disease is uncomplicated or complicated and according to the morphology of complicated disease itself (1,4). Nonoperative management has become a credible option for carefully selected uncomplicated appendicitis, while complicated cases may require appendectomy, drainage, antibiotics, or individualized combinations of these approaches depending on perforation, abscess, phlegmon, physiological instability, and patient factors (1-5). Preoperative severity classification is therefore not merely prognostic; it can influence the choice of treatment pathway.

The Saudi evidence provides a useful but incomplete foundation for such risk stratification. Alfehaid et al. showed that total and direct bilirubin were independently associated with complicated appendicitis, with direct bilirubin producing somewhat greater specificity and discrimination (10). Al Amri et al. demonstrated a significant relationship between NLR and complicated disease (11). Alghamdi et al. found symptom duration, Alvarado score, and CT changes to be independently informative (12), while Shirah et al. linked prolonged prehospital symptoms to perforation in a much larger Saudi cohort (13). SIRS has also been associated with advanced appendiceal disease (14). These findings are complementary rather than mutually exclusive. They suggest that Saudi patients exhibit the same broad clinical pathway identified internationally: temporal progression, intensification of inflammatory response, biochemical perturbation, and eventual anatomical evidence of complicated disease.

A crucial limitation, however, is that these Saudi predictors were generally studied in separate cohorts. Consequently, it remains unclear which retain independent value when entered simultaneously into a single model. Bilirubin may appear highly informative in a model without CRP or detailed CT characteristics but become less informative once those variables are included. NLR may show strong association in unadjusted analyses yet contribute little incremental discrimination after CRP, lymphocyte count, or anatomical features are incorporated. This phenomenon is evident internationally. In the NoCtApp study, bilirubin was significant on univariable testing but was not retained in the adjusted model (9). In Liang et al., WBC and lymphocytes entered the final nomogram, but not every inflammatory marker considered in candidate analyses (7). Le et al. retained CRP rather than NLR in a model containing physical examination and CT features (8). Thus, statistical association should not be equated with incremental predictive usefulness.

The evidence concerning symptom duration deserves similarly nuanced interpretation. Saudi and Pakistani studies demonstrate independent associations between delayed presentation and complicated disease (12,15), while recent scoring approaches have also retained duration as a predictor (6,9). These findings support rapid identification and referral of suspected appendicitis. However, contemporary thinking increasingly recognizes that uncomplicated and complicated appendicitis may not always represent obligatory sequential stages of a single time-dependent process (5). Some patients may have biological or anatomical factors that predispose to rapid progression, while others may experience limited inflammation without perforation despite longer symptom

duration. Appendicolith, anatomy, host response, age, and comorbidity may modify the relationship between time and severity.

This distinction is important when interpreting the large Saudi perforation study. Its observation that perforation was associated more strongly with prehospital symptom duration than with in-hospital operative delay supports efforts to reduce delayed presentation and referral (13), but it should not be interpreted as evidence that every short delay after hospital assessment independently causes perforation. Updated guidance indicates that appropriately selected uncomplicated appendicitis can tolerate limited operative delay without increased adverse outcomes, reinforcing the need to differentiate prehospital disease evolution from controlled in-hospital scheduling after appropriate assessment (1).

The growing evidence around appendicolith provides another important point of convergence. Appendicolith has been associated with complicated disease in Pakistan and multiple recent international studies (6-8,15,20). It also has implications beyond initial severity classification. Contemporary reviews indicate that appendicolith is associated with greater likelihood of failure or additional intervention during nonoperative treatment of uncomplicated appendicitis (5). Sohail et al. further examined patients who initially had non-perforated appendicitis with appendicolithiasis and found that older age, obesity, and hyponatremia were associated with subsequent in-hospital perforation (27). Thus, appendicolith may function both as an anatomical severity marker and as a modifier of treatment prognosis.

CT findings as a group appear among the strongest contemporary predictors because they directly visualize local pathology. The Saudi finding of marked association with cecal and ileal thickening (12) is conceptually consistent with more recent evidence showing the importance of periappendiceal free fluid, extraluminal air, wall defects, abscess, fat stranding, appendiceal enlargement, and adjacent inflammatory changes (6,8,21,26). These findings suggest that models using only laboratory parameters may sacrifice important anatomical information where CT is readily available.

Nevertheless, universal dependence on CT may not be optimal. Radiation exposure, cost, resource availability, workload, and differences in imaging practice remain relevant. The NoCtApp model demonstrates that reasonably strong discrimination can be achieved without sectional imaging (9). For Saudi Arabia, this raises the possibility that two related decision-support pathways may ultimately prove useful: a broadly available clinical-laboratory model for early triage and a second CT-enhanced model for patients undergoing cross-sectional imaging. Comparative evaluation could determine whether the additional discrimination produced by CT justifies its routine incorporation in particular risk groups.

The biomarker literature also requires careful interpretation. Saudi bilirubin evidence is compelling in demonstrating an association, but moderate AUCs and sensitivities show that bilirubin alone cannot reliably determine severity (10). Recent studies support the same conclusion. Hyperbilirubinemia is repeatedly associated with complicated disease, but models combining bilirubin with CRP, sodium, duration, or other markers tend to perform more effectively than bilirubin used alone (22-24). These findings strengthen the case for bilirubin as an adjunctive feature rather than a definitive diagnostic test.

The same applies to NLR. Saudi data demonstrate a strong association (11), and multiple international cohorts have reported elevated NLR in perforated or complicated appendicitis (18). Yet thresholds vary substantially, and NLR is affected by numerous inflammatory, infectious, physiological, and pharmacological factors. The inconsistent retention of NLR in adjusted models suggests that its principal value may lie in low-cost early stratification or in settings where CRP or advanced imaging is less accessible, rather than as a universal standalone discriminator.

CRP appears somewhat more consistently in recent multivariable modelling. It was retained in the NoCtApp model (9), in the Kobayashi score (20), in the Suzuki gangrenous appendicitis score (21), and in the recent internally validated clinical-CT model (8). This repeated appearance across differently structured models suggests that CRP captures clinically useful information regarding inflammatory burden. However, CRP is also time-dependent, and a low level early in disease does not exclude subsequent progression.

Emerging routine biomarkers such as sodium, albumin, glucose, and D-dimer offer potential additional information. Hyponatremia has been associated with complicated appendicitis across several adult studies (22,25,27). Albumin, glucose, and D-dimer were incorporated into the externally validated Liang model (7). Their usefulness may partly

arise from capturing dimensions of systemic illness distinct from local inflammatory markers. Conversely, broader inclusion of laboratory variables increases model complexity and the risk of overfitting. Model development must therefore balance statistical performance against parsimony and practical implementation.

The evolution from association studies toward prediction modelling also changes how research quality should be judged. Reporting a statistically significant P value for NLR, bilirubin, CRP, or symptom duration is insufficient for a clinically deployable prediction tool. High-quality model development requires transparent candidate-predictor selection, adequate events per parameter, appropriate handling of continuous variables, assessment of collinearity, missing-data management, model calibration, discrimination, internal validation, and preferably decision-curve analysis. External validation should then assess performance in geographically and clinically different populations.

Recent studies increasingly meet some of these requirements. Liang et al. externally validated their model and reported discrimination, calibration, and decision-curve findings (7). Le et al. used bootstrap internal validation and reported an optimism-corrected AUC together with calibration measures (8). The NoCtApp study separated development and validation datasets (9). These represent methodological advances over many earlier single-marker studies.

Yet even good international models cannot automatically solve the Saudi evidence gap. Prediction models can lose calibration when applied to populations with different baseline prevalence, demographic structure, disease definitions, laboratory practices, CT protocols, treatment thresholds, or referral pathways. Saudi Arabia also contains considerable regional variation in healthcare access and hospital type. A tertiary metropolitan surgical service may encounter a different disease spectrum from a peripheral or referral-based service. External validation within the country is therefore essential.

A future Saudi prediction study should begin by standardizing the outcome. Complicated appendicitis should be prospectively defined using reproducible operative and/or histopathological criteria, ideally mapped to an accepted severity classification. Gangrene, perforation, abscess, phlegmon, and generalized peritonitis should be reported separately as well as within the composite outcome so that models can be compared internationally and potentially tailored to clinically distinct forms of complicated disease.

Candidate predictors should then be selected before modelling on clinical and biological grounds. These may include age, comorbidity, symptom duration, temperature, guarding, rebound tenderness, Alvarado or AIR components, WBC, neutrophil and lymphocyte counts, NLR, CRP, total and direct bilirubin, sodium, albumin, glucose, and selected additional markers if routinely available. CT variables should be predefined rather than extracted opportunistically and could include appendicolith, maximal appendiceal diameter, mural disruption, periappendiceal fluid, abscess, extraluminal gas, fat stranding, cecal or ileal thickening, and other validated severity features.

Continuous predictors should preferably remain continuous during primary modelling rather than being prematurely converted into arbitrary cutoffs. Dichotomization simplifies bedside use but discards information and can reduce transportability. Thresholds can be derived subsequently if a simplified clinical score is required. Penalized or otherwise appropriately constrained regression methods may also reduce overfitting where candidate predictors are numerous.

The model should undergo robust internal validation and then geographical external validation. One practical design would derive the model in a multicenter cohort including several Saudi regions and reserve one or more geographically distinct hospitals for external validation. Alternatively, a temporal validation cohort could be followed by independent regional testing. Reporting should include AUC with confidence intervals, calibration intercept and slope, Brier score or equivalent overall-performance measures, sensitivity, specificity, positive and negative predictive values at clinically meaningful thresholds, and decision-curve analysis.

The purpose of such modelling should also be explicitly clinical. A risk tool is valuable only if it changes decisions. Potential applications include identifying patients who require expedited senior surgical review, prioritizing CT or operative access, distinguishing candidates for conservative management, selecting patients requiring broader antimicrobial coverage, or informing referral from facilities without surgical capacity. Decision thresholds should therefore be evaluated according to these intended uses rather than chosen solely to maximize the Youden index.

The current evidence additionally suggests an important healthcare-system research question. Saudi and Ethiopian findings indicate that prehospital and referral delays may contribute to complicated presentation (13,16). Future prospective research should therefore separate patient delay, referral delay, diagnostic delay, imaging delay, and post-diagnostic treatment interval. These are not interchangeable exposures. Identification of a modifiable system-level delay could have greater population impact than introduction of another laboratory biomarker.

The implications are particularly relevant for healthcare settings with variable access to CT and specialist surgical services. A parsimonious model using history, examination, and routine laboratories could support risk stratification before transfer, while a CT-enhanced model could refine probability after arrival at a definitive surgical center. Such staged risk assessment may prove more implementable than requiring the same diagnostic pathway at every level of care.

The available evidence also has implications for interpretation of negative results. A normal bilirubin, modest NLR, or absence of leukocytosis should not independently reassure clinicians when prolonged symptoms, peritoneal findings, CRP elevation, appendicolith, free fluid, wall defects, or other CT abnormalities suggest advanced disease. Conversely, isolated biochemical abnormalities without supportive clinical or anatomical findings should not automatically lead to classification as complicated appendicitis. Integrated interpretation remains essential.

Several limitations characterize the evidence synthesized in this review. Most Saudi and comparator studies are observational and many are retrospective. Single-center sampling creates vulnerability to case-selection and referral bias. Definitions of complicated appendicitis vary substantially. Candidate predictors and laboratory thresholds are inconsistent. Imaging is not uniformly obtained, and radiological interpretation may not be blinded to clinical information. Several studies report associations without complete model performance, calibration, or validation. Some analyses categorize continuous predictors at study-specific thresholds, reducing comparability and risking optimistic performance.

There is also heterogeneity in the reference standard. Operative appearance can differ from histopathological classification, particularly for gangrene or localized perforation. Composite outcomes may combine disease states with different management implications. A patient with localized abscess and a patient with diffuse fecal peritonitis are both commonly classified as complicated, although their physiology, urgency, and treatment pathways may differ markedly. Future prediction research should therefore consider both a primary standardized composite and clinically meaningful component outcomes.

An additional limitation is the relatively limited number of recent Saudi adult prediction studies compared with the rapidly expanding international literature. This makes international comparison essential but also prevents assuming direct transportability. The purpose of incorporating recent international models into the present review is consequently not to advocate their immediate implementation in Saudi practice, but to identify methodological and predictor patterns that can inform locally appropriate validation and model development.

Despite these limitations, the evidence is sufficiently consistent to support several conclusions. First, delayed presentation is an important recurrent risk feature, although it is not deterministic. Second, inflammatory biomarkers are useful but individually insufficient. Third, bilirubin has reproducible adjunctive value, including in Saudi adults, but moderate discrimination precludes standalone use. Fourth, appendicolith, appendiceal enlargement, free fluid, wall disruption, extraluminal gas, bowel or cecal inflammatory change, and related CT features provide important anatomical information. Fifth, contemporary integrated models achieve materially greater discrimination than most isolated biomarkers. Finally, the major unresolved Saudi research question is no longer whether individual markers are associated with complicated appendicitis, but how these predictors should be combined, validated, calibrated, and translated into clinically useful risk thresholds.

Future Research Priorities

Future Saudi research should prioritize a prospective multicenter adult cohort with standardized outcome definitions and predefined candidate predictors. Development should incorporate clinical history, examination findings, routinely available inflammatory and biochemical measures, and structured CT features. Candidate variables should include symptom duration, temperature, guarding, rebound tenderness, Alvarado or AIR components, WBC, differential counts, NLR, CRP, total and direct bilirubin, sodium, albumin, glucose, appendicolith,

appendiceal diameter, periappendiceal fluid, mural discontinuity, abscess, extraluminal air, fat stranding, and adjacent bowel inflammatory changes.

Predictor selection and modelling should avoid reliance solely on univariable P values. Continuous predictors should be modeled appropriately, missing data should be handled transparently, and optimism should be quantified through bootstrap or cross-validation procedures. A simplified clinical score can subsequently be derived if needed, but the statistical model should initially preserve available information.

External validation should be treated as mandatory before routine implementation. Validation should include hospitals from regions different from the derivation cohort and report both discrimination and calibration. The model should also be compared directly with established scores and simpler baselines to demonstrate incremental utility. Decision-curve analysis could identify probability thresholds at which the model provides greater net clinical benefit than treating every patient as high risk or low risk.

Future research should also evaluate whether the proposed tool improves patient-centered or system outcomes. Important endpoints include diagnostic delay, time to definitive management, negative appendectomy, conversion to open surgery, postoperative infection, intra-abdominal abscess, length of stay, readmission, intensive-care utilization, treatment failure after nonoperative management, and healthcare resource use. Without evidence of impact, a statistically accurate model may remain clinically irrelevant.

Finally, prospective studies should explicitly investigate patient- and system-level delay. Differentiating delayed symptom recognition from delayed first presentation, interfacility transfer, imaging access, surgical consultation, and treatment after diagnosis would permit targeted quality-improvement interventions. Such work may be particularly valuable because delayed presentation is one of the few recurrent predictors that potentially reflects modifiable healthcare behavior and organization rather than fixed patient biology.

CONCLUSION

Current evidence indicates that complicated acute appendicitis in adults cannot be reliably predicted by a single clinical finding, laboratory marker, or imaging feature. Prolonged symptom duration, peritoneal signs, higher clinical severity scores, CRP and other inflammatory abnormalities, hyperbilirubinemia, selected metabolic disturbances, appendicolith, increasing appendiceal diameter, periappendiceal fluid, and advanced CT inflammatory findings all contribute useful information, but their predictive value varies across populations and analytical models. Saudi studies provide important evidence for bilirubin, NLR, symptom duration, clinical scores, SIRS, and CT abnormalities, yet these factors have largely been investigated separately and have not been integrated into a geographically validated adult Saudi prediction framework. Recent international studies demonstrate that multidomain models combining clinical, laboratory, and CT information can achieve AUCs around 0.80-0.86 with increasingly rigorous internal or external validation. The principal research priority for Saudi Arabia is therefore not continued identification of isolated statistically significant biomarkers, but prospective development and external validation of a parsimonious, calibrated, clinically actionable prediction model capable of distinguishing patients at meaningful risk of complicated appendicitis across diverse Saudi healthcare settings.

REFERENCES

1. Podda M, Ceresoli M, de Simone B, Fugazzola P, Pata F, Balla A, et al. Diagnosis and treatment of acute appendicitis: 2025 edition of the World Society of Emergency Surgery Jerusalem Guidelines. JAMA Surg. 2026. doi:10.1001/jamasurg.2025.6218.

2. Diaz JJ, Ceresoli M, Herron TJ, Coccolini F. What you need to know: current management of acute appendicitis in adults. J Trauma Acute Care Surg. 2024. doi:10.1097/TA.0000000000004471.

3. Lamm R, Kumar SS, Collings AT, Haskins IN, Abou-Setta A, Narula N, et al. Diagnosis and treatment of appendicitis: systematic review and meta-analysis. Surg Endosc. 2023. doi:10.1007/s00464-023-10456-5.

4. Kumar SS, Collings AT, Lamm R, Haskins IN, Scholz S, Nepal P, et al. SAGES guideline for the diagnosis and treatment of appendicitis. Surg Endosc. 2024. doi:10.1007/s00464-024-10813-y.

5. Lee HG, Park I. Clinical outcomes and optimal indications for nonoperative management of acute appendicitis in adult patients: a comprehensive literature review. Ann Coloproctol. 2025. doi:10.3393/ac.2023.00192.0027.

6. Kaewlai R, Tongsai S, Teerasamit W, Wongsaengchan D, Noppakunsomboon N, Khamman P, et al. Validation of scoring systems for the prediction of complicated appendicitis in adults using clinical and computed tomographic findings. Insights Imaging. 2023. doi:10.1186/s13244-023-01540-4.

7. Liang Y, Sailai M, Ding R, Yimamu B, Kazi T, He M, et al. Predictive model for identification of gangrenous or perforated appendicitis in adults: a multicenter retrospective study. BMC Gastroenterol. 2024. doi:10.1186/s12876-024-03445-y.

8. Le KL, Le NK, Pham TN, Trinh MT, Tran MQ, Pham PC, et al. Clinical, laboratory, and computed tomography predictors of complicated acute appendicitis in adults selected for surgery: an internally validated cohort study. Medicine (Baltimore). 2026. doi:10.1097/MD.0000000000049367.

9. Strohäker J, Brüschke M, Feng YS, Beltzer C, Königsrainer A, Ladurner R. Predicting complicated appendicitis is possible without the use of sectional imaging—presenting the NoCtApp score. Int J Colorectal Dis. 2023. doi:10.1007/s00384-023-04501-x.

10. Alfehaid MS, Babiker AM, Alkharraz A, Alsaeed HY, Alzunaydi AA, Aldubaiyan AAR, et al. Elevated total and direct bilirubin are associated with acute complicated appendicitis: a single-center based study in Saudi Arabia. BMC Surg. 2023. doi:10.1186/s12893-023-02258-2.

11. Al Amri FS, Fihrah RS, Al Jabbar I, Alqahtani RM, Alnujaymi BM, Alshehri R, et al. Accuracy of neutrophil-to-lymphocyte ratio in predicting the severity of acute appendicitis: a single-center retrospective study. Cureus. 2023;15:e45923. doi:10.7759/cureus.45923.

12. Alghamdi AG, Alzhrani SM, Fayraq A, Alzahrani S. Predictive value of clinical and CT scan findings for complicated appendicitis: a retrospective analysis. Cureus. 2025. doi:10.7759/cureus.88948.

13. Shirah BH, Shirah HA, Alhaidari WA, Elraghi MA, Chughtai MA. Perforated appendix-delay in presentation rather than delay in the surgical intervention: retrospective database analysis of 2573 Saudi Arabian patients in 10 years. 2016.

14. Bukhari, Mirza. Complicated appendicitis is predicted by the presence of systemic inflammatory response syndrome. Saudi Surg J. 2020. doi:10.4103/ssj.ssj_6_20.

15. Khan et al. Factors associated with complicated appendicitis: view from a low-middle income country. Cureus. 2019. doi:10.7759/cureus.4765.

16. Bayissa B, Miressa F, Abulkadir A, Fekadu G. Predictors of complicated appendicitis among patients presented to public referral hospitals in Harari region, Eastern Ethiopia: a case-control study. SAGE Open Med. 2022. doi:10.1016/j.sipas.2022.100072.

17. Appraisal of various predictors for complicated appendicitis in adults: a comparative study between complicated and non-complicated appendicitis. 2021. doi:10.21203/RS.3.RS-303224/V1.

18. Uludağ S, Erginöz E. Effectiveness of pre-operative routine blood tests in predicting complicated acute appendicitis. Ulus Travma Acil Cerrahi Derg. 2022. doi:10.14744/tjtes.2021.13472.

19. Hamamcı E, Er S, Akın T, Bayazatli SM, Berkem H, Çetinkaya E. Can the modified systemic inflammation score predict complicated appendicitis? J Coll Physicians Surg Pak. 2023. doi:10.29271/jcpsp.2023.11.1240.

20. Kobayashi T, Hidaka E, Koganezawa I, Nakagawa M, Yokozuka K, Ochiai S, et al. Development of a scoring model based on objective factors to predict gangrenous/perforated appendicitis. BMC Gastroenterol. 2023. doi:10.1186/s12876-023-02767-7.

21. Suzuki T, Matsumoto A, Sugiki D, Akao T, Matsumoto H. Clinical prediction model for gangrenous appendicitis: a retrospective single-center study. Scand J Surg. 2025. doi:10.1177/14574969251319853.

22. Shuaib A, Alhamdan N, Arian H, Sallam MA, Shuaib A. Hyperbilirubinemia and hyponatremia as predictors of complicated appendicitis. Med Sci (Basel). 2022;10:36. doi:10.3390/medsci10030036.

23. Barut B, Ceylan C. Preoperative predictor laboratory markers for complicated appendicitis: a retrospective analysis of single center experience. Med Sci. 2023. doi:10.5455/medscience.2023.05.077.

24. Basukala S, Gurung S, Tamang A, Shrestha O, Devkota M, Thapa N, et al. Diagnostic value of biochemical markers in prediction of perforation in acute appendicitis: a cross-sectional study. Ann Med Surg. 2024. doi:10.1097/MS9.0000000000002547.

25. Symeonidis N, Pavlidis E, Psarras K, Stavrati K, Nikolaidou C, Marneri A, et al. Preoperative hyponatremia indicates complicated acute appendicitis. Surg Res Pract. 2022;2022:1836754. doi:10.1155/2022/1836754.

26. Akıncı Ö. Predictive value of sphericity index and other multidetector computed tomography findings in perforated acute appendicitis. Ulus Travma Acil Cerrahi Derg. 2023. doi:10.14744/tjtes.2023.45383.

27. Sohail AH, Hakmi H, Cohen K, Hurwitz JC, Brite J, Cimaroli S, et al. Predictors of in-hospital appendiceal perforation in patients with non-perforated acute appendicitis with appendicolithiasis at presentation. BMC Surg. 2023. doi:10.1186/s12893-023-02210-4.