Article

Not peer-reviewed version

Diagnostic Performance of Interleukin-6 (IL-6) and Membrane Glycoprotein Cluster of Differentiation64 (CD64) for Acute Appendicitis in Girls Presenting with Lower Abdominal Pain: A Case–Control Study

Eva Filo *, Vassileios Mouravas, Dimitrios Sfoungaris, Konstantina Kontopoulou, Asimina Fylaktou, Ioannis Valioulis *

Posted Date: 10 March 2026

doi: 10.20944/preprints202603.0822.v1

Keywords: appendicitis; interleukin-6 (IL-6); membrane glycoprotein cluster of differentiation-64 (CD64); abdominal pain; biomarkers; ROC; AUC; alvarado score

 doi:10.20944/preprints202603.0822.v1

Article

Diagnostic Performance of Interleukin-6 (IL-6) and Membrane Glycoprotein Cluster of Differentiation-64 (CD64) for Acute Appendicitis in Girls Presenting with Lower Abdominal Pain: A Case–Control Study

Eva Filo [1,*], Vassileios Mouravas [1], Dimitrios Sfoungaris [1], Konstantina Kontopoulou [2], Asimina Fylaktou [3] and Ioannis Valioulis [1,*]

[1] 1st Department of Pediatric Surgery, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[2] Laboratory of Microbiology, "G. Gennimatas" General Hospital, 54635 Thessaloniki, Greece
[3] Immunology Department, National Peripheral Histocompatibility Center, Hippocration General Hospital, Thessaloniki, Greece
* Correspondence: eva_filo@yahoo.com (E.F); Tel.: +30-6931-043-837 (E.F.); ivalioulis@auth.gr (I.V.) Tel.: +30-6946-260-288 (I.V.)

Abstract

Background: Acute appendicitis in girls presenting with lower abdominal pain represents a frequent diagnostic dilemma, given the overlap in clinical presentation with gynecological and non-surgical causes. This study aimed to evaluate the diagnostic performance of IL-6 and CD64 and to compare them with classical inflammatory markers and the Alvarado score. **Methods:** We conducted an observational case–control study over a three-year period (December 2022–December 2025) at the First University Paediatric Surgery Clinic (General Hospital of Thessaloniki "Georgios Gennimatas"). Consecutive girls aged ≤16 years presenting with lower abdominal pain were included. The primary outcome was the presence of appendicitis (yes/no), defined by the final clinical diagnosis and, where applicable, intraoperative and/or histopathological confirmation. Diagnostic performance was assessed using ROC curves/AUC with 95% confidence intervals estimated by the DeLong method. The prespecified primary model was a logistic regression including the Alvarado score and log1p(IL-6). **Results:** Of 74 initially assessed cases, one was excluded (appendiceal neuroendocrine tumour, NET G1), yielding a final sample of 73 girls: 37 with appendicitis and 36 without appendicitis. IL-6 was higher in the appendicitis group (median 19.41 vs 4.10 pg/mL) and showed moderate discrimination (AUC 0.696). CRP showed lower/borderline performance (AUC 0.595), whereas CD64 did not demonstrate useful discrimination (AUC 0.521). The Alvarado score had the highest discriminatory ability (AUC 0.885). Adding IL-6 to the Alvarado score did not materially improve the AUC in the common subset. **Conclusions:** IL-6 demonstrates moderate diagnostic performance as a standalone biomarker and may be useful as an adjunct, particularly when a clinical score is unavailable or unreliable. CD64 did not add diagnostic information in this setting. Larger, prespecified studies are required to identify clinically useful cut-offs.

Keywords: appendicitis; interleukin-6 (IL-6); membrane glycoprotein cluster of differentiation-64 (CD64); abdominal pain; biomarkers; ROC; AUC; alvarado score

1. Introduction

The term "vermiform appendix" has been traced back to the Egyptian civilization around 3000 BC, when organs were removed during mummification and stored in jars; some jars have been found bearing inscriptions interpreted as referring to an "intestinal worm". Owing to its apparent predisposition to inflammation, the appendix was long regarded as a vestigial organ [1]. The first

 doi:10.20944/preprints202603.0822.v1

recorded anatomical drawing of the appendix is attributed to Leonardo da Vinci in 1508 [2]. In 1886, the American physician Reginald Heber Fitz [1843–1913] introduced the term "appendicitis" in his monograph entitled "Perforating inflammation of the vermiform appendix; with special reference to its early diagnosis and treatment" [3]. Embryologically, development of the appendix begins during the fifth week of gestation and is closely linked to midgut development; by the eighth week of embryonic development, the appendix becomes macroscopically recognizable [4]. In topographic anatomy, the appendix arises from the posterior–medial aspect of the caecum, approximately 1.7 cm from the terminal ileum. Its base lies at the point of convergence of the taeniae coli, and the anterior taenia coli terminates at the base of the appendix [5].

Appendicitis is defined as inflammation of the vermiform appendix. Acute appendicitis is classified as "uncomplicated" or "complicated" by the European Association for Endoscopic Surgery [EAES]. Uncomplicated appendicitis is defined as inflammation in the absence of a phlegmon, free purulent fluid, abscess, or gangrene, whereas complicated appendicitis includes periappendiceal phlegmon with or without a pericaecal abscess, gangrene, or perforation [6]. Peak incidence occurs at 10–14 years in boys and 15–19 years in girls. The lifetime risk of developing acute appendicitis is estimated at 6.7% for women and 8.6% for men, while the lifetime risk of appendicectomy is reported as 23.1% for women and 12% for men [7].

During the acute inflammatory phase, typical histological features include mucosal ulceration, transmural neutrophil infiltration, perforation, and serositis; in chronic inflammation, lymphocytic infiltration predominates [8]. Regarding etiology in the pediatric population, acute appendicitis is most commonly attributed to lymphoid hyperplasia, characterised by excessive proliferation of lymphoid tissue within the appendix, leading to luminal obstruction, inflammation, and localised ischemia [9]. In the context of ongoing inflammation, particularly in the absence of timely intervention, complications may arise, including perforation, periappendiceal abscess, and peritonitis [10]. In such cases, the inflammatory cascade induces acute-phase mediators such as interleukin-6 [IL-6], tumour necrosis factor-α [TNF-α], and C-reactive protein [CRP], which have been associated with disease severity [11].

In clinical practice, the diagnosis of appendicitis in children relies on the clinical presentation, physical examination findings, and laboratory and imaging assessment. Several scoring systems that combine clinical features with laboratory values are available; the Pediatric Appendicitis Score and the Alvarado score are among the most widely used, with the Alvarado score being the most commonly applied in many settings [12]. In 1998, the World Health Organization defined a biomarker as "any substance, structure or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease" [13]. Serum biomarkers are objectively measurable and interpretable indicators that reflect physiological and pathological processes, as well as pharmacological responses to therapeutic interventions. They are used to identify, differentiate, and evaluate pathological states, to assess disease severity, and to guide the determination, monitoring, and prognostication of treatment response [14].

In a study published in BMC Surgery in 2006, Ulrich Sack and colleagues reported that white blood cell count, CRP, and IL-6 were directly associated with the severity of appendiceal inflammation in children. Identification of severe appendicitis was supported by IL-6 or CRP, but not by white blood cell count; nonetheless, CRP and IL-6 were considered complementary markers that may assist in identifying the need for prompt surgical management rather than serving as stand-alone diagnostic tests [11]. In a 2022 systematic review published in World Journal of Pediatrics, Arredondo Montero and colleagues concluded that the sensitivity and specificity of serum IL-6 for diagnosing uncomplicated acute appendicitis in the pediatric population are moderate, but appear higher for complicated appendicitis. They also noted an apparent association between serum IL-6 levels and the duration [in hours] of abdominal pain in children with acute appendicitis, and highlighted the need for further studies to evaluate this biomarker in distinguishing complicated from uncomplicated disease [15].

 doi:10.20944/preprints202603.0822.v1

Neutrophil CD64 expression has been investigated in recent years as a biomarker of infection and sepsis. It has features that support clinical applicability: under resting conditions, neutrophil CD64 expression is low, but following activation it increases markedly within a few hours. CD64 has been described as a promising sepsis biomarker, superior to CRP and potentially outperforming procalcitonin. Importantly, neutrophil CD64 appears to perform comparably across age groups, including adults, neonates, and infants [16].

Interleukin-6 [IL-6] was initially described under several names—B-cell stimulatory factor-2 [BSF-2], interferon-β2 [IFN-β2], hybridoma/plasmacytoma growth factor, macrophage–granulocyte inducer type 2, and hepatocyte stimulating factor [HGF]—reflecting its diverse biological activities. Following its cloning in 1986 and the characterization of its functional properties, these activities were attributed to a single cytokine, subsequently termed interleukin-6 [IL-6] [17]. IL-6 plays a central role in inflammation, acting as a key inducer of CRP, fibrinogen, and serum amyloid A, among many other mediators [17]. In a study of 137 children aged <15 years with appendicitis, Elliver and colleagues observed higher serum concentrations of IL-6 and IL-10 and lower concentrations of TNF-β in complicated appendicitis. Elevated serum IL-6 was associated with increased risk of complicated appendicitis, whereas serum IL-1α, IL-1β, IL-2, IL-10, IL-17A, and TNF-β were not similarly associated. The area under the ROC curve [AUC] for IL-6 was 0.75, indicating moderate ability to identify complicated acute appendicitis [18].

2. Materials and Methods

The objectives of this study were: (a) to evaluate the diagnostic performance of IL-6 and CD64 in discriminating appendicitis from non-appendicitis among girls presenting with lower abdominal pain; and (b) to assess whether IL-6 provides incremental diagnostic information beyond the Alvarado score and basic laboratory indices (WBC, CRP, and white cell differentials).

2.1. Study Design

We conducted an observational case–control study over a three-year period (December 2022 to December 2025) at the First University Paediatric Surgery Clinic, General Hospital of Thessaloniki "Georgios Gennimatas". Consecutive female pediatric cases presenting to the Emergency Department with suprapubic/lower abdominal pain were included, and all underwent a standardized clinical, laboratory, and imaging work-up.

2.1.1. Study population and Eligibility Criteria

Inclusion criteria were: (a) girls aged ≤16 years presenting with suprapubic pain; (b) absence of major comorbidities; and (c) completion of the initial diagnostic evaluation and follow-up until a final diagnosis was established. Exclusion criteria were: (a) oncology patients; (b) patients with major comorbidities or previous relevant lower abdominal surgery (as specified in the protocol); and (c) cases with an unsuitable blood sample for laboratory processing. In addition, one case operated on for suspected appendicitis but ultimately diagnosed with an appendiceal neuroendocrine tumour (NET G1) was excluded from the final analysis, as prespecified. (Figure 1).

 doi:10.20944/preprints202603.0822.v1

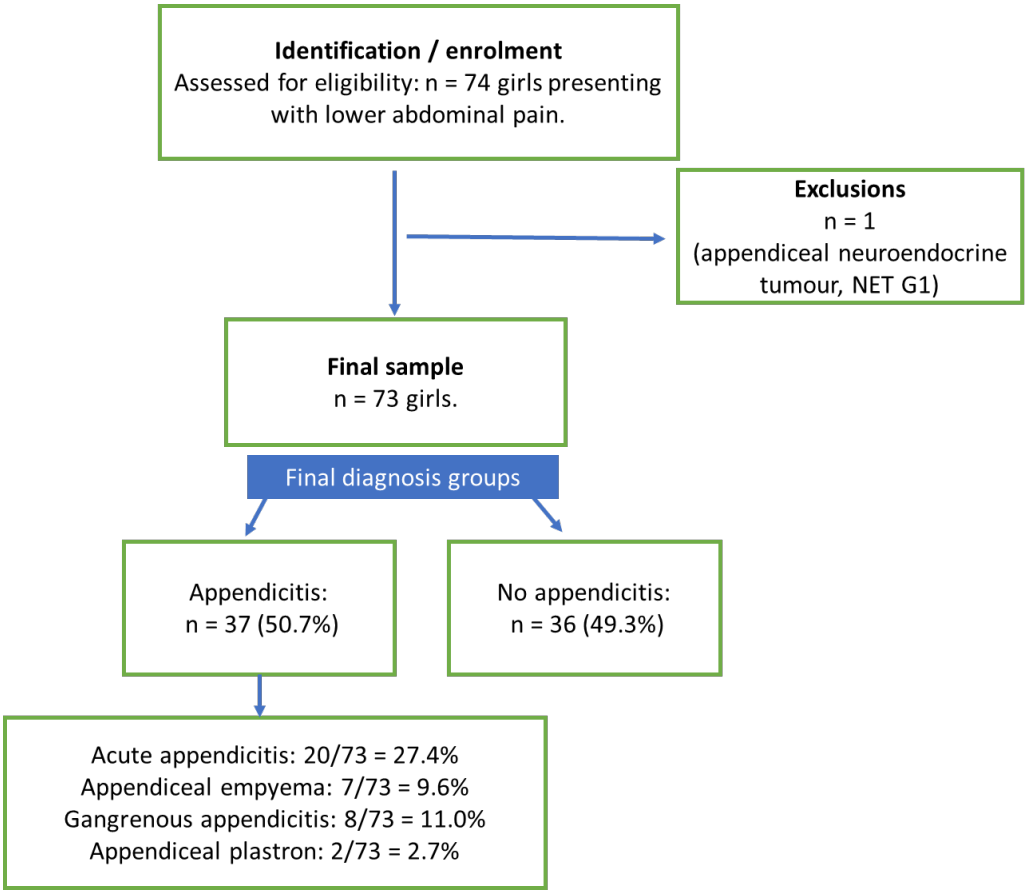

Figure 1. Study flow diagram.

2.1.2. Definition of Groups and Outcome

The primary diagnostic outcome was the presence of appendicitis (yes/no). The "appendicitis" group was defined based on the final clinical diagnosis and, where applicable, intraoperative and/or histopathological confirmation. The "non-appendicitis" group comprised other diagnoses (non-surgical and/or surgical causes other than appendicitis), as documented in the clinical records.

2.2. Variables and Indices

We recorded demographic and anthropometric data (age, weight, height, BMI), clinical signs/measurements (e.g., temperature), and routine laboratory parameters (e.g., WBC, lymphocyte percentage, CRP), as well as the biomarkers of interest: IL-6 (pg/mL) and CD64 (X-mean, arbitrary units). The Alvarado score was calculated according to the standard criteria.

2.3. Laboratory Measurements

IL-6 was measured using an immunochemiluminescent method on an automated immunoassay analyzer. CD64 was measured by flow cytometry at a collaborating laboratory (due to lack of availability at the reference hospital). Routine blood tests and CRP were performed using the laboratory's established automated methods (hematology/biochemistry assays and turbidimetry for CRP, where applicable).

2.4. Sample Size

The initial study design targeted 80% power using G*Power (version 3.1.9.7) for two independent groups (Wilcoxon–Mann–Whitney test), taking into account the rarity of certain surgical

 doi:10.20944/preprints202603.0822.v1

diagnoses specified in the protocol. For the present article, the final analyses were based on the available sample accrued during the recruitment period.

2.5. Statistical Analysis

Continuous variables are reported as mean ± SD when distributions were compatible with parametric approaches, or as median (IQR) when departures from normality were evident. Two-group comparisons used Welch's t-test or the Mann–Whitney U test, as appropriate. Effect sizes are reported as Hedges' g for mean differences and r for non-parametric comparisons, with standard interpretation by magnitude. Diagnostic performance of individual markers and models was evaluated using ROC curves and AUC, with 95% confidence intervals computed by the DeLong method [19].

The prespecified primary model was a logistic regression including the Alvarado score and log1p(IL-6), to address right-skewness in IL-6 and the presence of zero/very low values. Model performance was assessed using AUC in the common subset of observations with complete data for the included terms. Exploratory multivariable approaches without a clinical score were also examined, combining log1p(IL-6) with objective laboratory/clinical indicators (e.g., WBC, lymphocyte percentage, temperature, CD64), and evaluated via AUC. In addition, a clinically motivated IL-6 cut-off (≥7 pg/mL) was assessed, with sensitivity/specificity reported in the results. Analyses were performed using an available-case approach, with explicit reporting of the sample size (N) for each marker/model (e.g., differing N for IL-6, CD64, CRP, temperature). To assess the influence of extreme values, winsorization at the 99th percentile was applied to IL-6 and CD64, and selected ROC models were re-evaluated; this did not materially change the conclusions.

2.6. Software

Statistical analyses were implemented in Python 3.11.2 (main, Apr 28 2025, 14:11:48) [GCC 12.2.0] and automated via custom scripts. The following packages were used: pandas 2.2.3 (data management/cleaning), NumPy 1.24.0 (numerical computation), SciPy 1.14.1 (statistical tests/utilities), statsmodels 0.14.3 (logistic regression/Logit and calibration indices), scikit-learn 1.4.2 (ROC/AUC, Brier score), and matplotlib 3.7.5 (figures).

3. Results

Over the three-year study period (December 2022 to December 2025) at the First University Paediatric Surgery Clinic, General Hospital of Thessaloniki "Georgios Gennimatas", and in collaboration with the laboratory of General Hospital of Thessaloniki "Ippokrateio" for the analysis of the serum membrane glycoprotein cluster of differentiation-64 (CD64)—given that this assay could not be supported by our hospital—the following results were obtained. In total, 74 girls aged 4.5 to 15.5 years presented with suprapubic/lower abdominal pain and were enrolled. Thirty-seven cases were classified as appendicitis, including 20 with acute appendicitis, 7 with appendiceal empyema, 8 with gangrenous appendicitis, and 2 with appendiceal plastron. One patient in the appendicitis category who underwent surgery for suspected acute appendicitis was found on histopathological examination to have an appendiceal neuroendocrine tumour (NET G1) and was therefore excluded from the study.

3.1. Demographic and Anthropometric Characteristics

The final analytic sample comprised 73 girls presenting with suprapubic/lower abdominal pain. Of these, 37 (50.7%) were assigned to the appendicitis group and 36 (49.3%) to the non-appendicitis group. The distribution of participants across the two groups was essentially balanced, facilitating interpretation of between-group comparisons. Age was similarly distributed between groups. The median age was 13.0 years (IQR 11.0–14.0) in the appendicitis group and 12.0 years (IQR 9.38–14.0) in the non-appendicitis group, with no statistically significant difference ($p = 0.424$) and a very small

 doi:10.20944/preprints202603.0822.v1

effect size (r = −0.109). A similar pattern was observed for anthropometric parameters. Body weight, height, and body mass index (BMI) did not differ significantly between groups (p = 0.581, p = 0.390, and p = 0.168, respectively), and the corresponding effect sizes were small (Hedges' g = −0.129 for weight, g = 0.200 for height, and g = −0.333 for BMI). Regarding data completeness, BMI was unavailable in four cases; therefore, descriptive analyses for BMI were based on available observations (n = 36 in the appendicitis group and n = 33 in the non-appendicitis group). Overall, the demographic and anthropometric profiles indicate that the two groups were broadly comparable at baseline (Figure 2).

Table 1. This is a table with the variables.

Variable	Appendicitis (n = 37)	Non-appendicitis (n = 36)	p-value	Statistical test	Effect size
Age (years), median (IQR)	13.00 (11.00–14.00)	12.00 (9.38–14.00)	0.424	Mann–Whitney U	r = −0.109
Weight (kg), mean ± SD	45.97 ± 13.85	47.88 ± 15.39	0.581	Welch's t-test	g = −0.129
Height (cm), mean ± SD	151.05 ± 14.52	147.97 ± 15.89	0.390	Welch's t-test	g = 0.200
BMI (kg/m²), mean ± SD	19.87 ± 4.00 (n = 36)	21.27 ± 4.34 (n = 33)	0.168	Welch's t-test	g = −0.333

Note: Age is reported as median (IQR). Weight/height/BMI are reported as mean ± SD. BMI had missing values in 4 cases.

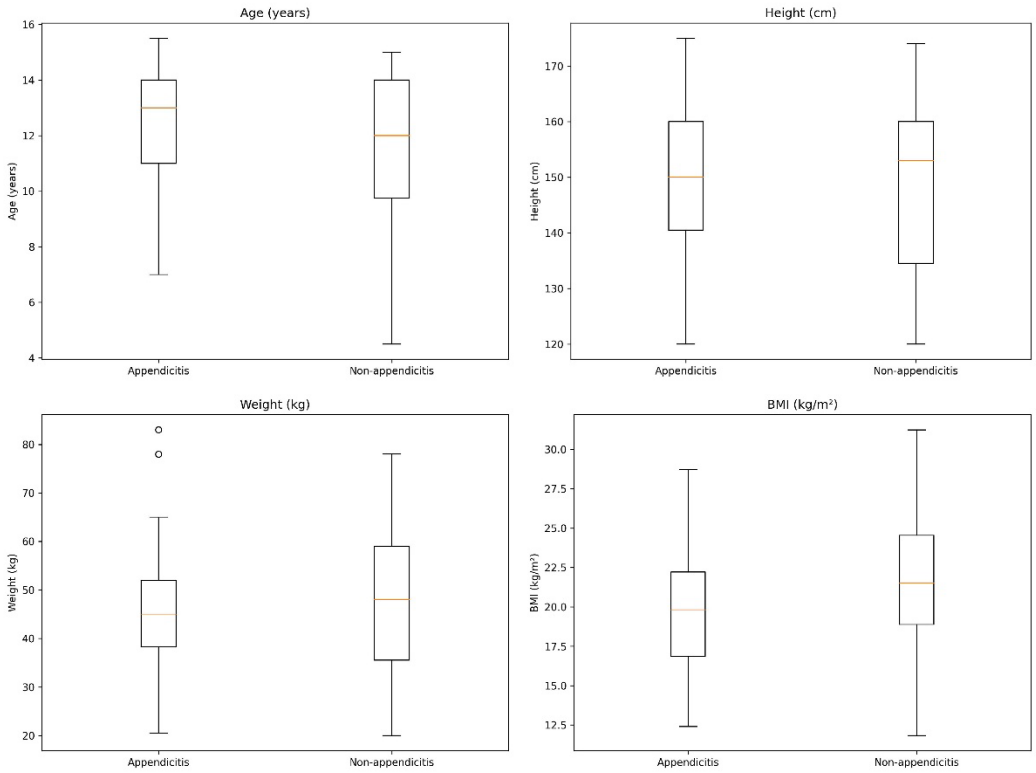

Figure 2. Boxplots of demographic and anthropometric characteristics by group.

 doi:10.20944/preprints202603.0822.v1

The clinical characteristics at presentation of girls with suprapubic/lower abdominal pain are shown in Table 4.2, comparing the appendicitis group (n = 37) with the non-appendicitis group (n = 36).

Table 2. Baseline clinical characteristics by final diagnosis.

Variable / Category	Appendicitis (n = 37)	Non-appendicitis (n = 36)	p-value (test)	Effect size
Pain duration (categories) (N = 37/35)			p = 0.703§	V = 0.202
– ≤12 h	13 (35.1%)	14 (38.9%)		
– >12–24 h	11 (29.7%)	8 (22.2%)		
– ≥7 days	2 (5.4%)	1 (2.8%)		
– >24–48 h	4 (10.8%)	7 (19.4%)		
– >48 h–≤7 days	7 (18.9%)	5 (13.9%)		
– Not recorded	0 (0.0%)	1 (2.8%)		
Temperature (N = 32/35)			p = 0.087§	V = 0.334
– Afebrile/no fever	13 (35.1%)	23 (63.9%)		
– <37.5°C	5 (13.5%)	5 (13.9%)		
– 37.5–37.9°C	6 (16.2%)	2 (5.6%)		
– ≥38.0°C	8 (21.6%)	5 (13.9%)		
– Not recorded	5 (13.5%)	1 (2.8%)		
Menarche (N = 34/31)			p = 0.708§	V = 0.097
– Yes	22 (59.5%)	21 (58.3%)		
– No	12 (32.4%)	10 (27.8%)		
– Not recorded	3 (8.1%)	5 (13.9%)		
Alvarado score (N = 37/36), median (IQR)	8.00 (6.00–9.00)	4.00 (3.00–6.00)	p < 0.001‡	r = 0.770

Note: Categorical variables are presented as n (%). For categorical variables, the header row also indicates the number of available observations per group (N = appendicitis/non-appendicitis). The "p-value (test)" column reports the p-value and the test as a symbol: § chi-square test, ¶ Fisher's exact test, † Welch's t-test, ‡ Mann–Whitney U test. Effect sizes: V = Cramer's V (categorical), g = Hedges' g (t-test), r = rank-biserial correlation (Mann–Whitney). "Not recorded" categories indicate missing documentation for that variable.

Haematological/biochemical parameters and biomarkers at admission are presented in Table 4.3. For each variable, descriptive statistics (mean ± SD or median [IQR]) were selected according to the within-group distribution, while between-group comparisons were performed using appropriate tests (Welch's t-test or the Mann–Whitney U test) and are accompanied by effect sizes. The distributions of these markers are illustrated in Figure 4.3.

Table 3. Haematological/biochemical parameters and biomarkers by final diagnosis.

Variable	**Appendicitis**	**Non-appendicitis**	**p-value (test)**	**Effect size**
WBC (×10³/μL)	15.00 (10.10–18.30) (N = 37)	9.75 (7.08–13.83) (N = 36)	p = 0.002‡	r = 0.428
Neutrophils (NE, %)	75.35 ± 12.08 (N = 37)	68.10 ± 13.27 (N = 36)	p = 0.017†	g = 0.566
Lymphocytes (LY, %)	16.67 ± 9.73 (N = 37)	21.76 ± 10.89 (N = 36)	p = 0.039†	g = –0.488

 doi:10.20944/preprints202603.0822.v1

Monocytes (MO, %)	7.00 (5.60–8.50) (N = 37)	7.05 (5.90–9.00) (N = 36)	p = 0.761‡	r = −0.042
Haematocrit (HCT, %)	38.94 ± 3.20 (N = 37)	38.33 ± 2.87 (N = 36)	p = 0.393†	g = 0.199
Haemoglobin (HGB, g/dL)	13.00 (12.30–13.50) (N = 37)	12.70 (12.20–13.70) (N = 36)	p = 0.611‡	r = 0.070
Platelets (PLT, ×10³/μL)	304.00 (251.00–333.00) (N = 37)	296.50 (262.50–326.50) (N = 36)	p = 0.830‡	r = −0.030
CRP (mg/dL)	0.60 (0.10–3.60) (N = 37)	0.10 (0.03–2.42) (N = 34)	p = 0.167‡	r = 0.190
IL-6 (pg/mL)	19.41 (3.33–33.13) (N = 34)	4.10 (1.88–13.00) (N = 34)	p = 0.006‡	r = 0.391
Neutrophil CD64 (X-mean, a.u.)	6.98 (5.22–12.47) (N = 34)	6.75 (4.12–14.15) (N = 35)	p = 0.769‡	r = 0.042

Note: Continuous variables are presented as mean ± SD († Welch's t-test; effect size g = Hedges' g) or as median (IQR) (‡ Mann–Whitney U test; effect size r = rank-biserial correlation), according to within-group normality. N corresponds to the number of available observations per group.

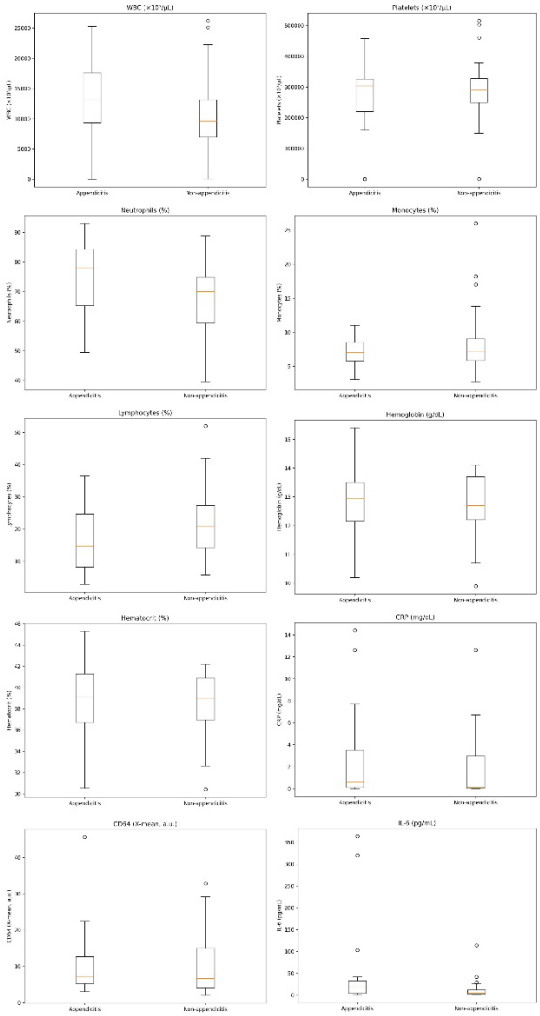

Figure 2. Boxplots of haematological/biochemical parameters and biomarkers at admission by final diagnosis (appendicitis vs non-appendicitis). Circles indicate outliers (values beyond 1.5×IQR from the quartiles).

 doi:10.20944/preprints202603.0822.v1

3.2. Exploratory Analysis of IL-6 and CD64

In this cohort of girls presenting with suprapubic/lower abdominal pain (N = 73), IL-6 showed moderate univariable discriminatory ability for distinguishing appendicitis from non-appendicitis (AUC = 0.696, N = 68), outperforming CRP (AUC = 0.595, N = 71) and clearly exceeding CD64, which did not provide clinically useful discrimination in this dataset (AUC = 0.521, N = 69).

Within multivariable models that did not incorporate a clinical score, IL-6 retained a consistent contribution when combined with objective laboratory/clinical indicators, with the best-performing specifications achieving overall moderate discrimination (e.g., log1p(IL-6) + WBC + LY%: AUC = 0.733, N = 68; log1p(IL-6) + WBC + temperature: AUC = 0.697, N = 63; log1p(IL-6) + CD64 + temperature: AUC = 0.692, N = 61). By contrast, when the Alvarado score was available, adding IL-6 did not meaningfully improve the discriminatory performance of the clinical model (AUC = 0.885 for Alvarado alone and AUC = 0.885 for Alvarado + log1p(IL-6) in the common subset), suggesting that the practical value of IL-6 in this setting lies primarily as an adjunct laboratory marker in scenarios where a clinical score is unavailable or cannot be reliably calculated, or when a standardised laboratory-based support of decision-making is required.

Sensitivity analyses addressing the influence of extreme values (winsorization at the 99th percentile for IL-6 and CD64) did not materially alter the conclusions (for example, AUC for WBC + log1p(IL-6) + CD64 was 0.732 before and 0.738 after winsorization), supporting the robustness of these findings.

Table 4. Haematological/biochemical parameters and biomarkers by final diagnosis.

Marker	N	AUC	95% CI (DeLong)
Alvarado score	73	0.885	0.811–0.959
WBC (×10³/μL)	73	0.714	0.594–0.834
IL-6 (pg/mL)	68	0.696	0.568–0.823
CRP (mg/dL)	71	0.595	0.461–0.729
CD64 (X-mean, a.u.)	69	0.521	0.382–0.660
Temperature (°C)	67	0.658	0.536–0.780

Table 5. Multivariable logistic regression models for the diagnosis of appendicitis.

Predictors	N	AUC (95% CI)	Odds ratios (OR, 95% CI)	p-values
Alvarado	73	0.885 (0.811–0.959)	Alvarado: 2.279 (1.585–3.277); Intercept: 0.008 (0.001–0.071)	Alvarado: <0.001; Intercept: <0.001
Alvarado + log1p(IL-6)	68	0.885 (0.805–0.964)	Alvarado: 2.244 (1.514–3.326); log1p(IL-6): 1.031 (0.604–1.760); Intercept: 0.008 (0.001–0.079)	Alvarado: <0.001; log1p(IL-6): 0.912; Intercept: <0.001

 doi:10.20944/preprints202603.0822.v1

WBC + log1p(IL-6) + CD64	65	0.732 (0.608–0.856)	WBC: 1.078 (0.971–1.196); log1p(IL-6): 1.767 (1.064–2.935); CD64: 0.963 (0.892–1.039); Intercept: 0.156 (0.033–0.743)	WBC: 0.159; log1p(IL-6): 0.028; CD64: 0.327; Intercept: 0.020
log1p(IL-6) + CD64 + Temperature	61	0.692 (0.555–0.829)	log1p(IL-6): 1.753 (1.020–3.014); CD64: 0.935 (0.857–1.020); Temperature: 2.025 (0.693–5.913); Intercept: 0.000 (0.000–214907.652)	log1p(IL-6): 0.042; CD64: 0.131; Temperature: 0.197; Intercept: 0.178
log1p(IL-6) + WBC + Temperature	63	0.697 (0.566–0.828)	log1p(IL-6): 1.428 (0.845–2.412); WBC: 1.081 (0.966–1.209); Temperature: 1.388 (0.577–3.341); Intercept: 0.000 (0.000–101480781.803)	log1p(IL-6): 0.183; WBC: 0.173; Temperature: 0.464; Intercept: 0.396
log1p(IL-6) + WBC + LY%	68	0.733 (0.610–0.855)	log1p(IL-6): 1.600 (0.987–2.593); WBC: 1.103 (0.977–1.245); LY%: 1.003 (0.939–1.072); Intercept: 0.096 (0.005–1.807)	log1p(IL-6): 0.056; WBC: 0.112; LY%: 0.918; Intercept: 0.118

Notes: Outcome: appendicitis (1) vs non-appendicitis (0). Each model was analysed on a complete-case basis; therefore, N varies across models. log1p(IL-6) = ln (IL-6 + 1). AUCs with 95% confidence intervals were estimated using the DeLong method. Units: WBC ×10³/μL, temperature °C, IL-6 pg/mL, CD64 as X-mean (a.u.), LY% lymphocyte percentage.

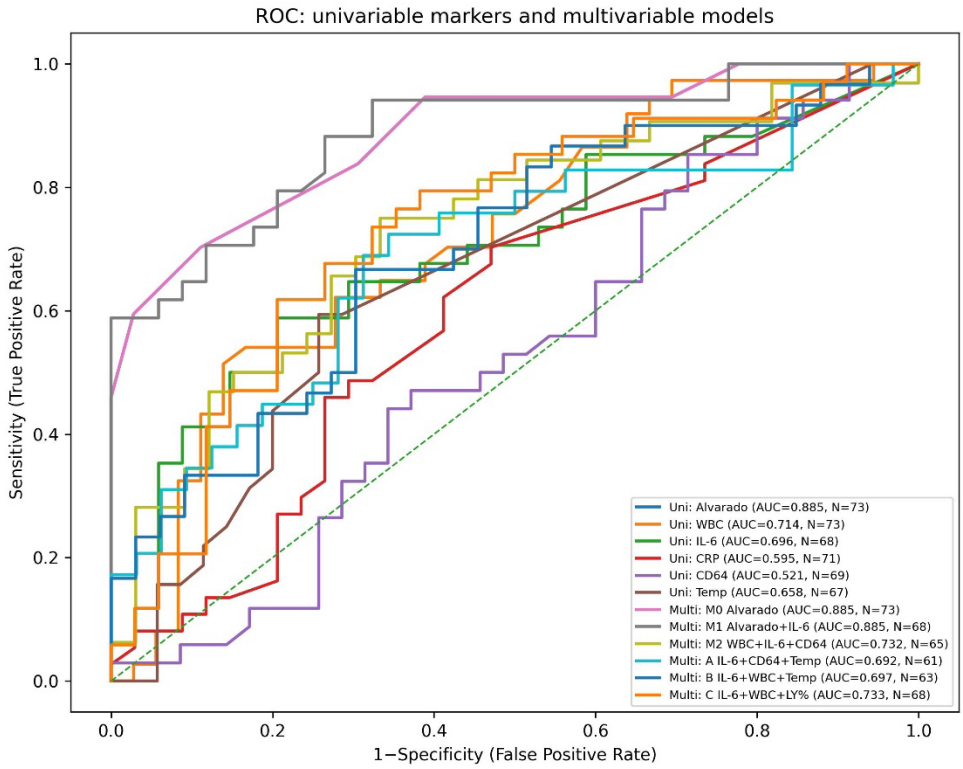

Figure 3. Receiver operating characteristic (ROC) curves for univariable markers and multivariable logistic regression models for diagnosing appendicitis. Univariable ROC curves are shown for the Alvarado score, white blood cell count (WBC), interleukin-6 (IL-6), C-reactive protein (CRP), neutrophil CD64 (X-mean), and temperature. Multivariable models include: M0 (Alvarado), M1 (Alvarado + log1p[IL-6]), M2 (WBC + log1p[IL-6] + CD64), M3 (log1p[IL-6] + CD64 + temperature), M4 (log1p[IL-6] + WBC + temperature), and M5 (log1p[IL-6] + WBC + lymphocyte percentage). The diagonal dashed line indicates no-discrimination performance. AUC values and sample sizes (N) for each curve are reported in the legend; 95% confidence intervals were estimated using the DeLong method. log1p(IL-6) denotes ln (IL-6 + 1).

 doi:10.20944/preprints202603.0822.v1

4. Discussion

The present study was informed by the rationale of the work by Reed JL et al. (2011) [19] in Academic Emergency Medicine, which examined the contribution of biomarkers to the differential diagnosis of surgical causes in girls presenting with suprapubic/lower abdominal pain. Lower abdominal pain is a common reason for presentation in paediatric and adolescent populations, and acute appendicitis remains a clinical challenge, particularly when the clinical spectrum overlaps with gynaecological or non-surgical causes.

IL-6 is a pleiotropic pro-inflammatory cytokine with a central role in the acute-phase inflammatory response, inducing the synthesis of acute-phase proteins such as CRP, serum amyloid A, and fibrinogen. In several clinical contexts, IL-6 increases earlier than other markers, which has led to its investigation as an early biomarker (e.g., in sepsis) [22,23]. This biological rationale supports the hypothesis that IL-6 may provide complementary information in acute appendicitis, especially at early stages or in cases with an equivocal clinical picture.

In the study cohort (December 2022 to December 2025), 37 girls had a final diagnosis of appendicitis and 36 did not. Univariable diagnostic performance indicated that the Alvarado score had the highest discriminatory ability, a finding consistent with its clinical utility as a structured tool for estimating disease probability. IL-6 demonstrated moderate discrimination, indicating that, as a standalone biomarker, it provides clinically relevant information but is insufficient on its own to support a confident diagnostic decision. Similarly, WBC showed moderate discrimination, consistent with the expected rise in leukocytosis during acute inflammation. By contrast, CRP performed less well, which is clinically plausible given that CRP may rise more slowly than earlier inflammatory markers. CD64 showed borderline discriminatory performance, without practically useful separation in the present sample.

The limited performance of CD64 in our study warrants cautious interpretation, as this marker has been primarily established in the context of bacterial infection and sepsis. CD64 (FcγRI) is expressed on monocytes and neutrophils and increases following cellular activation, making it an attractive infection biomarker [23]. However, its clinical application involves practical challenges (method/analyser requirements, lack of fully harmonised units, potential inter-laboratory variation and cut-offs). Moreover, in a clinical scenario such as acute appendicitis, a “sepsis-like” activation pattern may not be uniformly present across cases, particularly in early or uncomplicated episodes. Therefore, the borderline performance of CD64 in this cohort does not negate its broader biological relevance, but suggests that it did not add incremental diagnostic information in this specific population and measurement setting.

In relation to the international literature, several studies have shown that IL-6 tends to rise more markedly in complicated/perforated appendicitis and may perform better when the clinical question is discrimination between complicated and uncomplicated disease [25–28]. Associations between IL-6 and other indices (CRP, neutrophil measures, total WBC) have also been reported, as well as a potential relationship with symptom duration, which is clinically plausible given the dynamic nature of the inflammatory response [26,28]. In our dataset, we did not stratify by severity/complications of appendicitis; therefore, the observed performance of IL-6 primarily reflects discrimination of “appendicitis versus non-appendicitis”, rather than severity.

A central finding of this work is that, although IL-6 has biological and clinical plausibility as a marker of acute inflammation, its addition to the Alvarado score did not demonstrate meaningful incremental diagnostic value in this sample. In the common subset with complete data, improvement in model fit with IL-6 versus Alvarado alone was marginal and did not support a clear upgrade in performance. Practically, the Alvarado score already functions as a strong clinical classifier, leaving limited scope for measurable added value from a single biomarker. In addition, IL-6 likely overlaps with inflammatory and clinical features captured by the score (directly or indirectly), thereby reducing its independent contribution when the score is available. Nonetheless, this does not preclude targeted usefulness in specific scenarios, such as “intermediate” clinical probability (e.g., an Alvarado “grey zone”), where decision-making is more uncertain—provided that this is evaluated prospectively in larger, prespecified cohorts.

 doi:10.20944/preprints202603.0822.v1

Finally, several limitations should be considered. The sample size is limited for detailed subgroup analyses, there were missing values for certain markers (e.g., IL-6/CD64/CRP/temperature), and CD64 is subject to laboratory heterogeneity in terms of standardisation. Despite these constraints, the study provides clear, practice-relevant messages: the Alvarado score remains a central tool, IL-6 may be useful as an adjunct, and CD64 did not support clinically meaningful contribution in this setting.

5. Conclusions

In conclusion, IL-6 emerged as a biomarker with moderate discriminatory ability for the diagnosis of acute appendicitis in girls presenting with suprapubic/lower abdominal pain, supporting its potential to provide clinically useful information as a laboratory adjunct to the initial assessment. Its performance as a single marker was not sufficient for stand-alone diagnostic use; however, its biological relevance to the early inflammatory response and the observed performance in this cohort support its role as an ancillary tool, particularly in cases with an equivocal clinical presentation or when decision-making requires rapid strengthening of the probability of disease.

In multivariable analyses, adding IL-6 to the Alvarado score did not demonstrate a clear, measurable improvement in overall diagnostic performance in this sample. This likely reflects the already high discriminatory capacity of the clinical score and a degree of conceptual overlap between the biological information captured by IL-6 and features of the clinical inflammatory phenotype. Nevertheless, the findings suggest that laboratory-based combinations including IL-6 alongside objective haematological indices (such as WBC and lymphocyte percentage) can achieve moderate discrimination, providing an alternative approach when a clinical score is unavailable or when a standardised, measurement-based evaluation is preferred.

The practical value of IL-6 lies in the fact that, as a marker of early inflammatory activation, it can support diagnostic documentation through a simple blood draw and reduce uncertainty in selected clinical scenarios. Future studies with larger samples, prespecified evaluation in "intermediate" clinical probability categories (e.g., grey-zone clinical scores), and stratification by severity/complications are required to clarify when—and at what thresholds—IL-6 offers maximal clinical benefit, and whether it can be incorporated into standardised diagnostic pathways for paediatric and adolescent patients.

The knowledge gained after the completion of this study advances existing knowledge not only by utilizing these biomarkers in daily clinical practice so that there is a more immediate diagnosis but also helps future research in this area. In the future, it would be useful to utilize these biomarkers in countries with low socioeconomic levels so that they have more direct access to scientific laboratory documentation and diagnosis. Finally, the way seems to be opened for the search for other biomarkers for targeted diseases that concern young patients.

Author Contributions: "Conceptualization, E.F; methodology, E.F.; software, E.F.; validation, E.F.; formal analysis, E.F.; investigation, E.F.; resources, E.F.; data curation, E.F; writing—original draft preparation, E.F; writing—review and editing, E.F.; visualization, E.F.; supervision, E.F. and I.V.; project administration, E.F.; funding acquisition, E.F. All authors have read and agreed to the published version of the manuscript."

Funding: "This research received no external funding"

Institutional Review Board Statement: "The study was conducted in accordance with the Declaration of Helsinki, and approved by the Bioethics and Ethics Committee of the Medical School of the Aristotle University of Thessaloniki on the session 7 /20.6.2023 (approval code: 202/20.6.2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest

 doi:10.20944/preprints202603.0822.v1

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	Definition
AUC	Area under the curve
a.u.	Arbitrary units
BMI	Body mass index
BSF-2	B-cell stimulatory factor-2
CD64	Cluster of differentiation 64
CI	Confidence interval
CRP	C-reactive protein
EAES	European Association for Endoscopic Surgery
FcγRI	Fc gamma receptor I
HCT	Haematocrit
HGB	Haemoglobin
HGF	Hepatocyte stimulating factor
IFN-β2	Interferon beta-2
IL-6	Interleukin-6
IL-10	Interleukin-10
IQR	Interquartile range
log1p(IL-6)	Ln (IL-6 + 1)
LY%	Lymphocyte percentage
MO	Monocytes
NE	Neutrophils
NET	Neuroendocrine tumour
OR	Odds ratio
PAS	Pediatric Appendicitis Score
PLT	Platelets
ROC	Receiver operating characteristic
SD	Standard deviation
TNF-α	Tumour necrosis factor alpha
TNF-β	Tumour necrosis factor beta
WBC	White blood cell count

References

1. Bhasin, S.K.; Khan, A.B.; Kumar, V.; Sharma, S.; Saraf, R. Vermiform Appendix and Acute Appendicitis. *JK Science* 2007, 9, No. 4.
2. Herrod, P.J.J.; Kwok, A.T.; Lobo, D.N. Three Centuries of Appendicectomy. *World J. Surg.* 2023, 47, 928–936. https://doi.org/10.1007/s00268-022-06874-6
3. Constantin, M.; Petrescu, L.; Mătanie, C.; Vrancianu, C.O.; Niculescu, A.G.; Andronic, O.; Bolocan, A. The Vermiform Appendix and Its Pathologies. *Cancers* 2023, 15, 3872. https://doi.org/10.3390/cancers15153872
4. Skandalakis, J.E.; Skandalakis, P.N.; Skandalakis, L.J. *Χειρουργική Ανατομική και Τεχνική*; Ιατρικές Εκδόσεις Σιώκη: Athens, Greece, 2000.
5. Téoule, P.; Laffolie, J.; Rolle, U.; Reissfelder, C. Acute Appendicitis in Childhood and Adulthood. *Dtsch. Arztebl. Int.* 2020, 117, 764–774. https://doi.org/10.3238/arztebl.2020.0764
6. Zavras, N.; Vaos, G. Management of Complicated Acute Appendicitis in Children: Still an Existing Controversy. *World J. Gastrointest. Surg.* 2020, 12, 129–137. https://doi.org/10.4240/wjgs.v12.i4.129
7. Kooij, I.A.; Sahami, S.; Meijer, S.L.; Buskens, C.J.; te Velde, A.A. The Immunology of the Vermiform Appendix: A Review of the Literature. *Clin. Exp. Immunol.* 2016, 186, 1–9. https://doi.org/10.1111/cei.12821
8. Lotfollahzadeh, S.; Lopez, R.A.; Deppen, J.G. Appendicitis. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, USA, 2025. Available online: https://www.ncbi.nlm.nih.gov/books/NBK493193/ (accessed on 28 February 2026).
9. Sack, U.; Biereder, B.; Elouahidi, T.; Bauer, K.; Keller, T.; Tröbs, R.B. Diagnostic Value of Blood Inflammatory Markers for Detection of Acute Appendicitis in Children. *BMC Surg.* 2006, 6, 15. https://doi.org/10.1186/1471-2482-6-15

 doi:10.20944/preprints202603.0822.v1

10. Glass, C.C.; Rangel, S.J. Overview and Diagnosis of Acute Appendicitis in Children. *Semin. Pediatr. Surg.* 2016, 25, 198–203. https://doi.org/10.1053/j.sempedsurg.2016.05.001
11. Lagarou, E. *Biomarkers of Non-Alcoholic Fatty Liver Disease*. Diploma Thesis, Department of Biochemistry and Biotechnology, [Institution not stated], Greece, 2014.
12. Hung, S.K.; Lan, H.M.; Han, S.T.; Wu, C.C.; Chen, K.F. Current Evidence and Limitation of Biomarkers for Detecting Sepsis and Systemic Infection. *Biomedicines* 2020, 8, 494. https://doi.org/10.3390/biomedicines8110494
13. Arredondo Montero, J.; Bardají Pascual, C.; Bronte Anaut, M.; López-Andrés, N.; Antona, G.; Martín-Calvo, N. Diagnostic Performance of Serum Interleukin-6 in Pediatric Acute Appendicitis: A Systematic Review. *World J. Pediatr.* 2022, 18, 91–99. https://doi.org/10.1007/s12519-021-00488-z
14. Hoffmann, J.J. Neutrophil CD64 as a Sepsis Biomarker. *Biochem. Med. (Zagreb)* 2011, 21, 282–290. https://doi.org/10.11613/BM.2011.038
15. Kishimoto, T. Interleukin-6: Discovery of a Pleiotropic Cytokine. *Arthritis Res. Ther.* 2006, 8 (Suppl. 2), S2. https://doi.org/10.1186/ar1916
16. Elliver, M.; Salö, M.; Roth, B.; Ohlsson, B.; Hagander, L.; Gudjonsdottir, J. Associations between Th1-Related Cytokines and Complicated Pediatric Appendicitis. *Sci. Rep.* 2024, 14, 4613. https://doi.org/10.1038/s41598-024-53756-z
17. Reed, J.L.; Strait, R.T.; Kachelmeyer, A.M.; Byczkowski, T.L.; Ho, M.L.; Huppert, J.S. Biomarkers to Distinguish Surgical Etiologies in Females with Lower Quadrant Abdominal Pain. *Acad. Emerg. Med.* 2011, 18, 686–691. https://doi.org/10.1111/j.1553-2712.2011.01110.x
18. Cohen, S.B.; Wattiez, A.; Stockheim, D.; Seidman, D.S.; Lidor, A.L.; Mashiach, S.; Goldenberg, M. The Accuracy of Serum Interleukin-6 and Tumour Necrosis Factor as Markers for Ovarian Torsion. *Hum. Reprod.* 2001, 16, 2195–2197. https://doi.org/10.1093/humrep/16.10.2195
19. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* 1988, 44, 837–845.
20. Daponte, A.; Pournaras, S.; Hadjichristodoulou, C.; Lialios, G.; Kallitsaris, A.; Maniatis, A.N.; Messinis, I.E. Novel Serum Inflammatory Markers in Patients with Adnexal Mass Who Had Surgery for Ovarian Torsion. *Fertil. Steril.* 2006, 85, 1469–1472. https://doi.org/10.1016/j.fertnstert.2005.10.056
21. Röhrig, B.; du Prel, J.B.; Wachtlin, D.; Blettner, M. Types of Study in Medical Research: Part 3 of a Series on Evaluation of Scientific Publications. *Dtsch. Arztebl. Int.* 2009, 106, 262–268. https://doi.org/10.3238/arztebl.2009.0262
22. Tanaka, T.; Narazaki, M.; Kishimoto, T. IL-6 in Inflammation, Immunity, and Disease. *Cold Spring Harb. Perspect. Biol.* 2014, 6, a016295. https://doi.org/10.1101/cshperspect.a016295
23. Hung, S.K.; Lan, H.M.; Han, S.T.; Wu, C.C.; Chen, K.F. Current Evidence and Limitation of Biomarkers for Detecting Sepsis and Systemic Infection. *Biomedicines* 2020, 8, 494. https://doi.org/10.3390/biomedicines8110494
24. Nossuli, T.O.; Lakshminarayanan, V.; Baumgarten, G.; Taffet, G.E.; Ballantyne, C.M.; Michael, L.H.; Entman, M.L. A Chronic Mouse Model of Myocardial Ischemia-Reperfusion: Essential in Cytokine Studies. *Am. J. Physiol. Heart Circ. Physiol.* 2000, 278, H1049–H1055. https://doi.org/10.1152/ajpheart.2000.278.4.H1049
25. Lin, W.Y.; Lee, E.P.; Chen, C.Y.; Guo, B.C.; Lin, M.J.; Wu, H.P. Changes in Levels of Serum Cytokines and Chemokines in Perforated Appendicitis in Children. *Int. J. Mol. Sci.* 2024, 25, 6076. https://doi.org/10.3390/ijms25116076
26. Arredondo Montero, J.; Antona, G.; Rivero Marcotegui, A.; Bardají Pascual, C.; Bronte Anaut, M.; Ros Briones, R.; Fernández-Celis, A.; López-Andrés, N.; Martín-Calvo, N. Discriminatory Capacity of Serum Interleukin-6 between Complicated and Uncomplicated Acute Appendicitis in Children: A Prospective Validation Study. *World J. Pediatr.* 2022, 18, 810–817. https://doi.org/10.1007/s12519-022-00598-2
27. Elliver, M.; Salö, M.; Roth, B.; Ohlsson, B.; Hagander, L.; Gudjonsdottir, J. Associations between Th1-Related Cytokines and Complicated Pediatric Appendicitis. *Sci. Rep.* 2024, 14, 4613. https://doi.org/10.1038/s41598-024-53756-z

 doi:10.20944/preprints202603.0822.v1

28. Arredondo Montero, J.; Bardají Pascual, C.; Bronte Anaut, M.; López-Andrés, N.; Antona, G.; Martín-Calvo, N. Diagnostic Performance of Serum Interleukin-6 in Pediatric Acute Appendicitis: A Systematic Review. *World J. Pediatr.* 2022, 18, 91–99. https://doi.org/10.1007/s12519-021-00488-z
29. Di Mitri, M.; Parente, G.; Bonfiglioli, G.; Thomas, E.; Bisanti, C.; Cordola, C.; Vastano, M.; Cravano, S.; Collautti, E.; Di Carmine, A.; D'Antonio, S.; Gargano, T.; Libri, M.; Lima, M. IL-6 Serum Levels Can Enhance the Diagnostic Power of Standard Blood Tests for Acute Appendicitis. *Children* 2022, 9, 1425. https://doi.org/10.3390/children9101425
30. Fang, D.-H.; Fan, C.-H.; Li, J.; An, Q.; Yao, H.; Ji, Q.; Niu, G. Ratios of CD64 Expressed on Neutrophils, Monocytes, and Lymphocytes May Be a Novel Method for Diagnosis of Neonatal Sepsis. *J. Infect. Dev. Ctries* 2015, 9, 175–181. https://doi.org/10.3855/jidc.4992
31. Liang, J.; Xu, J.; You, L.; Yang, G.; Niu, G.; Pan, H.; Yu, S.; Mai, X.; Xu, J.; Zhang, F. Neutrophil CD64: A Potential Biomarker for the Diagnosis of Infection in Patients with Haematological Malignancies. *Hematology* 2021, 26, 970–975. https://doi.org/10.1080/16078454.2021.2003064
32. Cong, S.; Ma, T.; Di, X.; Tian, C.; Zhao, M.; Wang, K. Diagnostic Value of Neutrophil CD64, Procalcitonin, and Interleukin-6 in Sepsis: A Meta-Analysis. *BMC Infect. Dis.* 2021, 21, 384. https://doi.org/10.1186/s12879-021-06064-0
33. Unver, N.; McAllister, F. IL-6 Family Cytokines: Key Inflammatory Mediators as Biomarkers and Potential Therapeutic Targets. *Cytokine Growth Factor Rev.* 2018, 41, 10–17. https://doi.org/10.1016/j.cytogfr.2018.04.004