



# A risk score system for predicting complicated appendicitis and aid decision-making for antibiotic therapy in acute appendicitis

Tianlei Xu<sup>1</sup>, Qian Zhang<sup>1</sup>, Hongwei Zhao<sup>1</sup>, Yiting Meng<sup>2</sup>, Feng Wang<sup>1</sup>, Yuanxin Li<sup>1</sup>, Peng Zhang<sup>1</sup>

<sup>1</sup>Department of Gastrointestinal Surgery, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China;

<sup>2</sup>Department of Pathology, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China

**Contributions:** (I) Conception and design: T Xu, P Zhang; (II) Administrative support: H Zhao, Q Zhang; (III) Provision of study materials or patients: P Zhang, H Zhao, Q Zhang; (IV) Collection and assembly of data: T Xu, F Wang, Y Meng; (V) Data analysis and interpretation: T Xu, P Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

**Correspondence to:** Peng Zhang. No. 168 Litang Road, Changping District, Beijing 102218, China. Email: zpa01519@btch.edu.cn.

**Background:** Exclusive antibiotic therapy is a feasible treatment option for uncomplicated appendicitis, but the pre-treatment diagnosis of uncomplicated appendicitis is challenging. This study aimed to develop a risk score system to predict complicated appendicitis and aiding decision-making regarding antibiotic therapy for acute appendicitis.

**Methods:** The risk score system for predicting complicated appendicitis was constructed and validated by a surgical therapy cohort (n=543). Furthermore, we applied an independent antibiotic treatment cohort (n=169) to verify whether the risk score system could guide antibiotic treatment decision-making in patients with acute appendicitis (AA).

**Results:** A total of 543 patients were included in the surgical therapy cohort and was split into the primary (n=375) and validation (n=168) cohorts with repeated random sampling. In the primary cohort, multivariate analysis confirmed that periappendiceal fat stranding (PFS,  $P<0.001$ , OR =67.80), the C-reactive protein level (CRP  $\geq 38$  mg/L,  $P<0.001$ , OR =5.77) and the neutrophil-to-lymphocyte ratio (NLR  $\geq 7$ ,  $P<0.001$ , OR =3.51) were independent risk factors for complicated appendicitis. The PFS, CRP and NLR scores were 10.0, 4.0 and 3.0 points, respectively. Fourteen patients (3.7%, 14/375) and seven patients (4.2%, 7/168) with pathologically confirmed complicated appendicitis were classified as having uncomplicated appendicitis in the primary and validation cohorts based on the risk score system, respectively. In the independent antibiotic treatment cohort (n=169), the failure rate of antibiotic treatment was 49.2% and 5.3% for the risk score system predicted complicated AA and uncomplicated AA. Furthermore, the predictive accuracy of the risk score system for antibiotic treatment failure as measured by the area under the curve (AUC) was 0.823 (95% CI: 0.757–0.878).

**Conclusions:** We found that the proposed risk score system based on biological and CT features not only enables the accurate identification of complicated appendicitis patients before pre-treatment but also serves to guide antibiotic treatment decisions.

**Keywords:** Risk score; antibiotic therapy; uncomplicated appendicitis; complicated appendicitis

Submitted Jan 05, 2021. Accepted for publication Apr 22, 2021.

doi: 10.21037/apm-21-26

**View this article at:** <http://dx.doi.org/10.21037/apm-21-26>

## Introduction

Acute appendicitis (AA) is a common cause of abdominal pain and accounts for approximately 7.0–10.0% of emergency visits (1). AA is divided into uncomplicated appendicitis and complicated appendicitis according to the clinical manifestations or acute inflammatory processes, and complicated appendicitis refers to AA with perforation, gangrene, and abscess (2). Although appendectomy is a classic and standard treatment for AA, the overall complication rate after appendectomy is 8.2–31.4%, and the wound infection rate and incidence of pelvic abscess are 3.3–10.3% and 9.4%, respectively (3,4). Furthermore, the incidence of postoperative bowel obstruction is 2.8% (5).

Recently, a number of systematic reviews and meta-analyses of random controlled trials (RCTs) have suggested that antibiotic therapy is a feasible treatment option for uncomplicated appendicitis (6–8). Approximately 8.2–12.0% of patients treated with antibiotics experience treatment failure during their first hospitalization, and 19.2–22.6% of patients might need a second hospitalization for recurrence during the 1-year follow-up (7,8). In fact, some RCTs employed computerized tomography (CT) to assess whether appendicitis was complicated, whether CT can be used to accurately distinguish complicated appendicitis from uncomplicated appendicitis pre-treatment is still uncertain (9,10). Kim *et al.* summarized 23 articles on the CT features distinguishing complicated appendicitis from uncomplicated appendicitis and found that the CT features had a relatively high specificity but low sensitivity (11). Leeuwenburgh *et al.* found that ultrasonography and CT features incorrectly classified up to half of all patients with perforated appendicitis as having uncomplicated appendicitis (12). In one of the RCTs, although a preoperative CT assessment was used, 18% of the patients with complicated appendicitis were diagnosed with uncomplicated appendicitis prior treatment (10). Therefore, the accurate assessment of patients with complicated appendicitis prior treatment may improve the targeting of antibiotic treatment. However, the pre-treatment diagnosis of complicated appendicitis is challenging.

In this study, we attempted to investigate the factors that can distinguish complicated appendicitis from uncomplicated appendicitis and developed a risk score system to predict complicated appendicitis based on preoperative laboratory data and CT features before treatment. Furthermore, we applied an independent antibiotic treatment cohort to verify whether the risk score

system could guide antibiotic treatment decision-making in patients with AA. We present the following article in accordance with the TRIPOD reporting checklist (available at <http://dx.doi.org/10.21037/apm-21-26>).

## Methods

### Study population

#### Surgical therapy cohort (ST cohort)

A total of 713 consecutive patients who underwent surgical treatment for pathologically confirmed AA at Beijing Tsinghua Changgung Hospital (BTCH) between January 2016 and December 2019 were recruited. The inclusion criteria were age greater than 18 years, pathological confirmation of AA and performance of an abdominal pelvic CT scan before surgery. The exclusion criteria were age younger than 18 years, no CT scan before surgery, and pregnancy. All patients underwent three-port laparoscopic appendectomy as previously reported (13). All surgeries were performed by the same group of doctors, who had 3–5 years of experience with laparoscopic appendectomy and performed more than 50 laparoscopic appendectomies per year.

According to the inclusion and exclusion criteria, 543 patients who received surgical therapy (appendectomy) (ST cohort) were enrolled to develop the nomogram risk score and evaluate the predictive performance. The ST cohort dataset was split into the primary and validation cohorts with repeated random sampling until there were no significant differences ( $P$  value  $\geq 0.10$ ) between the two cohorts in any variables (*Table 1*). Ultimately, 375 patients were included in the primary cohort, and 168 patients were included in the validation cohort.

#### Antibiotic therapy cohort (AT cohort)

A total of 242 consecutive patients who were diagnosed with AA at BTCH between January 2016 and January 2019 were recruited. The inclusion criteria were the confirmation of AA based on the medical history, physical examination, laboratory inspection and abdominal pelvic CT scan; the refusal of surgical treatment; and administration of antibiotic therapy. The exclusion criteria were emergency surgical treatment and a diagnosis of appendiceal abscess prior to the initial treatment. According to the inclusion and exclusion criteria, we enrolled 169 patients who received antibiotic treatment. The definition of treatment failure was the formation of a pelvic or abdominal abscess that required

**Table 1** Baseline characteristics of the primary and validation cohorts

Characteristic	Primary cohort (n=375)	Validation cohort (n=168)	P value
Age, median (IQR), y	36 [28–52]	36 [29–52]	0.621
Male, No. (%)	188 (50.1)	76 (45.2)	0.308
Duration of abdominal pain, median (IQR), h	20.0 (11.0–34.5)	17.0 (12.0–33.0)	0.731
Shifting pain in right lower quadrant, No. (%)	218 (58.1)	110 (65.5)	0.108
Vomiting, No. (%)	173 (46.1)	87 (51.8)	0.229
Diarrhea, No. (%)	24 (6.4)	14 (8.3)	0.467
History of appendicitis, No. (%)	52 (13.9)	24 (14.3)	0.894
Peritonitis, No. (%)	231 (61.6)	109 (64.9)	0.502
TEMP, median (IQR), °C	37.30 (36.80–37.80)	37.20 (36.70–37.80)	0.852
WBC, median (IQR),	13.40 (10.86–16.10)	13.37 (11.37–15.59)	0.839
CRP, median (IQR), mg/L	11.00 (3.00–47.00)	13.00 (4.00–42.50)	0.637
NEUT, median (IQR)	83.30 (77.65–88.10)	84.25 (77.35–88.90)	0.515
NEUT%, median (IQR)	11.15 (8.64–13.75)	11.30 (9.01–13.37)	0.989
LY, median (IQR)	10.60 (6.85–15.60)	9.90 (6.50–16.38)	0.705
LY%, median (IQR)	1.40 (0.97–1.97)	1.40 (0.94–1.92)	0.581
PLT, median (IQR)	231.00 (196.00–262.50)	226.00 (192.75–276.25)	0.984
ALB, median (IQR)	45.50 (43.20–47.70)	45.20 (42.90–47.42)	0.25
NLR, median (IQR)	7.86 (5.00–12.95)	8.62 (4.74–13.61)	0.666
Appendix diameter, median (IQR), mm	10 [9–12]	10 [9–13]	0.425
Fecalith, No. (%)	139 (37.1)	67 (39.9)	0.566
Pelvic inflammatory, No. (%)	64 (17.1)	37 (22.0)	0.19
Periappendiceal fat infiltration, No. (%)	64 (17.1)	37 (22.0)	0.271
Time to operation, median (IQR), h	7.25 (4.50–12.12)	8.00 (5.00–12.50)	0.358

IQR, inter quartile range; TEMP, temperature; WBC, white blood cell count; NEUT, neutrophil count; NEUT%, neutrophil percentage; LY, lymphocyte count; LY%, lymphocyte percentage; PLT, platelet count; ALB, albumin level; NLR, neutrophil-to-lymphocyte ratio; CRP, C-reactive protein.

an operation within 30 days of antibiotic treatment.

All patients were treated with intravenous antibiotics after diagnosis; The conventional antibiotic treatment is the second or third generation cephalosporin + metronidazole or ornidazole. If the abdominal CT and laboratory tests indicate a more serious condition, the more aggressive antibiotic regimen will be selected. The therapeutic options were as follows: ertapenem 1 g + 0.9% NaCl 100 mL 1 time/day; ceftriaxone 2 g + 0.9% NaCl 100 mL 1 time/day and metronidazole or Austrian Nitrazole 0.5 g 2 times/day; cefuroxime 1.5 g + 0.9% NaCl 100 mL 2 times/day and metronidazole

or ornidazole 0.5 g 2 times/day; or levofloxacin 0.5 g 1 time/day and metronidazole or ornidazole 0.5 g 2 times/day (14,15). After 3 days of treatment, routine blood and CRP tests were performed. If the tests were normal, no further treatment was administered; otherwise, the infusion was continued to reduce inflammation until the tests returned with normal results. Imaging examinations were immediately reviewed if symptoms and signs worsened during treatment, and surgical drainage was performed if there was an abscess around the appendix; the decision to remove the appendix was made based on the intraoperative situation.

### *Pathological assessment*

The histopathological diagnosis was based on changes in and the density of neutrophils infiltrating the wall of the appendix. The severity of appendicitis can be divided into three histopathological conditions: mucosal, purulent, and gangrenous (16). The relevant conditions under the microscope are neutrophil-infiltrated mucosa/submucosa (mucosal AA); neutrophil-infiltrated mucosa, submucosa and muscularis mucosa (suppurative AA); and penetrating necrosis of the appendix wall and extensive mucosal ulcers. The definition of complicated appendicitis includes gangrenous and perforated appendicitis.

### *Data collection and CT findings*

This study was a retrospective study, and data on patients' medical history, physical examinations and laboratory results were collected retrospectively by a case manager on a structured case record form. The following clinical characteristics were obtained for this analysis: age, sex, duration of abdominal pain, presence of shifting pain in the right lower quadrant, vomiting, diarrhoea, peritonitis, temperature (TEMP), history of appendicitis, white blood cell (WBC) count, neutrophil (NEUT) count, neutrophil percentage (NEUT%), lymphocyte (LY) count, lymphocyte percentage (LY%), platelet (PLT) count, albumin (ALB) level, neutrophil-to-lymphocyte ratio (NLR), C-reactive protein (CRP) level and the time to operation. CT features associated with complicated appendicitis included appendix diameter, faecalith, pelvic inflammation, and periappendiceal fat stranding (PFS). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was reviewed and approved by the Ethics Committee of Beijing Tsinghua Changgung Hospital (20301-0-01). Informed consent for this retrospective analysis was waived.

### *Statistical analysis*

Statistical analyses were conducted using R software (version 3.5.1, <https://www.r-project.org/>) for Windows. For categorical variables, the P value was calculated using Fisher's exact test. Continuous variables are expressed as the means  $\pm$  standard deviations (SDs) for normally distributed data or as the medians and interquartile ranges (IQRs) for nonnormally distributed data. To improve the clinical applicability of the nomogram, continuous variables were

assessed by receiver operating characteristic (ROC) curves and the Youden index. A multivariable logistic regression model with the selected predictors was fitted to reduce the use of stepwise backward selection.

The nomogram was formulated based on the results of the multivariate logistic regression analysis with the rms package in R. The performance of the nomogram was measured by the area under the ROC curve (AUC) and the calibration curve. External validation was achieved by applying the nomogram to the validation cohort and evaluating the performance with similar statistics. The nomogram model was transformed into a clinically applicable scoring system called the nomogram risk score by nomogramEx package. The nomogram was used to calculate each patient's total score. Subsequently, a cutoff value analysis was performed in the primary cohort to demonstrate the ability of the nomogram risk score to select patients with complicated appendicitis. To explore whether the nomogram risk score could be used to guide antibiotic treatment, we obtained a nomogram risk score for each patient in the antibiotic treatment cohort. According to the cutoff value, patients were divided into the predicted complicated appendicitis group (risk score >11) and the predicted uncomplicated appendicitis group (risk score <11), and then the failure rates of antibiotic therapy were calculated for the predicted complicated appendicitis group and the uncomplicated appendicitis group.

## **Results**

### *Baseline characteristics of patients undergoing appendectomy*

The baseline characteristics of the primary cohort (n=375) and validation cohort (n=168) are summarized in *Table 1*. The baseline characteristics were similar between the primary and validation cohorts. In the primary cohort, 117 (31.2%) patients were pathologically confirmed complicated appendicitis, and 258 (69.8%) patients were pathologically confirmed uncomplicated appendicitis. The primary cohort population consisted of 188 men (188/375, 50.1%) and 76 women (187/375, 49.9%), with a median age of 36 years (IQR: 28–52 years). The median age in the complicated appendicitis group was older than that in the uncomplicated appendicitis group (43.0 vs. 35.0, P=0.005). There was no difference in the male to female ratio between the two groups (*Table 2*).

**Table 2** Univariate analysis of clinical and CT features in primary cohort (uncomplicated appendicitis vs. complicated appendicitis)

Characteristic	Uncomplicated appendicitis (n=258)	Complicated appendicitis (n=117)	P value
Age, median (IQR), y	35.00 (28.00–46.00)	43.00 (30.00–58.00)	0.005
Male, No. (%)	129 (50.0)	59 (50.4)	1
Duration of abdominal pain, median (IQR), h	17.00 (10.00–30.00)	26.00 (17.00–48.00)	<0.001
Shifting pain in right lower quadrant, No. (%)	134 (51.9)	84 (71.8)	<0.001
Vomiting, No. (%)	109 (42.2)	64 (54.7)	0.026
Diarrhea, No. (%)	15 (5.8)	9 (7.7)	0.5
History of appendicitis, No. (%)	39 (15.1)	13 (11.1)	0.336
Peritonitis, No. (%)	138 (53.5)	93 (79.5)	<0.001
TEMP, median (IQR), °C	37.10 (36.70–37.50)	37.60 (37.20–38.40)	<0.001
WBC, median (IQR)	12.88 (10.50–15.63)	14.51 (12.25–17.20)	<0.001
CRP, median (IQR)	7.00 (3.00–22.75)	52.00 (12.00–119.00)	<0.001
NEU, median (IQR)	81.85 (75.53–87.50)	85.00 (80.90–89.20)	<0.001
NE, median (IQR)	10.29 (8.33–13.15)	12.46 (10.09–14.69)	<0.001
LYM, median (IQR)	12.20 (7.73–17.25)	8.90 (5.50–12.20)	<0.001
LY, median (IQR)	1.45 (1.03–2.10)	1.29 (0.80–1.65)	0.001
PLT, median (IQR)	235.50 (200.00–264.00)	223.00 (193.00–259.00)	0.073
ALB, median (IQR)	45.90 (43.80–47.70)	44.70 (42.50–47.40)	0.007
NLR, median (IQR)	6.67 (4.38–11.35)	9.62 (6.62–16.19)	<0.001
Appendix diameter, median (IQR), mm	10.00 (9.00–12.00)	12.00 (10.00–14.00)	<0.001
Fecalith, No. (%)	87 (33.7)	52 (44.4)	0.05
Pelvic inflammatory, No. (%)	25 (9.7)	39 (33.3)	<0.001
Periappendiceal fat stranding, No. (%)	104 (40.3)	115 (98.3)	<0.001
Time to operation, median (IQR), h	7.50 (5.00–11.00)	9.00 (5.50–15.50)	0.011

CT, computed tomography; IQR, inter quartile range; TEMP, temperature; WBC, white blood cell count; NEUT, neutrophil count; NEUT%, neutrophil percentage; LY, lymphocyte count; LY%, lymphocyte percentage; PLT, platelet count; ALB, albumin level; NLR, neutrophil-to-lymphocyte ratio; CRP, C-reactive protein.

### *Nomogram construction and predictive performance*

In the univariate analysis, the factors associated with complicated appendicitis in the primary cohort were as follows: age, duration of abdominal pain, shifting pain in the right lower quadrant, vomiting, peritonitis, TEMP, WBC count, CRP level, NEUT count, NEUT%, LY count, LY%, PLT count, ALB level, the NLR, appendix diameter, pelvic inflammation and PFS. However, in the multivariate analysis, the data were compared between the uncomplicated appendicitis group and the complicated appendicitis group using inverse stepwise logistic regression

analysis, and PFS ( $P < 0.001$ , OR =67.80), a CRP level  $\geq 38$  mg/L ( $P < 0.001$ , OR =5.77) and an NLR  $\geq 7$  ( $P < 0.001$ , OR =3.51) were statistically significant independent risk factors (Table 3). The cut-offs for the CRP level and the NLR in the primary cohort were 38 and 7 mg/L, respectively (Figure 1).

The nomogram for predicting complicated appendicitis was constructed based on the results of the multivariate analysis in the primary cohort. The PFS, CRP and NLR scores were 10.0, 4.0 and 3.0, respectively (Figure 2). In the primary cohort, the predictive accuracy for complicated appendicitis as measured by the AUC value was 0.890

**Table 3** Multivariate analysis of the primary cohort

Characteristic	OR	95% CI	P value
Periappendiceal fat stranding	67.80	16.05–286.34	<0.001
CRP $\geq$ 38 mg/L	5.77	3.09–10.75	<0.001
NLR $\geq$ 7.0	3.51	1.90–6.51	<0.001

Values in parentheses are 95% CI. NLR, neutrophil-to-lymphocyte ratio; CRP, C-reactive protein; OR, odds ratio.

**Table 4** Univariate analysis of clinical and CT features in antibiotic therapy cohort (success vs. failure)

Characteristic	Success (n=129)	Failure (n=40)	P value
Age, median (IQR), y	32.00 (25.00–49.00)	43.00 (30.75–56.75)	0.002
Sex, male, No. (%)	64 (49.6)	17 (42.5)	0.363
TEMP, median (IQR), °C	36.90 (36.50–37.50)	38.00 (37.15–38.40)	<0.001
WBC, median (IQR), $10^9/L$	12.63 (10.49–14.59)	12.64 (11.32–16.68)	0.15
CRP $\geq$ 38 mg/L, No. (%)	33 (25.6)	25 (62.5)	<0.001
NLR $\geq$ 7, No. (%)	62 (48.1)	29 (72.5)	0.01
Appendix diameter, median (IQR), mm	9.00 (7.00–11.00)	11.50 (10.00–15.00)	<0.001
Fecalith, No. (%)	40 (31.0)	18 (45.0)	0.128
Periappendiceal fat stranding, No. (%)	42 (32.6)	34 (85.0)	<0.001
Cephalosporin, No. (%)	113 (87.6)	27 (67.5)	0.007
Antibiotic therapy time, median (IQR), day	3.00 (2.00–3.00)	2.00 (1.00–3.25)	<0.001
Nomogram risk score	3.00 (0.00–10.00)	14.00 (13.00–17.00)	<0.001

Abbreviations: TEMP, temperature; WBC, white blood cell count; NLR, neutrophil-to-lymphocyte ratio; CRP, C-reactive protein.

(95% CI: 0.854–0.920, *Figure 3A,B*). The nomogram risk score was calculated for each patient. The best cutoff value for the nomogram risk score was 11.0 points, and nomogram risk scores greater than 11 were predictive of complicated appendicitis. In the primary cohort, based on the risk score system, fourteen patients (3.7%, 14/375) with pathologically confirmed complicated appendicitis were classified as having uncomplicated appendicitis, and the sensitivity and specificity were 0.880 and 0.748, respectively. The calibration plot for the probability of complicated appendicitis showed a good correlation between the actual observed outcome and the outcome predicted by the nomogram (*Figure 3C*).

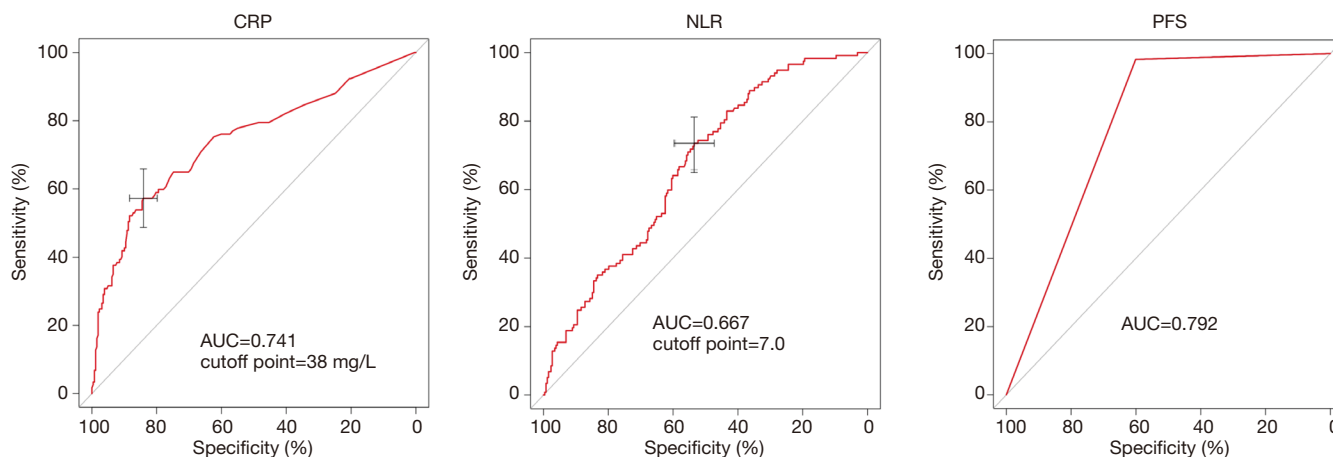
#### **Validation of the nomogram for the prediction of complicated appendicitis**

The nomogram was validated by the calibration plot and

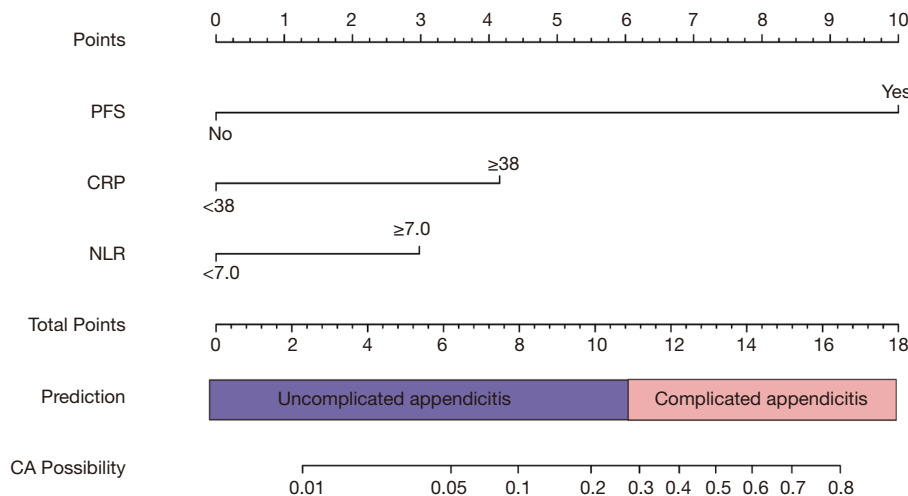
the AUC in an independent validation cohort of 168 patients. The AUC in the validation cohort was 0.890 (95% CI: 0.832–0.933, *Figure 3D,E*), which demonstrated that the model had good distinguishing ability. The calibration plot for the probability of having complicated appendicitis showed a good correlation between the actual observed outcome and the outcome predicted by the nomogram (*Figure 3F*). In the validation cohort, based on the risk score system, seven patients (4.2%, 7/168) with pathologically confirmed complicated appendicitis were classified as having uncomplicated appendicitis, and the sensitivity and specificity were 0.868 and 0.696, respectively.

#### **Prediction of the failure of antibiotic therapy during the first 30 days by the risk score system**

Of the 169 antibiotic treatment patients, 23.7% (40/169) underwent surgical drainage with or without appendectomy



**Figure 1** ROC and best of cut off value for CRP, NLR and PFS. NLR, neutrophil-to-lymphocyte ratio; CRP, C-reactive protein; PFS, periappendiceal fat stranding; AUC, area under the receiver-operating characteristic curve; ROC, receiver operating characteristic curve.



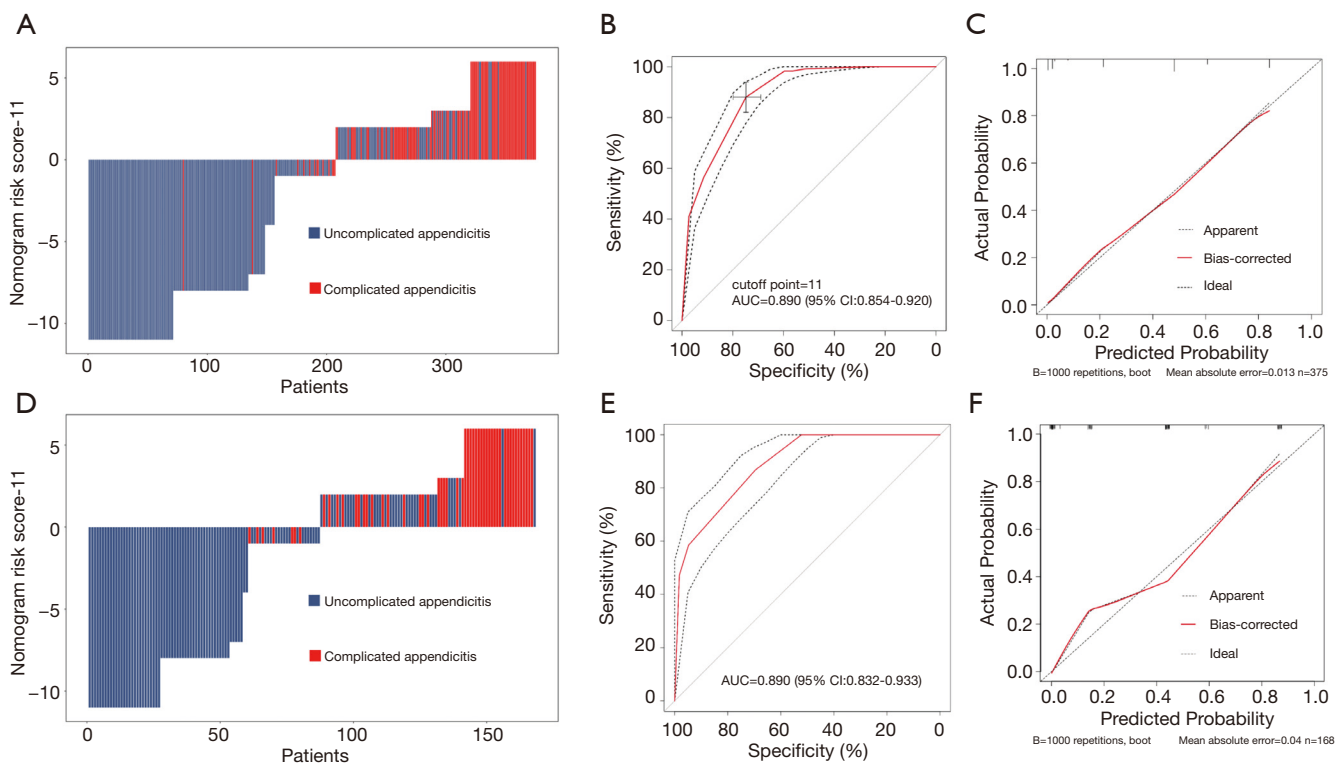
**Figure 2** Nomogram for complicated appendicitis based on biological and CT features. NLR, neutrophil-to-lymphocyte ratio; CRP, C-reactive protein; PFS, periappendiceal fat stranding; CA, complicated appendicitis.

due to the failure of antibiotic treatment. In the univariate analysis, the data were compared between the failure group and the success group and demonstrated that age, TEMP, CRP level, the NLR, appendix diameter, PFS, use of cephalosporin, antibiotic therapy time and the nomogram risk score were statistically significant risk factors (Table 4). Multivariate analysis showed that a nomogram risk score >11 points (P=0.001, OR =5.045) was a statistically significant independent risk factor (Table 5). When the patient’s nomogram risk score of was greater than 11, the patient has a high probability of complicated appendicitis, and the failure rate of antibiotic treatment was 49.2% (Figure 4A);

when the patient’s nomogram risk score was less than 11, the patient has a high probability of uncomplicated appendicitis, the failure rate of antibiotic treatment was only 5.3% (Figure 4A). Furthermore, the predictive accuracy of the nomogram risk score for antibiotic treatment failure as measured by the AUC was 0.823 (95% CI: 0.757–0.878, Figure 4B).

**Discussion**

In this study, we constructed a nomogram for the prediction of complicated appendicitis. This nomogram



**Figure 3** The distribution of patients with actual observed and predicted by nomogram (A), ROC curve with 95% CI and best of cut off value for nomogram risk score (B), and calibration of the nomogram (C) in primary cohort; The distribution of patients with actual observed and predicted by nomogram (D), ROC curve with 95% CI (E), and calibration of the nomogram (F) in validation cohort. 95% CI, 95 percent confidence interval; AUC, area under the receiver-operating characteristic curve; ROC, receiver operating characteristic.

**Table 5** Multivariate analysis of the antibiotic therapy cohort

Characteristic	OR	95% CI	P value
Nomogram risk score >11	5.045	1.901–13.922	0.001
TEMP, °C	4.064	1.550–10.654	0.004

TEMP, temperature.

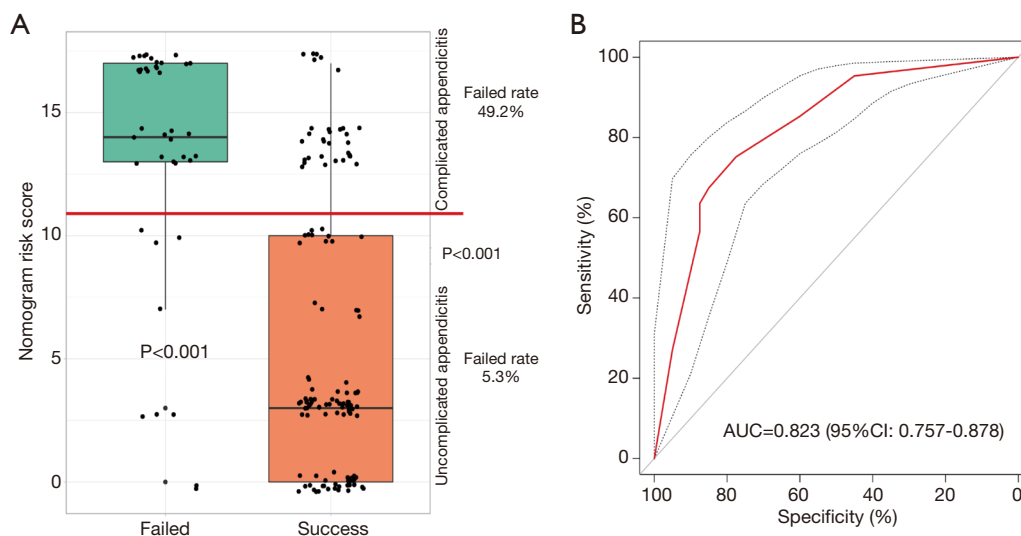
can be used to estimate the probability of complicated appendicitis based on the results of a logistic regression analysis. The nomogram includes one CT feature and two biological features that are routinely obtained. This model can accurately predict complicated appendicitis, and furthermore, this model can be used to guide decision-making regarding antibiotic therapy in AA.

AA is one of the most common abdominal emergencies in the world, and the cause of the condition is still unclear. To stratify patients based on the necessary clinical management, AA is divided into uncomplicated and complicated appendicitis; however, many patients still have

ambiguous diagnoses, which is one of the most challenging problems. Recently, several systematic reviews and meta-analyses have shown that most uncomplicated appendicitis patients can be treated with antibiotics first (6,7,11). Therefore, the World Society of Emergency Surgery (WES) guidelines recommend the use of antibiotic therapy as a safe alternative to surgery for patients with uncomplicated appendicitis and without appendicolith. Notably, the guidelines also indicate the possibility of treatment failure and the misdiagnosis of complicated appendicitis (16).

A recent meta-analysis by Podda *et al.* showed that the failure rate of antibiotic treatment and the recurrence rate at





**Figure 4** The nomogram risk score associated with failure and success in patients with antibiotic therapy and prediction of the failure of antibiotic therapy during the first 30 days by the nomogram risk score (A); ROC curve with 95% CI for nomogram risk score to predict the failure of antibiotic therapy (B). 95% CI, 95 percent confidence interval; AUC, area under the receiver-operating characteristic curve.; ROC: receiver operating characteristic.

the 1-year follow-up were 8.5% and 19.2%, respectively (7). However, the failure rate of antibiotic therapy in the Vons *et al.* study was 12% (14/120) (10). Notably, in this study, despite CT-scan assessment, 21 (18%) of 119 patients were unexpectedly identified as having complicated appendicitis during surgery (10). Salminen *et al.* showed that 4 (1.5%) of 237 patients were identified as having complicated appendicitis during surgery (9). This may be due to inaccurate CT scans or the progression of uncomplicated appendicitis to perforation.

Neither CT nor emergency MRI can be used to discriminate between nonperforated and perforated appendicitis. Leeuwenburgh *et al.* found that both methods incorrectly classified more than half of the patients with perforated appendicitis as having uncomplicated appendicitis (12). A systematic review and meta-analysis of 10 CT features (extraluminal appendicolith, abscess, appendiceal wall enhancement defect, extraluminal air, ileus, periappendiceal fluid collection, ascites, intraluminal air, intraluminal appendicolith and PFS) for the diagnosis of complicated appendicitis showed that PFS was the only feature that had a high sensitivity (94%; 95% CI: 86–98%), although it had a low specificity (40%; 95% CI: 23–60%) (11). Another study showed that the pooled sensitivity of the presence of any of the 10 CT features was higher than that

of individual assessments (92% *vs.* 64%;  $P<0.001$ ), although the pooled specificity was lower (43% *vs.* 76%;  $P<0.001$ ). In our study, the AUC, sensitivity and specificity of PFS were 79.2%, 98.7% and 59.7%, respectively. However, sensitivity, rather than specificity, should be given priority in the diagnosis of complicated appendicitis. Low specificity may lead to appendectomy in some patients who could be treated with antibiotics.

To improve the predictive ability, our nomogram includes clinical features and imaging features. Clinical features, especially when two or more are combined, have been suggested as being valuable in the diagnosis of complicated appendicitis. The NLR is a simple clinical inflammatory marker, and the NLR provides information about two different immune and inflammatory pathways, which may make it a good marker for predicting appendicitis and its severity (17–19). In our study, the optimal cutoff value for the NLR was 7, and an NLR greater than 7 indicated the presence of complicated appendicitis. Similar to other studies, the AUC, sensitivity, and specificity of the NLR for predicting complex appendicitis were 0.667, 0.735, and 0.535, respectively. In the early stage of AA, the sensitivity of CRP is low. CRP may be more sensitive for the detection of perforation of the appendix and the formation of an abscess, although the positive predictive value of traditional inflammatory markers is relatively lower (20). A systematic

review and meta-analysis demonstrated that the CRP level was more accurate (AUC =0.75, 95% CI: 0.71–0.78) than the WBC count and procalcitonin level (21). In our study, the AUC, sensitivity, and specificity of CRP for predicting complicated appendicitis were 0.741, 0.573, and 0.841, respectively.

Previously, Atema *et al.* constructed a model for predicting complicated appendicitis based on clinical and imaging characteristics (22). The model includes five clinical indicators (age, TEMP, duration of symptoms, WBC count, and CRP level) and three CT-based parameters (presence of extraluminal free air, periappendiceal fluid and appendicolith). Nonetheless, scores >6 points are not specific for complicated appendicitis; instead, other complex diseases, such as perforated diverticulitis and Crohn's disease, may also be indicated (22). These limitations and the time needed to calculate up to 22 points per patient may limit its daily use. In addition, the study lacked a validation group. Avanesov *et al.* constructed a model for predicting complicated appendicitis based on a combination of clinical and CT features called the appendicitis severity index (ASPI) (23). A score  $\geq 4$  points predicted complicated appendicitis with a positive predictive value of 92% and a negative predictive value of 83%. Again, the study lacked a validation group. In our study, evaluation of the nomogram showed that a score greater than or equal to 11 points indicated the presence of complicated appendicitis, and fourteen patients (3.7%, 14/375) and seven patients (4.2%, 7/168) with complicated appendicitis were classified as having uncomplicated appendicitis in the primary and validation cohorts, respectively. The false-negative classification of complicated appendicitis patients who then receive antibiotic treatment may result in an increased rate of treatment failure. In this study, we used an independent cohort who received antibiotic therapy to test whether this nomogram risk score could guide decision-making with regard to antibiotic therapy. When the nomogram risk score was more than 11 points, the failure rate of antibiotic treatment was 49.2%; in contrast, when the nomogram risk score was less than 11 points, the failure rate of antibiotic treatment was only 5.3%, which was lower than previously reported. This suggests that the nomogram risk score not only accurately predicts complicated appendicitis but can also be used to guide the use of antibiotic treatment in patients with AA.

Several limitations of this study need to be addressed. First, this study was retrospective and included a relatively limited number of patients from only one hospital.

Although the model was validated in an independent validation cohort, there was no external validation to ensure that the results support universal application. The second is the definition of complicated appendicitis in this study. The definition of complicated appendicitis is not yet clear. The definition of complicated appendicitis in this study included gangrenous and perforated appendicitis, which are determined based on postoperative pathological results. Patients with appendiceal abscesses or inflammatory masses were excluded because this subgroup of appendicitis requires special treatment.

## Conclusions

In conclusion, we found that the proposed nomogram risk score based on biological and CT features not only enables the accurate identification of complicated appendicitis patients before surgery but can also be used to guide management decisions.

## Acknowledgments

*Funding:* None.

## Footnote

*Reporting Checklist:* The authors have completed the TRIPOD reporting checklist. Available at <http://dx.doi.org/10.21037/apm-21-26>

*Data Sharing Statement:* Available at <http://dx.doi.org/10.21037/apm-21-26>

*Peer Review File:* Available at <http://dx.doi.org/10.21037/apm-21-26>

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/apm-21-26>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was reviewed and approved by the Ethics Committee of Beijing Tsinghua Changgung

Hospital (20301-0-01). Informed consent for this retrospective analysis was waived.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Cervellin G, Mora R, Ticinesi A, et al. Epidemiology and outcomes of acute abdominal pain in a large urban Emergency Department: retrospective analysis of 5,340 cases. *Ann Transl Med* 2016;4:362.
2. Bhangu A, Soreide K, Di Saverio S, et al. Acute appendicitis: modern understanding of pathogenesis, diagnosis, and management. *Lancet* 2015;386:1278-87.
3. National Surgical Research Collaborative. Multicentre observational study of performance variation in provision and outcome of emergency appendectomy. *Br J Surg* 2013;100:1240-52.
4. Jaschinski T, Mosch CG, Eikermann M, et al. Laparoscopic versus open surgery for suspected appendicitis. *Cochrane Database Syst Rev* 2018;11:CD001546.
5. Opaluwa AS. Bowel obstruction following appendectomy: what is the true incidence? *Ann Surg* 2010;251:991.
6. Huston JM, Kao LS, Chang PK, et al. Antibiotics vs. Appendectomy for Acute Uncomplicated Appendicitis in Adults: Review of the Evidence and Future Directions. *Surg Infect (Larchmt)* 2017;18:527-35.
7. Podda M, Gerardi C, Cillara N, et al. Antibiotic Treatment and Appendectomy for Uncomplicated Acute Appendicitis in Adults and Children: A Systematic Review and Meta-analysis. *Ann Surg* 2019;270:1028-40.
8. Sallinen V, Akl EA, You JJ, et al. Meta-analysis of antibiotics versus appendectomy for non-perforated acute appendicitis. *Br J Surg* 2016;103:656-67.
9. Salminen P, Paajanen H, Rautio T, et al. Antibiotic Therapy vs Appendectomy for Treatment of Uncomplicated Acute Appendicitis: The APPAC Randomized Clinical Trial. *JAMA* 2015;313:2340-8.
10. Vons C, Barry C, Maitre S, et al. Amoxicillin plus clavulanic acid versus appendectomy for treatment of acute uncomplicated appendicitis: an open-label, non-inferiority, randomised controlled trial. *Lancet* 2011;377:1573-9.
11. Kim HY, Park JH, Lee YJ, et al. Systematic Review and Meta-Analysis of CT Features for Differentiating Complicated and Uncomplicated Appendicitis. *Radiology* 2018;287:104-15.
12. Leeuwenburgh MM, Wiezer MJ, Wiarda BM, et al. Accuracy of MRI compared with ultrasound imaging and selective use of CT to discriminate simple from perforated appendicitis. *Br J Surg* 2014;101:e147-55.
13. Buckley FP 3rd, Vassaur H, Monsivais S, et al. Single-incision laparoscopic appendectomy versus traditional three-port laparoscopic appendectomy: an analysis of outcomes at a single institution. *Surg Endosc* 2014;28:626-30.
14. Bai N, Sun C, Wang J, et al. Ertapenem versus ceftriaxone for the treatment of complicated infections: a meta-analysis of randomized controlled trials. *Chin Med J (Engl)* 2014;127:1118-25.
15. Talan DA, Saltzman DJ, DeUgarte DA, et al. Methods of conservative antibiotic treatment of acute uncomplicated appendicitis: A systematic review. *J Trauma Acute Care Surg* 2019;86:722-36.
16. Di Saverio S, Podda M, De Simone B, et al. Diagnosis and treatment of acute appendicitis: 2020 update of the WSES Jerusalem guidelines. *World J Emerg Surg* 2020;15:27.
17. Sahbaz NA, Bat O, Kaya B, et al. The clinical value of leucocyte count and neutrophil percentage in diagnosing uncomplicated (simple) appendicitis and predicting complicated appendicitis. *Ulus Travma Acil Cerrahi Derg* 2014;20:423-6.
18. Eddama M, Fragkos KC, Renshaw S, et al. Logistic regression model to predict acute uncomplicated and complicated appendicitis. *Ann R Coll Surg Engl* 2019;101:107-18.
19. Hajibandeh S, Hajibandeh S, Hobbs N, et al. Neutrophil-to-lymphocyte ratio predicts acute appendicitis and distinguishes between complicated and uncomplicated appendicitis: A systematic review and meta-analysis. *Am J Surg* 2020;219:154-63.
20. Ozozan OV, Vural V. High C-reactive protein level as a predictor for appendiceal perforation. *Ulus Travma Acil Cerrahi Derg* 2020;26:63-6.
21. Yu CW, Juan LI, Wu MH, et al. Systematic review and meta-analysis of the diagnostic accuracy of procalcitonin, C-reactive protein and white blood cell count for suspected acute appendicitis. *Br J Surg* 2013;100:322-9.

22. Atema JJ, van Rossem CC, Leeuwenburgh MM, et al. Scoring system to distinguish uncomplicated from complicated acute appendicitis. *Br J Surg* 2015;102:979-90.
23. Avanesov M, Wiese NJ, Karul M, et al. Diagnostic prediction of complicated appendicitis by combined clinical and radiological appendicitis severity index (APSI). *Eur Radiol* 2018;28:3601-10.

**Cite this article as:** Xu T, Zhang Q, Zhao H, Meng Y, Wang F, Li Y, Zhang P. A risk score system for predicting complicated appendicitis and aid decision-making for antibiotic therapy in acute appendicitis. *Ann Palliat Med* 2021;10(6):6133-6144. doi: 10.21037/apm-21-26