Royal College of Surgeons
ADVANCING SURGICAL CARE

GENERAL SURGERY

Ann R Coll Surg Engl 2018; **00:** 1–12
doi 10.1308/rcsann.2018.0152

Logistic regression model to predict acute uncomplicated and complicated appendicitis

MMR Eddama[1,2], KC Fragkos[2], S Renshaw[2], M Aldridge[3], G Bough[2], L Bonthala[2], A Wang[2], R Cohen[1,2]

[1]Division of Surgery and Interventional Science, University College London, London, UK
[2]Department of Colorectal Surgery, University College london Hospital, London, UK
[3]Department of Surgery, Lister Hospital Stevenage, Stevenage, UK

ABSTRACT

INTRODUCTION While patients with acute uncomplicated appendicitis may be treated conservatively, those who suffer from complicated appendicitis require surgery. We describe a logistic regression equation to calculate the likelihood of acute uncomplicated appendicitis and complicated appendicitis in patients presenting to the emergency department with suspected acute appendicitis.

MATERIALS AND METHODS A cohort of 895 patients who underwent appendicectomy were analysed retrospectively. Depending on the final histology, patients were divided into three groups; normal appendix, acute uncomplicated appendicitis and complicated appendicitis. Normal appendix was considered the reference category, while acute uncomplicated appendicitis and complicated appendicitis were the nominal categories. Multivariate and univariate regression models were undertaken to detect independent variables with significant odds ratio that can predict acute uncomplicated appendicitis and complicated appendicitis. Subsequently, a logistic regression equation was generated to produce the likelihood acute uncomplicated appendicitis and complicated appendicitis.

RESULTS Pathological diagnosis of normal appendix, acute uncomplicated appendicitis and complicated appendicitis was identified in 188 (21%), 525 (59%) and 182 patients (20%), respectively. The odds ratio from a univariate analysis to predict complicated appendicitis for age, female gender, $\log_2$ white cell count, $\log_2$ C-reactive protein and $\log_2$ bilirubin were 1.02 (95% confidence interval, CI, 1.01, 1.04), 2.37 (95% CI 1.51, 3.70), 9.74 (95% CI 5.41, 17.5), 1.57 (95% CI 1.40, 1.74), 2.08 (95% CI 1.56, 2.76), respectively. For the same variable, similar odds ratios were demonstrated in a multivariate analysis to predict complicated appendicitis and univariate and multivariate analysis to predict acute uncomplicated appendicitis.

CONCLUSIONS The likelihood of acute uncomplicated appendicitis and complicated appendicitis can be calculated by using the reported predictive equations integrated into a web application at www.appendistat.com. This will enable clinicians to determine the probability of appendicitis and the need for urgent surgery in case of complicated appendicitis.

ACKNOWLEDGEMENTS This work would have not been completed without the help of: Clarissa Y. M. Carvallho, Consultant Anaesthetist at Guy's and St Thomas's Hospital; Bried O'Brien, the Head of Urgent Care Transformation at University College London Hospital; and Guang's Wu, Web Application Developer. Their contribution to building the web application allowed this work to materialise into clinical use to benefit patients.

KEYWORDS

Acute appendicitis – Complicated appendicitis – Right iliac fossa pain – Acute abdomen – Emergency surgery – Diagnostic strategy

Accepted 7 July 2018

CORRESPONDENCE TO

Mohammad Eddama, E: eddama@doctors.org.uk

Introduction

Worldwide, appendicitis remains one of the most common causes of acute abdominal pain in adults.[1] Despite the recent advancement in healthcare provision, thousands of people around the world still suffer from significant morbidity because of appendicitis.[2] One key factor in improving patient outcomes from appendicitis is to ensure that the diagnosis and management are instigated in a timely manner.[3]

It appears that all causes of appendicitis eventually lead to a final common pathway of intraluminal bacterial invasion to the appendiceal wall.[4] Subsequently, inflammation can progress from acute intraluminal inflammation to gangrene and perforation.[5] The severity of pathological changes are reflected clinically. If the appendiceal inflammation is contained within the appendix, patients may present to the emergency department with localised signs and symptoms and may be systemically well. This group of patients may be diagnosed with acute uncomplicated appendicitis (AUA).

Depending on their clinical presentation, age and other comorbidity, patients with AUA may be treated conservatively with antibiotics or offered surgery.[6] However, if patients are clinically unwell, complicated appendicitis (described as intra- or extraluminal pus, necrosis, gangrene or perforation) may be suspected. These cases may require immediate surgical intervention, since conservative treatment is unlikely to be effective. Therefore, in patients presenting with suspected appendicitis, careful clinical assessment and diagnosis should aim to promptly distinguish between patients who suffer from AUA and complicated appendicitis.

Currently, the most common diagnostic strategy for acute appendicitis includes clinical assessment and biochemical blood markers. Computed topography (CT) and ultrasonography have reduced the rate of negative appendicectomy.[7,8] However, the risk of ionising radiation exposure and unavailability of out-of-hours imaging services in some hospitals limits their routine use.[9] In some patients, imaging tests may also delay the delivery of surgical intervention unnecessarily.[10]

Clinical scoring systems, such as Alvarado score and the appendicitis inflammatory response (AIR) score, have acquired popularity.[11,12] These scoring systems enable clinicians to stratify the risk of appendicitis into low, moderate and high. The Alvarado score is most useful in predicting the absence of appendicitis, with a sensitivity of 94–99%,[13] but it lacks specificity. It is also less accurate in children[14] and tends to over predict the presence of acute appendicitis in women.[13] Similarly, the AIR score uses clinical parameters and C-reactive protein (CRP) level to predict appendicitis. Indeed, AIR has an improved discriminating power for ruling appendicitis in or out; it is more specific in patients with moderate risk and more reliable in children than the Alvarado score.[15] Bilirubin has been described as an additional marker to CRP and white cell count (WCC) in predicting the diagnosis of complicated appendicitis.[16] There is a need for a scoring system that incorporates recent markers, such as bilirubin, able to discriminate between AUC and complicated appendicitis. In this study we aim to describe a logistic regression equation that calculates the probability of AUA and complicated appendicitis.

Materials and methods

Patient population

A cohort of 1076 patients who underwent appendicectomy was identified retrospectively from two different NHS hospitals in the UK: 851 patients from the University College London Hospital (UCLH) and 225 patients from Lister General Hospital (LGH). The study periods include April 2012 to April 2015 and September 2012 to June 2013, from UCLH and LGH, respectively. From those, 895 patients were included in the final analysis (Fig 1). Regulatory approval was granted by the site institutional review board as a service evaluation. Patient data were anonymised. The exclusion criteria were patients who underwent elective appendicectomy, patients who underwent other surgical procedures such as hysterectomy, and patients who had no

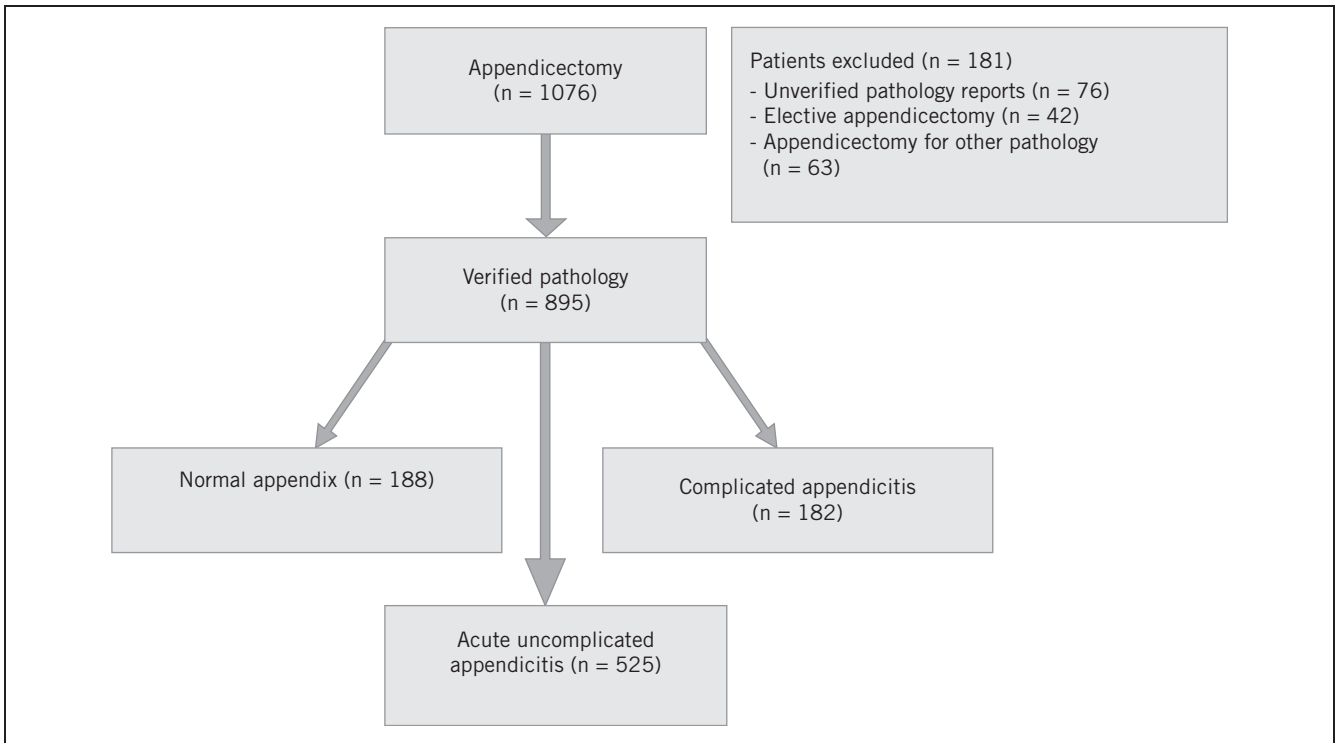

Figure 1 Flowcharts of patients who were included and excluded from the study.

pathology results verified by a consultant pathologist. We extracted data on patient demographics, preoperative biochemical markers and final histology. Data were extracted from the electronic patient record system ('clinical data repository' or CDR).

Definitions of comparison groups

Depending on the outcome of the final pathology, patients were classified into three groups: normal appendix, AUA and complicated appendix (Fig 1). We used the results from patients with normal pathology as a reference category. Normal pathology was identified when the microscopic appearance described no signs of inflammation. The diagnosis of AUA was defined when the appendix was suppurative/phlegmonous with microscopic appearance of transmural inflammation, ulceration or thrombosis and intramural pus. Complicated appendix was defined as gangrenous (transmural inflammation with necrosis), perforated (visible perforation), transmural inflammation with intramural or extramural pus (with or without perforation). Depending on the degree of inflammation, other diagnoses including adenocarcinoma of the appendix, neuroendocrine tumours and parasitic infection were either excluded or included within the respective analysis group.

Histological diagnosis was obtained retrospectively and performed as part of the standard clinical care. Data were collected and interpretation verified several times by members of the research team to ensure accuracy. Statistical analysis of the data was undertaken by a statistician who was blinded to the outcome code to avoid bias. The sample size exceeded the rule-of-thumb of 10 observations per independent variable and therefore this study is statistically powered.

Statistical analysis

Data were analysed using Statistical Package for Social Sciences, version 22, and GraphPad Prism version 6. For inference statistics, t-test was used to analyse continuous data and chi-square test was used to analyse categorical data. Univariate as well as multivariate regression analyses were performed. Logistic regression was used to determine the odds ratio (OR) for the independent predictors. Diagnostics for the goodness of fit was also employed. Receiver operating characteristic (ROC) curve and area under the curve (AUC) analyses were used to establish the diagnostic accuracy and a cut-off point of the preoperative blood levels of WCC, CRP and bilirubin for identifying AUA and complicated appendix. The level of statistical significance was set at 0.05 for all test procedures.

Results

Patient characteristics

Table 1 summarises the main patients' characteristics and highlights the difference between the groups. Patients with complicated appendix are significantly ($P < 0.01$) older

Table 1 Baseline characteristics of patients.

Characteristic	Group		
	Normal appendix	AUA	Complicated appendicitis
Age (years), mean (SD)[a]	28 (14)	29 (14)	33 (17)
Sex *n* (%):[b]			
Male	61 (32)	308 (59)	99 (54%)
Female	127 (68)	217 (41)	83 (46%)
ASA score *n* (%):			
I	54 (84)	234 (87)	86 (86)
II	6 (13)	27 (10)	12 (12)
III	< 5	7 (2)	< 5
IV	< 5	< 5	< 5
Admission to surgery waiting time (hours), mean (SD)[b]	29 (24)	21 (22)	17 (14)
Length of surgical procedure (minutes) mean (SD)	62 (37)	64 (28)	71 (31)
Length of hospital stay (days) mean (SD)[c]	2.7 (1.9)	2.9 (3)	3.7 (4)
WCC ($x10^9$/l) mean (SD)[b]	10.6 (0.7)	12.8 (0.2)	14.5 (0.4)
CRP (mg/l) mean (SD)[b]	27 (4)	39 (3)	88 (7)
Bilirubin (mg/dl) mean (SD)[b]	11 (1)	14 (0.5)	18 (1.6)

ASA, American Society of Anesthesiologists; AUA, acute uncomplicated appendicitis; CRP, C-reactive protein; WCC, white cell count.
[a] $P < 0.01$
[b] $P < 0.0001$
[c] $P < 0.05$

than patients with normal appendix as well as those with AUA. There were significantly more males in the AUA and complicated appendix groups than among the normal appendix group ($\chi^2_{[2]}$ 38.6, $P < 0.0001$). The duration of surgery in the patients with complicated appendix was significantly ($P < 0.01$) increased in comparison with those with AUA or normal appendix. This was still significant after being adjusted for the method of surgical access. Laparoscopic appendicectomy converted to open was significantly ($P < 0.0001$) longer (mean 92 minutes, standard deviation, SD, 39 minutes) than procedures completed laparoscopically (mean 68 minutes, SD 31 minutes) or with an open appendicectomy (mean 55 minutes, SD 25 minutes).

Correlation of biochemical blood markers

In patients with AUA, WCC is significantly positively correlated with both CRP (r 0.13, 95% confidence interval, CI, 0.04–0.21, $P < 0.01$) and bilirubin (r 0.16, 95% CI 0.07–0.25, $P < 0.001$). Similarly, CRP is significantly positively correlated with bilirubin (r 0.25, 95% CI 0.16–0.33, $P < 0.0001$). On the other hand, there was no significant correlation between WCC, CRP and bilirubin values at presentation in patients with normal appendix: WCC and CRP (r 0.06, 95% CI –0.09 to –0.20); white cell count and bilirubin (r 0.07, 95% CI –0.09 to –0.22): and CRP and bilirubin (r 0.12, 95% CI –0.03 to –0.28). Furthermore, there was no significant correlation between WCC, CRP and bilirubin values at presentation in patients with complicated appendix: WCC and CRP (r 0.003, 95% CI –0.15 to –0.15); WCC and bilirubin (r –0.02, 95% CI –0.18 to –0.13) and CRP and bilirubin (r 0.07, 95% CI –0.08 to –0.22) (Fig 2).

Predicting acute uncomplicated and complicated appendicitis with logistic regression

We hypothesised that five variables (age, gender, WCC, CRP and bilirubin) could function as predictors of AUA and complicated appendix, with the normal appendix being the reference category. Initial univariate modelling with CRP, WCC and bilirubin did not produce satisfactory models owing to lack of fit with the Hosmer–Lemeshow test,[17] and inappropriate links (as tested by Stata's linktest package).[18] Hence, the logarithms (with base 2) of CRP, WCC and bilirubin were used as predictors and produced satisfactory fits and links.

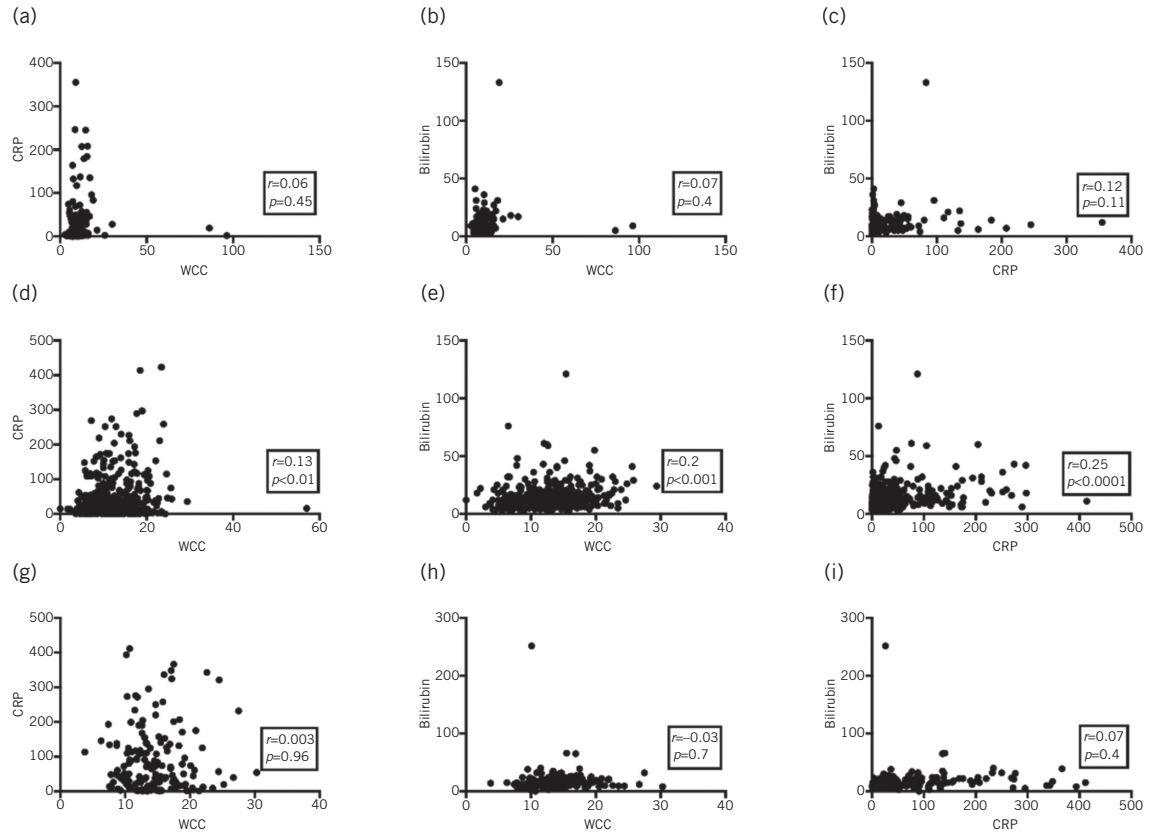

Figure 2 Correlation between white cell count (WCC), C-Reactive protein (CRP) and bilirubin showing statistically significantly positive correlation in the prediction of acute uncomplicated appendicitis (D, E and F), but not normal (A, B and C) or complicated appendicitis (G, H and I).

Prediction of acute uncomplicated appendicitis

Univariate and multivariate logistic regression analyses were performed. The likelihood ratios from the univariate analysis models demonstrated a statistically significant ability of gender, $\log_2$ WCC, $\log_2$ CRP and $\log_2$ bilirubin to distinguish patients with normal appendix from AUA. The likelihood ratio from the multivariate logistic regression model containing all the predictors were also statistically significant ($\chi^2_{[5]} = 84.65$, $P < 0.0001$), indicating that the model was able to distinguish between patients with normal appendix and AUA. A stepwise regression approach (backward procedure, based on *P*-value of predictor removed) indicated that that age could be removed. The Bayesian information criterion (BIC) was reduced in the model without age and the R-squares were not largely different between models and the likelihood ratio indicated that both models were not different from each other. Overall, the final model with age excluded was preferred as more parsimonious. The final model equation was:

$$P = \frac{1}{1 + e^{-(-3.664 + 0.144 \times \log_2 CRP + 0.260 \times \log_2 Bilirubin + 0.882 \times \log_2 WCC + 0.869 \times [0 \text{ if } female | 1 \text{ if } male])}}$$

The model as a whole explained between 14.3% (Cox–Snell R-square) and 20.9% (Nagelkerke R-square) of the variance in the pathology diagnosis and correctly classified 76% of cases (Fig 3). The Pearson chi-square goodness-of-fit test was non-significant indicating satisfactory fit (Table 3).

As shown in Table 2, of four (gender, $\log_2$ WCC, $\log_2$ CRP, $\log_2$ bilirubin) variables with statistically significant odds ratios (OR) to predict AUA, three remain statistically significant in the multivariate analysis and $\log_2$ bilirubin was borderline significant. Males are 2.4 times more likely to have AUA than females (OR 2.38). The model suggests that there is a 2.4 times increase in the likelihood of AUA for every doubling of WCC (OR 2.42). Similarly, for every doubling of the CRP the likelihood of AUA increases by 16% (OR 1.16).

Collinearity was tested with the coldiag package in STATA.[19] The condition number was 11.19, indicating no collinearity, as discussed above. Although the correlations are significant between the predictors, their association is weak, thus explaining the lack of collinearity in our model. The discriminatory power of the model is acceptable (area under the ROC curve 0.755). Ideally, there should be minimal overlap between the estimated prediction for normal appendix and AUA to indicate good discrimination between cases. However as shown in Fig 4 there is an overlap in between normal appendix and AUA and the jittered outcome does not have the more density observation in the edges of each section of the graph.

Prediction of complicated appendicitis

Univariate and multivariate logistic regression analysis were again performed. The likelihood ratios from the univariate analysis models demonstrated a statistically significant ability of gender, age, $\log_2$ WCC, $\log_2$ CRP and $\log_2$ bilirubin to distinguish patients with normal appendix from CA. The likelihood ratios from the multivariate logistic regression model containing all the predictors were also statistically significant $\chi^2_{[5]} = 130.107$, $P < 0.0001$, indicating that the model was able to distinguish between patients with normal appendix and CA. A stepwise regression approach (backward procedure, based on p-value of predictor removed) indicated that that $\log_2$ bilirubin could be removed. The BIC was reduced in the model without $\log_2$ bilirubin and the R-squares were not largely different between models and the likelihood ratios indicated that

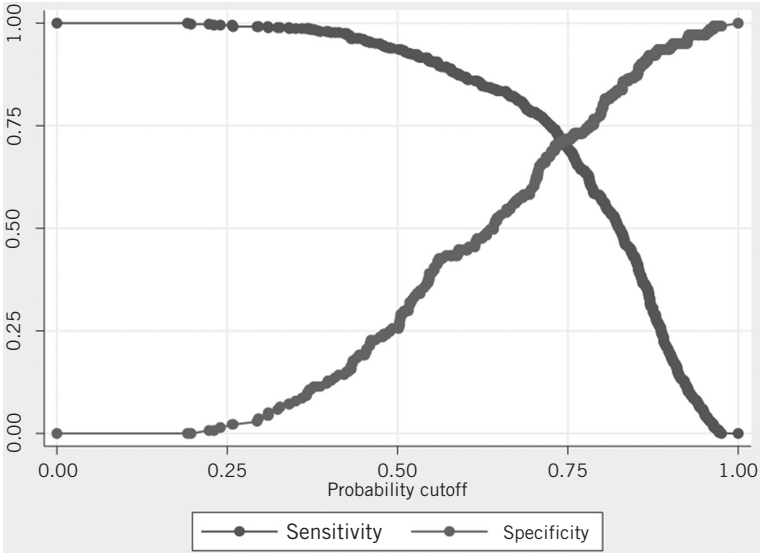

Figure 3 Plot of sensitivity and specificity against probability cut off from the fitted model for acute uncomplicated appendicitis.

Table 2 Univariate and multivariate logistic regression model showing the output for objective independent variable to predict the likelihood of acute uncomplicated and complicated appendicitis.

Appendicitis type[a]	Variable	Univariate analysis		Multivariate analysis	
		OR (95% CI)	LR (df)	Step 1 OR (95% CI)	Step 2 OR (95% CI)
Acute uncomplicated	Age	1.01 (0.99–1.02)[b]	χ^2 (1) = 0.77, P = 0.382	1.01 (0.99–1.02)	(removed)
	Sex (ref. female)	2.78 (1.93–4.00)[b]	χ^2 (1) = 31,56, P < 0.001	2.41 (1.55–3.73)[b]	2.38 (1.54–3.69)[b]
	Log_2 WCC	2.98 (2.16–4.11)[b]	χ^2 (1) = 48.57, P < 0.001	2.43 (1.70–3.48)[b]	2.42 (1.69–3.46)[b]
	Log_2 CRP	1.20 (1.11–1.29)[b]	χ^2 (1) = 22.62, P < 0.001	1.15 (1.05–1.26)[c]	1.16 (1.06–1.26)[c]
	log_2 bilirubin	1.76 (1.40–2.21)[b]	χ^2 (1) = 25.47, P < 0.001	1.29 (1.00–1.68)	1.29 (1.00–1.68)
Complicated	Age	1.02 (1.01–1.04)[b]	χ^2 (1) = 10.85, P = 0.001	1.03 (1.01–1.05)[c]	1.03 (1.01–1.05)[c]
	Sex (ref. female)	2.37 (1.51–3.70)[b]	χ^2 (1) = 14.50, P < 0.001	1.74 (0.91–3.32)	1.88 (1.01–3.48)[d]
	Log_2 WCC	9.74 (5.41–17.5)[b]	χ^2 (1) = 86.84, P < 0.001	5.58 (2.89–10.75)[b]	5.84 (3.05–11.17)[b]
	Log_2 CRP	1.57 (1.40–1.74)[c]	χ^2 (1) = 86.76, P < 0.001	1.44 (1.26–1.64)[b]	1.46 (1.28–1.66)[b]
	Log_2 bilirubin	2.08 (1.56–2.76)[b]	χ^2 (1) = 29.67, P < 0.001	1.14 (0.81–1.62)	(removed)

CI, confidence interval; CRP, C-reactive protein; df, degrees of freedom; LR, likelihood ratio; OR, odds ratio; WCC, white cell count.
[a] Normal appendix is the reference category.
[b] $P < 0.001$.
[c] $P < 0.01$
[d] $P < 0.05$

Table 3 Model fit criteria.

Criterion	Acute uncomplicated appendicitis			Complicated appendicitis		
	Step 2	Step 1	Difference (step 2 – step1)	Step 2	Step 1	Difference (step 2 – step1)
Likelihood ratio (degrees of freedom)	83.869 (4)	84.650 (5)	0.780 (1)	132.233 (3)	130.107 (5)	2.126 (2)
P-value for likelihood ratio	< 0.0001	<0.0001	0.377	< 0.0001	< 0.0001	0.345
McFadden's R^2 (%)	13.5	13.6	–0.1	32.1	33.9	–1.8
McFadden's adjusted R^2 (%)	11.5	11.3	0.2	29.7	30.2	–0.5
Maximum Likelihood R^2 (Cox–Snell) (%)	14.3	14.4	–0.1	35.9	37.5	–1.5
Cragg–Uhler's (Nagelkerke) R^2 (%)	20.9	21.1	–0.2	47.9	50.0	–2.1
Akaike information	1.012	1.014	–0.002	0.974	0.967	0.007
Bayes information	–2856.83	–2851.31	–5.52	–1383.24	–1264.68	–118.56
Hosmer–Lemeshow goodness-of-fit test *P*-value	0.042	0.098	–	0.1415	0.0604	–
Pearson χ^2 goodness-of-fit test *P*-value	0.056	0.068	–	0.1871	0.5588	–
Correctly classified cases	75.96%	76.88%	–	80.81%	80.51%	–

both models were not different from each other. Overall, the final model with log_2 bilirubin excluded was preferred as more parsimonious. The final model equation was:

$$P = \frac{1}{1 + e^{-(-8.814 + 0.364 \times \log_2 CRP + 1.768 \times \log_2 WCC + 0.025 \times \text{age} + 0.647 \times [0 \text{ if } female \;\; 1 \text{ if } male])}}$$

The model as a whole explained between 35.9% (Cox–Snell R-square) and 47.9% (Nagelkerke R-square) of the variance in the pathology diagnosis and correctly classified 80.8% of cases (Fig 5). The Hosmer–Lemeshow and Pearson χ^2 goodness-of-fit tests were non-significant indicating satisfactory fit (Table 3).

Table 2 also shows the odds ratios for the prediction of complicated appendix at presentation. In the univariate analysis, all the predictors demonstrated a statistically significant OR. Except for log_2 bilirubin, the multivariate model demonstrated statistically significant odds ratios for age, gender, log_2 WCC and log_2 CRP. For every one-year

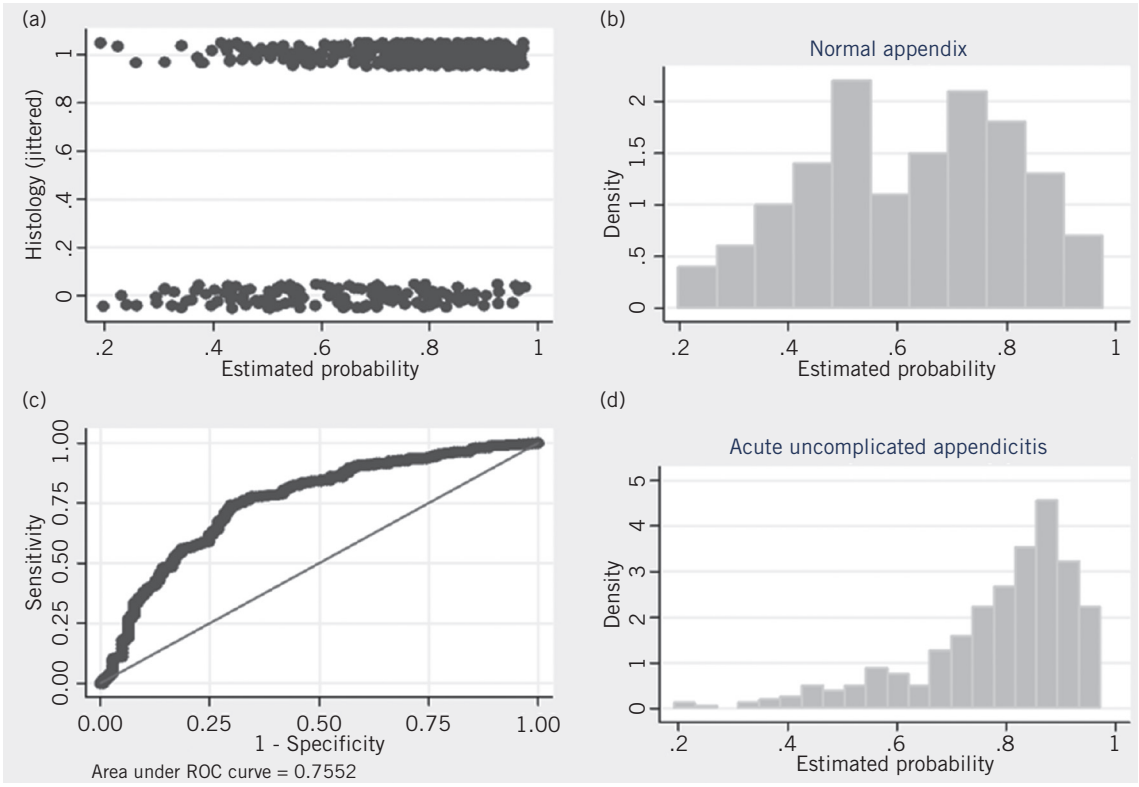

Figure 4 A) Plot of jittered outcome versus estimated probabilities from the fitted model for acute uncomplicated appendicitis (AUA). B) Histogram of estimated probabilities from the fitted model for AUA for the case of normal appendix. C) Receiver operating characteristic curve from the fitted model for AUA. D) Histogram of estimated probabilities from the fitted model for AUA for the case of AUA.

increase in the patient's age the likelihood of complicated appendix increases by 3.0% (OR 1.03). Males were almost 1.9 times more likely to present with complicated appendix than females. The model suggests that there is a 5.8 times increase in the likelihood of complicated appendix for every doubling of WCC (OR 5.84). Similarly, for every doubling of the CRP the likelihood of CA increases by 46% (OR 1.46).

Collinearity was not an issue in this regression model since the predictors were the same as for AUA. The discriminatory power of the model is excellent (area under the ROC curve 0.862) with minimal overlap in the case of normal and complicated appendix (Fig 6).

Cut-off points for the diagnosis of acute uncomplicated and complicated appendicitis

AUC and ROC analyses were performed to assess the sensitivity and specificity of WCC, CRP and bilirubin. Cut-off levels were chosen with priority given to higher sensitivity for WCC and higher specificity for CRP and bilirubin (Fig 7). For the prediction of AUA with normal appendix being the reference category, AUC for WCC, CRP and bilirubin were 0.70, 0.66, and 0.64, respectively (Fig 7a). Cut-off value for WCC at 9×10^9/l demonstrated a sensitivity of 78%, whereas CRP and bilirubin were more specific at levels of 14.2 mg/l (68% specificity) and 12.5 mg/dl (72% specificity), respectively.

For the prediction of complicated appendix with normal appendix being the reference category, AUC for WCC, CRP and bilirubin were 0.800, 0.800, and 0.710, respectively (Fig 7b). Cut-off points to predict complicated appendicitis at presentation for WCC of 10.7×10^9/l demonstrated 80% sensitivity. Cut-off point for the CRP level of 40 mg/l was associated with 80% specificity, whereas a bilirubin level of 14.5 mg/dl was associated with a 76% specificity (Table 4).

Comparison of present models with models from the literature

In 2015, Chambers *et al.*[20] published a logistic regression model predicting gangrenous/perforated appendicitis (complicated appendix in our study) compared with a normal appendix or gangrenous/perforated appendicitis compared with inflamed appendicitis compared with a normal

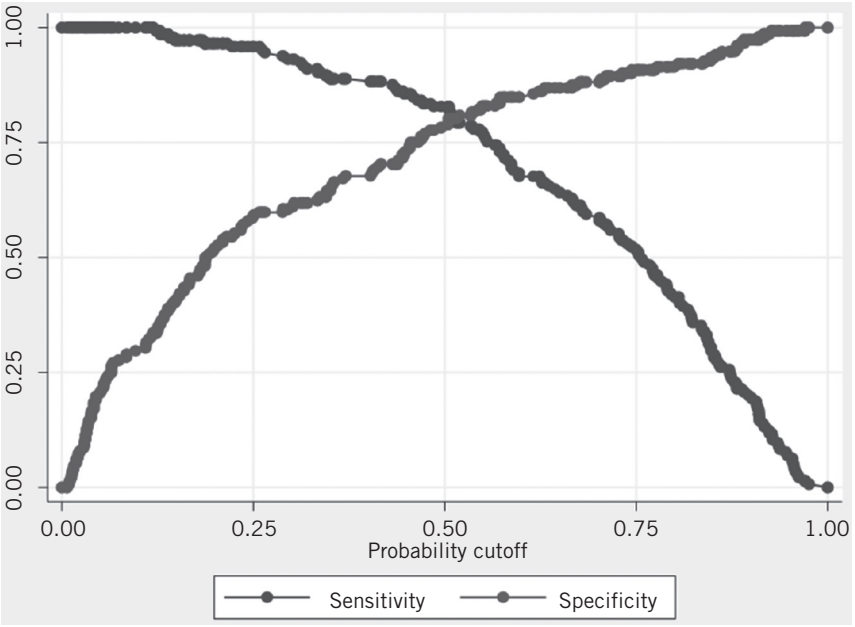

Figure 5 Plot of sensitivity and specificity against probability cut off from the fitted model for acute uncomplicated appendicitis.

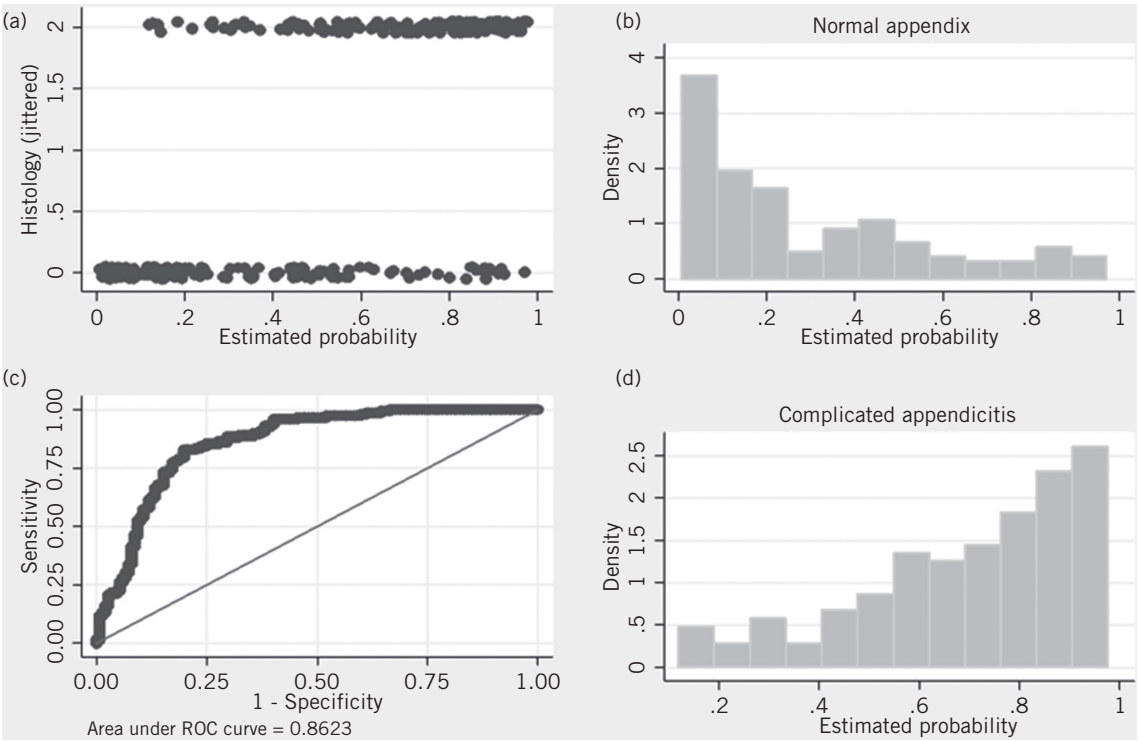

Figure 6 A) Plot of jittered outcome versus estimated probabilities from the fitted model for complicated appendicitis. B) Histogram of estimated probabilities from the fitted model for complicated appendicitis for the case of normal appendix. C) Receiver operating characteristic (ROC) curve from the fitted model for complicated appendicitis. D) Histogram of estimated probabilities from the fitted model for complicated appendicitis for the case of complicated appendicitis.

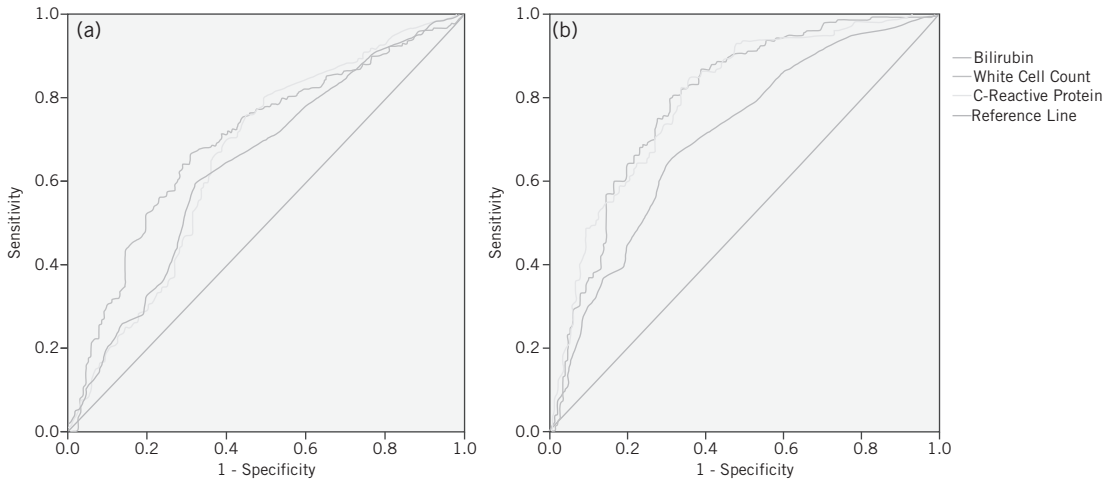

Figure 7 Acute uncomplicated appendicitis and receiver operating characteristic analysis of white cell count (green), C-reactive protein (yellow) and bilirubin (blue) for acute uncomplicated (A) and complicated appendicitis (B).

Table 4 Area under the curve and cut-off values of receiver operating characteristic curve for prediction of acute uncomplicated and complicated appendicitis at presentation. *p < 0.0001. Reference category is normal appendix.

Appendicitis type	Measure	WCC	CRP	Bilirubin
Acute uncomplicated	AUC (95% CI)	0.700 ± 0.02 (0.65–0.75)[a]	0.660 ± 0.03 (0.61-0.71)[a]	0.640 ± 0.03 (0.59-0.69)[a]
	Cut-off point	9 x 10^9/l	14.2 mg/l	12.5 mg/dl
	Sensitivity (%)	78	53	50
	Specificity (%)	50	68	72
Complicated	AUC (95% CI)	0.800 ± 0.03 (0.75-0.85)[a]	0.800 ± 0.3 (0.75-0.85)[a]	0.710 ± 0.3 (0.65-0.76)[a]
	Cut-off point	10.7 × 10^9/l	40 mg/l	14.5 mg/dl
	Sensitivity (%)	80	60	51
	Specificity (%)	70	80	76

AUC, area under the curve; CI, confidence interval; CRP, C-reactive protein; OR, odds ratio; WCC, white cell count.
[a]$P < 0.001$

appendix. Their models were complicated appendix compared with normal appendix (model 1):

$$P = \frac{1}{1 + e^{-(-3.59 + 0.006 \times \text{CRP} + 0.024 \times \text{Bilirubin} + 0.121 \times \text{WCC})}}$$

and complicated appendix compared with inflamed appendix (model 2):

$$\text{P} = \frac{1}{1 + e^{-(-2.77 + 0.005 \times \text{CRP} + 0.061 \times \text{Bilirubin} + 0.211 \times \text{WCC})}}$$

Age and gender were not included in their equation.

We compared our respective models with those published by Chambers *et al*.[20] When comparing model 1 with our respective model, the ROC curves were very similar (Fig 8) and the AUC from ROC curves were not statistically different (Eddama model AUC 0.861 vs Chambers model AUC 0.845, $P = 0.193$). When comparing model 2 with our respective model, the ROC curves were different (Fig 9) and the AUC from ROC curves were statistically different (Eddama model AUC 0.718 vs Chambers model AUC 0.641, $P = 0.002$). The increased predictive ability can possibly be attributed to the inclusion of demographics age and gender and the logarithms of WCC, CRP and bilirubin.

Discussion

Discrimination between AUA and complicated appendix is important in the management of patients presenting with suspected acute appendicitis. There are two reasons why this distinction is important:

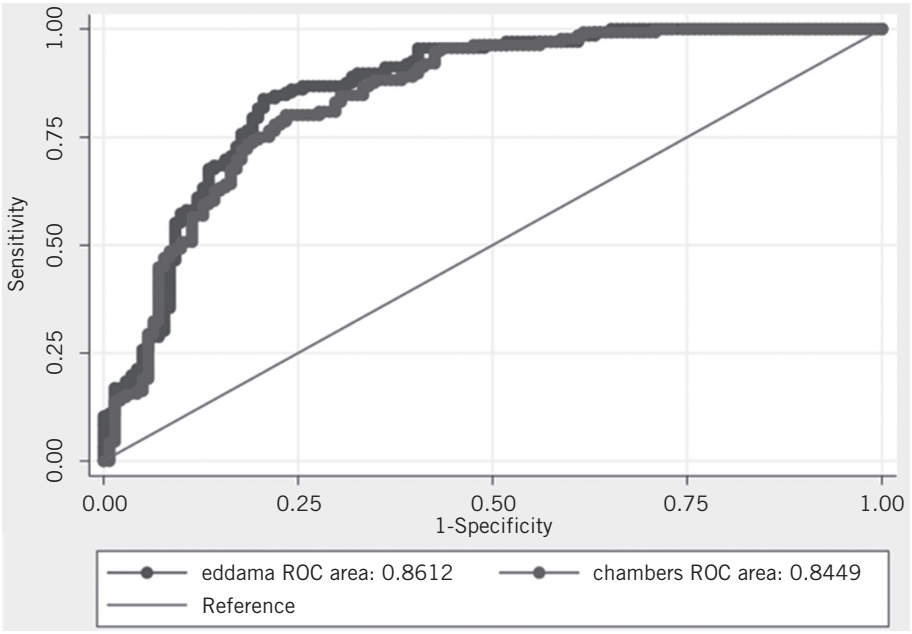

Figure 8 ROC curves for complicated compared with normal appendicitis from the logistic regressions of Eddama and Chambers, which have similar predictive ability (ROC, receiver operating characteristic).[20]

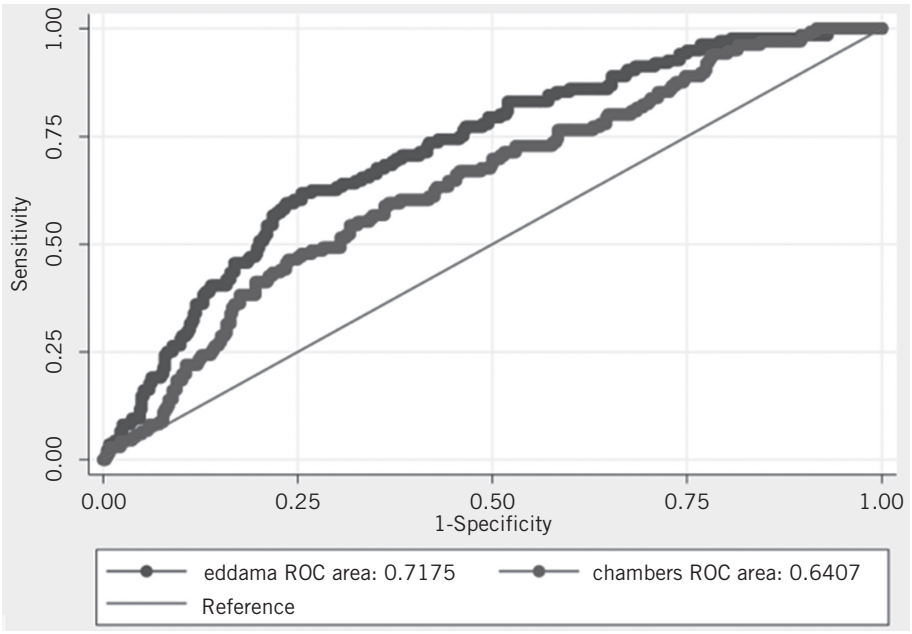

Figure 9 ROC curves for complicated appendicitis compared with acute uncomplicated appendicitis from the logistic regressions of Eddama and Chambers.[20] Eddama's model has slightly better predictive ability (ROC, receiver operating characteristic).

> The most common presentation of appendicitis, AUA, can safely be treated conservatively with antibiotics.[6,21]

> The less common presentation of appendicitis, complicated appendix, almost always requires surgical intervention.[22]

Thus, there is a need to foresee a trajectory towards demarcating the two categories of patients. The most significant finding of this study is an equation to calculate the percentage likelihood of AUA and complicated appendix. In a univariate and multivariate regression analysis, odds ratio for age, gender, WCC, CRP and bilirubin were significant predictors of AUC and complicated appendix. The percentage of likelihood generated from these variables can add an objective measure to the clinical diagnosis.

Our results show that when patients present to the emergency department with suspected appendicitis, males are more likely to have appendicitis on histology than females. Furthermore, the likelihood of complicated appendicitis increases with age. Other studies described a higher incidence of appendicitis in male than females presenting with right iliac fossa pain.[7] Indeed, the likelihood of AUA and complicated appendix is higher in females and increases with age.[7,23] We found that WCC, CRP and bilirubin individually, and in combination, are useful predictors of AUA and complicated appendix. Increased bilirubin levels have been previously shown to predict appendiceal perforation.[24–26] In a univariate analysis, we have demonstrated that raised bilirubin significantly increases in AUA and complicated appendix. However, in a multivariate analysis including age, gender, WCC and CRP, bilirubin was not a significant predictor. This is consistent with a previously published meta-analysis, which showed that bilirubin should be used as a predictor of complicated appendix, but only in conjunction with other predictors.[27] As described previously, this study has further emphasized that WCC and CRP remain significant predictors of both AUA and complicated appendix in a univariate and multivariate regression analysis.[28]

The natural history of appendicitis remains unclear. Although the progression from AUA to complicated appendix remains the strongest hypothesis,[22] other hypotheses marked distinct pathological mechanisms characterising both AUA and complicated appendix.[4] Different microbiological components of AUA and complicated appendix have been described.[29] Moreover, the inflammatory response of patients with appendicitis varies depending on patients demographics and other immunological factors.[30] In this study, we demonstrated a pattern of increase in the level of biochemical markers including WCC, CRP and bilirubin among the patient categories. For example, patients with a normal appendix had lower WCC in comparison with patients with AUA, whereas patients with complicated appendix had the highest WCC levels. Similarly, CRP and bilirubin have the highest means in the complicated appendix category. This finding supports the belief that complicated appendix is a natural progression of untreated AUA.[31]

We also found that WCC has a higher sensitivity of 80% in both acute uncomplicated and complicated appendicitis, whereas CRP and bilirubin tend to have higher specificity, particularly for complicated appendicitis. This is consistent with previously reported results,[32] which indicated similar specificities and sensitivities of the diagnostic values of WCC, CRP and bilirubin in acute appendicitis.

Another interesting finding in this study is the positive correlation between WCC, CRP and bilirubin in the AUA group but not in normal or complicated appendix. Although not specifically verified, this pattern of correlation between the variables may provide clues to distinguish between patients categories. For example, a patient with an isolated significant raise in their CRP, but normal WCC and bilirubin may have an alternative diagnosis. Furthermore, the difference in correlation between the variables in the categories may be an indicator of a different inflammatory response in patients with AUC compared with those who suffer a complicated appendix.

Although similar studies and scoring systems have been developed for the management of suspected appendicitis, such as Alvarado and AIR scores, we consider that this study has offered a percentage likelihood that may distinguish between patient categories. The equation described by the multivariate logistic regression is based on objective measures and has been adapted into a web application for emergency clinicians to use during the assessment of patients with suspected appendicitis. Using a larger sample size collected from the web application, the equation can be validated.

The strength of this study can be summarised. The study included a large number (895) of patients, a sample size valid to produce statistically powered logistic regression. The model exclusively used objective variables and therefore can only add a numerical validity to the clinical suspicion. To avoid selection bias, all cases that underwent emergency appendicectomy and final pathology of the appendix within the study period were included. Patient characteristics and outcome data were extracted from the hospital electronic record retrospectively and confirmed by four different authors (SR, GB, LNB and AW) to ensure accuracy. The immediate perioperative blood results were used to emphasise relevance to clinical judgment. The pathology reports were verified by a consultant pathologist, as a part of healthcare provision and not for the purpose of this study, which makes the outcome variables credible and less likely to be biased.

Despite the strength of this study, there are a few limitations that could not be avoided. The population does not represent the entire population of patients presenting with right iliac fossa pain, as the patients in this study are only those who underwent appendicectomy. This study is retrospective and lacks randomisation, so inherent biases associated with allocation and interpretation of independent and dependent variables may exist. Furthermore, it is possible that some patients have been treated conservatively or re-presented to other hospitals. Also, the use of WCC, CRP and bilirubin to predict the probability of AUA and complicated appendix is subject to available resources that provide the level of these markers.

The progress from this work is made possible by the building of a web application (www.appendistat.com) for clinicians. This has two objectives: first, to create a database at which the logistic regression model can be validated and improved; and second, to facilitate the use of our model in a clinical setting, whereby these variables can be

easily inputted into the web application that is enabled on desktops, tablets and smart-phone devices. In conjunction with the clinical assessment, emergency clinicians can use this tool to support their decisions on the management of patients with right iliac fossa pain and suspected appendicitis. Specifically, conservative management of patients with a low percentage likelihood of AUA and complicated appendix, and immediate surgical intervention for those with high percentage likelihood of complicated appendix can be advocated. Validation of the data from the web application will enable us to generate cut-off-points of the percentage likelihood. This would increase the sensitivity and specificity of this assessment tool.

References

1. Stewart B, Khanduri P, McCord C *et al*. Global disease burden of conditions requiring emergency surgery. *Br J Surg* 2014; **101(1)**: e9–e22.
2. Lozano R, Naghavi M, Foreman K *et al*. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380(9859)**: 2,095–2,128.
3. Velanovich V, Satava R. Balancing the normal appendectomy rate with the perforated appendicitis rate: implications for quality assurance. *Am Surg* 1992; **58(4)**: 264–269.
4. Bennion RS, Baron EJ, Thompson JE, Jr *et al*. The bacteriology of gangrenous and perforated appendicitis – revisited. *Ann Surg* 1990; **211(2)**: 165–171.
5. Carr NJ. The pathology of acute appendicitis. *Ann Diagn Pathol* 2000; **4(1)**: 46–58.
6. Varadhan KK, Neal KR, Lobo DN. Safety and efficacy of antibiotics compared with appendicectomy for treatment of uncomplicated acute appendicitis: meta-analysis of randomised controlled trials. *BMJ* 2012; **344**: e2156.
7. Collaborative S, Cuschieri J, Florence M *et al*. Negative appendectomy and imaging accuracy in the Washington State Surgical Care and Outcomes Assessment Program. *Ann Surg* 2008; **248(4)**: 557–563.
8. Raja AS, Wright C, Sodickson AD *et al*. Negative appendectomy rate in the era of CT: an 18-year perspective. *Radiology* 2010; **256(2)**: 460–465.
9. Wagner PL, Eachempati SR, Soe K *et al*. Defining the current negative appendectomy rate: for whom is preoperative computed tomography making an impact? *Surgery* 2008; **144(2)**: 276–282.
10. Pickhardt PJ, Lawrence EM, Pooler BD, Bruce RJ. Diagnostic performance of multidetector computed tomography for suspected acute appendicitis. *Ann Intern Med* 2011; **154(12)**: 789–796.
11. Alvarado A. A practical score for the early diagnosis of acute appendicitis. *Ann Emerg Med* 1986; **15(5)**: 557–564.
12. Andersson M, Andersson RE. The appendicitis inflammatory response score: a tool for the diagnosis of acute appendicitis that outperforms the Alvarado score. *World J Surg* 2008; **32(8)**: 1,843–1,849.
13. Ohle R, O'Reilly F, O'Brien KK *et al*. The Alvarado score for predicting acute appendicitis: a systematic review. *BMC Med* 2011; **9**: 139.
14. Kulik DM, Uleryk EM, Maguire JL. Does this child have appendicitis? A systematic review of clinical prediction rules for children with acute abdominal pain. *J Clin Epidemiol* 2013; **66(1)**: 95–104.
15. de Castro SM, Unlu C, Steller EP *et al*. Evaluation of the appendicitis inflammatory response score for patients with acute appendicitis. *World J Surg* 2012; **36(7)**: 1,540–1,545.
16. Farooqui W, Pommergaard HC, Burcharth J, Eriksen JR. The diagnostic value of a panel of serological markers in acute appendicitis. *Scand J Surg* 2015; **104 (2)**: 72–78.
17. Lemeshow S, Hosmer DW. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982; **115(1)**: 92–106.
18. Pregibon D. Goodness of link tests for generalized linear models. *Appl Stat* 1980; **29(1)**: 15.
19. Belsley DA, Kuh E, Welsch RE. Detecting and assessing collinearity. In: Belsley DA, Kuh E, Welsch RE, eds. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Hoboken, NJ: Wiley; 2005. pp. 85–191.
20. Chambers AC, Bismohun SL, Davies H *et al*. Predictive value of abnormally raised serum bilirubin in acute appendicitis: a cohort study. *Int J Surg* 2015; **13**: 207–210.
21. Mason RJ. Appendicitis: is surgery the best option? *Lancet* 2011; **377(9777)**: 1,545–1,546.
22. Temple CL, Huchcroft SA, Temple WJ. The natural history of appendicitis in adults: a prospective study. *Ann Surg* 1995; **221(3)**: 278–281.
23. McCartan DP, Fleming FJ, Grace PA. The management of right iliac fossa pain - is timing everything? *Surgeon* 2010; **8(4)**: 211–217.
24. Estrada JJ, Petrosyan M, Barnhart J *et al*. Hyperbilirubinemia in appendicitis: a new predictor of perforation. *J Gastrointest Surg* 2007; **11(6)**: 714–718.
25. Sand M, Bechara FG, Holland-Letz T *et al*. Diagnostic value of hyperbilirubinemia as a predictive factor for appendiceal perforation in acute appendicitis. *Am J Surg* 2009; **198(2)**: 193–198.
26. Atahan K, Ureyen O, Aslan E *et al*. Preoperative diagnostic role of hyperbilirubinaemia as a marker of appendix perforation. *J Int Med Res* 2011; **39(2)**: 609–618.
27. Giordano S, Paakkonen M, Salminen P, Gronroos JM. Elevated serum bilirubin in assessing the likelihood of perforation in acute appendicitis: a diagnostic meta-analysis. *Int J Surg* 2013; **11(9)**: 795–800.
28. Panagiotopoulou IG, Parashar D, Lin R *et al*. The diagnostic value of white cell count, C-reactive protein and bilirubin in acute appendicitis and its complications. *Ann R Coll Surg Engl* 2013; **95(3)**: 215–221.
29. Baron EJ, Bennion R, Thompson J *et al*. A microbiological comparison between acute and complicated appendicitis. *Clin Infect Dis* 1992; **14(1)**: 227–231.
30. Rivera-Chavez FA, Wheeler H, Lindberg G *et al*. Regional and systemic cytokine responses to acute inflammation of the vermiform appendix. *Ann Surg* 2003; **237(3)**: 408–416.
31. Korner H, Sondenaa K, Soreide JA. Perforated and non-perforated acute appendicitis: one disease or two entities? *Eur J Surg* 2001; **167(7)**: 525–530.
32. Emmanuel A, Murchan P, Wilson I, Balfe P. The value of hyperbilirubinaemia in the diagnosis of acute appendicitis. *Ann R Coll Surg Engl* 2011; **93(3)**: 213–217.