Ther Adv Gastroenterol

2021, Vol. 14: 1–11

DOI: 10.1177/17562848211053114

Machine learning for prediction of intra-abdominal abscesses in patients with Crohn's disease visiting the emergency department

Asaf Levartovsky, Yiftach Barash, Shomron Ben-Horin, Bella Ungar, Shelly Soffer, Marianne M. Amitai, Eyal Klang and Uri Kopylov

Abstract

Background: Intra-abdominal abscess (IA) is an important clinical complication of Crohn's disease (CD). A high index of clinical suspicion is needed as imaging is not routinely used during hospital admission. This study aimed to identify clinical predictors of an IA among hospitalized patients with CD using machine learning.

Methods: We created an electronic data repository of all patients with CD who visited the emergency department of our tertiary medical center between 2012 and 2018. We searched for the presence of an IA on abdominal imaging within 7 days from visit. Machine learning models were trained to predict the presence of an IA. A logistic regression model was compared with a random forest model.

Results: Overall, 309 patients with CD were hospitalized and underwent abdominal imaging within 7 days. Forty patients (12.9%) were diagnosed with an IA. On multivariate analysis, high C-reactive protein (CRP) [above 65 mg/l, adjusted odds ratio (aOR): 16 (95% CI: 5.51–46.18)], leukocytosis [above 10.5 K/μl, aOR: 4.47 (95% CI: 1.91–10.45)], thrombocytosis [above 322.5 K/μl, aOR: 4.1 (95% CI: 2–8.73)], and tachycardia [over 97 beats per minute, aOR: 2.7 (95% CI: 1.37–5.3)] were independently associated with an IA. Random forest model showed an area under the curve of 0.817 ± 0.065 with six features (CRP, hemoglobin, WBC, age, current biologic therapy, and BUN).

Conclusion: In our large tertiary center cohort, the machine learning model identified the association of six clinical features (CRP, hemoglobin, WBC, age, BUN, and biologic therapy) with the presentation of an IA. These may assist as a decision support tool in triaging CD patients for imaging to exclude this potentially life-threatening complication.

Keywords: abscess, Crohn complications, machine learning

Received: 8 June 2021; revised manuscript accepted: 23 September 2021.

Introduction

Patients with Crohn's disease (CD) often manifest with complications such as stricturing bowel obstruction and penetration due to fistula or abscess formation. The CD incidence is increasing over the last decades;[1–3] thus, more patients with CD are presenting to the emergency department (ED) with abdominal symptoms. The lifetime risk of intra-abdominal abscesses (IAs) in patients with CD is up to 30%;[4–6] The IA may often lead to serious complications.[7,8] Hypoalbuminemia, smoking, steroid therapy, and prior surgery are risk factors for IA development in patients with CD.[9–11] Several clinical and laboratory findings should raise the suspicion of an IA. These include abdominal pain, fever, and elevated inflammatory markers such as

Correspondence to:
Asaf Levartovsky
Department of Gastroenterology, Sheba Medical Center, Tel Hashomer, Tel-Aviv University, Israel.
alavra@gmail.com

Yiftach Barash
Eyal Klang
Department of Diagnostic Imaging, Sheba Medical Center, Tel Hashomer, Tel-Aviv University, Israel
DeepVision Lab, Sheba Medical Center, Tel Hashomer, Israel

Shomron Ben-Horin
Bella Ungar
Uri Kopylov
Department of Gastroenterology, Sheba Medical Center, Tel Hashomer, Tel-Aviv University, Israel

Shelly Soffer
DeepVision Lab, Sheba Medical Center, Tel Hashomer, Israel; Internal Medicine B, Assuta Medical Center, Ben-Gurion University of the Negev, Ashdod, Israel

Marianne M. Amitai
Department of Diagnostic Imaging, Sheba Medical Center, Tel Hashomer, Tel-Aviv University, Israel

C-reactive protein (CRP), thrombocytosis, and leukocytosis.[12] A multidisciplinary treatment approach is required, including antibiotics, percutaneous drainage, and surgery when necessary. The surgical management of these patients can be difficult and result in significant complications such as colostomy and a subsequent prolonged hospitalization. Therefore, an early diagnosis of an IA is crucial.

The diagnosis is based on imaging modalities, preferably computed tomography (CT), magnetic resonance enterography (MRE), or ultrasound (US). The CT has the advantage of being readily available and guidance during abscess percutaneous drainage, however exposes to more radiation than MRE. Cumulative effective doses of radiation is higher in patients with inflammatory bowel disease (IBD), especially in CD, mainly due to repeated exposures during diagnostic procedures.[13] Large cumulative doses of radiation may have serious ramifications, especially for younger patients with CD.[14,15] In addition, imaging is not routinely used during hospital admission. Thus, a high index of clinical suspicion is needed to diagnose an IA.

In view of these circumstances, a decision support tool can assist in identifying patients with CD who would benefit most from imaging in order to diagnose an IA. In the last decade, machine learning methods have become common among the medical community.[16,17] The combination of advances in machine learning methods with increasingly available digital healthcare data sets enables the medical community to have improved decision-making and better predictive models. Previous machine learning studies evaluated models for different ED outcomes, for example, prediction of mortality, intensive care unit (ICU) admission, and resource utilization.[18–20] A commonly used machine learning model is random forest. Random forest is an ensemble learning method commonly used for classification problems. The model operates by constructing multiple decision trees at training. The output of the random forest is selected by a majority vote of all the trees. Random forests correct the overfitting common to decision trees. As a nonlinear method, it often outperforms linear models, when higher order relationships exist in the data. We compared a linear logistic regression model with a nonlinear random forest model.

In this study, we reviewed a large cohort of patients with CD referred to a large tertiary hospital and determined the incidence of IA diagnosed within a week from admission. We aimed to identify clinical and laboratory predictors of an IA using machine learning techniques.

Materials and methods

We created an electronic data repository of all patients with CD who visited the ED of the Sheba medical center, a tertiary medical center, during the years 2012–2018. Data included demographic, clinical, and laboratory variables, as well as free-text imaging reports. For this study, we searched the data for CD patients with an IA who were diagnosed within a week from admission.

We included patients with CD aged 16 years and above, who presented with complaints that can possibly be attributed to CD or an IA (abdominal pain, diarrhea, vomiting, and fever). We excluded patients without these complaints or those not related to CD such as trauma, cardiac and respiratory symptoms, and neurological issues.

We excluded patients who were discharged from the ED and patients who did not undergo abdominal imaging (CT or MRE) within a week from admission. In addition, patients presenting within 30 days from an abdominal surgery were excluded. To avoid redundancy from repeated visits of the same patient, we included only the first ED visit for each patient.

An IA was diagnosed based on imaging reports of clinicians from the hospital's radiology department. The initial reports were revised by senior abdominal radiologists as those reports were eventually the standard for the diagnosis.

Retrieved data included demographics, IBD details (disease extent, behavior, extraintestinal manifestations, and therapy), comorbidities, clinical characteristics, vital signs measurements at admission (pulse, blood pressure, fever), and laboratory values (hemoglobin, white blood cells, albumin, CRP, platelets, and creatinine). The CRP levels are considered normal in our medical center if below 5 mg/l.

The primary analysis was of patients diagnosed with an IA within 7 days from admission compared with patients who were not diagnosed with an IA. The comparison included statistical analysis of clinical and laboratory factors between the groups. Subsequently, a secondary analysis of machine learning models was performed.

Data analysis methods

Machine learning models were trained to predict the presence of abscess. Data preprocessing included median imputation of missing values. We have compared a logistic regression model with a random forest model. Due to the sample size, to obtain statistical significance, for each experiment we have performed 100 random splits of 80% training and 20% testing, and averaged the results.

A regularized logistic regression model (ridge regression) was implemented using the scikit-learn library, including hyperparameters (l2 regularization, fit intercept, intercept scaling, class weight, and random state). Random forest hyperparameters included 200 estimators, with "gini" split. Data balancing techniques were not employed. For comparison, we have also evaluated the predictive power of CRP alone. In order for the comparison to be accurate in terms of the split between the training set and the testing set, we used CRP as a single feature in the logistic model, using the same data split as described.

We have used recursive feature elimination (RFE) to establish the optimal number of features in the models. The RFE was conducted by continuously fitting the model on a subset of features. In each round, all the remaining features in the subset were ranked. This was done by computing the area under the receiver operating characteristic (ROC) curve (AUC) for all the features in the subset other than the ranked feature. The least important feature was than discarded from the subset. This process was repeated until all the features were eliminated. The AUCs for each subset of features were recorded. The subset with the highest AUC was noted.

Sensitivity analysis was not performed.

The final model's metrics included AUC, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy.

The AUC was used as a performance measure for machine learning models, as this is a common benchmark used for this task.[21] The ROC curve is plotted with sensitivity *versus* (1 – specificity) in which sensitivity is on the y-axis and (1 – specificity) is on the x-axis. The AUC measures the calculated area under the ROC curve. It tells how much the model is capable of distinguishing between the positive and negative classes. The AUC ranges between 0 and 1, with higher values representing better performance.

Programming was done with Python (version 3.6.5 64bits).

Statistical analysis

Continuous variables were presented as medians and interquartile ranges (IQRs) and categorical variables as frequency and percentage. The association between IA presentation and categorical variables was studied using chi-square test and Fisher's exact test. The association between IA presentation and continuous variables was assessed using Mann–Whitney *U* test. Variables that were significantly associated with readmission at the univariate analysis were included in the multivariable analysis. For multivariate analysis, all variables with $p < 0.1$ were included in the model. A ROC analysis was performed for evaluation of the presence of an IA by clinical and laboratory parameters.

All statistical tests were two-sided, and a p value <0.05 was considered statistically significant. Statistical analysis was performed using the SPSS v23 statistical software (Armonk, NY, USA).

This article was prepared in accordance to the STROBE statement.[22]

Study ethics and patient consent

This study was carried out in accordance with the ethical guidelines of the Declaration of Helsinki. The study was approved by the institutional review board of the Sheba Medical Center. Because this was a retrospective analysis, no informed consent was obtained.

Results

Overall, 4507 ED visits of 1556 patients with CD occurred in the study period. We excluded 2124 visits due to complaints and symptoms not attributed to CD. An additional 1046 visits were

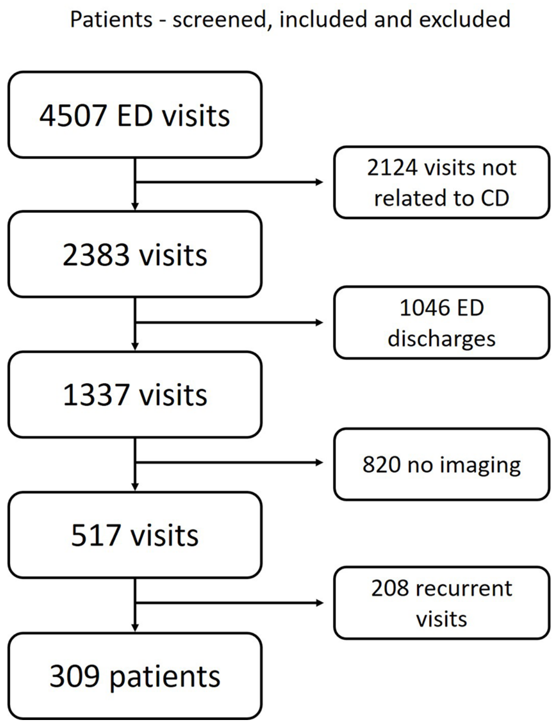

Figure 1. Flow chart showing the screened, included, and excluded patients.
CD, Crohn's disease; ED, emergency department.

excluded for being discharged from the ED, and 820 for not undergoing abdominal imaging (CT or MRE) within a week from admission. Of the 517 remaining patient visits, we included only the first ED visit for each patient and excluded cases with insufficient data. The remaining study population included 309 patients with CD who visited the ED with relevant complaints, were hospitalized, and underwent an abdominal imaging within a week from admission (Figure 1).

The median age of the study population was 37 (IQR: 26–50) years. A total of 167 patients (54%) were males. Forty-eight patients (15.5%) were currently treated with biologic therapy. Table 1 presents the background characteristics of the whole study population.

IA on imaging

Of the 309 patients in the study population, 40 patients (12.9%) had an IA on abdominal imaging during their admission (39 CT scans, one MRE). The median size of the largest abscess was 32 (IQR: 20.5–44.5) mm, as 20 patients (50%) had an abscess larger than 3 cm. Four patients (10%) were current or past smokers and seven patients (17.5%) underwent an abdominal surgery in their past history. On CD diagnosis, the majority of patients had a nonpenetrating and nonstricturing disease and ileal involvement, 72.5% and 50%, respectively.

The IBD and abscess characteristics of patients with an IA on presentation are summarized in Table 2.

The patients with an IA were younger compared with the 269 patients without an IA [median: 30 (IQR: 23–42.5) years *versus* 38 (IQR: 28–51) years, $p = 0.004$]. Tachycardia [heart rate > 100 beats per minute (bpm)] (p value = 0.012), leukocytosis (serum white blood cells > 11 K/μl) (p value = 0.009), and thrombocytosis (serum platelets > 440 K/μl) (p value = 0.017) were associated with an IA on imaging 7 days from admission. Neutrophil-to-lymphocyte ratio (NLR) was also higher in patients with an IA compared with patients without an IA [median: 9.2 (IQR: 4.8–17.14) *versus* 5.9 (IQR: 3.7–11.1), $p = 0.007$].

Table 3 shows the clinical and laboratory findings in both groups.

On multivariate analysis, high CRP values [above 65 mg/l, adjusted odds ratio (aOR): 16 (95% confidence interval, CI: 5.51–46.18)], leukocytosis [above 10.5 K/μl, aOR: 4.47 (95% CI: 1.91–10.45)], high NLR [above 7.7, aOR: 3.1 (95% CI: 1.56–6.3)], high platelet count [above 322.5 K/μl, aOR: 4.1 (95% CI: 2–8.73)], age [younger than 22.5 years, aOR 2.71 (95% CI 1.15–6.3)], and higher heart rate [over 97 bpm, aOR: 2.7 (95% CI: 1.37–5.3)] were independently associated with an IA.

A nomogram was computed to demonstrate the incidence of an IA based on CRP (Figure 2). The CRP was the variable most strongly associated with an IA among the variables mentioned above and six cutoff values were determined. As viewed in the figure, the PPV of an IA in patients with CD was greater as CRP cutoff values were higher. The PPV of an IA on abdominal imaging was lowest when examining all patients with a CRP above 5 mg/l (14.6% of patients), whereas incidence was greatest for a CRP cutoff above 85 mg/l (29.4% of patients). The NPV of an IA for CRP below cutoffs of 5 and 85 was 100% and 95.2%, respectively. The nomogram model demonstrated significant performance for each cutoff ($p = 0.01$).

In terms of clinical outcomes, all the 40 patients with an IA were treated with intravenous antibiotics during the hospitalization. Nineteen patients (47.5%) were treated surgically and these included 14 procedures of incision and drainage and five partial colectomies. Nine patients (22.5%) needed escalation in therapy during the upcoming gastroenterology clinic follow-up [mostly anti-TNFα (tumor necrosis factor α) induction].

Machine learning prediction model

Figure 3(a) and (b) presents the results of the RFE experiments of the logistic regression and random forest models. Overall, random forest and logistic regression showed similar performance. A logistic regression model showed an AUC of 0.816 ± 0.065 with seven features [CRP, hemoglobin, white blood cells (WBC), age, amylase, current biologic therapy, and current immunomodulatory therapy]. Top random forest model showed an AUC of 0.817 ± 0.065 with six features [CRP, hemoglobin, WBC, age, current biologic therapy, and blood urea nitrogen (BUN)]. For comparison, CRP alone achieved an AUC of 0.765 ± 0.071 ($p < 0.001$).

Supplementary Table 1 presents the averaged metrics table (sensitivity, specificity, PPV, NPV, and accuracy) of the random forest model for different cutoff values corresponding to specificities of 30%, 60%, and 90%.

Discussion

In a tertiary care center, there is a high frequency of referrals for patients with uncontrolled CD complicated disease manifestations. The IA, a major complication of CD, is diagnosed through abdominal imaging and, in most cases, abdominal CT. It is crucial to identify the patients presenting to the ED with an IA because their presence can have a significant role on disease management. In a study from 2013, clinical management was modified in 80.6% of CD visits based on the findings of abdominal CT. The decisions concerned mainly about hospitalization and discharge from the ED.[23]

As previously mentioned, cumulative effective dose of radiation is high in patients with IBD,

Table 1. Characteristics of total study population.

Study population, *N* = 309	
Male gender, *n* (%)	167 (54)
Age, *n* (IQR)	37 (26–50)
No prior hospitalization, *n* (%)	214 (69.3)
Comorbidities, *n* (%)	
Hypertension	18 (5.8)
Dyslipidemia	13 (4.2)
Hypothyroidism	6 (1.9)
COPD	3 (0.97)
Diabetes mellitus	5 (1.6)
CD extent at diagnosis, *n* (% of 226 patients)	
L1 (ileal)	127 (56.2)
L2 (colonic)	17 (7.5)
L3 (ileo-colonic)	82 (36.3)
Perianal disease	94 (41.6)
CD behavior at diagnosis, *n* (% of 283 patients)	
B1 (nonstricturing and nonpenetrating)	110 (38.9)
B2 (stricturing)	57 (20.1)
B3 (penetrating)	116 (41)
Current IBD treatment, *n* (%)	
Immunomodulators	47 (15.2)
Biologics	48 (15.5)
ASA	34 (11)
Steroids	35 (11.3)
Past biologic treatment, *n* (%)	
Infliximab	113 (36.6)
Adalimumab	122 (39.5)
Vedolizumab	53 (17.1)
Ustekinumab	45 (14.6)
Current or past smoker, *n* (%)	88 (28.5)
EIM, *n* (%)	84 (27.9)

ASA, 5-aminosalicylic acid; CD, Crohn's disease; COPD, chronic obstructive pulmonary disease; EIM, extra-intestinal manifestation; IBD, inflammatory bowel disease; IQR, interquartile range.

Table 2. IBD and abscess characteristics of patients with an IA on presentation.

Number of patients	40
Male, *n* (%)	21 (52.5)
Female, *n* (%)	19 (47.5)
CT, *n* (%)	39 (97.5)
MRE, *n* (%)	1 (2.5)
Age at CD diagnosis, years, median (IQR)	22.5 (17.25–29)
Current or past smoker, *n* (%)	4 (10)
Previous abdominal surgery, *n* (%)	7 (17.5)
EIM, *n* (%)	5 (12.5)
Abscess size, mm, median (IQR)	32 (20.5, 44.5)
Small (⩽30 mm)	20 (50)
Large (>30 mm)	20 (50)
Number of abscess, *n* (%)	
1	28 (70)
2	5 (12.5)
3 or more	7 (17.5)
CD extent at diagnosis, *n* (%)	
L1 (ileal)	20 (50)
L2 (colonic)	4 (10)
L3 (ileo-colonic)	16 (40)
Perianal disease	3 (7.5)
CD behavior at diagnosis, *n* (%)	
B1 (nonstricturing and nonpenetrating)	29 (72.5)
B2 (stricturing)	7 (17.5)
B3 (penetrating)	4 (10)
B2 + B3	2 (4.8)
Previous therapy, *n* (%)	
Immunomodulators	14 (35)
Biologics	15 (37.5)
ASA	18 (45)
Previous biologic therapy, *n* (%)	
Infliximab	18 (45)
Adalimumab	15 (33.3)
Vedolizumab	6 (15)
Ustekinumab	6 (15)

ASA, 5-aminosalicylic acid; CD, Crohn's disease; CT, computed tomography; EIM, extra-intestinal manifestation; IA, intra-abdominal abscess; IBD, inflammatory bowel disease; IQR, interquartile range; MRE, magnetic resonance enterography.

especially in CD, due to the incidence of abdominal CT scans. An important component in the clinician's work is the attempt to screen patients who could avoid further radiation among patients with CD. Thus, better decision-making tools are needed in order to select the patients who can benefit from abdominal imaging.

We evaluated clinical and laboratory factors that were associated with the presence of an IA on abdominal imaging, in the ED or during the first week of hospitalization. Patients with an IA were younger than control patients, a finding also demonstrated in a recent nationwide study.[24] We did not find an association between sex or comorbidities to an IA.

In this study, CRP levels were significantly higher in patients presenting with an IA compared with patients without an IA (114.21 *versus* 38.2 mg/l, respectively). Previous studies reported similar findings, as high CRP levels were associated with acute findings on imaging such as an IA, obstruction, or fistula in patients with CD.[25] In addition, CRP plays a role in IA formation during biologic treatment. High CRP levels measured 14 weeks after infliximab initiation can predict abscess formation.[26] As CRP is a widely used and inexpensive laboratory marker, it can allow a better selection of patients who could avoid abdominal imaging and unnecessary radiation with a high NPV. Desmond *et al.* sampled CRP assays from 147 patients with CD, a median of 2 days before abdominal imaging. Patients with normal CRP were significantly less likely to have penetrating or large bowel luminal disease.[27] Govani *et al.* created a model for patients with CD that predicted complications on CT with a miss rate of 0.8%. The model was then applied on two tertiary care centers using a decision tool including erythrocyte sedimentation rate (ESR) in addition to CRP.[12,28] In our study, 14.6% of patients with a CRP value above 5 mg/l had an IA. Incidence was reasonably higher for greater CRP values (29.4% of patients above CRP 85 mg/l). The NPV of an IA for low CRP values showed similar results to the studies mentioned above.

Additional laboratory markers may indicate an inflammatory process due to an undiagnosed IA. Such markers include white blood cells and platelets that take part in the systemic inflammation process. Studies from 2013 and 2014 found leukocytosis (above 10,000–12,000) as an independent predictor

Table 3. Characteristics of patients with an IA and control.

Variable	Non-abscess population ($n = 269$)	IA population ($n = 40$)	p value
Male gender	54.3%	52.5%	0.83
Age (IQR)	38 (28–51)	30 (23–42.5)	**0.004**
Days hospitalized (IQR)	4 (3–7)	6.5 (4–11)	**0.005**
Clinical features			
Pulse (IQR)	91 (77.25–103)	99.5 (88.25–114.75)	**0.004**
Fever	36.9 ± 2.2	37 ± 0.86	0.23
Systolic blood pressure (IQR)	117 (108–127)	111.5 (105–122)	0.07
Laboratory values (IQR)			
WBC	10.76 (8.13–14.31)	13.67 (10.7–16.64)	**0.003**
NLR	5.9 (3.67–11.12)	9.2 (4.87–17.14)	**0.007**
Hemoglobin	12.96 (11.65–14.24)	11.86 (10.4–12.43)	<**0.001**
Platelets	299 (231.5–377.5)	364 (303–486. 5)	**0.001**
Creatinine	0.79 (0.68–0.97)	0.8 (0.64–1.01)	0.9
LDH	193 (154.75–250)	162 (128–227.5)	**0.016**
Sodium	137 (135–138.75)	136 (134–138)	0.06
Urea	24.5 (19–31.75)	23.5 (19.25–28)	0.33
CRP	38.2 (11.96–92.8)	114.21 (85–150.2)	<**0.001**

CRP, C-reactive protein; IA, intra-abdominal abscess; IQR, interquartile range; LDH, lactate dehydrogenase; NLR, neutrophil-to-lymphocyte ratio; WBC, white blood cells.
The p value <0.05 was considered statistically significant (highlighted in bold).

of urgent findings on abdominal CTs, most commonly an IA.[29,30] The NLR, a prognostic factor in infectious diseases, can be used to predict the diagnosis of an IA complicating CD.[31] This factor was also used for the prediction of an IA in our study. We demonstrated on multivariate analysis that leukocytosis, thrombocytosis, and NLR were independently associated with an IA in patients with CD. Due to the potentially infectious nature of CD complications, there is a direct association between the levels of these laboratory markers and the presence of an IA. A recent machine learning–based approach for predicting clinically actionable findings on abdominal CT demonstrated the significance of an elevated NLR and these findings.[32]

Patients with CD can present to the ED with anemia, a result of malabsorption and inflammation. This sign can be especially noticeable during flares, and this study shows the association of anemia with an IA [hemoglobin: 12.96 (IQR: 11.65–14.24) in control *versus* 11.86 (IQR: 10.4–12.43) in patients with an IA, $p < 0.001$]. Tachycardia, a physiological marker mostly attributed to pain and inflammation, was associated with the presence of an IA in our study (above 97 bpm). This clinical sign can predict urgent clinical findings in abdominal CT.[29] An additional physiologic sign, fever, can be attributed to an IA formation.[31] In our study, however, there was no significant association between fever and the presence of an IA.

The main objective of this study was to predict the presence of an IA in patients with CD presenting to the ED. A machine learning algorithm

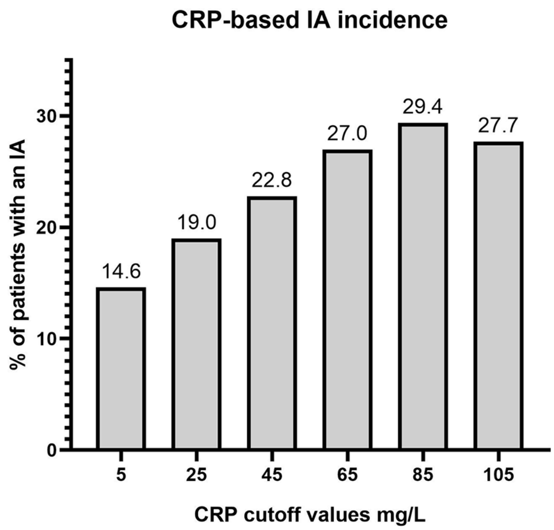

Figure 2. Nomogram demonstrating the PPV of an IA in patients with CD, based on CRP cutoff values (above 5, 25, 45, 65, 85, or 105 mg/l). The PPV of an IA was greater as CRP cutoff values were higher. The CRP above 5 mg/l had a PPV of 14.6%, whereas 29.4% for a CRP cutoff above 85 mg/l.
CD, Crohn's disease; CRP, C-reactive protein; IA, intra-abdominal abscess; PPV, positive predictive value.

was used to establish better clinical-decision guidance regarding the need for abdominal imaging. The machine learning model in our study predicted the presence of an IA based on six clinical markers. The AUC was 0.813, and the accuracies were 39%, 63%, and 85%, for fixed specificities of 30%, 60%, and 90%, respectively. Most of the results obtained from this analysis are consistent with the findings discussed above (CRP, hemoglobin, white blood count, and age). Biologic treatment is the mainstay therapy today in IBD; however, this therapy holds increased risk of infection, including abscesses. Moreover, the presence of an abscess is an absolute contraindication to anti-TNF therapy.[33] This might explain the findings in our study demonstrating an association between biologic therapy and the presence of an IA. Yet, an additional explanation might be that this population *a priori* has a more complicated and uncontrolled disease and therefore needed medical care in the ED. The BUN is a representative of hydration status that is affected by inflammation. Raised levels can be explained due to inflammation conditions such as an IA.

In this study, we compared two machine learning models. First, we trained a logistic regression model. Logistic regression is a well-established model. The model is easy to implement and easy to interpret and does not require significant resources. Second, we trained a random forest model. Random forest is an ensemble machine learning algorithm in which boosting is used to create multiple decision trees. Together, the ensemble produces superior results. As a nonlinear method, this aggregation of decision trees often outperforms linear models, when higher order relationships exist in the data.

The AUCs were similar between logistic regression and random forest models. This is perhaps due to having strong linear low-order correlations between the data and outcome (e.g. inflammatory markers). Because logistic regression is simple and easily interpretable, it should be considered for the task.

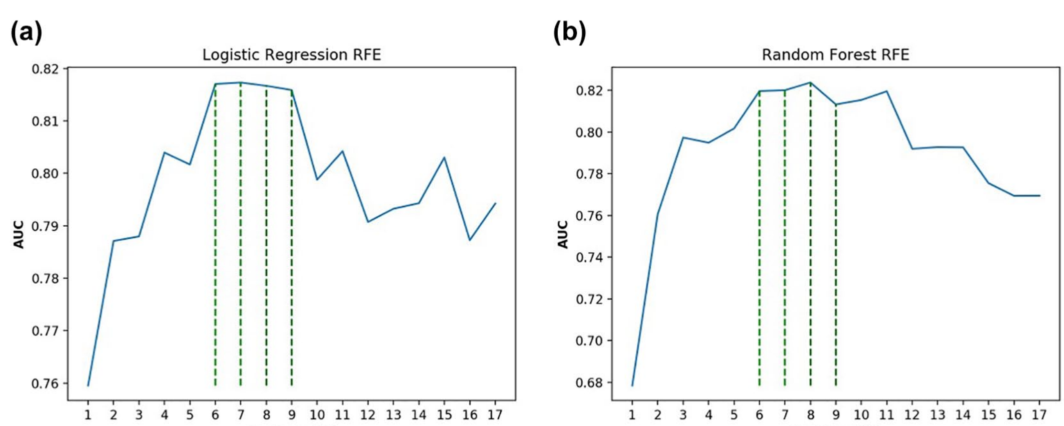

Figure 3. AUC graphs of RFE experiments of the (a) logistic regression and (b) random forest models.
AUC, area under the receiver operating characteristic curve; RFE, recursive feature elimination.

A machine learning decision support system as described in this study may be implemented in the emergency room EHR (electronic health record). This will help the physician when making clinical decisions such as ordering imaging tests, and discharging or admitting the patient. Although human intuition and knowledge is extremely important, the accumulated clinical data continue to grow. Thus, statistical tools as described may improve patient care.

This study has several limitations. First is the retrospective nature of the data collection from more than 4500 ED visits. Second, we included only patients who underwent abdominal imaging. Therefore, a selection bias is introduced as these patients *a priori* had a high clinical suspicion of an IA. This was, however, done for the purpose of implementing the machine learning algorithm. A larger cohort of all patients with CD, including visits without imaging, may further be required. Third, we excluded patients who were discharged from the ED under the assumption that they were presenting an uncomplicated disease. Some of these patients underwent abdominal imaging and discharged due to the absence of findings on imaging. Fourth, various gastrointestinal radiologists reviewed the imaging scans, therefore a potential interobserver variability. The likelihood of a gastrointestinal radiologist not detecting and differentiating an abscess through abdominal imaging, however, is low. Fifth, like many machine learning models, random forest outputs may be considered as a "black box," which may limit its acceptance by clinicians. Finally, this was a single-center study; therefore, further external validation with multicenter data is necessary before clinical implementation.

The strengths of this study include the investigation of a large cohort of patients with CD, being conducted in a tertiary center setting, and the proof-of-concept results supporting a role for machine learning in this clinical setting.

In conclusion, in this study we aimed to predict the presence of an IA. This was done in a large tertiary center cohort using a machine learning model. We identified clinical and laboratory features associated with an IA that were used by the algorithm to create a combined predictive model. Such a decision support tool may assist better triaging patients with CD for imaging in order to exclude this potentially life-threatening complication and avoid unnecessary radiation exposure.

Acknowledgements
Eyal Klang and Uri Kopylov contributed equally.

Author contributions
Conceptualization – UK; Formal analysis – AL, YB, SS, EK; Investigation – AL, YB, EK; Writing original draft – AL, EK, UK; Writing review & editing – AL, YB, SBH, BU, SS, MMA, EK, UK; All authors have read and agree to the published version of the manuscript.

Conflict of interest statement
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs
Asaf Levartovsky https://orcid.org/0000-0003-1235-1927

Shelly Soffer https://orcid.org/0000-0002-7853-2029

Uri Kopylov https://orcid.org/0000-0002-7156-0588

Supplemental material
Supplemental material for this article is available online.

References
1. Danese S and Fiocchi C. Etiopathogenesis of inflammatory bowel diseases. *World J Gastroenterol* 2006; 12: 4807–4812.

2. Molodecky NA, Soon IS, Rabi DM, *et al*. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012; 142: 46–54.e42.

3. Ng WK, Wong SH and Ng SC. Changing epidemiological trends of inflammatory bowel disease in Asia. *Intest Res* 2016; 14: 111–119.

4. Greenstein AJ, Sachar DB, Greenstein RJ, *et al*. Intraabdominal abscess in Crohn's (ileo) colitis. *Am J Surg* 1982; 143: 727–730.

5. Yamaguchi A, Matsui T, Sakurai T, *et al*. The clinical characteristics and outcome of

intraabdominal abscess in Crohn's disease. *J Gastroenterol* 2004; 39: 441–448.

6. Keighley MRB, Eastwood D, Ambrose NS, *et al*. Incidence and microbiology of abdominal and pelvic abscess in Crohn's disease. *Gastroenterology* 1982; 83: 1271–1275.
7. Poritz LS and Koltun WA. Percutaneous drainage and ileocolectomy for spontaneous intraabdominal abscess in Crohn's disease. *J Gastrointest Surg* 2007; 11: 204–208.
8. Louis E, Collard A, Oger AF, *et al*. Behaviour of Crohn's disease according to the Vienna classification: changing pattern over the course of the disease. *Gut* 2001; 49: 777–782.
9. Agrawal A, Durrani S, Leiper K, *et al*. Effect of systemic corticosteroid therapy on risk for intra-abdominal or pelvic abscess in non-operated Crohn's disease. *Clin Gastroenterol Hepatol* 2005; 3: 1215–1220.
10. Picco M. Tobacco consumption and disease duration are associated with fistulizing and stricturing behaviors in the first 8 years of Crohn's disease. *Am J Gastroenterol* 2003; 98: 363–368.
11. Huang W, Tang Y, Nong L, *et al*. Risk factors for postoperative intra-abdominal septic complications after surgery in Crohn's disease: a meta-analysis of observational studies. *J Crohns Colitis* 2015; 9: 293–301.
12. Govani SM, Guentner AS, Waljee AK, *et al*. Risk stratification of emergency department patients with Crohn's disease could reduce computed tomography use by nearly half. *Clin Gastroenterol Hepatol* 2014; 12: 1702–1707.e3.
13. Chatu S, Subramanian V and Pollok RCG. Meta-analysis: diagnostic medical radiation exposure in inflammatory bowel disease. *Aliment Pharmacol Ther* 2012; 35: 529–539.
14. Brenner DJ, Doll R, Goodhead DT, *et al*. Cancer risks attributable to low doses of ionizing radiation: assessing what we really know. *Proc Natl Acad Sci USA* 2003; 100(Suppl. 2): 13761–13766.
15. Newnham E, Hawkes E, Surender A, *et al*. Quantifying exposure to diagnostic medical radiation in patients with inflammatory bowel disease: are we contributing to malignancy? *Aliment Pharmacol Ther* 2007; 26: 1019–1024.
16. Klang E. Deep learning and medical imaging. *J Thorac Dis* 2018; 10: 1325–1328.
17. Kourou K, Exarchos TP, Exarchos KP, *et al*. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J* 2015; 13: 8–17.
18. Barash Y, Soffer S, Grossman E, *et al*. Alerting on mortality among patients discharged from the emergency department: a machine learning model. *Postgrad Med J*. Epub ahead of print 3 December 2020. DOI: 10.1136/postgradmedj-2020-138899.
19. Klug M, Barash Y, Bechler S, *et al*. A gradient boosting machine learning model for predicting early mortality in the emergency department triage: devising a nine-point triage score. *J Gen Intern Med* 2020; 35: 220–227.
20. Klang E, Kummer BR, Dangayach NS, *et al*. Predicting adult neuroscience intensive care unit admission from emergency department triage using a retrospective, tabular-free text machine learning approach. *Sci Rep* 2021; 11: 1381.
21. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit* 1997; 30: 1145–1159.
22. vonElm E, Altman DG, Egger M, *et al*. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007; 335: 806–808.
23. Israeli E, Ying S, Henderson B, *et al*. The impact of abdominal computed tomography in a tertiary referral centre emergency department on the management of patients with inflammatory bowel disease. *Aliment Pharmacol Ther* 2013; 38: 513–521.
24. Jeong S-H, Choi JS, Kim JW, *et al*. Clinical features of intra-abdominal abscess and intestinal free-wall perforation in Korean patients with Crohn's disease: results from the CONNECT study. *J Clin Med* 2020; 10: 116.
25. Yarur AJ, Mandalia AB, Dauer RM, *et al*. Predictive factors for clinically actionable computed tomography findings in inflammatory bowel disease patients seen in the emergency department with acute gastrointestinal symptoms. *J Crohns Colitis* 2014; 8: 504–512.
26. Yoneno K, Hisamatsu T, Matsuoka K, *et al*. Risk and management of intra-abdominal abscess in Crohn's disease treated with infliximab. *Digestion* 2014; 89: 201–208.
27. Desmond AN, O'Regan K, Malik N, *et al*. Selection of symptomatic patients with

Crohn's disease for abdominopelvic computed tomography: role of serum C-reactive protein. *Can Assoc Radiol J* 2012; 63: 267–274.

28. Govani SM, Waljee AK, Kocher KE, *et al*. Validation of a tool predicting important findings on computed tomography among Crohn's disease patients. *United European Gastroenterol J* 2017; 5: 270–275.
29. Jung YS, Park DI, Hong SN, *et al*. Predictors of urgent findings on abdominopelvic CT in patients with Crohn's disease presenting to the emergency department. *Dig Dis Sci* 2015; 60: 929–935.
30. Kerner C, Carey K, Baillie C, *et al*. Clinical predictors of urgent findings on abdominopelvic CT in emergency department patients with Crohn's disease. *Inflamm Bowel Dis* 2013; 19: 1179–1185.
31. Khoury T, Daher S, Massarwa M, *et al*. A validated score assessing the risk of an intra-abdominal abscess in patients with Crohn's disease presenting at the emergency department. *J Crohns Colitis* 2019; 13: 1131–1137.
32. Konikoff T, Goren I, Yalon M, *et al*. Machine learning for selecting patients with Crohn's disease for abdominopelvic computed tomography in the emergency department: AI selects CD patients for CT. *Dig Liver Dis*. Epub ahead of print 9 July 2021. DOI: 10.1016/j.dld.2021.06.020.
33. Dignass A, Van Assche G, Lindsay JO, *et al*. The second European evidence-based consensus on the diagnosis and management of Crohn's disease: current management. *J Crohns Colitis* 2010; 4: 28–62.