Risk Stratification of Emergency Department Patients With Crohn's Disease Could Reduce Computed Tomography Use by Nearly Half

Shail M. Govani,[*,‡] Amanda S. Guentner,[‡] Akbar K. Waljee,[*,‡,§] and Peter D. R. Higgins[*,‡]

*Division of Gastroenterology, ‡Department of Internal Medicine, University of Michigan, Ann Arbor; and §Veterans Affairs Center for Clinical Management Research, Ann Arbor, Michigan

BACKGROUND & AIMS: Computed tomography (CT) is a useful tool for assessing disease activity and excluding complications in patients with Crohn's disease (CD). However, excessive radiation increases risk for malignancy. We aimed to identify automatable algorithms with high negative predictive values for significant CT findings in patients with CD who present at the emergency department.

METHODS: We conducted a retrospective review of a tertiary center's medical records to identify adults diagnosed with CD who presented from 2000 through 2011. Logistic regression was used to model complications (perforations, abscesses, or other serious findings) and inflammation.

RESULTS: There were 1095 visits made by 613 individuals that included a CT scan within 24 hours of arrival. The average number of CT scans was 1.8 (range, 1–31). Complications of CD were observed in 16.8% of CT scans, inflammation in 54.5%, and new/worse findings in 67.2%. On the basis of 10-fold cross-validation, the area under the receiver operating characteristic curve value for the complications model was 0.80 (95% confidence interval, 0.74–0.86) and for the inflammation model was 0.71 (95% confidence interval, 0.68–0.74). Scanning only patients with model-predicted complications would reduce scans by 43.0%, with a miss rate of 0.8% (4 of 491).

CONCLUSIONS: Patients presenting to the emergency department with CD are frequently assessed by CT. However, no significant findings are observed in 32.8%, and only 17% have complications from CD. We created models to identify patients not likely to have significant findings from CT with high negative predictive values; these could aid physicians in avoiding CT scans for many patients. Studies are needed to validate these models beyond a single center.

Keywords: Inflammatory Bowel Disease; Imaging; Disease Progression; Diagnosis; IBD.

Computed tomography (CT) scan use has increased dramatically during the last 20 years as access has increased.[1] Imaging modalities, including CT, have important roles in the care of patients with Crohn's disease (CD), allowing clinicians to assess the extent of disease and the presence of penetrating complications. As CT scans have become easier to obtain, younger patients with CD may become exposed to large cumulative doses of radiation, with 11% exposed to more than 50 mSv, a level associated with increased risk of malignancy.[2,3] Approximately 30% of this radiation exposure occurs in the emergency department setting, and 75% of it is due to CT scans.[4] Young age at diagnosis, history of penetrating disease (fistulas and abscesses), history of multiple abdominal surgeries, and use of intravenous steroids and infliximab have been associated with higher cumulative radiation doses among patients with CD.[5]

We aimed to use logistic regression to develop algorithms that would predict the probability of complications or inflammation detected by CT scan in patients with CD presenting to the emergency department with abdominal symptoms. Ideal algorithms would have high negative predictive values, giving physicians the confidence to forgo CT scanning in patients with low risk for complications and inflammatory disease activity.

Methods

After receiving institutional review board approval, the electronic records database at the University of Michigan was queried for patients older than the age of 18 years with a diagnosis of CD by Institutional Classification

Abbreviations used in this paper: AEC, absolute eosinophil count; ALC, absolute lymphocyte count; AMC, absolute monocyte count; AST, aspartate aminotransferase; CD, Crohn's disease; CI, confidence interval; CRP, C-reactive protein; CT, computed tomography; ESR, erythrocyte sedimentation rate; MCH, mean corpuscular hemoglobin; MPV, mean platelet volume.

1542-3565/$36.00
http://dx.doi.org/10.1016/j.cgh.2014.02.036

of Diseases-9th Revision code 555.x who visited the emergency department between 2000 and 2011. Demographics, sex, age at emergency department visit, and laboratory results (obtained within 24 hours of the emergency department visit) were also electronically abstracted. The charts were then manually reviewed (S.M.G. and A.S.G.) to determine medications, chief complaint, whether a CT scan occurred within 24 hours of arrival, and, if so, the CT findings. Review of CT findings was performed by reviewers blinded to the laboratory results. Patients who did not actually have CD on review of records, did not undergo CT scan of the abdomen and pelvis with intravenous and oral contrast, or presented for a trauma complaint were excluded (Figure 1).

Obstruction was defined by the presence of a transition point requiring nasogastric tube decompression or surgical intervention. Patients were only classified as having appendicitis if they underwent surgery. Inflammation was defined by the presence of mucosal enhancement and/or increased vascular markings. Isolated wall thickening was not considered to be evidence of inflammation. Malignancies were only classified as a finding if they were new/unexpected findings. Urolithiasis was defined if there was no other explained cause of pain and/or findings of complications related to stones.

The 2 outcomes modeled were the presence of new or worsening complications (perforation, abscess, appendicitis, new malignancy, pyelonephritis, urolithiasis, diverticulitis, pancreatitis, or other serious findings) that would change clinical decision-making and active intestinal inflammation. Statistical analysis was performed by using SAS 9.3 software (SAS Inc, Cary, NC). Descriptive comparisons were performed by using either the Student t test for continuous values or the χ^2 square test for categorical values.

Logistic regression was used to model the outcomes. Candidate models were developed with the best subset selection method in SAS by using all predictors with P values $<.05$. All laboratory predictors were considered candidate variables in the process. Models were built with complete case analysis. Each visit was treated as a separate encounter so that repeat visits with the same patient were included in the model building process. The final model for each outcome was chosen for its superior negative predictive value and sensitivity. Because C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) are missing in approximately 50% of the visits, we developed separate models for the 2 outcomes both with and without these predictors. The best model for each outcome is presented in this text, whereas the other models can be found in the Supplementary Material.

In modeling outcomes when an external validation cohort is not available, a 10-fold cross-validation is preferred. In this approach, the data are divided into 10 subsets of size N/10, and models are trained on a pooled group of 9 subsets and tested on the 10th subset. This is repeated 9 additional times (by using a different subset for testing accuracy each time), and the internally validated mean accuracy is reported in the results section.[6] By performing 10-fold cross-validation, the effects of repeated measures in model building and test accuracy are minimized. The area under the receiver operating characteristic curve and associated confidence interval (CI) values were obtained by the cross-validation process.

Results

The initial data query resulted in 2875 emergency department visits for 1281 individuals (Supplementary Figure 1). After excluding patients without CD and those presenting with a traumatic complaint, there were 2472 emergency department visits for 1011 individuals. The average number of visits was 2.4, with a maximum of 56. Of these individuals, 613 had at least 1 CT scan (maximum, 31 in 1 individual). There were a total of

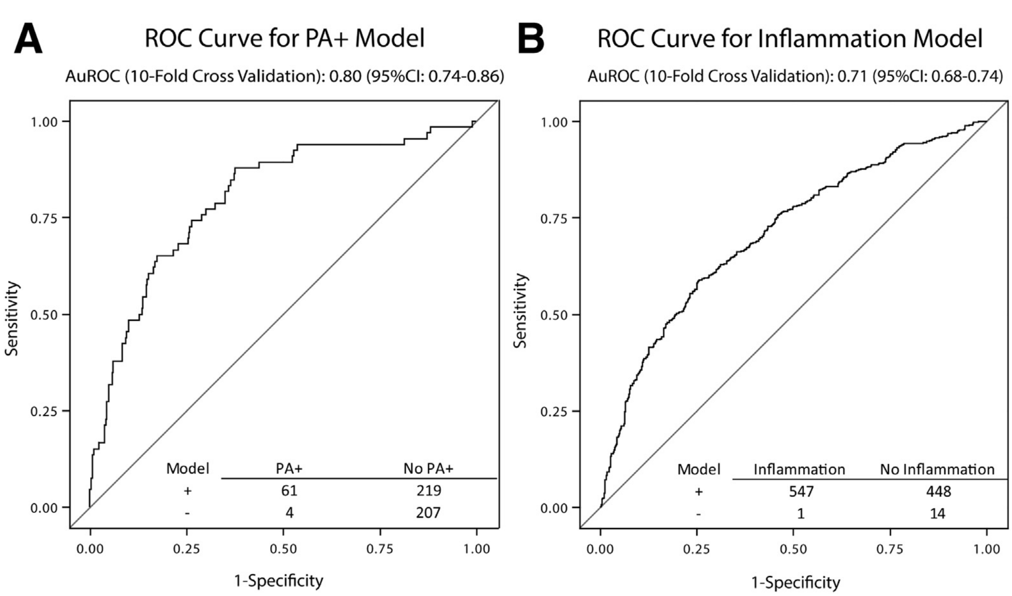

Figure 1. Receiver operating characteristic curve for individual models. This figure depicts sensitivity vs 1-specificity for the model depicting perforation, abscess, or other serious outcome (PA+) (*A*) and inflammation (*B*). The *bottom right corner* of each figure shows a classification table for each outcome compared with the prediction by the model at the defined cutoffs (6% for PA+ and 8% for inflammation). AuROC, area under the receiver operating characteristic curve.