A Novel Radiomics Nomogram for the Prediction of Secondary Loss of Response to Infliximab in Crohn’s Disease

Yueying Chen
Hanyang Li
Jing Feng
Shiteng Suo
Qi Feng
Jun Shen

1Division of Gastroenterology and Hepatology, Key Laboratory of Gastroenterology and Hepatology, Ministry of Health, Inflammatory Bowel Disease Research Center, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai Institute of Digestive Disease, Shanghai, 200127, People’s Republic of China; 2Department of Radiology, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, 200127, People’s Republic of China

Purpose: The prediction of the loss of response (LOR) to infliximab (IFX) is crucial for optimizing treatment strategies and shifting biologics. However, a secondary LOR is difficult to predict by endoscopy due to the intestinal stricture, perforation, and fistulas. This study aimed to develop and validate a radiomic nomogram for the prediction of secondary LOR to IFX in patients with Crohn’s disease (CD).

Patients and Methods: A total of 186 biologic-naive patients diagnosed with CD between September 2016 and June 2019 were enrolled. Secondary LOR was determined during week 54. Computed tomography enterography (CTE) texture analysis (TA) features were extracted from lesions and analyzed using LIFEx software. Feature selection was performed by least absolute shrinkage and selection operator (LASSO) and ten-fold cross validation. A nomogram was constructed using multivariable logistic regression, and the internal validation was approached by ten-fold cross validation.

Results: Predictors contained in the radiomics nomogram included three first-order and five second-order signatures. The prediction model presented significant discrimination (AUC, 0.880; 95% CI, 0.816–0.944) and high calibration (mean absolute error of = 0.028). Decision curve analysis (DCA) indicated that the nomogram provided clinical net benefit. Ten-fold cross validation assessed the stability of the nomogram with an AUC of 0.817 and an accuracy of 0.819.

Conclusion: This novel radiomics nomogram provides a predictive tool to assess secondary LOR to IFX in patients with Crohn’s disease. This tool will help physicians decide when to switch therapy.

Keywords: secondary loss of response, texture analysis, infliximab, Crohn’s disease, radiomics, prediction model

Introduction
Crohn’s disease (CD) is a recurrent inflammatory disease that affects the entire digestive tract, accompanied by extraintestinal manifestations. The prevalence has increased worldwide, and CD reduces the quality of life of patients. Anti-tumor necrosis factor (anti-TNF) agents have been considered the first-line therapy with improved mucosal healing and clinical remission rate. However, studies calculated that about 50–54% of patients with a loss of response (LOR) to infliximab (IFX) therapy, which was often driven by immunogenicity problems. The early assessment of LOR is crucial for physicians to select a new treatment for patients with CD. Recent studies have identified that the patient and disease characteristics, drug factors,
Radiologic images are composed of numerous voxels with individual gray-level values in accordance with the signal strength of the structure. Texture analysis (TA) is an emerging technology of radiomics, which includes the extraction, analysis, and quantification of lesion characteristics by analyzing the gray-scale values of regions of interest (ROIs) and is comprised of four steps: image acquisition, image segmentation, feature extraction, and informatics analysis. TA results include first-order and second-order features. Several studies have proposed that TA is a potentially useful technique to predict pathology characteristics, therapeutic effects, and prognoses of various tumors. Computerized tomography texture analysis (CTTA) is a TA based on CT and was first applied to evaluate the prognosis of patients with head and neck cancer. TA has also emerged as a predictor of patient outcomes in colorectal cancer (CRC) and a tool to distinguish malignant and benign breast tumors. Although TA has been widely used in tumor differentiation and prognosis prediction, few studies have focused on its use in inflammatory bowel disease. Furthermore, computed tomography enterography (CTE) is a convenient method to achieve a transluminal diagnosis and to monitor CD with high spatial resolution; however, there is little evidence to support CTE-based TA to predict the LOR to IFX in patients with CD. Our research aimed to identify the predictive value of CTE-based TA and construct a nomogram predicting LOR to IFX therapy in patients with CD, which may provide a useful tool to guide shifting the medical strategies for patients with CD.

Patients and Methods
Patient Characteristics
We reviewed the medical records of patients with CD treated by the Division of Gastroenterology and Hepatology at Renji Hospital between September 2016 and June 2019. CD was diagnosed based on the European Crohn’s and Colitis Organisation guidelines. We excluded 162 patients who were treated with IFX before admission, those who underwent maintenance therapy without a dose of 5–10 mg/kg, and those who did not undergo CTE prior to IFX therapy. A total of 186 biologic-naive patients were enrolled in our study. We initiated IFX therapy, which was maintained with 5–10 mg/kg of IFX every 8 weeks. Baseline characteristics and CTE features were collected retrospectively from a database of CT scans prior to treatment in our hospital. This study was approved by the Institutional Review Board of Shanghai Jiaotong University School of Medicine, Renji Hospital Ethics Committee (KY2017-129). Patients were not required to sign informed consent owing to the retrospective study, and their private information would be strictly protected in medical record system. This study was conducted in accordance with the Declaration of Helsinki to protect patient data confidentiality.

Outcomes and Definitions
A Crohn’s disease activity index (CDAI) reduction of more than 70 points or a decrease of Simplified Crohn’s Disease Activity Index (SCDAI) of more than 2 points from baseline was considered a clinical response to IFX therapy. Secondary LOR was defined by a multi-disciplinary team of experienced experts at our institution during week 54, including a relapse of CD after an initial response, such as CDAI or SCDAI rising again to the baseline level or even higher, and the requirement of an alternative biological agent, corticosteroid, or surgery.

CTE Imaging Technique
CTE examinations were performed using the conventional protocol. Patients fasted for 12 h prior to the scan and received 1,500 mL of polyethylene glycol solution (Wanghe Pharma, Shenzhen, China) orally as follows: a 500 mL bolus was administered at 45 min, 30 min, and 15 min prior to the CTE. CT scans were performed using a 64-detector computed tomography (CT) scanner (GE Medical System, Milwaukee, WI, USA), and patients underwent unenhanced and contrast-enhanced CT examinations in the supine position from the diaphragm to the perineum during a single breath-hold. Contrast-enhanced scanning was performed using a tube current of 228 mAs and voltage of 120 kV, and intravenous contrast (Lopamiro 370, Bracco Sine, Shanghai, China; 1.5mL/kg) at a speed of 3 mL/s. Scanning was performed with 40 mm collimation, 1.375:1 pitch, and a reconstruction thickness of 1.0 mm. CT images of the enteric phase were obtained 70 s after administration of the contrast agent and were retrieved for image feature extraction.
Imaging Analysis

ROIs were manually delineated by two clinical radiologists with over 13 years of experience in CT (SST and FQ) and two well-experienced gastroenterologists (CCY and SJ) within two weeks before IFX therapy, and we extracted one 2D ROIs for each scan. The consistency and repeatability of measurement between researchers were evaluated by intraclass correlation coefficient (ICC) with 0.905 (P<0.001). Before IFX therapy, the ROI was selected in the intestinal segment with the most active inflammation in each patient. Imaging features of active inflammation were defined as follows: inflamed bowel wall with ulceration on the intraluminal surface, bowel wall hyper-enhancement with moderate or severe thickening, bowel wall hyperenhancement with a halo sign (not due to intramural fat deposition), and inflammation penetrating the wall with or without an active fistula, a sinus tract, or an inflammatory mass in the adjacent mesentery.25

Imaging analysis was performed using Local Image Features Extraction (LIFEx) version 5.10 software (http://www.lifexsoft.org).26 Images were resampled to an isotropic 1 mm$^3$ voxel size. CT features were calculated manually using a 2D segmentation technique based on 400 discrete

---

![Flow Diagram](https://www.dovepress.com/)

**Figure 1** The flow diagram of study.
values with absolute scale bounds from −1,000 to 3,000 Hounsfield units. Twenty-four features were extracted, including conventional and histogram-based parameters and second and higher-order textural features. In particular, the following families of second and higher-order features were extracted: the grey-level co-occurrence matrix (GLCM), the grey-level run-length matrix (GLRLM), the neighborhood grey-level different matrix (NGLDM), and the grey-level zone length matrix (GLZLM) (*Supplementary Table 1*).

**Development and Validation of the Prediction Model**

Z-Score transformation was used to normalize the variables of different orders of magnitude to a common reference, and to improve the comparability of data. The equation was $x^* = \frac{(x-\mu)}{\sigma}$ (x: sample value; μ: population mean; σ: population standard deviation). Spearman correlation analysis was applied to reduce the dimension, for the highly correlated features (correlation coefficient > 0.9), the one showed a higher coefficient with other features was excluded. Least absolute shrinkage and selection operator (LASSO) and ten-fold cross-validation were applied in feature selection by determining the best λ, and the variables with nonzero coefficients were chosen. The prediction model and the nomogram were developed by a multivariate logistic regression analysis via the backward stepwise technique, and the best-fit model was determined with minimal Akaike information criterion (AIC), the workflow for developing the nomogram is presented in Figure 1. Ten-fold cross validation was performed to internally validate the nomogram. The discrimination of the nomogram was evaluated by receiver operating characteristic (ROC) analysis and presented as the area under the curve (AUC). The accuracy of the nomogram was assessed with a calibration curve, and decision curve analysis (DCA) was used to illustrate the clinical utility of the predictive nomogram.

**Statistical Analysis**

Continuous variables are presented as mean (standard deviation) or median (interquartile range), and categorical variables are presented as number or percentage. A chi-square test was used to compare categorical variables. The t-test or Mann–Whitney U-test was used to compare continuous variables. The reliability analysis was assessed by ICC with two-way random model. Spearman correlation coefficients more than 0.9 were considered very highly correlated. Statistical analyses were performed using SPSS version 25.0 and R version 3.6.3 statistical software, with a statistical significance of $P < 0.05$.

**Results**

**Baseline Characteristics and Feature Selection**

This study included 186 patients with CD who underwent IFX induction and maintenance therapy. Patient baseline characteristics, including sex, age, BMI, smoking history, surgical history, and simplified Crohn’s disease activity index are shown in Table 1. The incidence of secondary LOR was 23.12% (n = 43). The mean BMI of patients who experienced secondary LOR was significantly lower than that of patients who responded to IFX therapy ($p=0.07$). The 39 features identified by CTE imaging were normalized using the Z-Score due to obvious differences in

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Response ($n=143$)</th>
<th>Loss of Response ($n=43$)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (male/female, n)</td>
<td>60/83</td>
<td>20/23</td>
<td>0.60</td>
</tr>
<tr>
<td>Age (mean ±sd), year</td>
<td>29.57 (±9.17)</td>
<td>29.35 (±10.14)</td>
<td>0.89</td>
</tr>
<tr>
<td>Duration (median, IQR), month</td>
<td>36.00 (24.00–60.00)</td>
<td>36.00 (24.00–54.00)</td>
<td>0.961</td>
</tr>
<tr>
<td>BMI (mean ±sd)</td>
<td>20.89 (±4.46)</td>
<td>19.85 (±2.89)</td>
<td>0.07</td>
</tr>
<tr>
<td>Smoking (yes/no, n)</td>
<td>6/137</td>
<td>3/40</td>
<td>0.46</td>
</tr>
<tr>
<td>Surgery (yes/no, n)</td>
<td>49/94</td>
<td>13/30</td>
<td>0.63</td>
</tr>
<tr>
<td>Upper GI disease (yes/no, n)</td>
<td>50/143</td>
<td>14/43</td>
<td>0.605</td>
</tr>
<tr>
<td>CRP (median, IQR), mg/L</td>
<td>5.25 (1.09–18.40)</td>
<td>7.87 (1.87–19.40)</td>
<td>0.481</td>
</tr>
<tr>
<td>ESR (median, IQR), mm/h</td>
<td>13.00 (7.00–31.00)</td>
<td>15.00 (7.00–28.00)</td>
<td>0.963</td>
</tr>
<tr>
<td>SCDAI (median, IQR)</td>
<td>2.00 (1.00–3.00)</td>
<td>2.00 (1.00–4.00)</td>
<td>0.14</td>
</tr>
</tbody>
</table>

**Abbreviations:** BMI, body mass index; GI, gastrointestinal; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; SCDAI, Simplified Crohn’s Disease Activity Index; sd, standard deviation; IQR, interquartile range.
values. The threshold of Spearman correlation coefficient was set to 0.9, after filtering, there were 24 features remained (Supplementary Table 2). LASSO and ten-fold cross validation identified the optimum $\lambda = 0.042$. The 24 candidate features were reduced to the following 8 features with nonzero coefficients in 186 patients (93:4 ratio): CONVENTIONAL_HUstd (reflects the standard deviation value in the ROIs), HISTO_Energy (reflects the uniformity of the distribution), HISTO_Skewness (reflects the asymmetry of the grey-level distribution in the histogram), GLCM_Energy (reflects the uniformity of grey-level voxel pairs), GLZLM_ZP (measures the homogeneity of the homogeneous zones), NGLDM_Contrast (reflects the intensity difference between neighbouring regions.), GLZLM_LZLGE (reflects the distribution of the long homogeneous zones with low grey-levels), and GLCM_Contrast (reflects the local variations in the GLCM) (Figure 2).

**Figure 2** Texture features selection using the LASSO and ten-fold cross-validation. (A) Optimal parameter ($\lambda$) selection in LASSO model used ten-fold cross-validation via minimum criteria. Dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 SE of the minimum criteria (the 1-SE criteria). A $\lambda$ of 0.042 with log ($\lambda$) = -3.170 was chosen. (B) LASSO coefficient profiles of the 24 radiomics features. A coefficient profile plot was generated versus the selected log($\lambda$) value using ten-fold cross-validation, where optimal $\lambda$ resulted in 8 features with nonzero coefficients.
Nomogram Development and Validation

The model based on the 8 features with nonzero coefficients was constructed using backward stepwise multivariate logistic regression. The model, including these 8 features, had a good fit, as determined by minimum AIC (Table 2) and was presented as a nomogram (Figure 3). The predictive value of the model was calculated using the following formula:

\[
\text{logit} P = -1.832 - (0.631 \times \text{CONVENTIONAL_HUstd}) + (1.454 \times \text{HISTO_Energy}) + (1.041 \times \text{HISTO_Skewness}) - (1.490 \times \text{GLCM_Energy}) - (1.477 \times \text{GLZLM_ZP}) + (0.306 \times \text{NGLDM_Contrast}) + (0.311 \times \text{GLZLM_LZLGE}) + (0.692 \times \text{GLCM_Contrast})
\]

The performance of the nomogram is shown in Figure 4. The nomogram exhibited sense discrimination (AUC = 0.880), and the calibration curve revealed excellent accuracy (mean absolute error = 0.028), the Hosmer–Lemeshow test (P = 0.742) revealed that the predicted probability was close to the actual probability. The DCA showed that the predictive nomogram could be used to predict the second LOR rate and provide clinical net benefit if the risk threshold between 5% and 95%. Furthermore, ten-fold cross validation was performed to prove the robustness of this model, with a mean AUC of 0.817, and accuracy of 0.819 (Supplementary Table 3). A pair of comparisons for the loss and gain of response of the nomogram for risk assessment of LOR is illustrated in Figure 5. The following mean values were found in patients with LOR: CONVENTIONAL_HUstd = 24, HISTO_Energy = 18, HISTO_Skewness = 21, GLCM_Energy = 42, GLZLM_ZP = 57, NGLDM_Contrast = 0.7, GLZLM_LZLGE = 7.5, GLCM_Contrast = 16 (Figure 5A), which sums to 186.2, suggesting that the probability of LOR to IFX therapy was above 90%. In patients who responded well to IFX treatment, the score of the nomogram showed that the possibility was less than 50% (Figure 5B).

Discussion

This study developed and validated a radiomics signature-based nomogram to predict secondary LOR to IFX in

---

Table 2 Multivariate Regression Analyses of the Prediction Model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>OR (95% CI)</th>
<th>S.E.</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONVENTIONAL_HUstd</td>
<td>-0.631</td>
<td>0.532 (0.231–0.939)</td>
<td>0.425</td>
<td>0.138</td>
</tr>
<tr>
<td>HISTO_Energy</td>
<td>1.454</td>
<td>4.280 (2.160–8.483)</td>
<td>0.349</td>
<td>0.000</td>
</tr>
<tr>
<td>HISTO_Skewness</td>
<td>1.041</td>
<td>2.833 (1.443–5.563)</td>
<td>0.344</td>
<td>0.002</td>
</tr>
<tr>
<td>GLCM_Energy</td>
<td>-1.490</td>
<td>0.225 (0.091–0.560)</td>
<td>0.465</td>
<td>0.001</td>
</tr>
<tr>
<td>GLZLM_ZP</td>
<td>-1.477</td>
<td>0.228 (0.095–0.546)</td>
<td>0.445</td>
<td>0.001</td>
</tr>
<tr>
<td>NGLDM_Contrast</td>
<td>0.306</td>
<td>1.358 (0.657–2.807)</td>
<td>0.370</td>
<td>0.409</td>
</tr>
<tr>
<td>GLZLM_LZLGE</td>
<td>0.311</td>
<td>1.364 (0.598–3.115)</td>
<td>0.421</td>
<td>0.461</td>
</tr>
<tr>
<td>GLCM_Contrast</td>
<td>0.692</td>
<td>1.998 (0.855–4.670)</td>
<td>0.283</td>
<td>0.110</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; OR, odds ratio; S.E., standard error.
patients with CD. The nomogram includes eight items of the radiomics signature extracted by texture analysis. The performance of this prediction nomogram showed adequate discrimination and significant calibration power, as well as a net clinical benefit.

LASSO and ten-fold cross validation were used to reduce 24 features extracted by LIFEx software to eight potential predictors based on their strength of association with outcome. This method combined the individual features into a radiomics signature. We used backward-stepwise multivariate logistic regression analysis to determine the predictive ability of each variable to prevent omitting a potential predictor. Due to the limited sample size, our nomogram was validated by ten-fold cross validation, and the results showed good AUC, as well as high accuracy. The results of the calibration curve and DCA indicate that our nomogram is effective and clinically relevant. Previous models built to predict the efficacy of IFX in patients with CD were mainly based on clinical characteristics and laboratory indicators. Billiet et al developed a matrix-based prediction model with an AUC of 0.80 using BMI, surgical history, and age IFX initiation.\(^8\) The predictive powers of interleukin-9 (AUC = 0.746) and fecal calprotectin (AUC = 0.87) in evaluating the efficacy of IFX therapy in patients with CD have also been confirmed by clinical studies.\(^{31,32}\) Additionally, magnetic resonance imaging (MRI) has been reported to be conducive for optimizing therapy for patients with CD during IFX treatment, and the MRI prediction represented an AUC of 0.7359 in assessing endoscopic inflammation at 1 year.\(^{33}\) In contrast, our noninvasive nomogram had an AUC of 0.893.

Texture analysis has been frequently applied to diagnose and evaluate the prognoses of tumors based on different pixel gray values and spatial relationships.\(^{34}\) One previous study reported that radiomics signatures from second-order differentiations could be used to differentiate neoplastic and non-neoplastic lesions in the colon with an AUC of 0.85.\(^{35}\) A nomogram integrating CT features selected by TA and clinical factors was reported to have a C-index of 0.778, and good calibration, which favored the prediction of metastasis in individual lymph nodes in patients with CRC.\(^{36}\) Models and nomograms using TA have also been established to predict the efficacy of various therapies. Kim et al reported a radiomics model consisting of GLCM and histogram-based features, combined with clinical factors that effectively evaluated the prognosis of transarterial chemoembolization in patients with hepatocellular carcinoma (HR = 19.88, \(P < 0.0001\)).\(^{37}\) Similarly, a model generated with gray-level correlation and variance signatures analyzed by TA was reported to effectively predict the recurrence of non-small cell lung cancer before stereotactic radiation therapy.\(^{38}\)
Although TA has been widely used in solid cancers, few studies have focused on the inflammatory diseases and cavernous organs due to diffuse lesions. Wang et al reported that CTTA-based classifications contributed to a model used to differentiate normal thickened colon wall, inflammatory bowel disease (IBD), and CRC with an ideal AUC.

As CTE is an effective technique for assessing disease activity in IBD, we speculated that TA incorporated with CTE may be a more precise, promising tool to determine the prognosis of patients with CD. Furthermore, various anti-TNF agents have been developed in recent years, and prediction models and nomograms are necessary to assist in selecting the best treatment strategy. Therefore, we used noninvasive radiomics features as a convenient biomarker for the prediction of LOR to IFX therapy. Although endoscopy is an accurate method for evaluating the prognosis of CD, it cannot be used in patients with intestinal obstructions or internal fistulas. In addition, while endoscopic scanning can identify mucosal lesions, CD mainly affects multiple layers of the intestine. CTE is a useful tool to identify transluminal lesions. To the best of our knowledge, our study is the first to apply TA in the prognostic prediction of IFX in IBD, providing clinical evidence supporting a change in therapy, and reducing the inefficient use of medical resources.

However, our study has several limitations. First, an external validation by a multicenter study is needed to improve the reliability of our radiomics nomogram in clinical applications. Second, our radiomics nomogram does not include clinical or laboratory features. Third, we used the 2D rather than 3D segmentation technique to select the ROI as it is less challenging and takes less time. Further research focusing on the differences between the performances of 2D and 3D TA in patients with CD is required. In addition, serum biomarkers and nutritional status also reflect the response to IFX, and may be confounding factors that predict the effectiveness of IFX.

Our nomogram was constructed using features with predictive power prior to therapy to avoid the omission of important predictors. Future studies with larger patient populations are required to verify the relationship between the efficacy of IFX therapy and radiomics signatures.
Conclusion
We developed an effective novel radiomics nomogram consisting of eight radiographic features to predict the LOR to IFX treatment in patients with CD. TA based on CTE is a significant noninvasive biomarker that can accurately identify patients with CD who have LOR to IFX therapy and can be used to determine the optimal medical therapy on a patient-by-patient basis.

Data Sharing Statement
The data underlying this article will be shared on reasonable request to the corresponding author, Jun Shen.

Acknowledgments
Supported by grants from National Natural Science Foundation of China (No. 81770545, and 81701746) and MDT Project of Clinical Research Innovation Foundation, Renji Hospital, School of Medicine, Shanghai Jiaotong University (PYI-17-003).

Disclosure
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References


