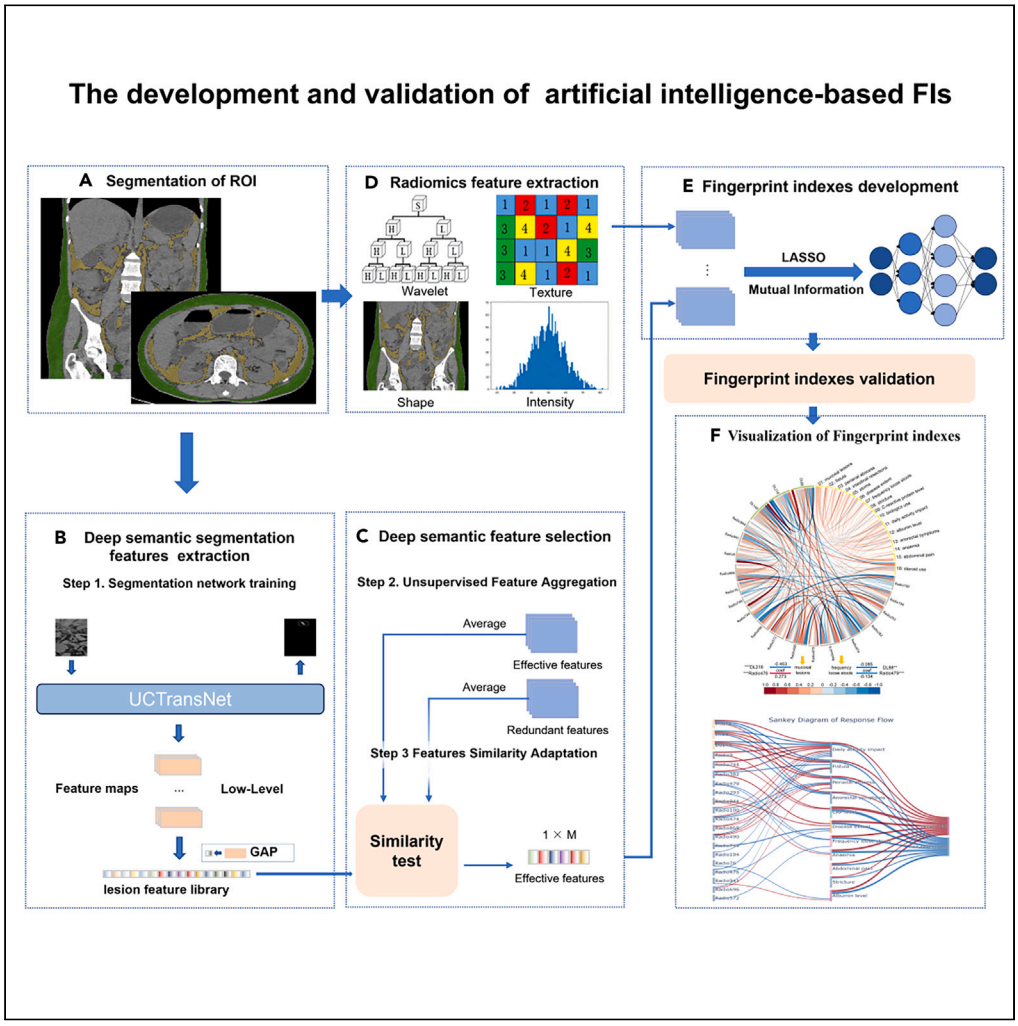
iScience

CellPress
OPEN ACCESS

Article

AI-based fingerprint index of visceral adipose tissue for the prediction of bowel damage in patients with Crohn's disease



Xuehua Li, Cicong Hu, Haipeng Wang, ..., Ren Mao, Bingsheng Huang, Shi-Ting Feng

lxueh@mail.sysu.edu.cn (X.L.)
shili101028@126.com (L.S.)
maor5@mail.sysu.edu.cn (R.M.)
huangb@szu.edu.cn (B.H.)
fengsht@mail.sysu.edu.cn (S.-T.F.)

Highlights

Visceral adipose tissue-fingerprint index (VAT-FI) can assess bowel damage in CD patients

VAT-FI outperforms conventional fat metrics in evaluating bowel damage

VAT-FI is an alternative to Lémann index, minimizing the need for extensive clinical data

Li et al., iScience 27, 111022
October 18, 2024
https://doi.org/10.1016/j.isci.2024.111022

iScience

CellPress
OPEN ACCESS

Article

AI-based fingerprint index of visceral adipose tissue for the prediction of bowel damage in patients with Crohn's disease

Xuehua Li,[1,12,*] Cicong Hu,[1,2,12] Haipeng Wang,[3,12] Yuqin Lin,[4,12] Jiaqiang Li,[5,12] Enming Cui,[6] Xiaozhao Zhuang,[7] Jianpeng Li,[8] Jiahang Lu,[9] Ruonan Zhang,[1] Yangdi Wang,[1] Zhenpeng Peng,[1] Canhui Sun,[1] Ziping Li,[1] Minhu Chen,[10] Li Shi,[11,*] Ren Mao,[10,*] Bingsheng Huang,[3,*] and Shi-Ting Feng[1,13,*]

SUMMARY

The fingerprint features of visceral adipose tissue (VAT) are intricately linked to bowel damage (BD) in patients with Crohn's disease (CD). We aimed to develop a VAT fingerprint index (VAT-FI) using radiomics and deep learning features extracted from computed tomography (CT) images of 1,135 CD patients across six hospitals (training cohort, *n* = 600; testing cohort, *n* = 535) for predicting BD, and to compare it with a subcutaneous adipose tissue (SAT)-FI. VAT-FI exhibited greater predictive accuracy than SAT-FI in both training (area under the receiver operating characteristic curve [AUC] = 0.822 vs. AUC = 0.745, *p* = 0.019) and testing (AUC = 0.791 vs. AUC = 0.687, *p* = 0.019) cohorts. Multivariate logistic regression analysis highlighted VAT-FI as the sole significant predictor (training cohort: hazard ratio [HR] = 1.684, *p* = 0.012; testing cohort: HR = 2.649, *p* < 0.001). Through Shapley additive explanation (SHAP) analysis, we further quantitatively elucidated the predictive relationship between VAT-FI and BD, highlighting potential connections such as Radio479 (wavelet-HLH-first-order standard deviation)-Frequency loose stools-BD severity. VAT-FI offers an accurate means for characterizing BD, minimizing the need for extensive clinical data.

INTRODUCTION

Crohn's disease (CD) is a progressively destructive inflammatory bowel disease that causes transmural inflammation. Over time, this inflammation can result in the accumulated bowel damage (BD), manifesting as various intestinal complications such as strictures, fistulas, and abscesses.[1] The incidence of BD in CD patients can reach 50% during their disease course.[2] Severe BD substantially contributes to long-term disability, impacting individuals' lives beyond only intestinal symptoms.[3] The active monitoring of BD with an attempt to improve the long-term prognosis of patients is increasingly becoming a pivotal objective in CD management.

The Lémann index[1] and the disease severity index (DSI)[4] have been proposed as clinical BD measuring tools. Recently, DSI has been shown to be a more sensitive tool for measuring BD compared to the Lémann index.[5] However, the utilization of DSI heavily relies on extensive clinical data collected throughout the prolonged course of a patient's illness. Obtaining accurate data becomes challenging when patients seek treatment at multiple health care facilities. The assessment of BD urgently requires a more efficient method.

Emerging evidence suggests the involvement of visceral adipose tissue (VAT) in CD pathogenesis. According to our previous studies[6–8] and those of other centers,[9–11] VAT hypertrophy is significantly associated with increased intestinal fibrosis/inflammation severity, the progression of intestinal stenosis and penetrating lesions, and the need for surgery. The VAT and intestines have recently been considered as

[1]Department of Radiology, The First Affiliated Hospital, Sun Yat-Sen University, 58 Zhongshan II Road, Guangzhou 510080, People's Republic of China
[2]Department of Radiology, The First Affiliated Hospital of Wenzhou Medical University, Nanbaixiang, Ouhai District, Wenzhou 325000, People's Republic of China
[3]Medical AI Lab, School of Biomedical Engineering, Medical School, Shenzhen University, Block A2, Lihu Campus of Shenzhen University, 1066 Xueyuan Avenue, Shenzhen 518000, People's Republic of China
[4]Department of Radiology, The First Affiliated Hospital of Shantou University Medical College, 57 Changping Road, Shantou 515000, People's Republic of China
[5]Department of Radiology, The First People's Hospital of Foshan, No.81, Lingnan Dadao north, Foshan 528000, People's Republic of China
[6]Department of Radiology, Jiangmen Central Hospital, Guangdong Medical University, 23 Beijie Haibang Street, Jiangmen 529030, People's Republic of China
[7]Department of Radiology, Hainan General Hospital (Hainan Affiliated Hospital of Hainan Medical University), No.19 Xiuhua Road, Xiuying District, Haikou 570311, People's Republic of China
[8]Department of Radiology, Affiliated Dongguan People's Hospital, Southern Medical University, No. 78 Wandao Road, Dongguan 523000, People's Republic of China
[9]Medical Imaging Department, The First Affiliated Hospital, Kunming Medical University, Xi Chang Road 295th, Kunming 650000, People's Republic of China
[10]Department of Gastroenterology, The First Affiliated Hospital, Sun Yat-Sen University, 58 Zhongshan II Road, Guangzhou 510080, People's Republic of China
[11]Department of Radiology, The Third Affiliated Hospital of Guangzhou Medical University, 63 Duobao Road, Guangzhou 510150, People's Republic of China
[12]These authors contributed equally
[13]Lead contact
*Correspondence: lxueh@mail.sysu.edu.cn (X.L.), shili101028@126.com (L.S.), maor5@mail.sysu.edu.cn (R.M.), huangb@szu.edu.cn (B.H.), fengsht@mail.sysu.edu.cn (S.-T.F.)
https://doi.org/10.1016/j.isci.2024.111022

a whole,[12] suggesting that damage occurring in the intestines can be "synchronously mapped" to the surrounding VAT. Consequently, VAT has potential as a prospective predictor of BD.

Computed tomography (CT) imaging allows for the precise capture of both macroscopic and microscopic features of VAT. Our previous study[6] confirmed that AI technology, such as radiomics, can thoroughly decipher intricate VAT fingerprint features on CT images, surpassing conventional fat metrics (e.g., VAT/subcutaneous adipose tissue [SAT] area ratio). Therefore, AI-based VAT fingerprint signatures could characterize BD.

Herein, we conducted a retrospective study involving six centers to develop and validate a CT-based VAT fingerprint index (FI) for accurately and conveniently quantifying BD severity, with DSI serving as a reference standard for BD assessment. This FI, which incorporates radiomics and deep learning features into a machine learning algorithm, served as a comprehensive representation of BD-associated changes. We also visualized this FI to highlight its ability to accurately quantify BD.

RESULTS

Patient characteristics

Ultimately, 1,135 patients were included, with 600 in training cohort and 535 in testing cohorts. The patients' characteristics are shown in Table S3. The training cohort comprised 433 patients with severe BD and 167 with nonsevere BD, while the total test cohort included 302 patients with severe BD and 233 with nonsevere BD. The intraclass correlation coefficients (ICCs) for DSI assessment were $\geq$0.80, indicating good inter- and intra-observer consistency. No differences in traditional adipose tissue imaging indices were found between patients with severe and nonsevere BD in training and testing cohorts (p = 0.06–0.89; Table S4) except for body mass index ($p < 0.05$).

Development, validation, visualization, and interpretation of VAT-FI

We selected 1,128 radiomics features with ICCs $\geq$0.80 and 512 deep learning features as candidate features for constructing VAT-FI. Through least absolute shrinkage and selection operator (LASSO) and mutual information methods, 25 radiomics features and 23 deep learning features with high correlation and significant weight coefficients were eventually selected to form VAT-FI for evaluating BD. The formula was as follows:

$$Prediction\ probability\ = \frac{1}{1+e^{F(x)}}, F(x) = -3.120 \times f(\boldsymbol{\beta}) + 1.352$$

where $F(x)$ is decision function of logistic regression model and $\boldsymbol{\beta}$ represents the weight vector of selected features. The VAT-FI utilized a threshold of 0.484 to distinguish BD severity, categorizing patients into two subgroups: severe ($\geq$0.484) and nonsevere (<0.484) BD.

The VAT-FI accurately predicted severe BD in both the training cohort (area under the receiver operating characteristic curve [AUC] = 0.822, $p < 0.001$) and the total test cohort (AUC = 0.791, $p < 0.001$; Figure 1A), but the differences were not significant (DeLong test, p = 0.439). Its AUC values in five testing cohorts were 0.778 ($p < 0.001$), 0.801 (p = 0.006), 0.758 (p = 0.015), 0.778 (p = 0.003), and 0.800 (p = 0.021) (Figure 1B; Table 1). Hosmer-Lemeshow test results for both cohorts were x^2 = 8.437 (p = 0.392) and x^2 = 11.610 (p = 0.169), indicating that VAT-FI was a good fit (Figure 1C). Moreover, the diagnostic accuracy of VAT-FI did not significantly differ across the CT scanner conditions (DeLong test, all $p > 0.050$; Table S5).

The Shapley additive explanation (SHAP) values and distributions of all 48 features used to develop VAT-FI are shown in Table S6, the top twenty fingerprint features were visualized in SHAP plots in descending order (Figures 2A and 2B). Specifically, the top three features were Radio293, Radio479, and Radio944. An increase in Radio944 or a decrease in Radio479 or Radio293 significantly increased the probability of severe BD. Also, SHAP plots reveal that deep learning features (DL149 and DL88) can contribute to BD prediction. The distributions of these features in the groups with and without severe BD are shown in Figures 2C and 2D. Additionally, Figures 2E and 2F show the individual SHAP force plots for severe and nonsevere BD patients to visualize how features in VAT-FI influenced prediction.

Subsequently, a chord diagram was constructed to display the intricate correlations between the top 20 important FI features and the 16 DSI items (Figure 3A; Table S7), highlighting the significant feature pairs like DL218-Mucosal lesions, Radio476-Mucosal lesions, DL88-Frequency loose stools, and Radio479-Frequency loose stools. Following a rigorous statistical process to identify features with significant intrinsic relationships, a Sankey diagram was constructed to visually map the connections among the "VAT-FI features," "DSI items," and "BD severity" (Figure 3B). VAT-FI features, such as DL218, DL88, Radio476, and Radio479, exhibited frequent correlations with DSI items. These various pathways, such as Radio479-Frequency loose stools-BD severity, showed that the VAT-FI features correlated with DSI items and then predicted BD severity.

Development, validation, visualization, and interpretation of the SAT-FI

To better investigate the role of VAT, we included SAT as a comparison. Following the same feature extraction and selection strategy as VAT-FI, 25 radiomics features and 23 deep learning features were combined for building the SAT-FI. The SAT-FI outcome was generated using the same strategy as VAT-FI. The formula for evaluating the BD severity is as follows:

$$Prediction\ probability\ = \frac{1}{1+e^{F(x)}}, F(x) = 0.916 \times f(\boldsymbol{\beta}) + 1.458$$

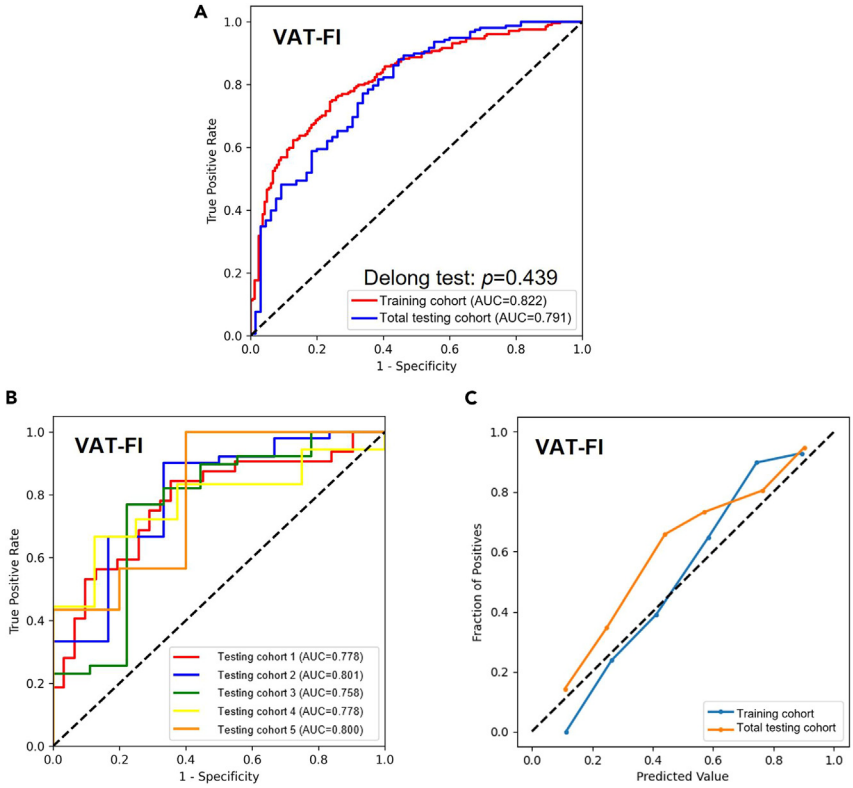


Figure 1. Predictive performance of VAT-FI in training and testing cohorts

(A–C) Plots show the receiver operating characteristic (ROC) curve of VAT-FI in the training cohort (A) and testing cohorts (B). Plot (C) shows the calibration curves for the VAT-FI in the training cohort and the total testing cohort. VAT, visceral adipose tissue; AUC, area under the receiver operating characteristic curve; FI, fingerprint index.

The optimal threshold for SAT-FI that distinguishes patients into the severe and nonsevere BD subgroups is 0.553. In the training cohort, SAT-FI achieved an AUC of 0.745 ($p < 0.001$) for predicting BD. Its predictive performance decreased in the total test cohort (AUC = 0.687; $p < 0.001$) and test cohorts (AUC = 0.698, 0.765, 0.516, 0.639, 0.646, respectively; $p = 0.002$–0.895; Table 1). No significant differences were observed between training and total test cohorts (DeLong test $p = 0.248$; Figure S3A) or among different test cohorts (Figure S3B; Table 1). The Hosmer-Lemeshow test yielded x^2 values of 3.994 ($p = 0.858$) in the training cohort and 6.792 ($p = 0.559$) in the total test cohort (Figure S3C).

The top 20 fingerprint features of SAT-FI were visualized in SHAP plots (Figures S4A and S4B). The distributions of important features are depicted using raincloud plots in Figures S4C and S4D. SHAP force plots for severe and nonsevere BD patients illustrate specific impacts on predictions (Figures S4E and S4F). Details of 48 features are available in Table S8.

Comparison of performance between VAT-FI and SAT-FI

The performance of VAT-FI was significantly better than that of SAT-FI in the training cohort (AUC = 0.822 vs. AUC = 0.745, $p = 0.019$; Figure 4A) and total test cohort (AUC = 0.791 vs. AUC = 0.687, $p = 0.019$; Figure 4B). Decision curve analysis indicated a greater net benefit of VAT-FI than of SAT-FI (Figures 4C and 4D). The threshold curve graph demonstrated that the VAT-FI predictions were more accurate than the SAT-FI predictions in reflecting the BD severity (Figure S5).

Identifying independent predictors of BD

Among the candidate factors, VAT-FI and erythrocyte sedimentation rate (ESR) were identified as significant factors ($p < 0.050$) through univariate logistic regression analysis and were subsequently included in the multivariate logistic regression analysis (Table 2). Ultimately, VAT-FI was identified as the sole independent predictor of BD (training cohort: hazard ratio [HR] = 1.496, p=0.039; total test cohort: HR = 2.434, $p = 0.003$). The difference in the SAT-FI score (training cohort: HR = 0.306, $p = 0.708$; total test cohort: HR = 2.265, $p = 0.091$) did not reach statistical significance.

Table 1. Predictive performance of FIs between the training cohort and total testing cohorts in predicting bowel damage

Model	Dataset	AUC (95% CI)	Accuracy	Sensitivity	Specificity	*p* values
VAT-FI	Training cohort	0.82 (0.79, 0.85)	0.75	0.78	0.72	<0.01
	Total testing cohort	0.79 (0.76, 0.83)	0.75	0.80	0.62	<0.01
	Testing cohort 1	0.78 (0.74, 0.82)	0.70	0.50	0.90	<0.01
	Testing cohort 2	0.80 (0.75, 0.85)	0.88	0.96	0.17	<0.01
	Testing cohort 3	0.76 (0.71, 0.81)	0.38	0.26	0.89	0.02
	Testing cohort 4	0.78 (0.74, 0.82)	0.54	0.15	0.92	<0.01
	Testing cohort 5	0.80 (0.76, 0.84)	0.68	0.71	0.50	0.02
SAT-FI	Training cohort	0.75 (0.71, 0.78)	0.71	0.60	0.81	<0.01
	Total testing cohort	0.69 (0.62, 0.73)	0.44	0.29	0.80	<0.01
	Testing cohort 1	0.70 (0.65, 0.74)	0.68	0.44	0.94	<0.01
	Testing cohort 2	0.77 (0.72, 0.81)	0.28	0.20	1.00	<0.01
	Testing cohort 3	0.52 (0.45, 0.59)	0.33	0.26	0.67	0.89
	Testing cohort 4	0.64 (0.60, 0.68)	0.50	0.23	0.77	0.22
	Testing cohort 5	0.65 (0.60, 0.69)	0.71	0.71	0.75	0.30

p values represent the significance level of AUC compared to random chance (AUC = 0.50).
AUC, area under the receiver operating characteristic curve; CI, confidence interval; VAT, visceral adipose tissue; SAT, subcutaneous adipose tissue; FI, fingerprint index.

DISCUSSION

In this study, VAT-FI accurately distinguished severe BD patients from nonsevere BD patients, demonstrating its robustness and generalizability across six medical centers. Using SHAP plots and Sankey diagram, we effectively visualized VAT-FI features and their importance and elucidated how VAT-FI could be used to evaluate BD. Moreover, VAT-FI outperformed SAT-FI and conventional fat metrics in assessing BD. Thus, our study highlighted the potential of VAT-FI as a precise surrogate biomarker for DSI in CD patients, enabling the assessment of long-term cumulative BD.

The effectiveness of VAT-FI in accurately identifying individuals with severe BD stems from the intimate connections between BD and VAT.[13] The fingerprint features (i.e., radiomics and deep learning features) of VAT facilitated the characterization of these connections of pathological alterations, thereby accurately reflecting the cumulative degree of disease burden in CD intestines. In this study, we employed SHAP plots to visualize the importance value and contribution direction of each factor in the VAT-FI and further utilized SHAP force plots to elucidate the contributions of these factors to the prediction of BD. More innovatively, despite the inherent complexity of comprehending these VAT-FI features, we employed Sankey diagrams as a visual aid to elucidate the intricate interplay between these fingerprint features and their associations with intestinal lesions or clinical factors, ultimately facilitating a comprehensive understanding of the workflow of this VAT-FI for predicting BD. Among the various pathways identified, several VAT-FI features, such as DL149 and DL88, effectively captured the distinct characteristics of fistulae and perianal abscess, thereby substantially contributing to advancing the diagnosis of severe BD. Radio749 and Radio476 may be correlated with nonsevere BD through disease extent. Therefore, VAT-FI has the potential to reveal important variations in VAT information among patients with different DSI levels and accurately evaluate their BD.

Moreover, the predictive efficacy of VAT-FI was superior to that of SAT-FI and conventional fat measures. The cellular morphology and composition and gene expression profiles in VAT exhibit variations in CD patients compared to healthy individuals,[14] whereas SAT shows similar morphological and molecular characteristics between CD patients and healthy controls.[15] Compared to SAT, VAT in CD patients exhibited pathological changes that were more consistent with intestinal lesions, resulting in a greater predictive efficacy of VAT-FI over SAT-FI. This result was consistent with our prior study, which demonstrated that VAT also outperformed SAT in predicting intestinal disease progression in CD patients.[6] Additionally, our study revealed no significant correlations between conventional fat metrics and DSI. The correlation between these conventional fat metrics and CD varies across different studies.[7,10,16,17] The inconsistency brought to light the constraints of traditional fat metrics, which predominantly offer information based on quantity. The inclusion of deep learning and radiomics features in our VAT-FI greatly enhanced the comprehensive extraction of multidimensional information from VAT, allowing for a more thorough assessment than conventional fat indicators.

Our study has several distinct advantages. First, an FI using radiomics and deep learning features was developed and validated to assess the correlation between VAT and BD. Furthermore, the use of an AI-based FI allowed us to extract a broader range of features for enhancing VAT-FI performance. Additionally, our study addressed the necessity for a comprehensive understanding of VAT-FI by effectively visualizing fingerprint features through SHAP plots and Sankey diagrams. Finally, this multicenter study included a substantial number of patients from various cities in China. The uniformity in the proportion of patients with severe BD across centers indicated the representative nature of the study population. Each center demonstrated consistent model performance, further confirming the generalizability and applicability of VAT-FI.

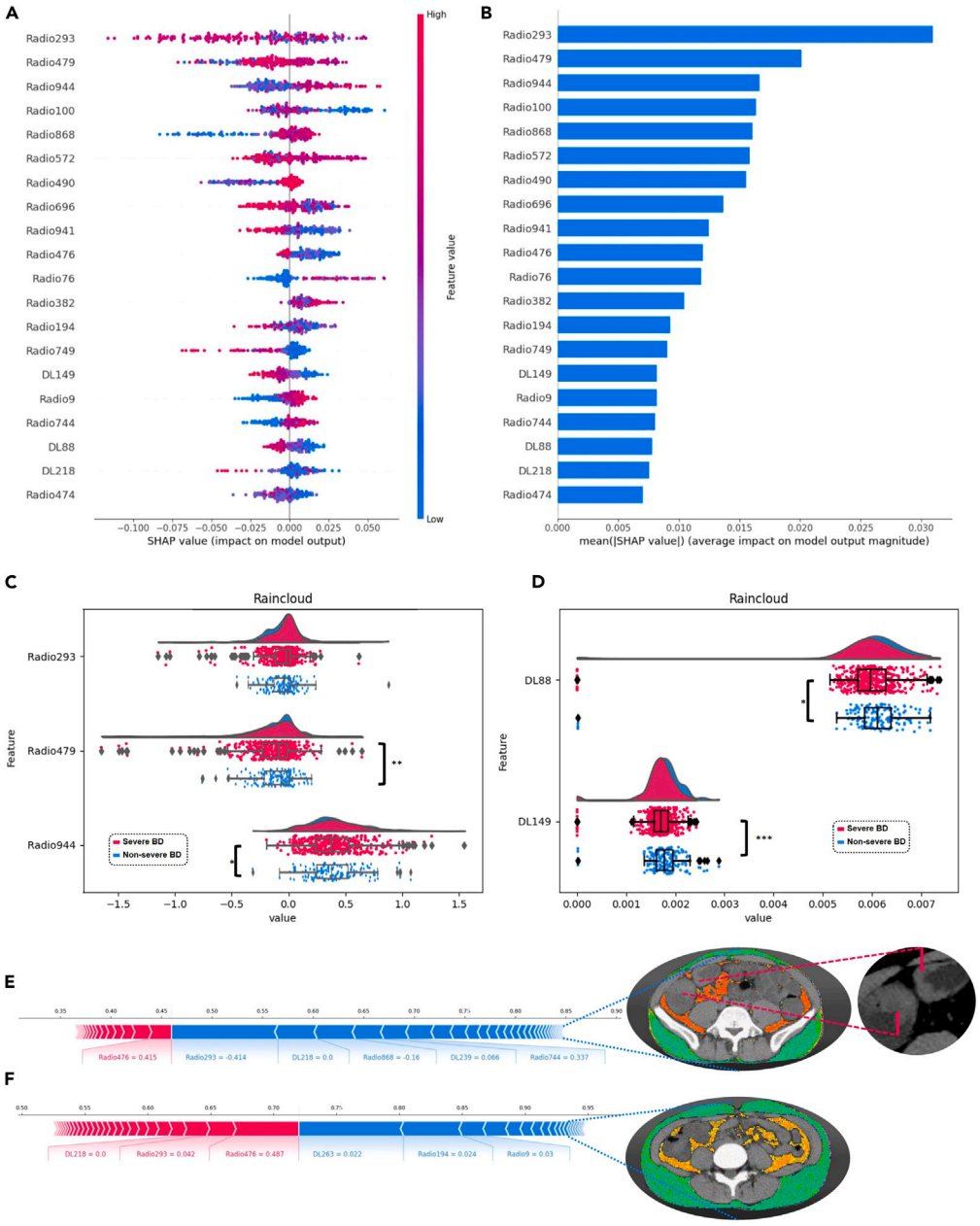


Figure 2. The SHAP plots and raincloud plots of VAT-FI

(A) The SHAP plot illustrates the most influential features on prediction and the distribution of each feature's impact on model, which includes a series of plots where each dot corresponds to an individual. The colors represent feature values for numeric features: red for larger values and blue for smaller. The line's thickness, composed of individual dots, reflected the count of examples for specific values. A negative SHAP value (extending to the left) indicates reduced risk of severe BD, while a positive value (extending to the right) indicates increased risk of severe BD.

(B) The weights of variables importance. The graph displays mean absolute value of the SHAP values for features in VAT-FI.

(C and D) The raincloud plots show the distribution of radiomics features (C) and deep learning features (D) between groups with (red color) and without (blue color) severe BD.

(E and F) SHAP individual force plots for two representative CD patients with severe (E; a 23-year-old female patient) or nonsevere BD (F; a 21-year-old male patient). The results show the impact of risk factor for prediction. Red features (left) indicate factors raising BD risk, and blue ones denote risk-reducing factors. Arrow lengths (red and blue) reflect SHAP values for each specific prediction feature. The heatmap images highlight their discrepancy in VAT between two patients with and without severe BD, whereas SAT differences were minimal. Error bars represent the 95% confidence intervals. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Radio293: wavelet-LHH-first-order standard deviation; Radio479: wavelet-HLH-first-order standard deviation; Radio944: log-sigma-3-0-mm-3D-first-order standard deviation. VAT, visceral adipose tissue; SAT, subcutaneous adipose tissue; FI, fingerprint index; Radio, radiomics feature; DL, deep learning feature; BD, bowel damage; SHAP, Shapley additive explanation.

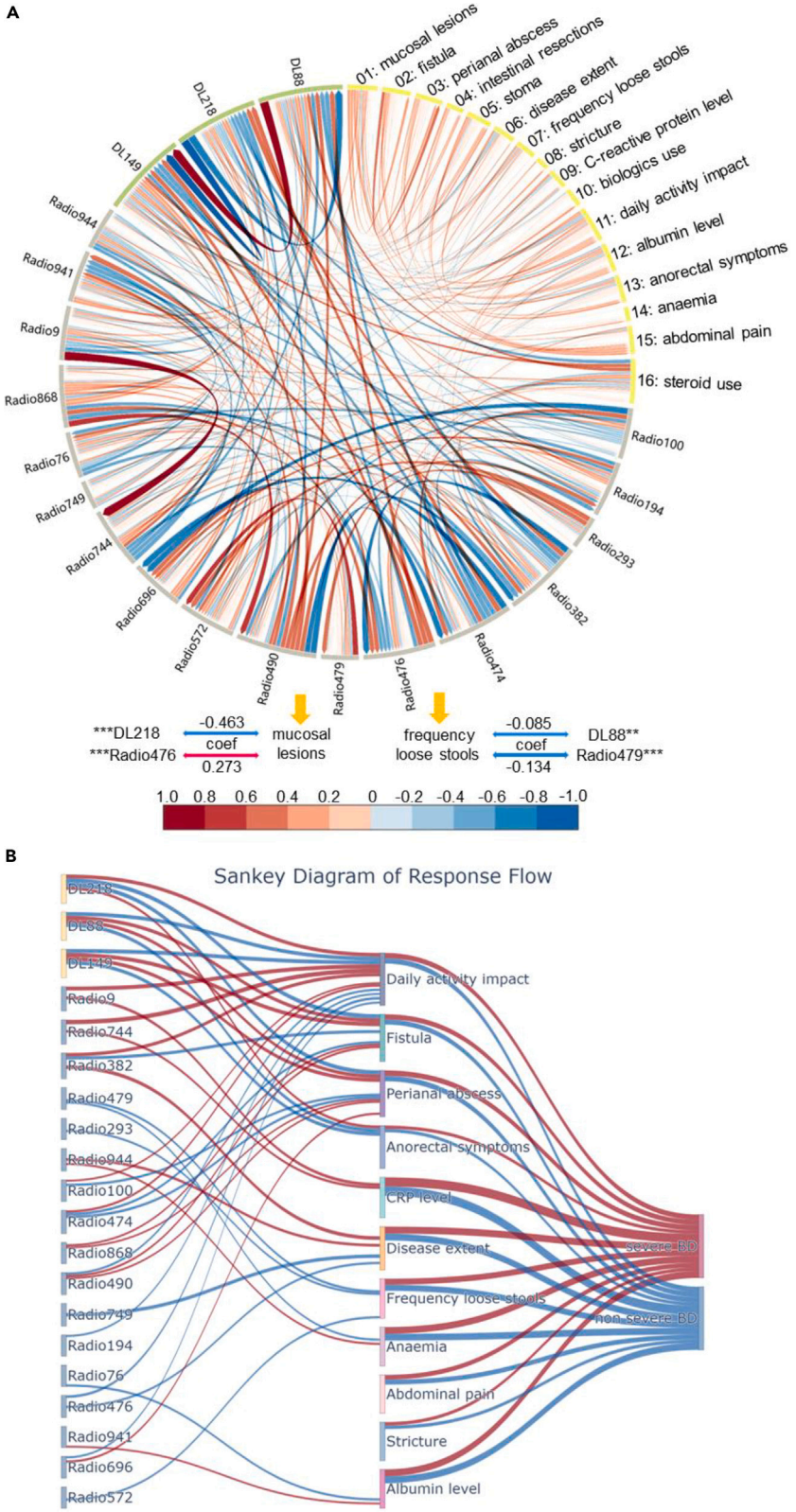

A
01: mucosal lesions
02: fistula
03: perianal abscess
04: intestinal resections
05: stoma
06: disease extent
07: frequency loose stools
08: stricture
09: C-reactive protein level
10: biologics use
11: daily activity impact
12: albumin level
13: anorectal symptoms
14: anaemia
15: abdominal pain
16: steroid use
DL88
DL218
DL149
Radio944
Radio941
Radio9
Radio868
Radio76
Radio749
Radio744
Radio696
Radio572
Radio490
Radio479
Radio476
Radio474
Radio382
Radio293
Radio194
Radio100
***DL218
***Radio476
-0.463
coef
0.273
mucosal lesions
frequency loose stools
-0.085
coef
-0.134
DL88**
Radio479***
1.0 0.8 0.6 0.4 0.2 0 -0.2 -0.4 -0.6 -0.8 -1.0
B
Sankey Diagram of Response Flow
DL218
DL88
DL149
Radio9
Radio744
Radio382
Radio479
Radio293
Radio944
Radio100
Radio474
Radio868
Radio490
Radio749
Radio194
Radio76
Radio476
Radio941
Radio696
Radio572
Daily activity impact
Fistula
Perianal abscess
Anorectal symptoms
CRP level
Disease extent
Frequency loose stools
Anaemia
Abdominal pain
Stricture
Albumin level
severe BD
non-severe BD

Figure 3. The chord diagram and Sankey diagram of VAT-FI

(A) Chord diagram illustrates pairwise correlations among top 20 fingerprint features of VAT and DSI items, with significant feature pairs and their coefficients shown below the chord graph. The red bars indicate positive correlations, while the blue ones represent negative correlations. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. (B) Sankey diagram illustrates the putative pathways connecting "fingerprint features of VAT" with "DSI items" and ultimately influencing "BD severity." The left nodes represent deep learning and radiomics features, middle nodes represent DSI items, and right nodes indicate BD severity. For (A) and (B), red lines indicate positive correlation, while blue lines indicate negative correlation. The thickness of lines presents strength of correlation, with thicker lines indicating a stronger correlation. (CRP, C-reactive protein; DSI, disease severity index; VAT, visceral adipose tissue; BD, bowel damage; DL, deep learning feature; Radio, radiomics feature.).

Our VAT-FI can offer a streamlined approach to characterize BD in patients with CD in clinical practice. Nevertheless, the clinical application of this diagnostic tool still has potential limitations. Firstly, its efficacy is significantly contingent upon the quality of the CT images, requiring clear imaging to ensure accurate feature extraction. Secondly, in cases where patients exhibit conditions that can distort the adipose tissue profile, such as the presence of ascites, the influence of hormone therapy, or cardiac/hepatic/renal insufficiency, the utility of this tool for evaluation may be compromised. Consequently, these factors necessitate a cautious approach and careful consideration when employing the tool in clinical practice.

In conclusion, our VAT-FI offers a precise means of assessing BD severity. It outperforms the SAT-FI and conventional fat metrics in terms of predictive efficiency and generalizability. It is expected to replace complex scoring systems such as DSI and Lémann index (LI). This novel approach simplifies the clinical BD assessment and eliminates the need to acquire and analyze extensive clinical data, making it a valuable tool in clinical practice.

Limitations of the study

However, there were several limitations in this study. First, the specific pathophysiological importance of fingerprint features remains unclear because most CD patients are not undergoing surgery, and VAT samples cannot be obtained for analysis. Future research should explore the intricate relationship between fingerprint features and VAT specimens, potentially utilizing metabolic fingerprints for this purpose.[18–21] Second, we did not include fingerprint features of the intestine in the analysis. In AI analysis of images, target segmentation and extraction are

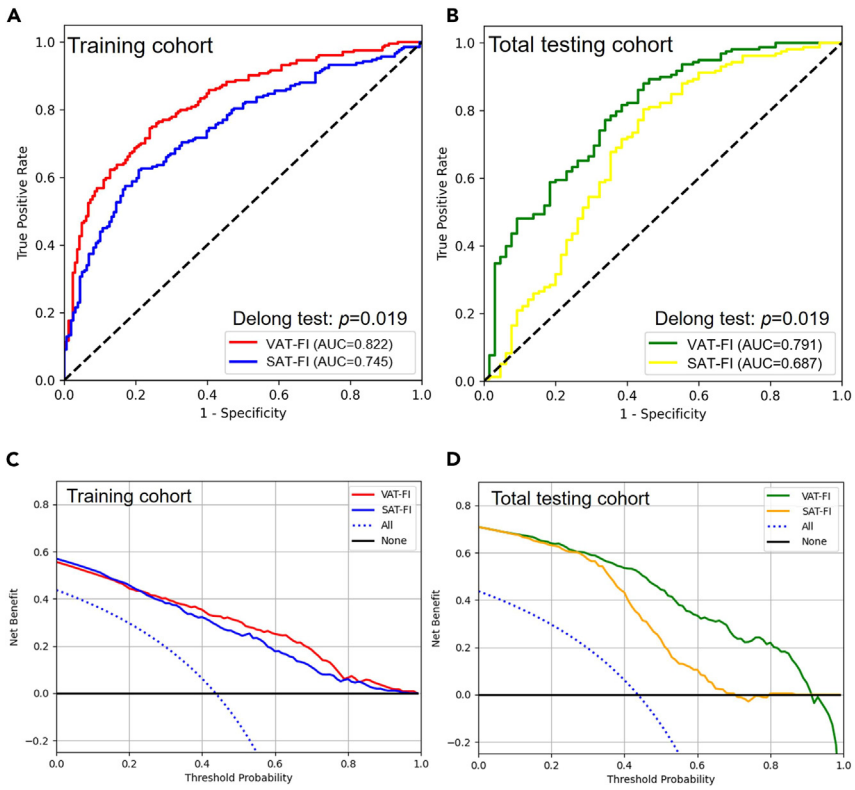


Figure 4. Comparison of the predictive performance between VAT-FI and SAT-FI

(A–D) ROC curves of FIs in training cohort (A) and total testing cohort (B). Decision curves of FIs in training cohort (C) and total testing cohort (D). AUC, area under the receiver operating characteristic curve; FI, fingerprint index; ROC, receiver operating characteristic; SAT, subcutaneous adipose tissue; VAT, visceral adipose tissue.

Table 2. Univariate and multivariate logistic regression analysis of clinical, laboratory, and imaging factors

	Training cohort		Total testing cohort	
Characteristics	Coefficients	*p* values	Coefficients	*p* values
Univariate logistic regression analysis				
Clinical factors				
Sex	−0.46	0.29	−0.95	0.06
Age	−0.01	0.56	−0.02	0.19
Smoking	−0.33	0.47	0.30	0.62
Disease duration (months)	<0.01	0.49	0.01	0.23
Body mass index (kg/m^2)	>−0.01	0.96	−0.01	0.86
Laboratory factors				
Erythrocyte sedimentation rate (mm/h)	0.04	<0.01	0.02	0.02
Adipose tissue metrics				
VAT volume	−0.03	0.57	0.02	0.41
SAT volume	0.01	0.76	>−0.01	0.86
Visceral to subcutaneous fat volume ratio	−0.59	0.86	1.35	0.74
Visceral to subcutaneous fat area ratio at L3	0.22	0.91	0.32	0.88
Visceral to subcutaneous fat area ratio at L4	1.40	0.42	−1.33	0.54
Radiomic models				
VAT-FI	1.50	0.04	2.43	<0.01
SAT-FI	0.31	0.71	2.27	0.09
Multivariate logistic regression analysis				
ESR	0.06	<0.01	−0.02	0.01
VAT-FI	1.68	0.01	2.65	<0.01

In both univariate and multivariate logistic regression analysis, $p < 0.050$ was considered statistically significant. The analyzed variables were VAT (visceral adipose tissue), SAT (subcutaneous adipose tissue), and FI (fingerprint index).

typically needed. In CT images, automatic segmentation and extraction of VAT can be easily achieved due to its distinctive fat density; however, manually outlining the intestine with soft tissue density can be time-consuming. Considering that VAT from a single CT scan at initial diagnosis or during follow-up has the potential to reflect cumulative BD, selecting VAT rather than the infected intestine for AI analysis was a more efficient approach, aligning with one of the objectives of our study.

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to the lead contact, Prof. Shi-Ting Feng (fengsht@mail.sysu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- All original code has been deposited at GitHub and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This study was supported by the National Key R&D Program of China (2023YFC2507300), Key Area Research and Development Program of Guangdong Province (2023B1111040003), National Natural Science Foundation of China (82070680, 82270693, 62371303, 82471948, 82271958, 82170537, 82222010, 82072002), Guangdong Basic and Applied Basic Research Foundation (2023B1515020070, 2023A1515011097, 2023A1515011304, 2023A1515010388), Fundamental Research Funds for the Central Universities Sun Yat-sen University (24ykqb003),and SKY Imaging Research Fund of the Chinese International Medical Foundation (Z-2014-07-2301). The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

AUTHOR CONTRIBUTIONS

X.L., L.S., R.M., B.H., and S.-T.F. were responsible for the overall study design. X.L., C.H., H.W., Y.L., Jiaqiang Li, R.Z., and Y.W. supervised the data collection. X.L., C.H., H.W., Y.L., Jiaqiang Li, R.Z., and Y.W. performed data analysis. X.L., C.H., H.W., Y.L., and Jiaqiang Li completed manuscript drafting. E.C., X.Z., Jianpeng Li, J. Lu, Z.P., C.S., Z.L., M.C., L.S., B.H., R.M., and S.-T.F. were responsible for manuscript editing. All authors read, discussed, and approved the final version of the manuscript. All authors had full access to the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis as well as the decision to submit for publication.

DECLARATION OF INTERESTS

All authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Ethics statement
  - Patients
- METHOD DETAILS
  - Assessment of BD
  - Computed tomography enterography protocol
  - The development and validation of the artificial intelligence-based FIs
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.111022.

Received: April 22, 2024
Revised: June 30, 2024
Accepted: September 20, 2024
Published: September 28, 2024

REFERENCES

1. Pariente, B., Mary, J.Y., Danese, S., Chowers, Y., De Cruz, P., D'Haens, G., Loftus, E.V., Jr., Louis, E., Panés, J., Schölmerich, J., et al. (2015). Development of the Lémann Index to Assess Digestive Tract Damage in Patients With Crohn's Disease. Gastroenterology *148*, 52–63.e3.
2. Peyrin-Biroulet, L., Loftus, E.V., Colombel, J.F., and Sandborn, W.J. (2010). The Natural History of Adult Crohn's Disease in Population-Based Cohorts. Am. J. Gastroenterol. *105*, 289–297.
3. Peyrin-Biroulet, L., Cieza, A., Sandborn, W.J., Kostanjsek, N., Kamm, M.A., Hibi, T., Lémann, M., Stucki, G., and Colombel, J.F. (2010). Disability in inflammatory bowel diseases: developing ICF Core Sets for patients with inflammatory bowel diseases based on the International Classification of Functioning, Disability, and Health. Inflamm. Bowel Dis. *16*, 15–22.
4. Siegel, C.A., Whitman, C.B., Spiegel, B.M.R., Feagan, B., Sands, B., Loftus, E.V., Jr., Panaccione, R., D'Haens, G., Bernstein, C.N., Gearry, R., et al. (2018). Development of an index to define overall disease severity in IBD. Gut *67*, 244–254.
5. Qiu, Y., Zhou, L., and Lu, B. (2023). Validation of disease severity index for predicting complicated disease in Crohn's disease: A comparison study with Lemann index. Dig. Liver Dis. *56*, 635–640.
6. Li, X., Zhang, N., Hu, C., Lin, Y., Li, J., Li, Z., Cui, E., Shi, L., Zhuang, X., Li, J., et al. (2023). CT-based radiomics signature of visceral adipose tissue for prediction of disease progression in patients with Crohn's disease: A multicentre cohort study. EClinicalMedicine *56*, 101805.
7. Li, X.H., Feng, S.T., Cao, Q.H., Coffey, J.C., Baker, M.E., Huang, L., Fang, Z.N., Qiu, Y., Lu, B.L., Chen, Z.H., et al. (2021). Degree of Creeping Fat Assessed by Computed Tomography Enterography is Associated with Intestinal Fibrotic Stricture in Patients with Crohn's Disease: A Potentially Novel Mesenteric Creeping Fat Index. J. Crohns Colitis *15*, 1161–1173.
8. Meng, J., Mao, Y., Zhou, J., Chen, Z., Huang, S., Wang, Y., Huang, L., Zhang, R., Shen, X., Lv, W., et al. (2022). Mesenteric abnormalities play an important role in grading intestinal fibrosis in patients with Crohn's disease: a computed tomography and clinical marker-based nomogram. Therap. Adv. Gastroenterol. *15*, 17562848221122504.
9. Erhayiem, B., Dhingsa, R., Hawkey, C.J., and Subramanian, V. (2011). Ratio of Visceral to Subcutaneous Fat Area Is a Biomarker of Complicated Crohn's Disease. Clin. Gastroenterol. Hepatol. *9*, 684–687.e1.
10. Bryant, R.V., Schultz, C.G., Ooi, S., Goess, C., Costello, S.P., Vincent, A.D., Schoeman, S., Lim, A., Bartholomeusz, F.D., Travis, S.P.L., and Andrews, J.M. (2019). Visceral Adipose Tissue Is Associated With Stricturing Crohn's Disease Behavior, Fecal Calprotectin, and Quality of Life. Inflamm. Bowel Dis. *25*, 592–600.
11. Büning, C., von Kraft, C., Hermsdorf, M., Gentz, E., Wirth, E.K., Valentini, L., and Haas, V. (2015). Visceral Adipose Tissue in Patients with Crohn's Disease Correlates with Disease Activity, Inflammatory Markers, and Outcome. Inflamm. Bowel Dis. *21*, 2590–2597.
12. Coffey, J.C., Byrnes, K.G., Walsh, D.J., and Cunningham, R.M. (2022). Update on the mesentery: structure, function, and role in disease. Lancet Gastroenterol. Hepatol. *7*, 96–106.
13. Eder, P., Adler, M., Dobrowolska, A., Kamhieh-Milz, J., and Witowski, J. (2019). The Role of Adipose Tissue in the Pathogenesis and Therapeutic Outcomes of Inflammatory Bowel Disease. Cells *8*, 628.
14. Kredel, L.I., and Siegmund, B. (2014). Adipose-tissue and intestinal inflammation - visceral obesity and creeping fat. Front. Immunol. *5*, 462.
15. Zulian, A., Cancello, R., Micheletto, G., Gentilini, D., Gilardini, L., Danelli, P., and Invitti, C. (2012). Visceral adipocytes: old actors in obesity and new protagonists in Crohn's disease? Gut *61*, 86–94.
16. Van Der Sloot, K.W.J., Joshi, A.D., Bellavance, D.R., Gilpin, K.K., Stewart, K.O., Lochhead, P., Garber, J.J., Giallourakis, C., Yajnik, V., Ananthakrishnan, A.N., et al. (2017). Visceral Adiposity, Genetic Susceptibility, and Risk of Complications Among Individuals with Crohn's Disease. Inflamm. Bowel Dis. *23*, 82–88.
17. Schineis, C.H.W., Pozios, I., Boubaris, K., Weixler, B., Kamphues, C., Margonis, G.A., Kreis, M.E., Strobel, R.M., Beyer, K., Seifarth, C., et al. (2024). Role of visceral fat on postoperative complications and relapse in patients with Crohn's disease after ileocecal resection: Is it overrated? Int. J. Colorectal Dis. *39*, 20.

18. Chen, X., Shu, W., Zhao, L., and Wan, J. (2023). Advanced mass spectrometric and spectroscopic methods coupled with machine learning for *in vitro* diagnosis. View *4*, 20220038.
19. Zhu, Y., and Girault, H.H. (2023). Algorithms push forward the application of MALDI–TOF mass fingerprinting in rapid precise diagnosis. View *4*, 20220042.
20. Wang, L., Zhang, M., Pan, X., Zhao, M., Huang, L., Hu, X., Wang, X., Qiao, L., Guo, Q., Xu, W., et al. (2022). Integrative Serum Metabolic Fingerprints Based Multi-Modal Platforms for Lung Adenocarcinoma Early Detection and Pulmonary Nodule Classification. Adv. Sci. *9*, 2203786.
21. Chen, Y., Xu, W., Zhang, W., Tong, R., Yuan, A., Li, Z., Jiang, H., Hu, L., Huang, L., Xu, Y., et al. (2023). Plasma metabolic fingerprints for large-scale screening and personalized risk stratification of metabolic syndrome. Cell Rep. Med. *4*, 101109.
22. Swaminathan, A., Fan, D., Borichevsky, G.M., Mules, T.C., Hirschfeld, E., Frampton, C.M., Day, A.S., Siegel, C.A., and Gearry, R.B. (2022). The disease severity index for inflammatory bowel disease is associated with psychological symptoms and quality of life, and predicts a more complicated disease course. Aliment. Pharmacol. Ther. *56*, 664–674.
23. Wang, H., Cao, P., Wang, J., and Zaiane, O.R. (2022). Uctransnet: rethinking the skip connections in u-net from a channel-wise perspective with transformer. In Proceedings of the AAAI Conference on Artificial Intelligence, *36*Proceedings of the AAAI Conference on Artificial Intelligence, pp. 2441–2449.
24. Shamshad, F., Khan, S., Zamir, S.W., Khan, M.H., Hayat, M., Khan, F.S., and Fu, H. (2023). Transformers in medical imaging: A survey. Med. Image Anal. *88*, 102802.
25. Huang, B., Tian, J., Zhang, H., Luo, Z., Qin, J., Huang, C., He, X., Luo, Y., Zhou, Y., Dan, G., et al. (2021). Deep Semantic Segmentation Feature-Based Radiomics for the Classification Tasks in Medical Image Analysis. IEEE J. Biomed. Health Inform. *25*, 2655–2664.
26. MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, *1*Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, pp. 281–297.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw and analyzed data	This paper	N/A
Software and algorithms		
UctransNet	Wang H et al.[14]	https://github.com/McGregorWwww/UCTransNet
DSFR	Huang B et al.[15]	https://github.com/Rahysea/DSFR_MTM_Radiology
ITK-SNAP	Version 3.6.0	https://www.itksnap.org/
SPSS	Version 20	https://www.ibm.com/products/spss-statistics
PyRadiomics	Version 3.1.0	https://pyradiomics.readthedocs.io/en/v3.1.0/installation.html
OpenCV-Python	Version 4.8.0	https://pypi.org/project/opencv-python/
Pytorch	Version 2.0.1	https://pytorch.org/
Python	Version 3.9.0	https://www.python.org/
Custom computer code	This paper	https://github.com/MedcAILab/DSFR_CD_VAT

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Ethics statement

This multicenter observational study was approved by the institutional ethics review board of The First Affiliated Hospital of Sun Yat-Sen University (2022-046)). Informed consent was waived by the review board.

Patients

This multicenter observational study, encompassing 1148 CD patients from six tertiary referral centers from August 2010 to March 2023. Among the total testing cohort, 70.7% were male and the median age was 30 $\pm$ 11 years. Among the training cohort, 69.5% were male and the median age was 31 $\pm$ 12 years. No significant differences in gender and age were observed between patients with and without disease progression. The inclusion criteria were as follows: (A) CD patients with comprehensive data, including clinical information, abdominal and pelvic CT images, laboratory data, and endoscopic examination; and (B) laboratory data obtained within two weeks and endoscopy within one month of the CT scan. The exclusion criteria were as follows: (A) severe abdominal/pelvic effusion, subcutaneous edema, or metallic artifacts that impeded accurate identification of adipose tissue on CT images; (B) concomitant diseases such as malignant tumors, endocrine disorders, or cardiac/hepatic/renal insufficiency that could impact the distribution or pathophysiological characteristics of adipose tissue; and (C) received corticosteroid treatment within 3 months before CT scan (Figure S1). The clinical and laboratory data were obtained from electronic medical records. Endoscopic images with reports were used to retrospectively analyze the mucosal lesions.

METHOD DETAILS

Assessment of BD

The DSI[4] was used to assess the long-term cumulative BD in patients with CD. This evaluation utilized biological markers, endoscopic and/or imaging findings, clinical indicators, history of CD-related complications, and subjective impact on patients' lives[4] (Table S1). Based on specialist opinion, the attributes of IBD that are most important for overall disease severity have been defined, as shown in the following table. The relative importance of each attribute was adapted to a 100-point scale. During the evaluation, a gastroenterologist conducted the clinical and endoscopic assessments, while a radiologist evaluated the radiological contents. Their evaluation results were used for subsequent data analysis. An additional gastroenterologist or radiologist independently assessed the corresponding items at two different times at 3-month intervals. To evaluate the inter/intraobserver agreement, the intraclass correlation coefficient (ICC) was used. A DSI>23 and DSI $\leq$ 23 indicated severe and nonsevere BD, respectively.[22]

Computed tomography enterography protocol

All patients with Crohn's disease were required to fast and ingest an oral polyethylene glycol electrolyte solution for 6–8 hours, followed by ingestion of 1600–2000 mL of a 2.5% mannitol solution 1 hour before the image acquisition. Plain and dual phase contrast-enhanced CTs were performed from the dome of the liver to the symphysis pubis using one of the following CT scanners shown in the Table S2, where the scanning parameters are also provided. After the plain scanning, 2 mL/kg of a non-ionic contrast agent ([Ultravist; Schering, Berlin, Germany] or

[Ioversol; Hengrui, Jiangsu, China]) was injected intravenously in the right antecubital vein at a rate of 3.5 mL/s using a dual-head power injector. The upper limit of the contrast dose was set at 150 mL for every patient. Forty millilitres of 0.9% saline was administered immediately after the injection of contrast agent at the same rate. The arterial phase scan was started 28 seconds post-injection, and the venous phase scan was initiated 70 seconds post-injection.

The development and validation of the artificial intelligence-based FIs

Volumes of interest (VOI) segmentation and CT-based radiomics feature extraction

All patients underwent CT examinations utilizing CT scanners across six centers (Table S2). To obtain the CT-based VOIs, we trained the segmentation framework[23] with 15 samples from the training set that exhibited fine delineation results. These samples, including three-dimensional VOI regions of the VAT and SAT, were manually contoured by a Radiologist using open-source medical imaging software (ITK-SNAP, version 3.6.0, www.itkSnap.org). A total of 1128 radiomics features with an ICC greater than 0.80 were extracted from the corrected VOIs that encompassed the VAT or SAT, using the PyRadiomics software (https://pyradiomics.readthedocs.io/en/v1.0/installation.html) (Figure 1B). To evaluate the inter-/intra-observer consistency of the extracted features, 30 patients were randomly selected from the training cohort. Another radiologist (Radiologist B) conducted the VOI inspection and correction under the same conditions, and the same features were extracted. Then, after 3 months, Radiologist A re-examined and corrected the VOIs of the same 30 patients under the same conditions, and the same features were extracted again. The ICCs were computed using a two-way random-effects model to determine the inter-/intra-observer consistency. Features with ICCs $\geq$ 0.80 were selected indicating good consistency.

After the training of segmentation model with corrected VOIs, the remaining 1120 samples in the training cohort and total testing cohort were also segmented automatically by this model to obtain the VOIs of VAT and SAT. In summary, our extracted texture features encompassed a comprehensive set of descriptors: 14 Shape-Based features, 18 First-Order Statistics features, 24 Grey Level Co-occurrence Matrix features, 16 Grey Level Run Length Matrix features, 16 Grey Level Size Zone Matrix features, 14 Grey Level Dependence Matrix features, 5 Neighborhood Grey Tone Difference Matrix features, 742 Wavelet Filter Features, and 279 LoG Filter Features. This comprehensive feature set offered a robust foundation for our analysis and extended our understanding of the image characteristics under consideration.

The extraction and selection of deep learning features

We employed UCtransNet,[23] a transformer-based segmentation network that excels at capturing comprehensive global context and leverage low-level CNN features[24] and construct a deep learning feature library to access a set of 512 deep learning features for each patient.[25] The procedure of that consists of two parts.

The first step involved the trained network that is capable of efficiently extracting semantic information from medical images with the dataset comprising CT images annotated for intestinal fat and corresponding VOIs. Subsequently, A deep learning feature library was then constructed based on the trained network. Each image in the dataset, along with its corresponding segmentation result, underwent preprocessing operations such as cropping and scaling before being input to the network (Figure 1C). The deep learning feature maps obtained from network are transformed into corresponding feature vectors using global average pooling. These feature vectors were then averaged to get access to a fixed-size feature vector, which formed the known deep learning feature library for subsequent feature similarity calculations.

Next, an unsupervised feature aggregation algorithm (K-means clustering algorithm[26]) was applied to divide the features obtained in the previous step, into two classes: effective features (assumed to contain more features of lesion layers) and redundant features (Figure 1D). Then, the similarity between the two unsupervised generated clusters and the known lesion feature database generated in Figure 1C was computed to select the optimal combination of effective features (Figure 1D).

Feature selection and FI development for radiomics analysis

The final step involved utilizing LASSO and mutual information (MI) to screen the optimal subset of features to construct logistic regression models named VAT-FI and SAT-FI. Each model's features underwent three rounds of LASSO selection——initial selection of radiomics features (a), initial selection of deep learning features (b), and further selection of features combining radiomics and deep learning (c). The tuning parameter α of LASSO was set within a predefined search range, specifically between 0.0001 and 0.1. In the initial selection phase, α values were kept consistent for both (a) and (b) to streamline the search process. In the further selection phase (c), the range was maintained at a certain value, and α was re-searched. In summary, α was optimized simultaneously, and the random grid search approach was employed to optimize each model. Selection of the most appropriate hyperparameter values was guided by the mean squared error observed in the training cohort. Furthermore, we incorporated MI, an information theory-based approach, for feature selection and evaluation. MI quantifies the correlation or dependency between two random variables, considering non-linear relationships rather than solely linear ones. MI measures the mutual information gain between two random variables, indicating how much the predictive power of one variable increases when we know the power of the other variable. We calculated the mutual information between each feature and the target variable, then ranked or selected features based on their mutual information values. Ultimately, we chose 25 radiomics features and 23 deep learning features based on this criterion.

Ultimately, we utilized cross-validation and a grid search to optimize the FI parameters to develop the best FIs. In the testing phase, the trained model was applied to testing cohorts for BD prediction. We assessed FI stability through stratified analysis across testing cohorts and subgroups categorized based on the type of CT scanners, ensuring consistent performance across diverse conditions.

Visualized correlations between BD and FIs

To interpret the performance of FIs, we utilized the SHapley Additive exPlanations (SHAP) method to quantify features' impact on FI predictions. The feature distributions in the FIs were visualized using raincloud plots.

To further investigate how VAT-FI predicted BD, we calculated Pearson correlation coefficients between features and DSI items, visualizing them through a chord diagram. This analysis used t-test to determine the statistical significance of the Pearson correlation coefficients. To mitigate the risk of false-positives inherent in multiple comparisons, we employed Benjamini–Hochberg procedure to control the false discovery rate for p values less than 0.05. The key features identified by SHAP were correlated with DSI parameters and BD severity using Pearson correlation. Their relationships were elucidated in a Sankey diagram.

Measurement of conventional adipose tissue parameters

To evaluate the predictive performance of FIs, we compared their efficacy with that of conventional metrics commonly employed to characterize adipose tissue[6] (Figure S2). The VAT volume and SAT volume (mm3) were calculated using open-source radiomic toolkit (https://pyradiomics.readthedocs.io/en/v1.0/installation.html). The VAT/SAT volume ratio was determined by dividing the VAT volume by the SAT volume. VAT and SAT areas (mm2) at the levels of L3 and L4 vertebrae were measured using OpenCV-Python toolkit (https://pypi.org/project/opencv-python/).

QUANTIFICATION AND STATISTICAL ANALYSIS

The clinical factors were analyzed through Student's t test, Mann−Whitney U test, or Chi−square test. The efficiency of FIs was evaluated through ROC analysis and Hosmer–Lemeshow test for goodness-of-fit. The AUC was compared against reference standards (AUC=0.50) and pairwise compared using DeLong test. Decision curve analysis was used to compare the net benefit between FIs. Additionally, candidate parameters were tested using univariate logistic regression analysis. Factors with a p value less than 0.05 in this analysis were selected for multivariate logistic regression analysis to identify significant predictors of BD. Statistical analysis was conducted using SPSS software (version 20; https://www.ibm.com/cn-zh/analytics/SPSS statistics software) or Python (version 3.9; https://www.python.org/).