

Derivation and Internal Validation of a Clinical Prediction Tool to Predict Nonalcoholic Fatty Liver Disease in Patients With Crohn's Disease

Scott McHenry, MD,*[Ⓞ] Ankita Tirath, MD,[†] Richard Tsai, MD,[‡] Yeshika Sharma, MD,[§] Avegail G. Flores, MD,*
Nicholas O. Davidson, MD, DSc,* Kathryn J. Fowler, MD,^{‡,¶} Matthew A. Ciorba, MD,*^{||} and
Parakkal Deepak, MD*^{||}[Ⓞ]

Background: Crohn's disease (CD) patients have more than double the risk of nonalcoholic fatty liver disease (NAFLD) compared with the general population after considering traditional risk factors. NAFLD remains underappreciated because routine imaging and liver biochemistries are neither sensitive nor specific for the diagnosis. Here we developed a Clinical Prediction Tool for NAFLD in CD (CPN-CD) using readily accessible parameters to diagnose NAFLD, as determined by magnetic resonance proton density fat fraction (PDFF).

Methods: A total of 311 consecutive CD patients who underwent magnetic resonance enterography from June 1, 2017, to May 31, 2018, were screened for NAFLD, defined as a PDFF >5.5% after excluding other liver diagnoses. CPN-CD was derived using binary multivariate logistic regression and internally validated with a 10-fold cross-validation. CPN-CD was compared with the Hepatic Steatosis Index (HSI) by the C-statistic and categorical Net Reclassification Improvement (NRI).

Results: CPN-CD included age, sex, ethnicity/race, serum alanine aminotransferase, body mass index, known cardiometabolic diagnoses, CD duration, and current use of azathioprine/6-mercaptopurine. At <20% risk, NAFLD could be excluded with a sensitivity of 86% (negative predictive value, 86%). At ≥50% risk, NAFLD was diagnosed with a specificity of 87% (positive predictive value, 75%). CPN-CD exhibited good discrimination (C-statistic 0.85) compared with fair discrimination of the HSI (C-statistic, 0.76). CPN-CD was superior to the HSI by net reclassification improvement (+0.20; $P < 0.001$) and decision curve analysis.

Conclusions: CPN-CD outperforms HSI in detecting NAFLD in patients with CD. Future directions include external validation, outcome validation, and testing generalizability to patients with ulcerative colitis.

The findings of this article are also available as a playable video in the HTML version of this article.

Key Words: nonalcoholic fatty liver disease, Crohn's disease, clinical prediction tool

INTRODUCTION

With the introduction of treatment paradigms that can induce clinical, endoscopic, and mucosal disease remission, the care of patients with Crohn's disease (CD) has increasingly focused on health maintenance and chronic disease management.¹ This is highlighted by the growing burden of obesity

in this population, which is paradoxical as those patients are known to exhibit malabsorption, undernutrition, and wasting.^{2,3} The impact of obesity is expected to compound the burden of nonalcoholic fatty liver disease (NAFLD), one of the earliest appreciated extra-intestinal manifestations of inflammatory bowel diseases (IBDs).^{4,5}

Received for publications September 17, 2019; Editorial Decision December 3, 2019.

From the *Division of Gastroenterology, Washington University in Saint Louis, St. Louis, Missouri, USA; [†]Department of Medicine, Washington University in Saint Louis, St. Louis, Missouri, USA; [‡]Department of Radiology, Mallinckrodt Institute of Radiology, Washington University in Saint Louis, St. Louis, Missouri, USA; [§]Division of Gastroenterology, University of Cincinnati, Cincinnati, Ohio, USA; [¶]Department of Radiology, University of California at San Diego, San Diego, California, USA; ^{||}Washington University Inflammatory Bowel Diseases Center, Washington University in Saint Louis, St. Louis, Missouri, USA

Supported by: S.M. is supported by an Institutional National Research Service Award (T32-DK007130-45). N.O.D. is supported by grants DK-119437, HL-38180, and DK-112378. P.D. is supported by a Junior Faculty Development Award from the American College of Gastroenterology. K.J.F. declares research grant support from General Electric, Bayer, and Pfizer. M.A.C. is supported by DK109384, a Crohn's and Colitis Foundation Daniel H. Present Senior Research Award (Ref. 370763), and by philanthropic support from the Givin' it All for Guts Foundation (<https://givinitallforguts.org>) and the Lawrence C. Pakula MD IBD Research Innovation and Education Fund. The work performed in this paper was additionally supported by grants provided by the National Institutes of Health through the Washington University in Saint Louis' Digestive Disease Research Core (P30 DK-52574). Additional grant support for the REDCap database was provided by the Clinical and Translational Science Award (UL1 TR000448) and the Siteman Cancer Center Support Grant (P30-CA091842).

Conflicts of interest: There are no potential personal or financial conflicts of interest to declare.

Conflicts of interest: There are no potential personal or financial conflicts of interest to declare.

Address correspondence to: Parakkal Deepak, MBBS, MS, Division of Gastroenterology, John T. Milliken Department of Medicine, Washington University in St. Louis School of Medicine, St. Louis, MO 63110 (deepak.parakkal@wustl.edu).

© 2020 Crohn's & Colitis Foundation. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

doi: 10.1093/ibd/izz324

Published online 7 January 2020

Although NAFLD is already known to affect one-quarter of the world population,⁶ CD patients have an even higher burden of NAFLD than community-based or clinic-based controls despite being thinner and younger.⁷ The risk for NAFLD is more than doubled in CD after adjusting for classic risk factors such as body mass index (BMI) and ethnicity.⁷ As the hepatic manifestation of the metabolic syndrome, NAFLD patients are at risk for both liver cirrhosis and hepatocellular carcinoma.^{8,9}

Independent of these liver-related outcomes, the overall clinical course of NAFLD is shaped by the attendant risks of cardiometabolic complications such as diabetes mellitus, cerebrovascular disease, and coronary artery disease.¹⁰ These comorbidities are predicted to compound the increased risk of acute myocardial infarction and heart failure recently demonstrated in patients with IBD.¹¹ In the general population, longitudinal studies demonstrate that the resolution of hepatic steatosis is associated with a reduction in the incidence of diabetes mellitus; however, patients with worsening obesity may have more progressive hepatic fibrosis.^{12,13} This suggests that early recognition of NAFLD in patients with CD may likewise mitigate the development of cardiometabolic complications in this population.

Although an ultrasound-based screening program for NAFLD in patients with IBD has been proposed,¹⁴ this approach can only reliably detect hepatic steatosis if >30% of the liver is involved.^{15,16} An expected consequence of this screening approach is it would fail to recognize patients with subtle degrees of hepatic steatosis. Although this limitation can be circumvented with modified techniques such as the continuous attenuation parameter,^{17,18} this approach is not widely available. Furthermore, there persists an as-yet unproven cost-benefit of such an NAFLD screening program.

Recognizing these important limitations, we reasoned that the accuracy and cost-effectiveness of an NAFLD screening program in CD could be improved if patients are first screened by a clinical prediction tool, with a subset of at-risk patients then undergoing imaging to evaluate for hepatic steatosis and fibrosis. Accordingly, we developed and internally validated a Clinical Prediction Tool for NAFLD in CD (CPN-CD) coupling magnetic resonance proton density fat fraction (MR-PDFF) imaging with clinical predictors readily available to the physician.

METHODS

Ethical Statement

The protocol and conduct of this study conformed with the 1975 Declaration of Helsinki and the Health Insurance Portability and Accountability Act (HIPPA). A waiver of consent was obtained through the institutional review board at Washington University School of Medicine in St. Louis (IRB #201705093).

Patient Selection

This study represents a secondary analysis of a previously published analysis in patients with CD.⁷ In brief, we included all consecutive CD patients who underwent magnetic resonance enterography through our tertiary referral center's Inflammatory Bowel Disease network of clinics from June 1, 2017, to May 31, 2018 (BJC Health Care, Washington University Physicians, Saint Louis, MO, USA). Magnetic resonance (MR) enterography is standard of care in our practice for disease staging and to screen for clinically occult activity. These data were linked to the patient's electronic medical records, which then underwent a structured chart review using Research Electronic Data Capture (REDCap) to ensure data fidelity, security, and to provide a clear audit trail. Missing data were handled by multiple (5) imputations, with the results reported as the aggregate value, unless otherwise noted.

Subjects were included if a board-certified gastroenterologist had made the clinical diagnosis of CD. The exclusion criteria included the presence of a confounding IBD diagnosis (such as indeterminate colitis, microscopic colitis, ulcerative colitis) or the presence of an alternative etiology of liver disease. We excluded patients who consumed excess alcohol (defined as >3 drinks a day for men, >2 standard drinks a day for women, consistent with the guideline criteria for the diagnosis of NAFLD),¹⁹ chronic viral hepatitis (defined using serology regardless of treatment status), autoimmune liver disease (defined as a the clinical diagnosis of primary sclerosing cholangitis, primary biliary cholangitis, or autoimmune hepatitis), or a clinical diagnosis of an alternative metabolic/toxic liver disease. We did not exclude severe malnutrition (defined as a BMI <18.5 kg/m²), glucocorticoid use, methotrexate use, or total parenteral nutrition, as we wished to increase the generalizability of our model and potentially include these as predictors if appropriate.

Imaging and Fat Quantification

All imaging was performed on 1 of 9 1.5–3.0-T Siemens (Erlangen, Germany) MR scanners. During the study period, clinical MR enterography protocols had been modified to allow the calculation of proton density fat fraction (PDFF) maps. The additional sequence itself is a 2-dimensional low-flip-angle multi-echo proton density-weighted gradient-recalled echo sequence for fat and iron quantification, which is obtained in a single breath hold in <30 seconds. This method has excellent sensitivity/specificity compared with the histologic diagnosis of hepatic steatosis (C-statistic, 0.989).²⁰

To create the PDFF maps from the raw imaging sequence, the imaging files were transferred to a separate, secure local computer and processed using open-source OsiriX software.²¹ The mean PDFF was calculated using 1-cm² regions of interest (ROI) in 9 liver segments (separating 4A and 4B), taking care to avoid blood vessels, bile ducts, or any focal liver lesions. NAFLD was defined as a mean liver fat of >5.5%, which is the well-established

threshold from the Dallas Heart Study and represents the 95th percentile of metabolically normal, lean subjects.²²

Statistical Methods

Derivation of the risk prediction model

We identified candidate predictor variables based on literature review of known risk factors of NAFLD and an exploratory analysis for Crohn's-specific variables. As the CPN-CD was designed to be used before obtaining staging MR enterography or endoscopy, predictors obtained by those means (eg, Simple Endoscopic Score of Crohn's disease, Magnetic Resonance Index of Activity) were not included as candidate predictors. Median and interquartile range for continuous variables and frequency and proportions for categorical variables were calculated. Continuous variables were examined for the linearity assumption. BMI did not meet the linearity assumption, so variables were grouped into categories chosen in a data-driven fashion to have a linear relationship to the logit for NAFLD rather than by World Health Organization classification.

To derive the CPN-CD, binary multivariate logistic regression was performed for the outcome of NAFLD (liver fat >5.5%). Model selection was then achieved by a backwards selection, likelihood ratio method ($P_{\text{entry}} = 0.15$ and $P_{\text{stay}} = 0.10$). The model was assessed for overall performance by the Hosmer and Lemeshow test (a measure of fitting the data where a statistically significant value implies poor performance), Nagelkerke R^2 statistics (measure of the explained variance), discrimination by C-statistic (excellent ≥ 0.90 , good 0.80–0.89, fair 0.70–0.79, and poor < 0.70),²³ and calibration by the slope and intercept of the predicted vs observed plot. The CPN-CD^{LOGIT} was converted to the predicted probability of NAFLD by:

$$\text{Predicted Probability of NAFLD} = \frac{e^{CPN-CD (\text{Logit})}}{1 + e^{CPN-CD (\text{logit})}}$$

Three CPN-CD risk categories (low, intermediate, and high) were created in a data-driven fashion to maximize the sensitivity/negative predictive value for those at "low risk," to maximize the specificity/positive predictive value for those at "high risk," and to minimize the total number of subjects categorized as "intermediate risk." This was achieved by mapping the CPN-CD logit values from the receiver operating curve to give a sensitivity and specificity of ~90% at the lower- and higher-risk thresholds, respectively.

Model performance was internally validated with a 10-fold cross-validation, with the internally validated C-statistic reported as the apparent C-statistic minus the average optimism/overfitting. All statistical tests utilized SPSS, version 25 (IBM Corp., Armonk, NY, USA).

Comparison with the Hepatic Steatosis Index

Our findings were then compared with the previously published Hepatic Steatosis Index (HSI).²⁴ The HSI has previously

been used to assess the burden of NAFLD in IBD²⁵ and is an appropriate comparator as a model that may inform the development of a screening program, as it includes clinical predictors readily available to a treating gastroenterologist, compared with more complex models such as the Fatty Liver Index, which includes waist circumference and insulin concentration.²⁶

The CPN-CD was compared with the HSI in its native, recalibrated, and revised forms by the C-statistic and calibration plots. The strategy to recalibrate the model was to keep the same beta coefficients and to add a constant to best fit the data from our population while keeping the same risk category thresholds as the native model, and the strategy to revise the model was to keep the same predictor variables but to adjust both the beta coefficients and constant.²⁷

The CPN-CD was then compared with the HSI by its categorical Net Reclassification Improvement (NRI).²⁸ The NRI is the sum of the proportion of patients with NAFLD who were more correctly classified (eg, intermediate risk to high risk) and the proportion of patients with normal liver fat who were more correctly classified (eg, intermediate risk to low risk). A paired proportion (McNemar's) test was calculated to assess the statistical significance of the NRI. The native and recalibrated HSI were compared by the previously published low-score (<30), intermediate-score (30–35.9), and high-score (≥ 36) thresholds. As the HSI is not strictly a probability-based risk prediction model, the revised version of the HSI was split into 3 risk categories chosen in a data-driven fashion analogous to the CPN-CD to maximize the sensitivity and specificity at the low-risk and high-risk thresholds, respectively.

Decision curve analysis

Decision curve analysis was performed by comparing the net benefit for the CPN-CD for each decile of calculated risk with the net benefit of screening all patients for NAFLD regardless of calculated risk.²⁹ The net benefit of a risk prediction model is defined as:

$$\text{Net Benefit} = \left(\frac{\text{True Positives}}{N} \right) - \left(\frac{\text{False Positives}}{N} \right) \times \left(\frac{\text{Risk Probability}}{1 - \text{Risk Probability}} \right)$$

Decision curve analysis was also performed for the revised HSI if only obese (BMI >30 kg/m²) patients were screened and if all patients were screened.

Sensitivity analysis: NAFLD with an elevated Fibrosis-4 score

As the risk of liver-related outcomes in NAFLD is associated with the stage of hepatic fibrosis,³⁰ we performed a sensitivity analysis by testing the operating characteristics of the CPN-CD for the outcome of those with both NAFLD

(PDFF >5.5%) and at least intermediate risk of hepatic fibrosis. The later determination was based on the value of the Fibrosis-4 (FIB-4) score, which is a histologically validated score using the patient's age along with the alanine/aspartate aminotransferase and platelet concentrations.^{31,32} An FIB-4 threshold of >1.3 was used, as that is the division between low-risk and intermediate-risk advanced-stage fibrosis. This threshold was chosen as it is reasoned that patients with NAFLD and a FIB-4 with low risk of hepatic fibrosis (ie, <1.3) could be managed by the treating gastroenterologist without needing elastography or specific hepatology consultation.

RESULTS

After applying the inclusion/exclusion criteria, there were 311 subjects in the analytic cohort (90% non-Hispanic white, 50% female, 77% ileocolonic distribution, 30%/35%/35% inflammatory/stricturing/penetrating phenotype, median age 40 years, median CD duration 11 years, median BMI 26 kg/m²).⁷ In general, there were only rare missing data, with only 9 (3%) patients missing liver biochemistry results, which was supplemented by multiple imputation. Missing data to calculate clinical severity using the Harvey-Bradshaw Index (HBI) were more common (60 patients, 19%); however, as the HBI was not a risk factor for NAFLD and appeared to be missing at random, this variable was not imputed or included as a candidate predictor for the multivariate model.

As shown previously,⁷ NAFLD was present in 118 patients (38%), with univariate odds ratios (ORs) calculated (Table 1). The presence of hepatic steatosis was listed in the clinical MR report in only 11 patients (9% of those with NAFLD) and was a recognized problem in the treating gastroenterologist's note in only 9 patients (7% of those with NAFLD). Statistically significant univariate risk factors included traditional predictors such as age, race/ethnicity, alanine aminotransferase (ALT), recent weight loss, BMI, and known cardiometabolic diagnoses (dyslipidemia, hypertension, diabetes mellitus).

We identified several CD-specific univariate predictors including duration of disease, prior ileocolonic or small bowel resection, and current use of azathioprine. Interestingly, the subjective disease severity in the form of the HBI, the use of total parenteral nutrition, and the current use of corticosteroids, biologics, or methotrexate were not predictors for a diagnosis of NAFLD. This is true even if PDFF was used as a continuous outcome rather than the dichotomous outcome of having NAFLD.

The final multivariate risk factors and associated ORs are shown in Table 2. They were used to develop the CPN-CD, defined as:

$$\begin{aligned} \text{CPN CD}^{\text{logit}} = & 0.668 \text{ (if CD duration } \geq 15 \text{ years)} \\ & -1.996 \text{ (if black/African American)} \\ & +1.232 \text{ (if female and age } \geq 50 \text{ years)} \\ & +0.956 \text{ (if there is a clinical diagnosis of} \\ & \text{either dyslipidemia, hypertension,} \\ & \text{or diabetes mellitus)} \\ & +0.871 \text{ (if currently on azathioprine} \\ & \text{or 6-mercaptopurine)} \\ & +1.794 \text{ (if ALT } 20\text{--}39.9 \text{ IU/L)} \\ & +2.213 \text{ (if ALT } \geq 40 \text{ IU/L)} \\ & +1.062 \text{ (if BMI } 27.5\text{--}34.9 \text{ kg/m}^2\text{)} \\ & +1.884 \text{ (if BMI } 35\text{--}37/49 \text{ kg/m}^2\text{)} \\ & +2.768 \text{ (if BMI } \geq 37.5 \text{ kg/m}^2\text{)} \\ & -3.336. \end{aligned}$$

The CPN-CD has good discrimination for CD subjects with NAFLD, with a C-statistic of 0.85. The agreement between the observed and predicted probability of NAFLD was excellent, with a calibration slope of 0.98, calibration intercept of -0.040, and a nonsignificant Hosmer and Lemeshow test (chi-square 10.86; df 8; $P = 0.21$). The model explained 47% of the variability for NAFLD by the Nagelkerke R^2 statistic. After performing 10-fold cross-validation, the average optimism of the C-statistic was 0.016, show that the CPN-CD's optimism correct C-statistic 0.83.

The $\text{CPN-CD}^{\text{logit}}$ was converted to predicted probability of NAFLD, with the diagnostic operating characteristic of the CPN-CD for each 10% increase in calculated risk shown in Table 3. This was used to stratify 3 clinically useful risk categories (low <20%, intermediate 20–49%, and high $\geq 50\%$). For example, at a threshold of <20% vs $\geq 20\%$ risk, 40% of patients would be at low risk and NAFLD could be excluded with a sensitivity of 86% and a negative predictive value of 86%. At a threshold of <50% vs $\geq 50\%$ risk, 27% of patients would be at high risk and NAFLD could be diagnosed with a specificity of 87% and a positive predictive value of 75%.

Due to the observational nature of this study, a liver biopsy had been obtained for only 5 patients at an average of 4 years before the MR. Although clearly underpowered, the CPN-CD predicted a high risk for NAFLD in the 3 patients with biopsy-proven NAFLD, whereas the CPN-CD predicted an intermediate risk for NAFLD for the 2 patients who had normal histology.

Comparison of the CPN-CD and the Hepatic Steatosis Index

The HSI was calculated in native, recalibrated, and revised forms (see Supplementary Table 1 for relevant risk equations and Supplementary Table 2 for their diagnostic operating characteristics). For example, almost half of all patients would be classified as high risk by the native his, but this had only 65% specificity and a 56% positive predictive value.

TABLE 1. Population Characteristics and Univariate Risk Factors

		NAFLD (n = 118)	Normal Liver (n = 193)	Odds Ratio (95% CI)	P
Demographics and cardiometabolic risk factors					
Age, y	<40	38 (32)	117 (61)	Ref	Ref
	40–59.9	56 (48)	49 (25)	3.5 (2.1–6.0)	<0.001
	≥60	24 (30)	27 (14)	2.7 (1.4–5.3)	0.003
Black/African American		4 (3)	19 (10)	0.3 (0.1–1.0)	0.035
Female		63 (53)	93 (48)	1.2 (0.8–1.9)	0.373
ALT, IU/L	<20	20 (17)	107 (55)	Ref	Ref
	20–39.9	65 (55)	65 (34)	5.4 (3.0–9.6)	<0.001
	≥40	33 (28)	21 (11)	8.4 (4.1–17)	<0.001
≥5% weight loss over last 12 mo		9 (8)	34 (18)	0.4 (0.1–0.8)	0.013
BMI, kg/m ²	<27.5	41 (35)	141 (73)	Ref	Ref
	27.5–34.9	53 (45)	46 (24)	4.0 (2.3–6.7)	<0.001
	35–37.49	8 (7)	3 (2)	9.2 (2.3–36)	0.002
	≥37.5	16 (14)	3 (2)	18.3 (5.1–66)	<0.001
Dyslipidemia		20 (17)	18 (9)	2.0 (1.0–3.9)	0.046
Diabetes mellitus		15 (13)	2 (1)	13.9 (3.1–62)	<0.001
Coronary artery disease		7 (6)	13 (7)	0.9 (0.3–2.2)	0.779
Hypertension		45 (38)	28 (15)	3.6 (2.1–6.3)	<0.001
Crohn's-related risk factors					
CD duration ≥15 y		59 (50)	60 (31)	2.2 (1.4–3.6)	0.001
Harvey Bradshaw Index (n = 251)	≤3	37 (39)	72 (46)	Ref	Ref
	4–7	37 (39)	55 (35)	1.3 (0.7–2.3)	0.359
	≥8	20 (21)	30 (19)	1.3 (0.7–2.6)	0.460
Total parental nutrition (within 6 mo)		1 (1)	2 (1)	0.8 (0.1–9.0)	1.000
Montreal location	Ileal (L1)	8 (7)	16 (8)	1.1 (0.4–3.1)	0.861
	Colon (L2)	17 (14)	31 (16)	Ref	Ref
	Ileocolonic (L3)	93 (80)	146 (76)	1.2 (0.5–3.1)	0.593
Upper GI involved (L4)		5 (4)	16 (8)	0.5 (0.2–1.4)	0.167
Montreal phenotype	Inflammatory	31 (31)	60 (26)	Ref	Ref
	Stricturing	43 (36)	67 (35)	1.2 (0.7–2.2)	0.462
	Penetrating	44 (37)	66 (34)	1.2 (0.7–2.3)	0.387
Perianal disease		32 (27)	50 (26)	1.1 (0.6–1.8)	0.814
Surgical history	No prior surgery	41 (35)	100 (51)	Ref	Ref
	Ileocolonic resection	43 (36)	51 (26)	2.1 (1.2–3.5)	0.009
	Small bowel resection only	24 (20)	25 (13)	2.3 (1.2–4.6)	0.013
	Colonic resection only	10 (9)	17 (9)	1.4 (0.6–3.4)	0.412
Biologic therapy	None	45 (38)	69 (36)	Ref	Ref
	Anti-TNF	48 (41)	82 (43)	0.9 (0.5–1.5)	0.682
	Anti-integrin	12 (10)	15 (8)	1.2 (0.5–2.9)	0.636
	Anti-IL 12/23	13 (11)	26 (14)	0.8 (0.4–1.6)	0.496
Immunomodulator	None	50 (42)	111 (58)	Ref	Ref
	Azathioprine/6-MP	51 (43)	59 (31)	1.9 (1.2–3.2)	0.011
	Methotrexate	17 (14)	23 (12)	1.6 (0.8–3.3)	0.172
Corticosteroids	None	108 (92)	163 (85)	Ref	Ref
	Prednisone	6 (5)	16 (8)	0.6 (0.2–1.5)	0.250
	Budesonide	4 (3)	14 (7)	0.4 (0.1–1.3)	0.147

P values represent chi-square test results for these categorical variables. Abbreviations: 6-MP, 6 mercaptopurine; IL, interleukin; TNF, tumor necrosis factor.

TABLE 2. Clinical Prediction Tool to Detect Nonalcoholic Fatty Liver Disease in Crohn's Disease (CPN-CD)

Predictor Variable	Beta Coefficient	Odds Ratio (95% CI)	P
CD duration ≥15 y	0.668	2.0 (1.1–3.6)	0.032
Black/African American	–1.996	0.1 (0.01–0.6)	0.007
Postmenopausal female	1.232	3.4 (1.5–8.1)	0.005
Any cardiometabolic complications	0.956	2.6 (1.3–5.1)	0.006
Azathioprine or 6-MP	0.871	2.4 (1.3–4.5)	0.006
ALT, IU/L			
<20	Ref	Ref	Ref
20–39.9	1.794	6.0 (2.9–12)	<0.001
≥40	2.213	9.1 (3.8–22)	<0.001
BMI, kg/m ²			
<27.5	Ref	Ref	Ref
27.5–34.9	1.062	2.9 (1.6–5.4)	0.001
35–37.4	1.884	6.6 (1.4–30)	0.015
≥37.5	2.768	16 (3.5–73)	<0.001
Constant	–3.336		

Postmenopausal female was defined by the interaction of age and sex variables with females ≥50 years (coded as “1”) compared with either males (coded as “0”) or females aged <50 years (coded as “0”). Any cardiometabolic complication was coded as “0” or “1” for patients who were already known to have a clinical diagnosis of hypertension, dyslipidemia, impaired fasting glucose, diabetes mellitus, or coronary artery disease by manual review of the patient's problem list (without using any specific a priori screening program or laboratory values).

Abbreviations: 6-MP, 6 mercaptopurine; CI, confidence interval.

TABLE 3. Operating Characteristics at Each Decile of Predicted Risk for the CPN-CD

Risk Threshold	Subjects Above Risk Threshold (n = 311), No. (%)	NAFLD (n = 118)	Normal Liver (n = 193)	Sensitivity, %	Specificity, %	PPV, %	NPV, %	Net Benefit (Screen at Risk Threshold)	Net Benefit (Screen All)
>10%	232 (75)	113	119	96	38	49	94	0.320	0.310
>20%	186 (60)	101	85	86	56	54	86	0.256	0.224
>30%	160 (51)	94	66	80	66	59	84	0.211	0.113
>40%	124 (40)	85	39	72	80	69	82	0.190	–0.034
>50%	105 (34)	79	26	67	87	75	81	0.170	–0.241
>60%	83 (27)	69	14	59	93	83	79	0.154	–0.551
>70%	61 (20)	55	6	47	97	90	75	0.132	–1.069
>80%	41 (13)	40	1	34	99.5	98	71	0.116	–2.103
>90%	18 (6)	18	0	15	100	100	66	0.058	–5.206

Net benefit is the difference between the true-positive count proportion and the false-positive count proportion * (risk probability/1-risk probability).

Abbreviations: NPV, negative predictive value; PPV, positive predictive value.

The C-statistic (0.85) for the CPN-CD is numerically higher than the native/recalibrated (0.76) and revised (0.78) HSI (Fig. 1), but this did not reach statistical significance. The CPN-CD had superior discrimination by the NRI, with all comparisons having a *P* value <0.001 (Table 4). For example, 28% of all patients with NAFLD were more accurately classified as intermediate or high risk by the CPN-CD than the recalibrated HSI.

By decision curve analysis, there is a higher net benefit using the CPN-CD to triage patients for screening compared with using the HSI (revised) to triage, screening only obese (BMI >30 kg/m²), or simply screening all patients throughout the entire range of predicted risk (Fig. 2).

Sensitivity Analysis

A total of 38 (12% of the entire population) patients had both NAFLD (PDFF >5.5%) and an elevated FIB-4 score (≥1.3), where further evaluation (eg, elastography) would routinely be indicated. The diagnostic operating characteristics per each decile of predicted risk are shown in Supplementary Table 3. For example, a CPN-CD predicting a low risk (<20%) of NAFLD would exclude this important subgroup with a sensitivity of 90%, whereas a CPN-CD predicting a high risk (>50%) of NAFLD would identify this subgroup with a specificity of 71%.

For this subpopulation, the C-statistic for the CPN-CD (0.78) was numerically higher than the native/recalibrated

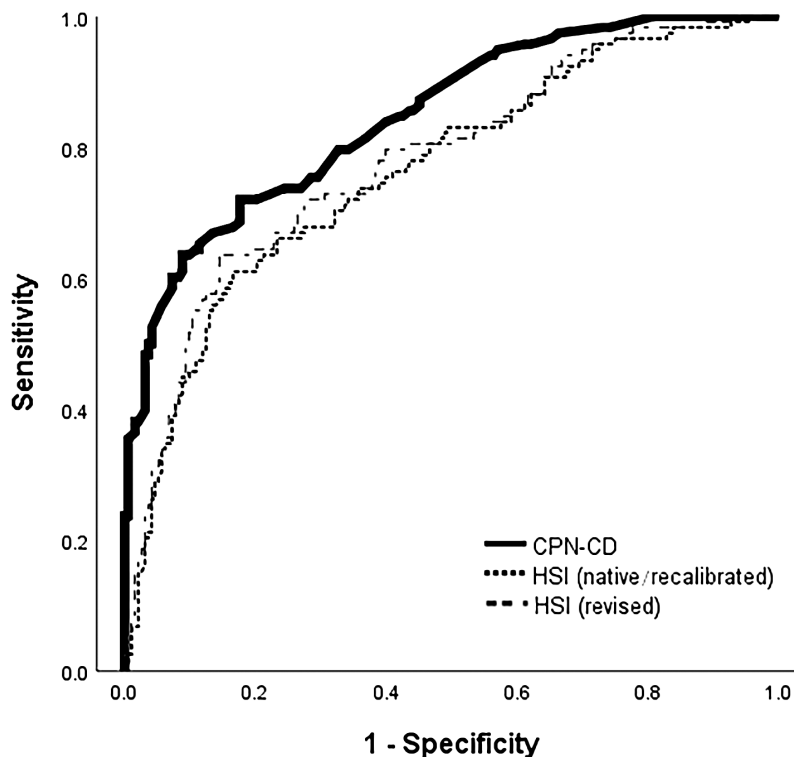


FIGURE 1. Receiver operating curve for CPN-CD and the HSI. The receiver operating curves for the risk model's native and recalibrated (where the beta coefficients are the same but with a different constant) versions are identical; however, the diagnostic operating characteristics differ due to risk category reclassification.

TABLE 4. CPN-CD Has Superior Discrimination Compared With the Hepatic Steatosis Index

	CPN-CD		HSI Index (Native)		HSI (Recalibrated)		HSI (Revised)		
C-statistic (95% CI)	0.848 (0.804–0.892)		0.763 (0.708–0.819)		0.763 (0.709–0.819)		0.779 (0.725–0.833)		
Net Reclassification Index	RI for NAFLD	Ref	Ref	–0.30	+0.08	–0.20	–0.28	–0.13	–0.19
	RI for normal		Ref		–0.38		+0.08		+0.06
	<i>P</i>		Ref		<0.001		<0.001		<0.001

The C-statistic represents the area under the receiver operating curve (plotting the sensitivity by the 1-specificity). The Net Reclassification Index is the sum of the proportion of patients with NAFLD who are reclassified correctly and the proportion of patients with a normal liver who are reclassified correctly. The NRI *P* value refers to the results of the McNemar test.

Abbreviations: CI, confidence interval; RI, reclassification improvement.

(0.62) and revised (0.67) HSI, but this did not reach statistical significance. The CPN-CD had statistically superior discrimination by the NRI approach (Supplementary Table 4). For example, 53% of patients with NAFLD with an elevated FIB-4 were more correctly classified as intermediate or high risk by the CPN-CD than the recalibrated HSI.

DISCUSSION

Here we report the derivation and internal validation of the Clinical Prediction Tool for NAFLD in Crohn's Disease

using the highly accurate magnetic resonance proton density fat fraction method as the means to diagnose hepatic steatosis. The model's discrimination remains very good and the predictors robust after internal validation; furthermore, this model significantly outperforms the Hepatic Steatosis Index in this patient population by the clinically relevant Net Reclassification Improvement measurement.

Although the CPN-CD has a numerically higher C-statistic (0.85), our study was underpowered to detect a statistically significant difference between either the native/recalibrated (0.76) or revised

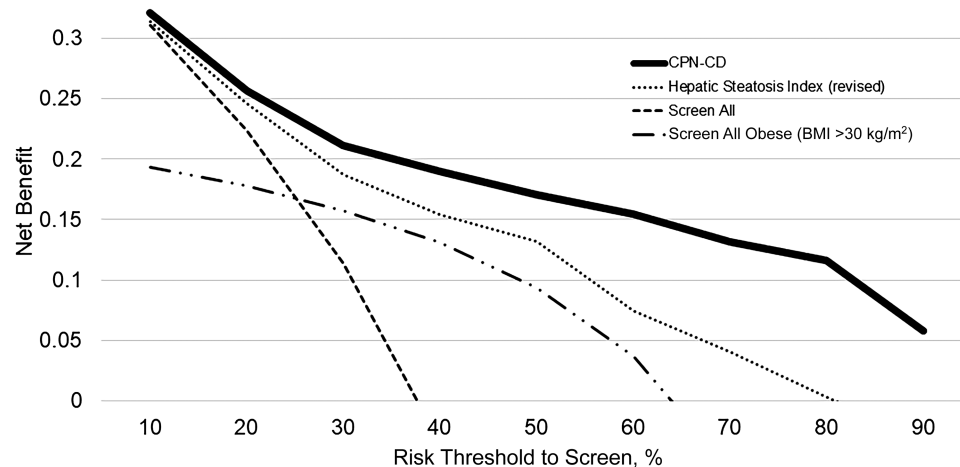


FIGURE 2. Decision curve analysis for CPN-CD, HSI, screening only obese patients, and screening all patients. Decision curve analysis is a method to assess the utility of a risk prediction tool to guide a screening program by plotting the risk threshold to screen against the net benefit. The net benefit of a risk model is equal to $\left(\frac{\text{True Positives}}{N}\right) - \left(\frac{\text{False Positives}}{N}\right) \times \left(\frac{\text{Risk Probability}}{1 - \text{Risk Probability}}\right)$.

(0.78) HSI. This conclusion is supported by the results of Bessissow et al., where the HSI had only fair discrimination in 62 IBD patients to detect ultrasonographic evidence of hepatic steatosis (0.74).²⁵

The CPN-CD was developed to inform the development of an NAFLD screening program in the Crohn's disease population. For example, in Crohn's patients with a risk of NAFLD $\geq 20\%$, the addition of PDFF sequences and elastography at the time of staging MR enterography has the potential to impact the care of these patients with only minimal additional cost/scanner time. Identification of otherwise healthy Crohn's patients with NAFLD should then prompt screening and treatment for prevalent cardiometabolic complications (diabetes mellitus, dyslipidemia, coronary artery disease), including weight loss strategies to mitigate additional incident disease. This is particularly relevant, as there appears to already be an increased risk of acute myocardial infarction,^{11,33} heart failure,¹¹ and diabetes mellitus³⁴ in patients with inflammatory bowel disease.

With regards to liver-related outcomes, identification of CD patients with NAFLD is also important, considering that the prevalence of hepatic fibrosis as detected by transient elastography has been reported to be as high as 12%.¹⁴ This prevalence agrees with our observation in this cohort using the FIB-4 score as a marker of hepatic fibrosis⁷; however, we acknowledge that the FIB-4 has not been validated against histology specifically in CD patients with NAFLD. Nevertheless, evidence of NAFLD with hepatic fibrosis would prompt referral to a hepatologist to help direct care and evaluation.

We are also aware of the limitations of these findings. Importantly, the CPN-CD is derived from a tertiary referral center of a primarily non-Hispanic white and black/African American population with small bowel and ileocolonic CD. The current findings will need to be examined in patients with ulcerative colitis, and those studies are currently in progress.

The limitations of race and ethnicity are also relevant, as the HSI was developed in a Korean population.²⁴ External validation to nontertiary centers with more straightforward disease is further warranted; however, it is important to note that many clinical features of CD were considered in this model without impacting the results. For example, the Montreal location (ileal, colonic, ileocolonic, upper gastrointestinal [GI]), Montreal phenotype (inflammatory, stenosing, penetrating), subjective clinical activity (Harvey Bradshaw Index), and biologic use did not appear to be risk factors in this population.

We acknowledge that the findings that some expected predictors were not acting as risk factors for NAFLD (eg, corticosteroids, methotrexate) could be a function of the study simply being underpowered; however, an alternative hypothesis is that they were not acting as risk factors because there is a unique pathology in this population. Although the scope of this manuscript was not to address the underlying pathophysiology, mechanistic studies to better explain the role of medications, enterokines, and dynamic fluctuations in gut inflammation are currently underway.

In conclusion, we have developed and internally validated the CPN-CD to identify patients with Crohn's disease who exhibit NAFLD. Future directions include external and outcome validation of the CPN-CD, while also performing cost-effectiveness analysis for a multitiered NAFLD screening program in this at-risk population.

SUPPLEMENTARY DATA

Supplementary data are available at *Inflammatory Bowel Diseases* online.

REFERENCES

1. Farraye FA, Melmed GY, Lichtenstein GR, et al. ACG clinical guideline: preventive care in inflammatory bowel disease. *Am J Gastroenterol*. 2017;112:241–258.

2. Singh S, Dulai PS, Zarrinpar A, et al. Obesity in IBD: epidemiology, pathogenesis, disease course and treatment outcomes. *Nat Rev Gastroenterol Hepatol*. 2017;14:110–121.
3. Moran GW, Dubeau MF, Kaplan GG, et al. The increasing weight of Crohn's disease subjects in clinical trials: a hypothesis-generating time-trend analysis. *Inflamm Bowel Dis*. 2013;19:2949–2956.
4. Chao CY, Battat R, Al Khoury A, et al. Co-existence of non-alcoholic fatty liver disease and inflammatory bowel disease: a review article. *World J Gastroenterol*. 2016;22:7727–7734.
5. McGowan CE, Jones P, Long MD, et al. Changing shape of disease: nonalcoholic fatty liver disease in Crohn's disease—a case series and review of the literature. *Inflamm Bowel Dis*. 2012;18:49–54.
6. Younossi Z, Anstee QM, Marietti M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. *Nat Rev Gastroenterol Hepatol*. 2018;15:11–20.
7. McHenry S, Sharma Y, Tirath A, et al. Crohn's disease is associated with an increased prevalence of nonalcoholic fatty liver disease: a cross-sectional study using magnetic resonance proton density fat fraction mapping. *Clin Gastroenterol Hepatol*. 2019;17:2816–2818.
8. Adams LA, Lymp JF, St Sauver J, et al. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. *Gastroenterology*. 2005;129:113–121.
9. Kanwal F, Kramer JR, Mapakshi S, et al. Risk of hepatocellular cancer in patients with non-alcoholic fatty liver disease. *Gastroenterology*. 2018;155:1828–1837.e2.
10. Franque SM, van der Graaff D, Kwanten WJ. Non-alcoholic fatty liver disease and cardiovascular risk: pathophysiological mechanisms and implications. *J Hepatol*. 2016;65:425–443.
11. Aniwaniwan S, Pardi DS, Tremaine WJ, et al. Increased risk of acute myocardial infarction and heart failure in patients with inflammatory bowel diseases. *Clin Gastroenterol Hepatol*. 2018;16:1607–1615.e1.
12. Yamazaki H, Tsuboya T, Tsuji K, et al. Independent association between improvement of nonalcoholic fatty liver disease and reduced incidence of type 2 diabetes. *Diabetes Care*. 2015;38:1673–1679.
13. Kim Y, Chang Y, Cho YK, et al. Obesity and weight gain are associated with progression of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2019;17:543–550.e542.
14. Saroli Palumbo C, Restellini S, Chao CY, et al. Screening for nonalcoholic fatty liver disease in inflammatory bowel diseases: a cohort study using transient elastography. *Inflamm Bowel Dis*. 2019;25:124–133.
15. Ryan CK, Johnson LA, Germin BI, et al. One hundred consecutive hepatic biopsies in the workup of living donors for right lobe liver transplantation. *Liver Transpl*. 2002;8:1114–1122.
16. Lee JY, Kim KM, Lee SG, et al. Prevalence and risk factors of non-alcoholic fatty liver disease in potential living liver donors in Korea: a review of 589 consecutive liver biopsies in a single center. *J Hepatol*. 2007;47:239–244.
17. Myers RP, Pollett A, Kirsch R, et al. Controlled Attenuation Parameter (CAP): a noninvasive method for the detection of hepatic steatosis based on transient elastography. *Liver Int*. 2012;32:902–910.
18. Eddowes PJ, Sasso M, Allison M, et al. Accuracy of fibroscan controlled attenuation parameter and liver stiffness measurement in assessing steatosis and fibrosis in patients with nonalcoholic fatty liver disease. *Gastroenterology*. 2019;156:1717–1730.
19. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of nonalcoholic fatty liver disease: practice guidance from the American Association for the Study of Liver Diseases. *Hepatology*. 2018;67:328–357.
20. Tang A, Tan J, Sun M, et al. Nonalcoholic fatty liver disease: MR imaging of liver proton density fat fraction to assess hepatic steatosis. *Radiology*. 2013;267:422–431.
21. Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging*. 2004;17:205–216.
22. Szczepaniak LS, Nurenberg P, Leonard D, et al. Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. *Am J Physiol Endocrinol Metab*. 2005;288:E462–E468.
23. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. *CJEM*. 2006;8:19–20.
24. Lee JH, Kim D, Kim HJ, et al. Hepatic steatosis index: a simple screening tool reflecting nonalcoholic fatty liver disease. *Dig Liver Dis*. 2010;42:503–508.
25. Bessissow T, Le NH, Rollet K, et al. Incidence and predictors of nonalcoholic fatty liver disease by serum biomarkers in patients with inflammatory bowel disease. *Inflamm Bowel Dis*. 2016;22:1937–1944.
26. Ruhl CE, Everhart JE. Fatty liver indices in the multiethnic United States National Health and Nutrition Examination Survey. *Aliment Pharmacol Ther*. 2015;41:65–76.
27. Steyerberg EW, Borsboom GJ, van Houwelingen HC, et al. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med*. 2004;23:2567–2586.
28. Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30:11–21.
29. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26:565–574.
30. Angulo P, Kleiner DE, Dam-Larsen S, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology*. 2015;149:389–97.e10.
31. Sterling RK, Lissen E, Clumeck N, et al; APRICOT Clinical Investigators. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology*. 2006;43:1317–1325.
32. Shah AG, Lydecker A, Murray K, et al; Nash Clinical Research Network. Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2009;7:1104–1112.
33. Choi YJ, Lee DH, Shin DW, et al. Patients with inflammatory bowel disease have an increased risk of myocardial infarction: a nationwide study. *Aliment Pharmacol Ther*. 2019;50:769–779.
34. Jess T, Jensen BW, Andersson M, et al. Inflammatory bowel disease increases risk of type 2 diabetes in a nationwide cohort study. *Clin Gastroenterol Hepatol*. In press.