Research Article

Laboratory based Non-Invasive Markers are Suboptimal in Detecting Advanced Fibrosis in Patients with Non-Alcoholic Steatohepatitis

Na Li[1*], Alexander Miller[2], Alice Hinton[3], Wei Chen[4], Khalid Mumtaz[1]

[1]Division of Gastroenterology, Hepatology and Nutrition, The Ohio State University Wexner Medical Center, Columbus, OH, USA

[2]Department of Internal Medicine, The Ohio State University Wexner Medical Center, Columbus, OH, USA

[3]Division of Biostatistics, College of Public Health, The Ohio State University, Columbus, OH, USA

[4]Department of Pathology, The Ohio State University Wexner Medical Center, Columbus, OH, USA

*Corresponding author: Na Li, Assistant Professor, Division of Gastroenterology, Hepatology and Nutrition, The Ohio State University Wexner Medical Center, 395 w 12th Ave, Rm 210F, Columbus, OH, 43210, USA

Abstract

Background and Aim: Hepatic fibrosis is a major determinant of clinical outcomes in patients with Non-Alcoholic Steatohepatitis (NASH). We aimed to investigate the diagnostic performance of non-invasive tests in detecting advanced fibrosis (F3-4) in a large NASH cohort from central Ohio, the United States.

Methods: Data of all patients with biopsy-proven NASH between 2014 and 2017 were collected. Diagnostic performance of aspartate aminotransferase (AST) to Platelets Ratio Index (APRI), Fibrosis-4 index (FIB-4) and NAFLD Fibrosis Score (NFS) were studied.

Results: A total of 284 NASH patients were included, 27.82% of whom had F3-4. The cohort was predominantly female (60.92%) and White (88.38%) with a mean age of 50±13 years. The most common comorbidities were obesity (77.11%) and type 2 diabetes (49.65%). There was a significant difference in NFS between fibrosis stage F0-2 and F3-4 (-0.43±1.99 and 0.30±2.28, p=0.01). The sensitivity of APRI <1, FIB-4 <1.3, NFS <-1.455 were 28%, 64%, and 73.33%, respectively. The specificity of APRI ≥2, FIB-4 ≥3.25, NFS ≥0.675 were 93.1%, 84.73%, 74.26%, respectively. The negative predictive value of all three models ranged between 72.59% and 77.72%, and the positive predictive values were consistently low (<40.38%). The area under receiver operator curves of APRI, FIB-4, and NFS were 0.52, 0.55, and 0.59, respectively. Diagnostic performance of these models appeared to be better in older (>35 year) and male population.

Conclusions: Overall APRI, FIB-4, NFS were suboptimal in detecting advanced fibrosis in our NASH cohort. Newer non-invasive tests with robust diagnostic accuracy are needed.

Introduction

Non-Alcoholic Fatty Liver Disease (NAFLD) is one of the most common chronic liver diseases around the world [1]. Non-alcoholic Steatohepatitis, a progressive form of NAFLD, promotes the development of liver fibrosis and cirrhosis. Multiple studies have demonstrated that stage of fibrosis is positively associated with all-cause and liver-related mortality [2,3].

Although liver biopsy remains the gold standard for staging of fibrosis, clinically it is not pragmatic nor necessary to biopsy

every patient with NASH. There are multiple reasons to consider a Non-Invasive Test (NIT) for diagnosing stage of fibrosis. These include improved patient's experience, reduced cost and biopsy-related complications, and improved access for point-of-care.

Current available NITs include laboratory-based scoring systems and imaging based testing such as elastography [4-8]. Fibrosis scoring models including AST to Platelets Ratio Index (APRI), Fibrosis-4 index (FIB-4) and NAFLD Fibrosis Score (NFS) have been shown to be potentially useful to rule out advanced fibrosis [4-7]. Area under Receiver Operator Curve (AUROC) for APRI, FIB-4, and NFS were reported between 0.77-0.84 [9]. However, many patients fall into the indeterminate zone for fibrosis assessment with the scoring models. Factors such as age, liver enzymes levels, prevalence of obesity, diabetes, and fibrosis may influence diagnostic accuracy of these scoring models [10,11]. In addition, different regions and practice (e.g. decision on liver biopsy) may also affect the sample selection of the NASH population. Imaging based tests such as elastography appears to be promising but are not readily available in primary care settings or small hospitals. Therefore, majority of facilities use laboratory based NITs despite their limitations.

The literature on the utility of NITs is growing all around the world. Majority of the reported studies are based on relatively small sample sizes and there is a need for larger studies on the utility of NITs for stage of fibrosis in NASH. It is also clinically relevant to test these NITs in region-specific NASH populations. Therefore, we aim to examine the diagnostic performance of three commonly used fibrosis scoring models including FIB-4, NFS, and APRI for advanced fibrosis in our NASH population from central Ohio, the United States.

Methods

This study was conducted at the Ohio State University, Wexner Medical Center (OSUWMC), Columbus, Ohio where patients from central Ohio are referred. We reviewed the records of all patients with biopsy-proven Steatohepatitis from 2014 to 2017. Patients who had history of excessive alcohol use or other competing liver etiologies were excluded. Excessive alcohol use among men was defined as consuming ≥21 standard drinks a week or ≥30 grams per day; and women consuming ≥14 drinks a week or ≥20 grams per day. Other liver etiologies including hepatitis B, hepatitis C, autoimmune hepatitis, hemochromatosis, alpha 1 antitrypsin deficiency, Wilson's disease, and history of liver transplant were excluded. We also excluded patients who had fatty liver disease due to chronic use of drugs (corticosteroids, methotrexate, tamoxifen), or total parenteral nutrition.

We collected clinical data including age, gender, race, body mass index, comorbidities (obesity, type 2 diabetes, dyslipidemia,

hypertension, hypothyroidism, obstructive sleep apnea, ischemic heart disease). We also collected information on history of bariatric surgery, history of alcohol use and smoking, and family history of liver and metabolic disorders. Laboratory data including Aspartate Aminotransferase (AST), Alanine Aminotransferase (ALT), total and indirect bilirubin, alkaline phosphatase, albumin, hemoglobin, white blood cell counts, platelet, creatinine, and International Normalized Ratio (INR) were collected closest to the visit for liver biopsy within 6 months window. Patients with more than 5% missing data were not included in analysis. These data included triglyceride, low-density lipoprotein, high-density lipoprotein, glucose, ferritin, iron saturation, anti-smooth antibody, and anti-mitochondrial antibody.

The Body Mass Index (BMI) was calculated using the formula: weight (kg)/height (m^2). The APRI was calculated as AST (U/L)/(upper limit of normal)/platelet count (x 10^9/L) x 100 [7]. The FIB-4 score was calculated according to the following formula: age x AST (U/L)/platelet count (x 10^9/L) x $\sqrt{\text{ALT}}$ (U/L) [4,5]. The NFS was calculated according to the following formula: -1.675 + 0.037 x age (years) + 0.094 x BMI (kg/m^2) + 1.13 x impaired fasting glycaemia or diabetes (yes=1, no=0) + 0.09 x AST/ALT ratio–0.013 x platelet (x 10^9/L)–0.06 x albumin (g/dL) [6]. We used literature-reported cut-offs of 1 and 2 for APRI, 1.3 and 3.25 for FIB-4, and -1.455 and 0.675 for NFS, respectively [5-7]. Specimens of liver pathology were fixed in formalin solution and stained with hematoxylin & eosin. Reticulin stain was used to assess stage of fibrosis. Mean length of liver biopsy sample was 20 mm with at least 11 portal tracts. All of the biopsies were reviewed by two experienced liver pathologists at the OSUWMC. Histological scoring of Nonalcoholic Steatohepatitis (NASH) and fibrosis were described according to the NAFLD Clinical Research Network criteria [12]. The Institutional Review Board of the OSUWMC approved the study.

Statistical Analysis

All statistical analyses were conducted using SAS 9.4 (SAS institute, Cary, NC). As the identification of patients with advanced fibrosis is of clinical importance, the patients were divided into two groups: patients with no/mild fibrosis (F0-2) and patients with advanced fibrosis (F3-4). Categorical variables were expressed as weighted frequency (percentage) and differences between groups were analyzed by χ^2 tests or Fisher exact tests in the case of small cell sizes. Continuous variables were expressed as mean ± SD and differences were analyzed with student's t tests or Wilcoxon rank-sum tests. Statistical significance was defined as p-value <0.05. The sensitivity, specificity, Positive Predictive Values (PPV), and Negative Predictive Values (NPV) for relevant cutoff values were calculated. AUROC with 95% Confidence Interval (CI) was calculated for each scoring model treated as a continuous variable.

Results

A total of 462 patients with liver biopsy-proven Steatohepatitis were identified at OSUWMC during the study period. After chart review, 284 patients met the inclusion criteria for NASH for analysis. Baseline characteristics of these patients are shown in Table 1. The mean age of patients was 50±13 years, mean BMI was 36.33±8.61 kg/m^2, and majority were females (60.92%) and White (88.38%). The most common comorbidity was obesity (77.11%), followed by type 2 diabetes (49.65%) and hypertension (37.68%). The prevalence of F0-2 and F3-4 was reported in 205 (72.18%) and 79 (27.82%) patients, respectively. Patients in the F0-2 group had higher platelet counts (215.42±78.03 vs 192.1±72.28 x10^9/L, p=0.02), lower serum glucose (138.15±63.48 vs 157.45±68.4 mg/dL, p=0.02) and lower INR (1.08±0.19 vs 1.16±0.35, p=0.01) as compared to patients in the F3-4 group. The mean NFS score for patients with F0-2 and F3-4 were -0.43±1.99 and 0.3±2.28, respectively, p=0.01. No significant differences in APRI and FIB-4 scores were found between the two groups.

	Full Cohort (n=284)		Fibrosis Stage 0-2 (n=205)		Fibrosis Stage 3-4 (n=79)		p-value
Demographics	*n*	*%*	*n*	*%*	*n*	*%*	
Age (year, mean, SD)	49.54	12.85	48.97	12.50	51.03	13.68	0.23
BMI (kg/m^2, mean, SD)	36.33	8.61	36.20	8.45	36.67	9.06	0.68
Gender							0.61
Female	173	60.92	123	60.00	50	63.29	
Male	111	39.08	82	40.00	29	36.71	
Race							0.08
White	251	88.38	184	89.76	67	84.81	
Black	21	7.39	11	5.37	10	12.66	
Other	12	4.23	10	4.88	2	2.53	
Comorbidities	*n*	*%*	*n*	*%*	*n*	*%*	
Obesity	219	77.11	157	76.59	62	78.48	0.73
Type 2 Diabetes	141	49.65	94	45.85	47	59.49	0.04
Dyslipidemia	61	21.48	41	20.00	20	25.32	0.33
Hypertension	107	37.68	72	35.12	35	44.30	0.15
Hypothyroidism	66	23.24	46	22.44	20	25.32	0.61
Sleep Apnea	30	10.56	20	9.76	10	12.66	0.48
Ischemic Heart Disease	20	7.04	15	7.32	5	6.33	0.77
Metabolic Disorder	78	27.46	51	24.88	27	34.18	0.12
Labs	*mean*	*SD*	*mean*	*SD*	*mean*	*SD*	
Bilirubin (mg/dL)	1.39	3.57	1.37	3.72	1.42	3.15	0.77
Albumin (g/dL)	4.00	0.67	4.04	0.67	3.88	0.67	0.05
Alkaline Phosphatase (U/L)	114.17	99.48	110.00	96.56	125.44	106.86	0.28
ALT (U/L)	82.51	73.66	87.51	78.88	69.07	55.59	0.05
AST (U/L)	66.01	53.77	68.79	55.07	58.54	49.69	0.05
Hemoglobin (g/dL)	13.34	1.90	13.45	1.87	13.08	1.98	0.14

	Full Cohort (n=284)		Fibrosis Stage 0-2 (n=205)		Fibrosis Stage 3-4 (n=79)		p-value
Platelet (x 10⁹/L)	208.97	77.07	215.42	78.03	192.10	72.28	0.02
Glucose (mg/dL)	143.48	65.31	138.15	63.48	157.45	68.40	0.02
Creatinine (mg/dL)	0.91	0.52	0.90	0.53	0.95	0.49	0.60
INR	1.10	0.25	1.08	0.19	1.16	0.35	0.01
Scores	*mean*	*SD*	*mean*	*SD*	*mean*	*SD*	
APRI	0.99	1.61	1.06	1.74	0.96	1.22	0.70
FIB-4	2.33	2.74	2.22	2.82	2.60	2.50	0.19
NFS	-0.22	2.10	-0.43	1.99	0.30	2.28	0.01
NASH: Non-Alcoholic Steatohepatitis; BMI: Body Mass Index; ALT: Alanine Aminotransferase; AST: Aspartate Aminotransferase; INR: International Normalized Ratio; APRI: AST to Platelets Ratio Index; FIB-4: Fibrosis-4 Index; NFS: Non-Alcoholic Fatty Liver Disease Fibrosis Score; SD: Standard Deviation							

Table 1: Baseline characteristics of the NASH patient cohort overall and by fibrosis stage (F0-2 vs F3-4).

The sensitivity, specificity, PPV, NPV to predict stage F3-4 fibrosis are shown in Table 2. We found that APRI with cutoffs of 1 and 2 had specificity of 70.44% & 93.10%, respectively but extremely low sensitivity (<28%). The sensitivity of FIB-4 <1.3 and NFS <-1.455 were 64% and 73.33%, respectively, with an unacceptably low specificity (<44.33%). The specificity of FIB-4 ≥3.5 and NFS ≥0.675 were 84.73% and 74.26%, respectively with low sensitivity (<42.67%). The NPVs for all three models ranged from 72.59% to 77.72%. The PPVs were consistently poor (<40.38%). AUROCs for APRI, FIB-4, and NFS were 0.52 (95% CI: 0.44, 0.60), 0.55 (95% CI: 0.47, 0.63), and 0.59 (95% CI: 0.52, 0.67), respectively in our cohort.

	F0-2	F3-4	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)
APRI			28.00%	70.44%	25.93%	72.59%	0.52 (0.44, 0.60)
<1	143	54					
≥1	60	21					
APRI			8.00%	93.10%	30.00%	73.26%	
≤2	189	69					
>2	14	6					
FIB-4			64.00%	44.33%	29.81%	76.92%	0.55 (0.47, 0.63)
<1.3	90	27					
≥1.3	113	48					
FIB-4			28.00%	84.73%	40.38%	76.11%	
≤3.25	172	54					
>3.25	31	21					
NFS			73.33%	29.21%	27.78%	74.68%	0.59 (0.52, 0.67)
<-1.455	59	20					
≥-1.455	143	55					

NFS			42.67%	74.26%	38.10%	77.72%	
≤0.675	150	43					
>0.675	52	32					
NASH: Non-Alcoholic Steatohepatitis; PPV: Positive Predictive Value; NPV: Negative Predictive Value; CI: Confidence Interval; APRI: Aspartate Aminotransferase to Platelets Ratio Index; FIB-4: Fibrosis-4 Index; NFS: Non-Alcoholic Fatty Liver Disease Fibrosis Score							

Table 2: Sensitivity, specificity, and positive and negative predictive values for identifying NASH patients with stage F3-4 fibrosis.

Subgroup analysis

We performed various sub-group analysis to identify a group of patients who may benefit more from NITs. To examine the impact of age on diagnostic performance of APRI, FIB-4, and NFS, we divided the patients into groups of age 18-35 years (n=41, 14.75%), 36-64 years (n=202, 72.66%), and ≥65 years (n=35, 12.59%). Advanced fibrosis (F3-4) was present in 24.39%, 26.24%, and 34.29% patients with age 18-35 years, 36-64 years, and ≥65 years, respectively. NFS between F0-2 and F3-4 in the three age groups were -1.59±2.32 and -1.63±1.61 (p=0.96), -0.38±1.86 and 0.22±2.09 (p=0.05), and 0.80±1.55 and 2.45±1.94 (p=0.01), respectively. No significant differences in APRI or FIB-4 scores were noted between F0-2 and F3-4 among any of the age groups. AUROCs of these scoring models increased with age particularly NFS showing 0.45 (95% CI: 0.25, 0.66), 0.58 (95% CI: 0.48, 0.67), and 0.74 (95% CI: 0.57, 0.91) in ages 18-35, 36-64, and ≥65, respectively (Table 3). Sensitivity, specificity, PPV, and NPV of APRI, FIB-4, and NFS are shown in Table 3. Overall, all three models had good specificity with high cutoff values (>90%) for identifying F3-4 fibrosis in NASH patients younger than 35 but had poor sensitivity (<50%). As age advances, there was improved test sensitivity at the cost of lower specificity.

Age 18-35 years	F0-2	F3-4	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)
APRI			20.00%	67.74%	16.67%	72.41%	0.36 (0.15, 0.57)
<1	21	8					
≥1	10	2					
APRI			10.00%	93.55%	33.33%	76.32%	
≤2	29	9					
>2	2	1					
FIB-4			20.00%	83.87%	28.57%	76.47%	0.39 (0.16, 0.62)
<1.3	26	8					
≥1.3	5	2					
FIB-4			10.00%	93.55%	33.33%	76.32%	
≤3.25	29	9					
>3.25	2	1					
NFS			50.00%	56.67%	27.78%	77.27%	0.45 (0.25, 0.66)
<-1.455	17	5					
≥-1.455	13	5					
NFS			10.00%	90.00%	25.00%	75.00%	
≤0.675	27	9					
>0.675	3	1					

Age 36-64 years	F0-2	F3-4	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)
APRI			30.19%	71.81%	27.59%	74.31%	0.50 (0.41, 0.59)
<1	107	37					
≥1	42	16					
APRI			9.43%	91.95%	29.41%	74.05%	
≤2	137	48					
>2	12	5					
FIB-4			64.15%	40.94%	27.87%	76.25%	0.54 (0.45, 0.64)
<1.3	61	19					
≥1.3	88	34					
FIB-4			26.42%	86.58%	41.18%	76.79%	
≤3.25	129	39					
>3.25	20	14					
NFS			71.70%	26.85%	25.85%	72.73%	0.58 (0.48, 0.67)
<-1.455	40	15					
≥-1.455	109	38					
NFS			39.62%	74.50%	35.59%	77.62%	
≤0.675	111	32					
>0.675	38	21					
Age ≥65 years	F0-2	F3-4	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)
APRI			25.00%	65.22%	27.27%	62.50%	0.54 (0.33, 0.75)
<1	15	9					
≥1	8	3					
FIB-4			100.00%	13.04%	37.50%	100.00%	0.58 (0.38, 0.78)
<1.3	3	0					
≥1.3	20	12					
FIB-4			50.00%	60.87%	40.00%	70.00%	
≤3.25	14	6					
>3.25	9	6					
NFS			100.00%	8.70%	36.36%	100.00%	0.74 (0.57, 0.91)
<-1.455	2	0					
≥-1.455	21	12					
NFS			83.33%	52.17%	47.62%	85.71%	
≤0.675	12	2					

>0.675	11	10					
NASH: Non-Alcoholic Steatohepatitis; PPV: Positive Predictive Value; NPV: Negative Predictive Value; CI: Confidence Interval; APRI: Aspartate Aminotransferase to Platelets Ratio Index; FIB-4: Fibrosis-4 Index; NFS: Non-Alcoholic Fatty Liver Disease Fibrosis Score. No patients had APRI >2 in age ≥65 years group.							

Table 3: Sensitivity, specificity, and positive and negative predictive values of APRI, FIB-4, and NFS for identifying NASH patients with stage F3-4 fibrosis among three age groups.

We also analyzed the diagnostic performance of APRI, FIB-4, and NFS based on normal vs elevated ALT (women: ≤30U/L, men: ≤45U/L) (Supplementary Table 1). No significant differences of scores were found between F0-2 and F3-4 for each NIT model. AUROC of NFS was 0.61 (95% CI: 0.46, 0.75) in patients with normal ALT and 0.53 (95% CI: 0.43, 0.63) in patients with elevated ALT. AUROCs of APRI and FIB-4 were similar in patients with normal and elevated ALT ranging 0.51-0.56. NPVs of APRI, FIB-4, and NFS were approximately 10% higher in patients with elevated ALT compared to patients with normal ALT.

Normal ALT	F0-2	F3-4	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)
APRI			16.67%	82.93%	41.67%	57.63%	0.51 (0.38, 0.65)
<1	34	25					
≥1	7	5					
APRI			3.33%	97.56%	50.00%	57.97%	
≤2	40	29					
>2	1	1					
FIB-4			70.00%	31.71%	42.86%	59.09%	0.56 (0.43, 0.70)
<1.3	13	9					
≥1.3	28	21					
FIB-4			40.00%	78.05%	57.14%	64.00%	
≤3.25	32	18					
>3.25	9	12					
NFS			83.33%	19.51%	43.10%	61.54%	0.61 (0.46, 0.75)
<-1.455	8	5					
≥-1.455	33	25					
NFS			63.33%	58.54%	52.78%	68.57%	
≤0.675	24	11					
>0.675	17	19					
Elevated ALT	**F0-2**	**F3-4**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC (95% CI)**
APRI			35.56%	67.28%	23.19%	78.99%	0.52 (0.43, 0.62)
<1	109	29					
≥1	53	16					
APRI			11.11%	91.98%	27.78%	78.84%	

≤2	149	40					
>2	13	5					
FIB-4			60.00%	47.53%	24.11%	81.05%	0.52 (0.43, 0.62)
<1.3	77	18					
≥1.3	85	27					
FIB-4			20.00%	86.42%	29.03%	79.55%	
≤3.25	140	36					
>3.25	22	9					
NFS			66.67%	31.68%	21.43%	77.27%	0.53 (0.43, 0.63)
<-1.455	51	15					
≥-1.455	110	30					
NFS			28.89%	78.26%	27.08%	79.75%	
≤0.675	126	32					
>0.675	35	13					
NASH: Non-Alcoholic Steatohepatitis; ALT: Alanine Aminotransferase; PPV: Positive Predictive Value; NPV: Negative Predictive Value; CI: Confidence Interval; APRI: Aspartate Aminotransferase to Platelets Ratio Index; FIB-4: Fibrosis-4 Index; NFS: Non-Alcoholic Fatty Liver Disease Fibrosis Score. Elevated ALT is defined as Women ≤30U/L, Men ≤45U/L.							

Supplementary Table 1: Sensitivity, specificity, and positive and negative predictive values for identifying NASH patients with stage F3-4 fibrosis based on normal and elevated ALT.

In addition, we analyzed the impact of gender on the diagnostic performance of APRI, FIB-4, and NFS for predicting advanced fibrosis. NFS between F0-2 group and F3-4 group were -0.91 ± 1.90 and 0.26 ± 2.62 in men (p=0.01), and -0.11 ± 2.00 and 0.37 ± 2.06 in women (p=0.17), respectively (Table 4). No significant differences of APRI or FIB-4 scores were found between the two groups based on gender. AUROC of NFS was higher in men (0.65, 95% CI: 0.52, 0.78) compared to that in women (0.55, 95% CI: 0.46-0.65). AUROCs of APRI and FIB-4 for men and women were similar. NPV of NFS at cutoff ≤0.675 to rule out F3-4 fibrosis was slightly better in men (82.28%) compared to that in women (74.56%). Similarly, FIB-4 also had higher NPV at cutoff 1.3 in men (80.43%) than that in women (74.65%).

Men	**F0-2**	**F3-4**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC (95% CI)**
APRI			27.59%	70.37%	25.00%	73.08%	0.54 (0.42, 0.66)
<1	57	21					
≥1	24	8					
APRI			3.45%	92.59%	14.29%	72.82%	
≤2	75	28					
>2	6	1					
FIB-4			68.97%	45.68%	31.25%	80.43%	0.59 (0.46, 0.72)
<1.3	37	9					
≥1.3	44	20					

			Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)
FIB-4			31.03%	85.19%	42.86%	77.53%	
≤3.25	69	20					
>3.25	12	9					
NFS			68.97%	41.25%	29.85%	78.57%	0.65 (0.52, 0.78)
<-1.455	33	9					
≥-1.455	47	20					
NFS			51.72%	81.25%	50.00%	82.28%	
≤0.675	65	14					
>0.675	15	15					
Women	**F0-2**	**F3-4**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC (95% CI)**
APRI			28.26%	70.49%	26.53%	72.27%	0.51 (0.41, 0.61)
<1	86	33					
≥1	36	13					
APRI			10.87%	93.44%	38.46%	73.55%	
≤2	114	41					
>2	8	5					
FIB-4			60.87%	43.44%	28.87%	74.65%	0.53 (0.42, 0.63)
<1.3	53	18					
≥1.3	69	28					
FIB-4			26.09%	84.43%	38.71%	75.18%	
≤3.25	103	34					
>3.25	19	12					
NFS			76.09%	21.31%	26.72%	70.27%	0.55 (0.46, 0.65)
<-1.455	26	11					
≥-1.455	96	35					
NFS			36.96%	69.67%	31.48%	74.56%	
≤0.675	85	29					
>0.675	37	17					

NASH: Non-Alcoholic Steatohepatitis; PPV: Positive Predictive Value; NPV: Negative Predictive Value; CI: Confidence Interval; APRI: Aspartate Aminotransferase to Platelets Ratio Index; FIB-4: Fibrosis-4 Index; NFS: Non-Alcoholic Fatty Liver Disease Fibrosis Score

Table 4: Sensitivity, specificity, and positive and negative predictive values for identifying NASH patients with stage F3-4 fibrosis between men and women.

Discussion

With the enormous global prevalence of NAFLD, it is imperative to develop non-invasive diagnostic tools to identify high-risk population. There are multiple studies addressing the role of laboratory-based scoring models for assessment of fibrosis stage especially advanced fibrosis in patients with NASH [5,6,13-15]. Majority of these studies are small comprising of sample size less than 200 [9]. The diagnostic performance of APRI, FIB-4, and NFS in our NASH cohort from central Ohio is consistent with but lower than other large studies [5,6,9,13]. The PPVs are consistently poor to detect advanced fibrosis in our cohort. The NPVs are acceptable but unsatisfactory around 75% for all three models regardless of previously published cutoff value used. NFS appears to have better diagnostic performance compared to FIB-4 or APRI.

The largest cohort reported the diagnostic performance of NITs is from the global phase 3 trials (STELLAR) including 3,202 biopsy-proven NASH patients 14. These trials were designed to include patients with significant fibrosis. The prevalence of F3-4 fibrosis in this cohort was 70.60% compared to average 24% in other large studies, and 27.82% in our study. This likely contributes to their high PPVs of NFS and FIB-4 (around 97%) due to increased pre-test probability but at the cost of lower NPVs (around 68%). In contrast, other studies including ours demonstrate higher NPVs of these NITs, indicating the clinical value of ruling out rather than ruling in the diagnosis of advanced fibrosis [5,6,13]. This is probably also true in the studies from communities where the estimated prevalence of advanced fibrosis is even lower than tertiary medical centers.

Most of the current NITs are developed in the NASH patient population between ages of 35 and 65 years. McPherson and others studied the effect of age on the performance of NITs for advanced fibrosis 10. In their study, the diagnostic accuracy of NFS and FIB-4 were low in patients younger than ≤35 years with AUROCs of 0.52 and 0.60, respectively, but improved with advancing age (0.81 in patients ≥65 years). Our study showed similar findings particularly with NFS. However, AUROCs are consistently lower for all three models. It is worth mentioning that our patient cohort included higher percentages of patients in both age groups 18-35 years (24%) and ≥65 years (31%) compared to those in McPherson's study, which reported approximately 11% in each group.

Our patient cohort is female-predominant which is different from the majority of other cohorts [5,6,13-15]. Therefore, we also analyzed the diagnostic performance of APRI, FIB-4, and NFS for advanced fibrosis based on gender. Interestingly, NFS demonstrated better performance in men (AUROC 0.65) than in women (AUROC 0.55) while FIB-4 and APRI were similar between genders. The impact of gender on performance of NITs has not been reported previously. The underlying reason remains unclear and could be related to the gender differences on NASH development and progression. Recent meta-analysis showed that women have a lower risk of non-alcoholic fatty liver disease, but a higher risk of advance fibrosis than men, especially after age 50 years [16]. In addition, differences may exist between genders regarding laboratory values and NASH related comorbidities such as diabetes and obesity [17,18].

A few factors may potentially contribute to overall lower diagnostic performance of NITs in our study compared to others in the literature. There are differences in the distribution of studied population including age, gender, comorbidities, and prevalence of fibrosis stages. This may suggest a regional difference of NASH populations. Selection bias may exist towards patients who undergo liver biopsy. Factors that could affect the pursuit of liver biopsy include local practice patterns, indications of biopsy, comorbid conditions, availability of treatment such as clinical trials in the local area, etc. In addition, substantial (~40%) sampling error may occur with biopsy that can result in disease severity being misclassified [19].

Our study has a few limitations. First, patients included in the study are from a tertiary academic center in Central Ohio. Therefore, the present findings may not be generalizable to other NASH populations. Second, this is a retrospective study and all data are collected from medical records. Given the significant difference of the diagnostic performance of NITs between our study and other published studies, we made every effort to ensure accurate data collection. All patient records were reviewed by two study authors separately. This has reduced our sample size by 57 cases without any major changes in the findings. Third, this is a cross-sectional study with laboratory test results collected closest to the time of liver biopsy within a six-month window period. We know that the laboratory test results that are used to calculate the fibrosis scores may fluctuate over time. Longitudinal studies assessing the value of these scoring models are needed to determine their utility in clinical practice. Despite these limitations, our study is in parallel with the other studies demonstrating suboptimal performance of these laboratory-based NITs, probably more useful ruling out rather than ruling in advanced fibrosis. Combinations or sequential use of other NITs particularly elastography has been suggested to improve the diagnostic value of NITs for advanced fibrosis [14,20,21].

In summary, the diagnostic accuracy of APRI, FIB-4, and NFS is suboptimal to predict advanced hepatic fibrosis in our NASH patient cohort from central Ohio, the United States. NFS has relatively better performance than FIB-4 or APRI. Age and gender appear to be affecting factors on performance besides regional differences of NASH populations. Clinicians should be aware of

the limitations of current NITs and apply them to clinical practice appropriately. There is a need for further studies to develop strong NITs to detect advanced fibrosis in patients with NASH.

References

1. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, et al. (2016) Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology 64: 73-84.

2. Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, et al. (2015) Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology 149: 389-397.e10.

3. Taylor RS, Taylor RJ, Bayliss S, Hagstrom H, Nasr P, et al. (2020) Association Between Fibrosis Stage and Outcomes of Patients With Nonalcoholic Fatty Liver Disease: A Systematic Review and Meta-Analysis. Gastroenterology 158: 1611-1625.e12.

4. Sterling RK, Lissen E, Clumeck N, Sola R, Correa MC, et al. (2006) Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 43: 1317-1325.

5. Shah AG, Lydecker A, Murray K, Tetri BN, Contos MJ, et al. (2009) Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 7: 1104-1112.

6. Angulo P, Hui JM, Marchesini G, Bugianesi E, George J, et al. (2007) The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. Hepatology 45: 846-854.

7. Wai CT, Greenson JK, Fontana RJ, Kalbfleisch JD, Marrero JA, et al. (2003) A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. Hepatology 38: 518-526.

8. Boursier J, Vergniol J, Guillet A, Hiriart JB, Lannes A, et al. (2016) Diagnostic accuracy and prognostic significance of blood fibrosis tests and liver stiffness measurement by FibroScan in non-alcoholic fatty liver disease. J Hepatol 65: 570-578.

9. Xiao G, Zhu S, Xiao X, Yan L, Yang J, et al. (2017) Comparison of laboratory tests, ultrasound, or magnetic resonance elastography to detect fibrosis in patients with nonalcoholic fatty liver disease: A meta-analysis. Hepatology 66: 1486-1501.

10. McPherson S, Hardy T, Dufour JF, Petta S, Romero-Gomez M, et al. (2017) Age as a Confounding Factor for the Accurate Non-Invasive Diagnosis of Advanced NAFLD Fibrosis. Am J Gastroenterol 112: 740-751.

11. McPherson S, Anstee QM, Henderson E, Day CP, Burt AD (2013) Are simple noninvasive scoring systems for fibrosis reliable in patients with NAFLD and normal ALT levels? Eur J Gastroenterol Hepatol 25: 652-658.

12. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, et al. (2005) Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 41: 1313-1321.

13. McPherson S, Stewart SF, Henderson E, Burt AD, Day CP (2010) Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. Gut 59: 1265-1269.

14. Anstee QM, Lawitz EJ, Alkhouri N, Wong VWS, Romero-Gomez M, et al. (2019) Noninvasive Tests Accurately Identify Advanced Fibrosis due to NASH: Baseline Data From the STELLAR Trials. Hepatology 70: 1521-1530.

15. Sumida Y, Yoneda M, Hyogo H, Itoh Y, Ono M, et al. (2012) Validation of the FIB4 index in a Japanese nonalcoholic fatty liver disease population. BMC Gastroenterol 12: 2.

16. Balakrishnan M, Patel P, Dunn-Valadez S, Dao C, Khan V, et al. (2021) Women Have a Lower Risk of Nonalcoholic Fatty Liver Disease but a Higher Risk of Progression vs Men: A Systematic Review and Meta-analysis. Clin Gastroenterol Hepatol 19: 61-71.e15.

17. Tramunt B, Smati S, Grandgeorge N, Lenfant F, Arnal JF, et al. (2020) Sex differences in metabolic regulation and diabetes susceptibility. Diabetologia 63: 453-461.

18. Nevill AM, Metsios GS (2015) The need to redefine age- and gender-specific overweight and obese body mass index cutoff points. Nutr Diabetes 5: e186.

19. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, et al. (2005) Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology 128: 1898-1906.

20. Petta S, Vanni E, Bugianesi E, Di Marco V, Camma C, et al. (2015) The combination of liver stiffness measurement and NAFLD fibrosis score improves the noninvasive diagnostic accuracy for severe liver fibrosis in patients with nonalcoholic fatty liver disease. Liver Int 35: 1566-1573.

21. Petta S, Wong VW, Camma C, Hiriart JB, Wong GLH, et al. (2017) Serial combination of non-invasive tools improves the diagnostic accuracy of severe liver fibrosis in patients with NAFLD. Aliment Pharmacol Ther 46: 617-627.