REVIEW ARTICLE

Histological assessment based on liver biopsy: the value and challenges in NASH drug development

Xiao-fei Tong[1], Qian-yi Wang[1], Xin-yan Zhao[1], Ya-meng Sun[1], Xiao-ning Wu[1], Li-ling Yang[1], Zheng-zhao Lu[1], Xiao-juan Ou[1], Ji-dong Jia[1] and Hong You[1]

Nonalcoholic steatohepatitis (NASH) is increasingly recognized as a serious disease that can lead to cirrhosis, hepatocellular carcinoma (HCC), and death. However, there is no effective drug to thwart the progression of the disease. Development of new drugs for NASH is an urgent clinical need. Liver biopsy plays a key role in the development of new NASH drugs. Histological findings based on liver biopsy are currently used as the main inclusion criteria and the primary therapeutic endpoint in NASH clinical trials. However, there are inherent challenges in the use of liver biopsy in clinical trials, such as evaluation reliability, sampling error, and invasive nature of the procedure. In this article, we review the advantages and value of liver histopathology based on liver biopsy in clinical trials of new NASH drugs. We also discuss the challenges and limitations of liver biopsy and identify future drug development directions.

Keywords: nonalcoholic steatohepatitis; liver biopsy; value and challenges; new drug development

INTRODUCTION

Nonalcoholic fatty liver disease (NAFLD) is a chronic liver disease related to metabolic disorders. The estimated global prevalence of NAFLD in the general population is ~25% [1]. Nonalcoholic steatohepatitis (NASH) is a serious form of NAFLD [2]. Although it has an insidious onset, it can eventually progress to potentially fatal condition, such as decompensated cirrhosis, hepatocellular carcinoma, and necessitate liver transplantation [3, 4]. The estimated global prevalence of NASH is 3%–5%, and the complications related to NASH impose a considerable morbidity and mortality burden [5, 6]. NAFLD is associated with an increased risk of cardiovascular disease, given the close association of NAFLD with metabolic syndrome, including obesity, hypertension, atherogenic dyslipidemia, and insulin resistance [7, 8]. In a recent study of biopsy-proven NAFLD patients, the fibrosis score was the only independent predictor of cardiovascular disease [9]. NAFLD is associated with an increased risk of developing extrahepatic cancers [10, 11]. In a Swedish biopsy-confirmed NAFLD cohort, patients had significantly increased overall mortality (15.3%), which was due primarily to the increased incidence of cancer [12].

Currently, there are no effective drugs approved for the treatment of NASH [13]. Therefore, research and development of novel drugs for treatment of NASH is a key imperative to reduce the disease burden and decrease the incidence of severe complications in the late stage of NASH. Although several drugs with different mechanisms have completed phase IIb/III clinical trials, the reported efficacy has been limited [14–17]. Owing to the complex pathogenetic mechanism of NASH and the considerable heterogeneity, histological evaluation based on liver biopsy is necessary for distinguishing NASH from NAFLD. Moreover, histological evaluation plays an important role in the research for development of new drugs for NASH [13]. To achieve a breakthrough in the development of NASH therapeutic drugs, the role of liver biopsy needs to be emphasized in the trial design.

Percutaneous fine-needle liver biopsy is a simple, reliable procedure that was introduced into clinical practice in the late 1950s [18]. Liver biopsy is considered as the "gold standard" for assessment of hepatic disease. The American Association for the Study of Liver Disease considers liver biopsy for establishing a diagnosis, staging of underlying liver disease, and directing management based on the underlying histology [19]. In the clinical practice and scientific research of NASH, liver biopsy provides much more essential information than the mere presence or absence of NASH. Liver biopsy remains the only diagnostic modality for identification of the cause, assessment of the classification, and prognostic evaluation.

Liver biopsy has been widely applied in the development of new NASH drugs. From the early PIVENS and FLINT studies to the several phase IIb/III studies carried out in recent years, liver biopsy has been used for patient screening and evaluation of therapeutic efficacy, which is a key link in the development of new drugs [20, 21]. The U.S. Food and Drug Administration (FDA), the European Medicines Agency (EMA), and the Chinese National Medical Products Administration (NMPA) recommend liver histological evaluation as the main inclusion criterion and primary efficacy/endpoint in the new NASH drug development. However, the failure of several phase III trials has raised concerns about the possible limitations of pathological assessment based on liver

biopsy in clinical trials and ways to improve the situation. In this article, we review the value of liver biopsy in the development of new NASH drugs and the associated challenges.

THE VALUE OF LIVER BIOPSY IN THE NEW NASH DRUG DEVELOPMENT

Liver biopsy used as NASH inclusion criteria in new drug development

Liver biopsy is recommended by drug regulatory authorities as the inclusion criteria for new NASH drug development. In1980, Ludwig et al. proposed the concept of NASH for the first time [22]. To date, histology is still an indispensable method to distinguish NASH from NAFLD. The main histologic features of NASH include steatosis, lobular inflammation, and ballooning hepatocytes [23]. The typical pattern of NASH fibrosis consists of delicate strands of collagen around the sinusoids and hepatocytes, which usually present as a "chicken wire" pattern involving the centrilobular region.

The first proposed grading and staging system for lesions of NASH in adults was the "Brunt system". The system enables semiquantitative evaluation of the typical pathological characteristics of NASH. Based on a comprehensive evaluation, the lesions are classified as mild, moderate, or severe [24]. NAFLD Activity Score (NAS) was proposed by NASH Clinical Research Net (NASH-CRN) in 2005 and is being currently used as the criteria for pathological diagnosis of NASH in clinical trials [25]. The NAS scoring system semiquantitatively evaluates steatosis (0–3), lobular inflammation (0–3), and ballooning (0–2). According to the NAS score, the disease can be classified into three categories: NASH, not NASH, and borderline. This system can be applied to the entire spectrum of NAFLD and can also be used in pediatric NALFD [26, 27]. The SAF/FLIP system was developed by the European FLIP Pathology Consortium and is currently used for clinical diagnosis [28, 29]. This system also covers the entire disease spectrum of NAFLD, however, it cannot be used in children. The SAF/FLIP and NASH-CRN system are consistent for the evaluation of steatosis (0–3) and fibrosis (F0–F4). However, the diagnosis criteria for lobular inflammation (0–2 in the SAF system) and ballooning (0–2 in the SAF system) are slightly different. In the SAF/FLIP system, the combination of the semi-quantitative values of these two components is regarded as "activity score"; therefore, a subscore is awarded for each component of the SAF (steatosis + activity + fibrosis), which is the primary advantage of this system [30].

Based on the important role of liver biopsy in the diagnosis of NASH, drug regulatory authorities (including the FDA, EMA, and NMPA) have approved histology as the main inclusion criteria for phase II/III clinical trials of NASH [31–33]. In 2018, the FDA issued a draft guideline for developing therapeutic drugs for NASH [31]. For the patient inclusion criteria of phase III development, FDA has accepted NASH activity score (NAS) ≥ 4 with at least 1 point each for inflammation and ballooning along with a NASH-CRN fibrosis score >stage 1 fibrosis but <stage 4 fibrosis. To determine the use of liver

biopsy findings as the entry criteria in the different phases of trials and the use of the NAS score, we retrieved 348 currently ongoing NASH-related clinical trials from ClinicalTrials.gov. A total of 188 trials use histological findings as the inclusion criteria, of which 79 trials use the NASH-CRN system as the diagnostic criteria, and 2 trials use the SAF/FLIP system (Table 1). In the early development stage, findings of non-invasive tests were the main inclusion criteria. However, histological evaluation is the major inclusion criterion in the phase of dose exploration and efficacy evaluation. For the NAS score, there is currently no uniform standard. A majority of the phase II–III trials, NAS score ≥4 has been used as the inclusion criterion (Table 1).

Liver biopsy is an accurate method to diagnose and differentiate NASH. The cardinal advantage of liver biopsy is that it allows for direct histological evaluation of the lesions, making it a reliable procedure for the diagnosis of NASH [34]. Clinical biochemical parameters such as serum aminotransferase can indicate disease activity to a certain extent but their levels may not always be consistent with the disease state. A recent study evaluated 534 adults with biopsy-proven NAFLD with normal aminotransferase levels; the prevalence of NASH with F2-F3 fibrosis and cirrhosis was 19% and 7%, respectively [35]. Normal aminotransferase levels do not necessarily indicate the absence of significant fibrosis or even cirrhosis; therefore, histopathological examination is more accurate method for diagnosis. Liver biopsy can also differentiate NAFLD from other chronic liver diseases. In an earlier study, among the 24 patients diagnosed with NASH before liver biopsy, the diagnosis was corrected in four patients after liver biopsy (three were diagnosed as normal/nonspecific and one was diagnosed as primary sclerosing cholangitis) [36]. In the study by Adams et al., 45 out of 51 (88%) NAFLD patients with positive autoantibodies qualified the diagnostic criteria for "probable" or "definite" autoimmune hepatitis (AIH). After liver biopsy, only four patients qualified the AIH diagnostic criteria [37]. Liver biopsy provides clinically meaningful information independent of purely clinical parameters. Therefore, use of liver biopsy as the diagnostic criterion for NASH is a more accurate method for the development of new NASH drugs.

Liver biopsy provides more morphological information than non-invasive tests. First, only histological examination can detect lobular inflammation and ballooning, which cannot be determined by current non-invasive diagnostic methods. Second, in the evaluation of steatosis, image-based MRI-proton density fat fraction (PDFF) can relatively accurately evaluate liver fat content. However, this test cannot reflect the distribution pattern of steatosis and the size of fat vesicles. Only liver biopsy can provide this important disease information. The distribution of steatosis was shown to be related with disease activity and correlate with fibrosis [38]. Diffuse microvesicular steatosis can be a manifestation of other etiologies including acute alcohol-induced, drug-induced liver disease, congenital metabolic disorder, and acute

Table 1. Number of NASH clinical trials in different phases and the corresponding inclusion criteria.

Phase of trials	Histological criteria for inclusion						Non-invasive tests for inclusion
	NAS ≥ 5 n (%)	NAS ≥ 4 n (%)	NAS ≥ 3 n (%)	SAF n (%)	Undefined n (%)	Total number of studies (Histology)	Total number of studies (Non-invasive)
I	0	3 (20.0)	1 (6.7)	0	11 (73.3)	15	60
II	6 (4.5)	40 (29.9)	13 (9.7)	1 (0.7)	74 (55.2)	134	70
III	1 (3.6)	8 (28.6)	4 (14.3)	1 (3.6)	14 (50.0)	28	7
IV	1 (9.1)	2 (18.2)	0	0	8 (72.7)	11	23
Total	8	53	18	2	107	188	160

fatty liver of pregnancy. Third, for the diagnosis of NASH fibrosis, previous studies showed a certain degree of correlation of image-based examination or serological markers with NASH fibrosis. However, non-invasive modalities do not allow for morphological characterization of fibrosis. Sun et al. established the P-I-R classification (Beijing classification) based on the morphology of the liver fibrosis septa to evaluate the antiviral therapeutic effect of chronic hepatitis B [39]. This new classification based on liver biopsy can help optimize therapeutic strategies, which cannot be achieved through non-invasive examinations. Therefore, compared with non-invasive tests, liver biopsy can provide more comprehensive disease information. Therefore, it is more suitable as inclusion criteria for new NASH drug development.

Liver biopsy used as therapeutic efficacy/endpoint in new NASH drug development

Liver biopsy is recommended by drug regulatory authorities as the therapeutic efficacy/endpoint for new NASH drug development. The ultimate goal of NASH treatment is to slow the progress, halt, or reverse disease progression and improve clinical outcomes. However, considering the slow progression of NASH, it will take a long time to observe the hard clinical endpoints of mortality and morbidity. Therefore, the drug regulatory authorities allowed the use of surrogate endpoints to accelerate the drug approval process. The ideal surrogate endpoint should not only reflect the long-term clinical benefit but also have the characteristics of flexibility and short-term availability to accelerate the new drug approval process.

Pathological evaluation of NASH based on liver biopsy meets the requirements of surrogate treatment endpoints for new drug development. Previous studies have shown a close relation between the pathological manifestations of NASH and prognosis. A recent cohort substudy analyzed 446 adult patients with NAFLD who underwent two liver biopsies over a period of 9 years. By comparing the findings of the two liver biopsies, the authors found that improvement or worsening of disease activity was associated with fibrosis regression or progression [40]. In several studies based on histological findings and long-term follow-up, the degree of fibrosis was a predictor of liver-related endpoint events [40–42]. A meta-analysis of 13 studies had a combined study population of 4428 patients with NAFLD, of which 2875 patients were reported to have NASH. The results showed an association of biopsy-confirmed fibrosis with the risk of mortality and liver-related morbidity [43]. A recent study developed a new machine learning (ML) algorithm based on the classical histological features and hepatic venous pressure gradient (HVPG) in patients with compensated cirrhosis NASH [44]. The ML HVPG score differentiated patients with normal, elevated HVPG (5.5–9.5 mmHg), and clinically significant portal hypertension (HVPG ≥ 10 mmHg), which were closely related to liver-related endpoint events. This new AI technology also demonstrated the association between histological features and clinical disease progression. Therefore, the close correlation between the histological manifestations of NASH and the clinical hard endpoints makes it a reasonable surrogate endpoint for NASH clinical trials.

Based on the advantages of pathological evaluation, several drug regulatory authorities recommend liver biopsy as the surrogate efficacy/endpoint in the new NASH drug development [31–33, 45]. The 2019 Liver Forum also provided recommendations for the endpoint selection of NASH trials [34]. In phase IIb/III clinical trials, the following criteria can be used as therapeutic endpoints: Resolution of steatohepatitis and no worsening of liver fibrosis based on the NASH-CRN fibrosis score; or improvement in liver fibrosis by ≥1 stage and no worsening of steatohepatitis; or both resolution of steatohepatitis and improvement in fibrosis. Resolution of steatohepatitis is defined as absence of signs of fatty liver disease or presence of isolated or simple steatosis without steatohepatitis and NAS scores of 0–1 for inflammation, 0 for ballooning, and any value for steatosis [31, 34].

New drug development with the efficacy of resolution of steatohepatitis. The resolution of NASH is used as the primary endpoint in many NASH therapeutic trials. Among the trials that are registered on Clinicaltrials.gov with results posted, the most commonly used criterion for NASH resolution is decrease in NAS score by two or more without worsening of fibrosis. A Phase II trial assessed the efficacy of silymarin in NASH patients for 12 months [46]. There were 4/27 (15%) in the 700 mg dose arm, 5/26 (19%) in 420 mg dose arm, and 3/25 (12%) in the placebo arm who reached the primary endpoint of improvement in NAS score by ≥2 points. Another phase II study also used the resolution of NASH as the primary endpoint to assess the efficacy of volixibat, an inhibitor of the apical sodium-dependent bile acid transporter [47]. On histological analysis, a greater proportion of participants in the placebo arm (38.5%) met the primary endpoint compared with the volixibat arm (30.0%). A pilot study of metformin for treatment used improvement in NAS by 3 points as the primary endpoint, and 30% of participants achieved a histological response [48].

In addition to NASH resolution as the primary endpoint, some studies have also used fibrosis improvement as a co-primary endpoint or secondary outcome. In the PIVENS study, vitamin E therapy was associated with a higher rate of improvement in NASH (improvement in NAS score by ≥2 points); however, neither vitamin E nor pioglitazone was associated with improvement in fibrosis scores which was the secondary endpoint [20]. In the FLINT study, 45.5% of participants in the obeticholic acid group reached the primary endpoint of improvement in NAS score by ≥2 and 35.3% of participants reached the secondary endpoint of improvement in fibrosis [21]. In the phase II trial of semaglutide, a higher percentage of patients in the semaglutide arm showed NASH resolution compared with the placebo arm. However, the trial did not show a significant difference in the percentage of patients with improvement in fibrosis stage (Table 2) [49].

New drug developments with the efficacy endpoint of improvement in liver fibrosis. Trials of new drugs that target fibrosis tend to use improvement of fibrosis as the primary endpoint. In the recently published data of combination therapy of selonsertib (SEL), cilofexor (CILO), and firsocostat (FIR) in NASH patients with bridging fibrosis and cirrhosis, the primary endpoint was a ≥1-stage improvement in fibrosis without worsening of NASH, which was achieved in 11% of placebo-treated patients versus 21% of patients treated with CILO/FIR [50]. The phase III trials of SEL in patients with NASH and bridging fibrosis (F3, STELLAR-3) or compensated cirrhosis (F4, STELLAR-4) used ≥1-stage improvement in fibrosis as the primary endpoint. Unfortunately, SEL did not show any antifibrotic effects in the two studies (Table 2) [14].

Advantages of liver biopsy over non-invasive tests as a surrogate endpoint in trials. Non-invasive diagnosis of NASH and the associated fibrosis is a contemporary research hot spot. Some non-invasive biomarkers have been used in clinical trials. MRI-PDFF has been used as the primary therapeutic endpoint in the early drug development stages such as the proof of concept studies, especially for drugs with the main target of reducing liver steatosis [33, 51]. For the evaluation of fibrosis, non-invasive tests such as controlled transient elastography (VCTE), FibroScan-AST score, and enhanced liver fibrosis test have been used as evaluation methods before liver biopsy to increase the success rate of screening [52, 53]. Magnetic resonance elastography has been used as a secondary endpoint in clinical trials because of its good performance in predicting fibrosis [54, 55].

However, irrespective of how good the predictive efficacy is, the area under the receiver operating characteristic curve (AUROC) will never reach 1.0. Although non-invasive imaging modalities offer some advantages, these do not provide in-depth characterization of the pathological changes. Furthermore, there is inadequate evidence to prove that non-invasive testing can

Table 2. Clinical trials in NASH patients with histological outcomes (phase II–III with results).

| Author [ref.] | Target | Treatment | Phase | Duration | Histological outcome measurement | | | |
| | | | | | NASH resolution without worsening of fibrosis | | Fibrosis improvement without worsening of NASH | |
					Drug (%)	Placebo (%)	Drug (%)	Placebo (%)
Navarro [46]	Antioxidant	Silymarin; 420 mg, 26; 700 mg, 27; PL, 25	2	48–50 week	NAS improvement ≥2 points 19 (420 mg)/15 (700 mg)	12	–	–
Unpublished	FXR agonist	Tropifexor (LJN452); 140 µg, 11; 200 µg, 11; PL, 11	2 (adaptive design)	48 week	NASH resolution as secondary outcome 10.5 (140 µg)/20.0 (200 µg)	7.1	Secondary outcome 28.9 (140 µg) 31.4 (200 µg)	26.2
Newsome [47]	Apical sodium-dependent Bile Acid Transporter inhibitor	Volixibat (SHP626); 5 mg, 11; 10 mg, 11; 20 mg, 8; PL, 13	2	48 week (terminated at 24 week)	NAS improvement ≥2 points 36.4 (5 mg)/18.2 (10 mg)/37.5 (20 mg)	38.5	–	–
Harrison [14]	ASK1 inhibitor	Selonsertib (GS-4997), F3:18 mg, 322; 6 mg, 321; PL 159 F4:18 mg, 354; 6 mg, 351; PL 172	3	48 week	–	–	F3:9.6 (18 mg) /12.1 (6 mg) F4: 14.4(18 mg)/12.8(6 mg)	F3: 13.2 F4: 12.8
Loomba [50]	FXR agonist; ACC inhibitor; ASK1 inhibitor	Selonsertib, 18 mg, 7; Firsocostat, 20 mg, 33; Cilofexor, 30 mg, 34; SEL + FIR, 71; SEL + CILO, 68; FIR + CILO, 67; PL, 38	2b	48 week	–	–	Selonsertib 28.6/ Firsocostat 12.1/ Cilofexor 11.8 SEL + FIR 15.5 SEL + CILO 19.1 FIR + CILO 20.9	10.5
Ratziu [16]	CCR2/CCR5 antagonist	Cenicriviroc (CVC), 150 mg, 145; PL, 144	2b	52 week	NAS improvement ≥2 points 15.9	18.8	Secondary outcome 20.0	10.4
Sven [77]	Pan-PPAR agonist	Lanifibranor (IVA337), 1200 mg, 83; 800 mg, 83; PL, 81	2b	24 week	SAF improvement ≥2 points 49.4 (1200 mg) 41.0 (800 mg)	27.2	Secondary outcome 42.2 (1200 mg) 27.7 (800 mg)	23.5
Newsome [49]	GLP-1 receptor agonist	Semaglutide, 0.1 mg, 57; 0.2 mg, 59; 0.4 mg, 56; PL, 58	2	72 week	NASH resolution 40.4 (0.1 mg) 35.6 (0.2 mg) 58.9 (0.4 mg)	17.2	Secondary outcome 49.1 (0.1 mg) 32.2 (0.2 mg) 42.9 (0.4 mg)	32.8
Neuschwander-Tetri [21]	FXR agonist	Obeticholic Acid, 25 mg,110; PL, 109	2	72 week	NAS improvement ≥2 points 45.5	21.1	Secondary outcome 35.3	19.4
Sanyal [20]	PPARγ agonist; Antioxidants	Pioglitazone, 30 mg, 80; Vitamin E, 800IU, 84; PL, 83	3	96 week	NAS ≤ 3, or NAS improvement ≥2 points Pioglitazone 33.8 Vitamin E 42.9	19.3	Secondary outcome Pioglitazone 44.3 Vitamin E 41.3	30.6
Loomba [48]	Insulin sensitizing agents	Metformin, 2000 mg, 23	2	48 week	NAS improvement ≥3 points 34.6	No	–	–

Table 2. continued

Author [ref.]	Target	Treatment	Phase	Duration	Histological outcome measurement		Fibrosis improvement without worsening of NASH	
					NASH resolution without worsening of fibrosis			
					Drug (%)	Placebo (%)	Drug (%)	Placebo (%)
Unpublished	PPARγ agonist	Pioglitazone, 30 mg, 18	2	48 week	NAS improvement ≥3 points 66.7	No	–	–
Unpublished	ACC inhibitor	Aramchol, 600 mg, 98; 400 mg, 101; PL, 48	2b	52 week	Secondary outcome NASH Resolution 16.7 (600 mg)/7.5(400 mg)	5	Secondary outcome 29.5 (600 mg)/21.3 (400 mg)	17.5
Sanyal [78]	SREBP inhibitor	Ethyl icosapentate, 2700 mg, 64; 1800 mg, 55; PL, 55	2b	52 week	NAS ≤ 3, or NAS improvement ≥2 points 32.7 (1800 mg) 31.3 (2700 mg)	32.7	–	–
Dasarathy [79]	Potent activators of PPARα and PPARγ agonist	Polyunsaturated fatty acids (Opti-EPA), (360 mg EPA and 240 mg DHA), 18; PL, 19	2	48 week	NAS improvement ≥2 points 44.4	47.4	–	–
Harrison [15]	Insulin sensitizer	MSDC-0602K, 62.5 mg, 84; 125 mg, 82; 250 mg, 86; PL, 74	2b	52 week	NAS improvement ≥2 points 29.8 (62.5 mg) 32.9 (125 mg) 39.5 (250 mg)	29.7	Secondary outcome 23.8 (62.5 mg) 28 (125 mg) 29.1 (250 mg)	
Ratziu [80]	Dual PPARα/δ agonist	Elafibranor (GFT505), 80 mg, 93; 120 mg, 91; PL, 92	2	52 week	Disappearance of NASH 23 (80 mg) 21 (120 mg)	17		

FXR farnesoid X receptor, ACC acetyl-CoA carboxylase, ASK1 apoptosis signal-regulating kinase 1, PPAR peroxisome proliferator-activated receptors, CCR2/CCR5 antagonist C–C chemokine receptors type 2 and 5 dual antagonist, GLP-1 glucagon-like peptide-1, SREBP sterol response element binding protein.

reflect the long-term prognosis of NASH and to predict clinical benefit with reliability and consistency. The current drug regulatory authorities and international guidelines have approved non-invasive testing as the main endpoint of the early IIa stage of drug development, according to the target of the drug [3, 33, 45]. In the pivotal (phase III) trials whose main purpose is to confirm efficacy, histology is still recognized as a relatively reliable efficacy endpoint. Non-invasive testing can be used as a secondary or exploratory endpoint in phase III trials [33, 56].

CHALLENGES OF LIVER BIOPSY IN NEW NASH DRUG DEVELOPMENT
Suboptimal reliability of pathological evaluation
Histological diagnosis obtained by liver biopsy is not perfect in all aspects. The reliability of pathological diagnosis is currently a major challenge. A recent study evaluated 678 biopsies from 339 patients conducted by three experienced hepatopathologists [57]. The inter-reader agreement was just fair to moderate (weighted kappa: steatosis 0.609, fibrosis 0.484, lobular inflammation 0.328, and ballooning 0.517). Although the same diagnostic criteria (NASH-CRN) were used, the agreement of pathological diagnosis was suboptimal. In the study by Pavlides et al., three pathologists evaluated biopsy specimens from 65 patients using NASH-CRN criteria; the kappa value for the fibrosis stage was 0.54. In another study, liver biopsy samples of 100 consecutive adult patients with suspected NAFLD were randomly assigned to four pathologists. Inter-observer agreement was acceptable for steatosis, lobular inflammation, and fibrosis, but not for hepatocyte ballooning (ICC: 0.012) [58]. The intra-observer agreement in this study was acceptable in all scales but was only 0.42 for lobular inflammation. Another study yielded similar results; the kappa value for lobular inflammation and ballooning was only 0.23–0.37 [59].

A high degree of discordance in liver histology interpretation may weaken the reliability of pathological evaluation and adversely affect clinical trials. In terms of screening of patients, pathological findings are the core inclusion criteria. An unreliable pathological evaluation may allow patients who do not meet histological criteria to be included and eligible patients to be excluded. In terms of efficacy evaluation, it is hard to accurately identify the true responders, which may attenuate the treatment effects. Ensuring the quality of liver biopsy specimens can help improve the consistency of evaluation. In addition, training of pathologists is another potential way to improve consistency. The FDA recommends establishing an adjudication committee of central pathologists to read baseline and posttreatment specimens together. The committee should consist of at least two experienced and trained pathologists to decide how each of the components will be interpreted [56].

Potential risks of invasive inspection and sample variability
Liver biopsy is an invasive procedure associated with potential risks such as pain and hemorrhage. The procedure is liable to cause patient discomfort. Therefore, liver biopsy may not be suitable for large-scale use, but is more suitable for use in patients with a high risk of progression [60]. Two or three liver biopsies are usually included in the design of phase IIb/III NASH clinical trials, which is also one of the main obstacles for patient enrollment. However, the overall incidence of adverse reactions is relatively low. The reported incidence of intraperitoneal hemorrhage is only 0.03%–0.7%, that of hemobilia and bile peritonitis is 0.006%–0.2% and 0.03%–0.22%, respectively [61]. Ultrasound-guided biopsy is associated with a lower incidence of complications and improved safety of liver biopsy, especially for patients with known specific mass lesions and history of intra-abdominal surgery [62].

Sample variability is another limitation of liver biopsy. Histologic lesions of NASH are unevenly distributed throughout the liver parenchyma, and 1 or 2 pieces of specimens may not be adequately representative of the entire liver. In the study by Ratziu et al., 51 patients with NAFLD underwent percutaneous liver biopsy with two samples collected to assess the sampling error of liver biopsy. The negative predictive value of a single biopsy for the diagnosis of NASH was at best 0.74. Six of seventeen patients with bridging fibrosis (35%) had only mild or no fibrosis on single sample reading [63]. In clinical trials, the influence of sampling variability will lead to bias in enrollment, disease classification, and efficacy evaluation.

Insufficient evaluation of dynamic changes in fibrosis
In the posttreatment liver biopsy, improvement in liver fibrosis by ≥1 stage is one of the main histological efficacy endpoints of NASH drugs. However, similar to the progress of fibrosis, the reversal of fibrosis is also a gradual and slow process. Therefore, within the limited treatment time of 48 to 72 weeks, the existing pathological staging system may not be sensitive enough to detect subtle changes in fibrosis. In addition, the fibrosis classification is merely a static assessment and does not suggest the dynamic changes in fibrosis morphology, which provides limited information about the direction of disease development. According to the "hepatic repair complex" concept proposed by Wanless et al., even a single biopsy can provide a sense of progression or regression [64]. For example, wide/broad, loosely aggregated collagen fibers and thin, densely compacted stroma can convey different disease information. In NASH clinical trials, we speculate that even in the absence of any posttreatment decrease in the overall fibrosis stage, change in the morphology of the fibrosis scar to a thin and dense form is also a sign of reversal. Recently, several phase III clinical trials that failed to meet the primary efficacy endpoint challenged the pathological evaluation system of NASH, which points to the need for a more nuanced and dynamic understanding of the disease beyond a mere static assessment of scar.

Weak evidence of the correlation between histological improvement and hard endpoints
Another concern about liver biopsy is whether the drug that has been approved by reaching the histological surrogate endpoint can truly achieve the hard endpoint in the long term. Currently, there is a paucity of data to assess whether improvement in histology can predict clinical endpoints. Recently, a study analyzed data from two clinical trials of NASH compensated cirrhosis patients. The results showed an association of cirrhosis regression with a lower risk of liver-related endpoint events compared with non-regression [65]. However, this study only included patients with compensated cirrhosis, the number of patients with fibrosis regression was relatively small, and the follow-up time was not long enough. Therefore, additional studies are required to justify the histological improvement and the hard clinical endpoints. The current FDA draft guidance endorses that premarketing trials for NASH which plan to evaluate the histology surrogate endpoint should ensure at least a 12–18 months period of treatment time [31]. Furthermore, the FDA's current thinking recommends efficacy evaluations of 2 or more years to confirm clinical benefit because of the subtle and slow changes in histologic features [56]. An investigational drug that is approved through an accelerated pathway based on the histologic efficacy endpoints requires a phase IV clinical outcomes trial to verify long-term benefits. The hard endpoints that truly reflect clinical benefit include the progression to cirrhosis, decompensatory events, Model for End-Stage Liver Disease score from ≤12 to ≥15, all-cause mortality, and need for liver transplantation. After completing pivotal trials and post-marketing trials, the product can be fully approved.

Inadequate identification of the complex heterogeneity of NASH
NAFLD is a multifactorial disease characterized by complex heterogeneity [66]. Several factors contribute to the heterogeneity

1206

of this disease, including, age, sex, genetic variants, type 2 diabetes mellitus, reproductive and hormonal status, metabolic health, and physical activity [67, 68]. These factors affect the progression of NAFLD disease through different mechanisms, thus the subphenotypes may have a distinct natural history and prognosis. The failure of several recent NASH trials has recently been attributed to disease heterogeneity [69]. Patients in different subgroups may show different responses to the drug. As mentioned above, histology is the main efficacy/endpoint for phase IIb/III clinical trials of NASH; however, the existing NASH histological criteria may not accurately render complex heterogeneity information. The potential relationship between NASH histological manifestations and dominant disease mechanisms is still not very clear [67, 68]. Thus, new NASH drug trials should take cognizance of clinical heterogeneity while designing the study. In the future, more research is required to explore the relationship between disease heterogeneity and histopathological characteristics and make better use of liver biopsy to identify the therapeutic effect in different subgroups of patients.

## THE FUTURE OF LIVER BIOPSY IN THE NEW NASH DRUG DEVELOPMENT

AI technology may improve the reliability of pathological evaluation

Recent advances in artificial intelligence technology has provided support for improving the reliability of pathological diagnosis. Second harmonic generation/two photon excitation fluorescence (SHG/TPEF) microscopy is a new technology that can fully quantify fibrosis and permit identification of individual collagen fibers without staining (Fig. 1). It can help precisely quantify the parameters of collagen fibers, such as the length, width, diameter, and the cross-linkages of the collagen fibers. A study established the Fibrosis-SHG index based on SHG/TPEF technology to evaluate the level of liver fibrosis in chronic liver disease [70]. Wang et al. used the SHG/TPEF technology for evaluation of NASH fibrosis [71]. They developed quantitation of fibrosis-related parameters

(q-FPs) and applied four main q-FPs to distinguish the F0–F4 stage of NASH fibrosis which showed a good diagnostic performance (AUROC 0.81–0.93). Chang et al. developed another fibrosis index (SHG B-index) comprising 14 unique SHG-based collagen parameters [72]. An SHG B-index score of >1.76 showed an overall accuracy of 98.5% in predicting the presence of bridging fibrosis. A recent study developed and validated qFIBS, a computational algorithm that quantifies key histological features of NASH. qFIBS included qFibrosis, qInflammation, qBallooning, and qSteatosis, which showed a satisfactory correlation with each respective component and good AUROC values (0.708–0.986) [73]. SHG/TPEF-based quantitative method can quantify the cardinal histological features of NASH independently of operators and experimental conditions. Therefore, this technology can allow for a more objective evaluation, especially when the diagnosis between different pathologists is inconsistent.

Good quality specimen may reduce sampling variability of pathological evaluation

Qualified liver biopsy specimens are the basis of pathological evaluation. Unqualified biopsy specimens affect the pathological evaluation, which in turn will affect enrollment screening and efficacy evaluation. Common factors that affect the quality of liver biopsy are the length and the integrity of the specimen, the number of portal areas sampled, and the clarity of staining. For pathological assessment, 25 mm is considered the ideal liver biopsy length to fully evaluate NASH lesions. However, 15 mm is also acceptable to balance the potential invasive injury and adequate evaluation [74]. For the width of the specimen, most studies recommend 16G cutting biopsy needles to provide a sufficient observation area. In addition, research centers need to avoid fragmentation of specimens as much as possible, as the integrity of specimens can also improve the reliability of diagnosis. For digital images, the sharpness of scanning is also a factor that can affect interpretation. Sponsors and investigators should deploy high-definition scanning equipment to ensure each specimen's clarity, and provide qualified scanned images for pathologists to read.
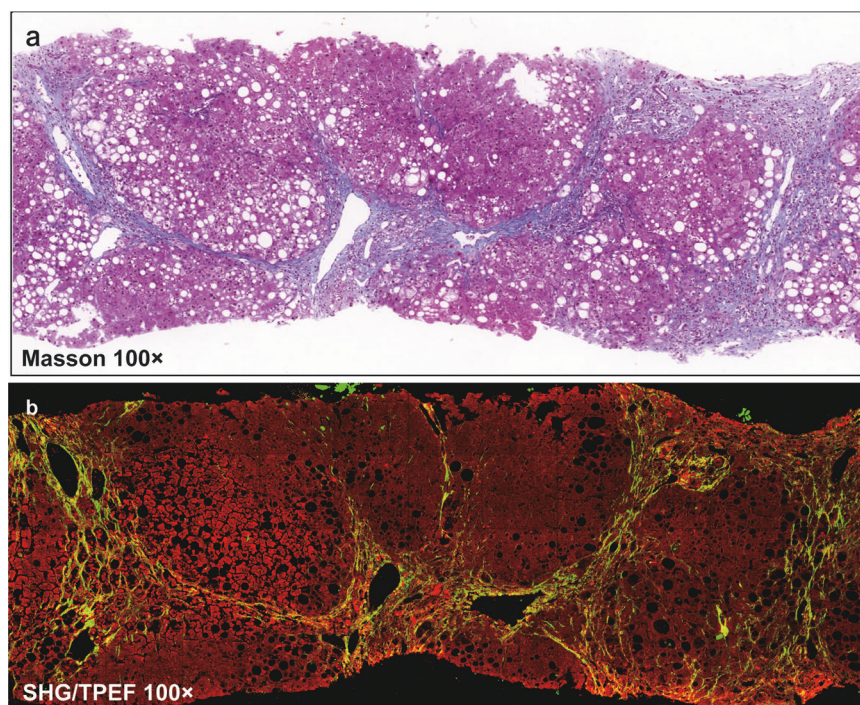
Fig. 1 Pathological images of Masson Trichrome staining and SHG/TPEF for liver specimen of a NASH patient with fibrosis. SHG/TPEF technology for quantification of NASH fibrosis. **a**. Masson Trichrome staining. **b**. SHG/TPEF technology for quantification of NASH.

Precise morphological assessment may improve identification of dynamic changes in fibrosis

The existing classification system categorizes fibrosis of NASH into five stages (0–4) [25]. Nevertheless, fibrosis/scarring in the same stage may show different morphology [64]. Future research should focus on the evaluation of precise morphological characteristics of fibrosis, to predict the direction of disease development. The "Beijing classification" provides a new potential idea for the evaluation of dynamic changes in fibrosis [39]. As mentioned above, this new fibrosis quality classification focuses on the morphological balance between progressive and regressive fibrosis/scarring. Based on the traditional fibrosis evaluation system, this new classification system categorizes patients into predominantly progressive (P), indeterminate (I), and predominately regressive (R), according to the morphology of fibrosis after antiviral therapy for chronic hepatitis B. A recent research of chronic hepatitis C used this new classification to assess the antiviral efficacy of direct-acting antiviral agents [75]. However, there are distinct differences between the patterns of fibrotic changes in viral hepatitis and NASH. Zone 3 pericellular fibrosis is the typical feature of NASH fibrosis [76]. Further studies are required to assess whether the "Beijing classification" can be used for the evaluation of NASH fibrosis (progress or reversal).

CONCLUSIONS

Histological evaluation based on liver biopsy is of great value in the development of new NASH drugs. It is used as the inclusion criteria, classification basis, and therapeutic efficacy/endpoint in clinical trials. Due to the close correlation between histological findings and long-term prognosis, liver biopsy can be used as a surrogate efficacy/endpoint in clinical trials of new drugs. Despite the inherent challenges in liver biopsy, including variable reliability of pathological evaluation, potential risk of invasive injury, sampling variability, insufficient evaluation of dynamic changes in fibrosis, inadequate identification of the complex heterogeneity, liver biopsy is still irreplaceable in the development of new NASH drugs. These challenges could be hopefully solved through new technology and good experimental design in the future.

REFERENCES

1. Younossi ZM. Non-alcoholic fatty liver disease: a global public health perspective. J Hepatol. 2019;70:531–44.
2. Singh S, Allen AM, Wang Z, Prokop LJ, Murad MH, Loomba R. Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. Clin Gastroenterol Hepatol. 2015;13:643–54.
3. Chalasani N, Younossi ZM, Lavine JE, Charlton M, Cusi K, Rinella M, 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;67:328–57.
4. European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO clinical practice guidelines for the management of non-alcoholic fatty liver disease. J Hepatol. 2016;64:1388–402.
5. Younossi ZM, Anstee QM, Marietti M, Hardy T, Henry L, Eslam M, et al. Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. Nat Rev Gastroenterol Hepatol. 2018;15:11–20.
6. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology. 2016;64:73–84.
7. Adams LA, Anstee QM, Tilg H, Targher G. Non-alcoholic fatty liver disease and its relationship with cardiovascular disease and other extrahepatic diseases. Gut. 2017;66:1138–53.
8. Targher G, Byrne CD, Tilg H. NAFLD and increased risk of cardiovascular disease: clinical associations, pathophysiological mechanisms and pharmacological implications. Gut. 2020;69:1691–705.
9. Henson JB, Simon TG, Kaplan A, Osganian S, Masia R, Corey KE. Advanced fibrosis is associated with incident cardiovascular disease in patients with non-alcoholic fatty liver disease. Aliment Pharmacol Ther. 2020;51:728–36.
10. Mantovani A, Petracca G, Beatrice G, Csermely A, Tilg H, Byrne CD, et al. Non-alcoholic fatty liver disease and increased risk of incident extrahepatic cancers: a meta-analysis of observational cohort studies. Gut. 2021. https://doi.org/10.1136/gutjnl-2021-324191.
11. Wang Z, Zhao X, Chen S, Wang Y, Cao L, Liao W, et al. Associations between nonalcoholic fatty liver disease and cancers in a large cohort in China. Clin Gastroenterol Hepatol. 2021;19:788–96.
12. Simon TG, Roelstraete B, Khalili H, Hagstrom H, Ludvigsson JF. Mortality in biopsy-confirmed nonalcoholic fatty liver disease: results from a nationwide cohort. Gut. 2021;70:1375–82.
13. Diehl AM, Day C. Cause, pathogenesis, and treatment of nonalcoholic steatohepatitis. N Engl J Med. 2017;377:2063–72.
14. Harrison SA, Wong VW-S, Okanoue T, Bzowej N, Vuppalanchi R, Younes Z, et al. Selonsertib for patients with bridging fibrosis or compensated cirrhosis due to NASH: results from randomized phase III STELLAR trials. J Hepatol. 2020;73:26–39.
15. Harrison SA, Alkhouri N, Davison BA, Sanyal A, Edwards C, Colca JR, et al. Insulin sensitizer MSDC-0602K in non-alcoholic steatohepatitis: a randomized, double-blind, placebo-controlled phase IIb study. J Hepatol. 2020;72:613–26.
16. Ratziu V, Sanyal A, Harrison SA, Wong VW, Francque S, Goodman Z, et al. Cenicriviroc treatment for adults with nonalcoholic steatohepatitis and fibrosis: final analysis of the phase 2b CENTAUR study. Hepatology. 2020;72:892–905.
17. Harrison SA, Abdelmalek MF, Caldwell S, Shiffman ML, Diehl AM, Ghalib R, et al. Simtuzumab is ineffective for patients with bridging fibrosis or compensated cirrhosis caused by nonalcoholic steatohepatitis. Gastroenterology. 2018;155:1140–53.
18. Menghini G. One-second needle biopsy of the liver. Gastroenterology. 1958;35:190–9.
19. Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD. American Association for the Study of Liver Disease. Liver Biopsy Hepatol. 2009;49:1017–44.
20. Sanyal AJ, Chalasani N, Kowdley KV, McCullough A, Diehl AM, Bass NM, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. N Engl J Med. 2010;362:1675–85.
21. Neuschwander-Tetri BA, Loomba R, Sanyal AJ, Lavine JE, Van Natta ML, Abdelmalek MF, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. Lancet. 2015;385:956–65.
22. Ludwig J, Viggiano TR, McGill DB, Oh BJ. Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. Mayo Clin Proc. 1980;55:434–8.
23. Kleiner DE, Makhlouf HR. Histology of nonalcoholic fatty liver disease and non-alcoholic steatohepatitis in adults and children. Clin Liver Dis. 2016;20:293–312.
24. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol. 1999;94:2467–74.
25. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology. 2005;41:1313–21.
26. Brunt EM. Nonalcoholic fatty liver disease: pros and cons of histologic systems of evaluation. Int J Mol Sci. 2016;17:97.
27. Nascimbeni F, Ballestri S, Machado MV, Mantovani A, Cortez-Pinto H, Targher G, et al. Clinical relevance of liver histopathology and different histological classifications of NASH in adults. Expert Rev Gastroenterol Hepatol. 2018;12:351–67.
28. Bedossa P, Consortium FP. Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. Hepatology. 2014;60:565–75.
29. Bedossa P, Poitou C, Veyrie N, Bouillot JL, Basdevant A, Paradis V, et al. Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. Hepatology. 2012;56:1751–9.

30. Nascimbeni F, Bedossa P, Fedchuk L, Pais R, Charlotte F, Lebray P, et al. Clinical validation of the FLIP algorithm and the SAF score in patients with non-alcoholic fatty liver disease. J Hepatol. 2020;72:828–38.

31. U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research. Noncirrhotic non-alcoholic steatohepatitis with liver fibrosis: developing drugs for treatment draft guidance for industry. 2018. https://www.fda.gov/media/119044/download.

32. National Medical Products Administration. Guidelines for clinical trials of drugs for the treatment of non-alcoholic steatohepatitis. [EB/OL] 2019-12-17 [2019-12-30]. 2019. http://www.nmpa.gov.cn/WS04/CL2138/372284.html.

33. Rinella ME, Tacke F, Sanyal AJ, Anstee QM. Report on the AASLD/EASL joint workshop on clinical trial endpoints in NAFLD. J Hepatol. 2019;71:823–33.

34. Cheung A, Neuschwander-Tetri BA, Kleiner DE, Schabel E, Rinella M, Harrison S, et al. Defining improvement in nonalcoholic steatohepatitis for treatment trial endpoints: recommendations from the liver forum. Hepatology. 2019;70:1841–55.

35. Gawrieh S, Wilson LA, Cummings OW, Clark JM, Loomba R, Hameed B, et al. Histologic findings of advanced fibrosis and cirrhosis in patients with nonalcoholic fatty liver disease who have normal aminotransferase levels. Am J Gastroenterol. 2019;114:1626–35.

36. Sorbi D, McGill DB, Thistle JL, Therneau TM, Henry J, Lindor KD. An assessment of the role of liver biopsies in asymptomatic patients with chronic liver test abnormalities. Am J Gastroenterol. 2000;95:3206–10.

37. Adams LA, Lindor KD, Angulo P. The prevalence of autoantibodies and autoimmune hepatitis in patients with nonalcoholic fatty liver disease. Am J Gastroenterol. 2004;99:1316–20.

38. Chalasani N, Wilson L, Kleiner DE, Cummings OW, Brunt EM, Unalp A,NASH Clinical Research Network. Relationship of steatosis grade and zonal location to histological features of steatohepatitis in adult patients with non-alcoholic fatty liver disease. J Hepatol. 2008;48:829–34.

39. Sun Y, Zhou J, Wang L, Wu X, Chen Y, Piao H, et al. New classification of liver biopsy assessment for fibrosis in chronic hepatitis B patients before and after treatment. Hepatology. 2017;65:1438–50.

40. Kleiner DE, Brunt EM, Wilson LA, Behling C, Guy C, Contos M, et al. Association of histologic disease activity with progression of nonalcoholic fatty liver disease. JAMA Netw Open. 2019;2:e1912565.

41. Dulai PS, Singh S, Patel J, Soni M, Prokop LJ, Younossi ZM, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: systematic review and meta-analysis. Hepatology. 2017;65:1557–65.

42. Hagstrom H, Nasr P, Ekstedt M, Hammar U, Stal P, Hultcrantz R, et al. Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. J Hepatol. 2017;67:1265–73.

43. Taylor RS, Taylor RJ, Bayliss S, Hagstrom H, Nasr P, Schattenberg JM, et al. Association between fibrosis stage and outcomes of patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis. Gastroenterology. 2020;15:1611–25.

44. Bosch J, Chung C, Carrasco-Zevallos OM, Harrison SA, Abdelmalek MF, Shiffman ML, et al. A machine learning approach to liver histological evaluation predicts clinically significant portal hypertension in NASH cirrhosis. Hepatology. 2021. https://doi.org/10.1002/hep.32087.

45. Human Medicines Development and Evaluation, European Medicines Agency. Report of the stakeholder interaction meeting on the development of medicinal products for chronic non-infectious liver diseases (PBC, PSC, NASH). 2019. https://www.ema.europa.eu/en/events.

46. Navarro AM, Belle SH, D'Amato M, Adfhal N, Brunt EM, Fried MW, et al. Silymarin in non-cirrhotics with non-alcoholic steatohepatitis: a randomized, double-blind, placebo controlled trial. PLoS One. 2019;14:e0221683.

47. Newsome PN, Palmer M, Freilich B, Sheikh MY, Sheikh A, Sarles H, et al. Volixibat in adults with non-alcoholic steatohepatitis: 24-week interim analysis from a randomized, phase II study. J Hepatol. 2020;73:231–40.

48. Loomba R, Lutchman G, Kleiner DE, Ricks M, Feld JJ, Borg BB, et al. Clinical trial: pilot study of metformin for the treatment of non-alcoholic steatohepatitis. Aliment Pharmacol Ther. 2009;29:172–82.

49. Newsome PN, Buchholtz K, Cusi K, Linder M, Okanoue T, Ratziu V, et al. A placebo-controlled trial of subcutaneous semaglutide in nonalcoholic steatohepatitis. N Engl J Med. 2021;384:1113–24.

50. Loomba R, Noureddin M, Kowdley KV, Kohli A, Sheikh A, Neff G, et al. Combination therapies including cilofexor and firsocostat for bridging fibrosis and cirrhosis attributable to NASH. Hepatology. 2021;73:625–43.

51. Caussy C, Reeder SB, Sirlin CB, Loomba R. Noninvasive, quantitative assessment of liver fat by MRI-PDFF as an endpoint in NASH trials. Hepatology. 2018;68:763–72.

52. Newsome PN, Sasso M, Deeks JJ, Paredes A, Boursier J, Chan W-K, et al. FibroScan-AST (FAST) score for the non-invasive identification of patients with non-alcoholic steatohepatitis with significant activity and fibrosis: a prospective derivation and global validation study. Lancet Gastroenterol Hepatol. 2020;5:362–73.

53. Vilar-Gomez E, Chalasani N. Non-invasive assessment of non-alcoholic fatty liver disease: clinical prediction rules and blood-based biomarkers. J Hepatol. 2018;68:305–15.

54. Jayakumar S, Middleton MS, Lawitz EJ, Mantry PS, Caldwell SH, Arnold H, et al. Longitudinal correlations between MRE, MRI-PDFF, and liver histology in patients with non-alcoholic steatohepatitis: analysis of data from a phase II trial of selonsertib. J Hepatol. 2019;70:133–41.

55. Ajmera VH, Liu A, Singh S, Yachoa G, Ramey M, Bhargava M, et al. Clinical utility of an increase in magnetic resonance elastography in predicting fibrosis progression in nonalcoholic fatty liver disease. Hepatology. 2020;71:849–60.

56. Anania FA, Dimick-Santos L, Mehta R, Toerner J, Beitz J. Nonalcoholic steatohepatitis: current thinking from the division of hepatology and nutrition at the Food and Drug Administration. Hepatology. 2020;73:2023–7.

57. Davison BA, Harrison SA, Cotter G, Alkhouri N, Sanyal A, Edwards C, et al. Suboptimal reliability of liver biopsy evaluation has implications for randomized clinical trials. J Hepatol. 2020;73:1322–32.

58. Pournik O, Alavian SM, Ghalichi L, Seifizarei B, Mehrnoush L, Aslani A, et al. Inter-observer and intra-observer agreement in pathological evaluation of non-alcoholic fatty liver disease suspected liver biopsies. Hepat Mon. 2014;14:e15167.

59. Gawrieh S, Knoedler DM, Saeian K, Wallace JR, Komorowski RA. Effects of interventions on intra- and interobserver agreement on interpretation of nonalcoholic fatty liver disease histology. Ann Diagn Pathol. 2011;15:19–24.

60. Lonardo A, Ballestri S. Perspectives of nonalcoholic fatty liver disease research: a personal point of view. Explor Med. 2020;1:85–107.

61. Shiha G, Ibrahim A, Helmy A, Sarin SK, Omata M, Kumar A, et al. Asian-Pacific Association for the Study of the Liver (APASL) consensus guidelines on invasive and non-invasive assessment of hepatic fibrosis: a 2016 update. Hepatol Int. 2017;11:1–30.

62. Tian G, Kong D, Jiang T, Li L. Complications after percutaneous ultrasound-guided liver biopsy: a systematic review and meta-analysis of a population of more than 12,000 patients from 51 cohort studies. J Ultrasound Med. 2020;39:1355–65.

63. Ratziu V, Charlotte F, Heurtier A, Gombert S, Giral P, Bruckert E, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. Gastroenterology. 2005;128:1898–906.

64. Wanless IR, Nakashima E, Sherman M. Regression of human cirrhosis. Morphologic features and the genesis of incomplete septal cirrhosis. Arch Pathol Lab Med. 2000;124:1599–607.

65. Sanyal AJ, Anstee QM, Trauner M, Lawitz EJ, Abdelmalek MF, Ding D, et al. Cirrhosis regression is associated with improved clinical outcomes in patients with nonalcoholic steatohepatitis. Hepatology. 2021. https://doi.org/10.1002/hep.32204.

66. Eslam M, Sanyal AJ, George J, Sanyal A, Neuschwander-Tetri B, Tiribelli C, et al. MAFLD: a consensus-driven proposed nomenclature for metabolic associated fatty liver disease. Gastroenterology. 2020;158:1999–2014.

67. Lonardo A, Arab JP, Arrese M. Perspectives on precision medicine approaches to NAFLD diagnosis and management. Adv Ther. 2021;38:2130–58.

68. Arrese M, Arab JP, Barrera F, Kaufmann B, Valenti L, Feldstein AE. Insights into nonalcoholic fatty-liver disease heterogeneity. Semin Liver Dis. 2021;41:421–34.

69. Ratziu V, Friedman SL. Why do so many NASH trials fail? Gastroenterology. 2020;18:S0016–5085. https://doi.org/10.1053/j.gastro.2020.05.046.

70. Gailhouste L, Grand YL, Odin C, Guyader D, Turlin B, Ezan F, et al. Fibrillar collagen scoring by second harmonic microscopy: a new tool in the assessment of liver fibrosis. J Hepatol. 2010;52:398–406.

71. Wang Y, Vincent R, Yang J, Asgharpour A, Liang X, Idowu MO, et al. Dual-photon microscopy-based quantitation of fibrosis-related parameters (q-FP) to model disease progression in steatohepatitis. Hepatology. 2017;65:1891–903.

72. Chang PE, Goh GBB, Leow WQ, Shen L, Lim KH, Tan CK. Second harmonic generation microscopy provides accurate automated staging of liver fibrosis in patients with non-alcoholic fatty liver disease. PLoS One. 2018;13:e0199166.

73. Liu F, Goh GBB, Tiniakos D, Wee A, Leow WQ, Zhao JM, et al. qFIBS: an automated technique for quantitative evaluation of fibrosis, inflammation, ballooning, and steatosis in patients with nonalcoholic steatohepatitis. Hepatology. 2020;71:1953–66.

74. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology. 2003;38:1449–57.

75. Huang R, Rao HY, Yang M, Gao YH, Wang J, Jin Q, et al. Histopathology and the predominantly progressive, indeterminate and predominately regressive score in hepatitis C virus patients after direct-acting antivirals therapy. World J Gastroenterol. 2021;27:404–15.

76. Wanless IR, Shiota K. The pathogenesis of nonalcoholic steatohepatitis and other fatty liver diseases: a four-step model including the role of lipid release and hepatic venular obstruction in the progression to cirrhosis. Semin Liver Dis. 2004;24:99–106.

77. Sven MF, Pierre B, Manal FA, Quentin MA, Elisabetta B, Vlad R, et al. A randomised, double-blind, placebo-controlled, multi-centre, dose-range, proof-of-concept, 24-

week treatment study of lanifibranor in adult subjects with non-alcoholic stea-
tohepatitis: Design of the NATIVE study. Contemp Clin Trials. 2020;98:106170.

78. Sanyal AJ, Abdelmalek MF, Suzuki A, Cummings OW, Chojkier M, Group EA. No
significant effects of ethyl-eicosapentanoic acid on histologic features of non-
alcoholic steatohepatitis in a phase 2 trial. Gastroenterology. 2014;147:377–84 e1.

79. Dasarathy S, Dasarathy J, Khiyami A, Yerian L, Hawkins C, Sargent R, et al. Double-
blind randomized placebo-controlled clinical trial of omega 3 fatty acids for the

treatment of diabetic patients with nonalcoholic steatohepatitis. J Clin Gastro-
enterol. 2015;49:137–44.

80. Ratziu V, Harrison SA, Francque S, Bedossa P, Lehert P, Serfaty L, et al. Elafibranor,
an agonist of the peroxisome proliferator-activated receptor-alpha and -delta,
induces resolution of nonalcoholic steatohepatitis without fibrosis worsening.
Gastroenterology. 2016;150:1147–59.