DR. BRENT A NEUSCHWANDER-TETRI (Orcid ID : 0000-0002-8520-7398)

Article type : Original

Defining Improvement in Nonalcoholic Steatohepatitis for Treatment Trial Endpoints: Recommendations from the Liver Forum

Amanda Cheung, MD[1]

Brent A. Neuschwander-Tetri, MD[2]

David E. Kleiner, MD, PhD[3]

Elmer Schabel, MD[4]

Mary Rinella, MD[5]

Stephen Harrison, MD[6]

Vlad Ratziu, MD, PhD[7]

Arun J. Sanyal, MD[8]

Rohit Loomba, MD, MHSc[9]

Sophie Jeannin Megnien, MD[10]

Richard Torstenson, MScPharm, PhD [11]

Veronica Miller, PhD[12]

On behalf of the Liver Forum Case Definitions Working Group

[1]Stanford University, cheungac@stanford.edu

[2]Saint Louis University, brent.tetri@health.slu.edu

[3]National Cancer Institute, kleinerd@mail.nih.gov

[4]Bundesinstitut für Arzneimittel und Medizinprodukte/ European Medicines Agency, Elmer.Schabel@bfarm.de *

[5]Northwestern University, Mary.Rinella@nm.org

[6]Pinnacle Clinical Research, sharrison@pinnacleresearch.com

[7]Hospital Pitié-Salpêtrière, vlad.ratziu@inserm.fr

[8]Virginia Commonwealth University, arun.sanyal@vcuhealth.org

[9]University of California, San Diego, roloomba@ucsd.edu

[10]Summit Clinical Research, sophie.megnien@mac.com

[11]Novo Nordisk Denmark, torstenson.richard@gmail.com

[12]University of California, Berkeley, veronicam@berkeley.edu

* The content of this paper represents the considerations and reflections of the authors, and not necessarily the views of the Food and Drug Administration, Bundesinstitut fuer Arnzeimittel und Medizinprodukte, or the European Medicines Agency

Corresponding author:

Brent A. Neuschwander-Tetri

3635 Vista Ave, GI/FDT-9

St. Louis, MO 63110

Phone: 314-577-8764

Fax: 314-577-8125

brent.tetri@health.slu.edu

Keywords: trial design, endpoints, biomarkers, nonalcoholic fatty liver disease, fibrosis

Abbreviations: American Association for the Study of Liver Diseases (AASLD), Controlled attenuation parameter (CAP), European Association for Study of the Liver (EASL), European Medicines Agency (EMA), Liver Inflammation and Fibrosis score (LIF), Magnetic resonance elastography (MRE), MR spectroscopy (MRS), Multiparametric MRI (MMRI), NAFLD activity score (NAS), Nonalcoholic fatty liver (NAFL), Nonalcoholic fatty liver disease (NAFLD), Nonalcoholic steatohepatitis (NASH), Proton-density-fat-fraction (PDFF), U.S. Food and Drug Administration (FDA), Vibration controlled transient elastography (VCTE)

Acknowledgement:

The authors thank Dr. Lara Dimick-Santos* for providing feedback and input on the development of this manuscript, as well as the entire Case Definitions Working Group (Appendix A).

Financial support:

The Forum for Collaborative Research receives unrestricted educational or research support from the following: Abbott Molecular; AbbVie; Advanced Biological Laboratories; Afimmune; Aicuris; Allergan; Altimmune; AMRA Medical; Arbutus Biopharma; Arrowhead Pharmaceuticals; Assembly Biosciences; Astellas; Astra-Zeneca; Biogen; Boeheringer Ingelheim; Bristol-Myers Squib; Celgene; Chimerix; Cocrystal Pharma, Inc.; Conatus Pharmaceuticals, Inc.; ConSynance; Covance; DDL Diagnostics; CTI Clinical Trial and Consulting Services; Cirius Therapeutics Inc.; ContraVir; Cymabay; DeuteRx; DiaPharma; DS Biopharma; Durect; Echosens; Enanta Pharmaceuticals, Inc.; Enyo Pharma; Exalenz; Ferring Pharmaceuticals; Fractyl; Galectin Therapeutics; Genentech; Genfit; Gilead Sciences; GlaxoSmithKline; Hepatitis B Foundation; HepQuant; HistoIndex Pte Ltd; Humedics GmbH; ICON Clinical Research; Immuron; Intercept Pharmaceuticals; Inventiva; Ironwood Pharmaceuticals; Janssen; LabCorp; Lilly; Madrigal Pharmaceuticals; Mallinckrodt Pharmaceuticals; MEDIAN Technologies; MediciNova; Medivir; Merck Laboratories; Microbiotix; Morphic Therapeutic; Mylan; NGM Biopharmaceuticals; Nimbus Therapeutics; Nitto Denko Technical Corp; Nordic Bioscience; Northsea Therapeutics; Novartis; Novo Nordisk Inc.; Nusirt; OWL Metabolomics; Perspectum Diagnostics; Pfizer; PPD Inc.; ProSciento; Quest Diagnostics; Raptor Pharmaceuticals; Resonance Health; Resoundant; Roche Molecular Systems; Roivant Sciences; RuiYi; SeraCare; Shire; Springbank Pharma; Somalogic; Takeda; TARGET PharmaSolutions; Teva Pharmaceuticals; ViiV Healthcare; VLVBio; Zafgen; and Zealand Pharma.

This work was also supported in part by RO1 DK 10596, and in part by the Intramural Research Program of the NIH, National Cancer Institute. These funding sources had no role in the analysis and interpretation of data, the writing of the report, or the decision to submit the article for publication.

Abstract

Identifying effective therapies for nonalcoholic steatohepatitis (NASH) with fibrosis is a pressing challenge, with 1-2% of the population in developed nations at risk of developing NASH cirrhosis and its complications. The design of NASH clinical therapeutic trials is hampered by the long period of minimally symptomatic disease that typically precedes the development of decompensated cirrhosis, and the accompanying uncertainties regarding the best pre-cirrhotic trial endpoints that reliably reflect a subsequent reduction in liver-related morbidity and mortality.

The Liver Forum is a multi-stakeholder organization comprised of academic, industry, and regulatory experts working from a regulatory science perspective to identify barriers, prioritize research, and identify solutions to accelerate therapeutic development for NASH. Prior work of The Liver Forum has focused on recommendations for disease definitions and baseline parameters to be implemented in clinical trials that are designed to assess disease status and prevent progression to cirrhosis, liver transplantation, hepatocellular carcinoma, and death.

The purpose of this summary is to review currently available clinical data to identify parameters that change in parallel with liver histology and are likely to reflect clinically meaningful reductions in the risk of developing cirrhosis and its complications. We review available data on exploratory histologic, blood-based and imaging pharmacodynamic biomarkers that may reflect meaningful treatment responses and provide recommendations regarding measurements to be considered in phase 2 and 3 trials as well as during post-marketing monitoring trials.

Nonalcoholic steatohepatitis (NASH) has a significant impact on lifespan and health care expenditures. Pharmacotherapy may be a reasonable approach for many patients, and numerous potentially effective therapeutic agents are under evaluation in phase 2 and 3 trials, the results of which are pending at this time (**Table 1**, www.clinicaltrials.gov).

Drug approval by regulatory agencies requires demonstration of clinically meaningful benefit to patients. In the United States and in Europe, approval requires demonstration of improvement in how a patient feels, functions, or survives.[(1)] Since NASH is typically asymptomatic until it has progressed to decompensated cirrhosis, a major challenge in drug development is to identify and validate surrogate markers that predict a reduction in the progression to hard outcomes (death, hepatocellular carcinoma, liver transplantation, cirrhosis and its complications). The ideal surrogate markers should predict future progression to hard outcomes even in the early stages of NASH and be sensitive to change in the context of treatment to reflect a reduction in the risk of progression.[(1)]

Summarized here are the data on currently available putative surrogate markers, measured in the context of clinical trials, with an emphasis on correlating changes in these biomarkers with changes in histological steatohepatitis or fibrosis and how these changes may predict outcomes. Areas of uncertainty and ambiguity are identified with recommendations on how such ambiguity might be resolved. This summary is the result of work by The Liver Forum, a multi-stakeholder collaboration of academic investigators, patient representatives, professional societies, industry stakeholders and regulatory authorities including representatives from the American Association for the Study of Liver Diseases (AASLD), the European Association for Study of the Liver (EASL), the U.S. Food and Drug Administration (FDA), and the European Medicines Agency (EMA). The Liver Forum has previously recommended baseline parameters to be measured in the context of clinical trials[(2)] and disease definitions to be used in clinical trials[(3)] with the goal of promoting comparability in trial design and data analysis. These disease definitions are aligned with those outlined in guidance documents

developed by the AASLD[4] and EASL.[5] The definitions and recommendations are nonbinding and do not represent official positions of the FDA or EMA.

Linguistic challenges in the field of NAFLD and NASH

The field of NASH struggles with linguistic and semantic issues that must be recognized to frame the discussion of endpoints. We discuss these issues here to acknowledge and highlight approaches taken to resolve them.

One issue has been the lack of an appropriate name for the type of nonalcoholic fatty liver disease (NAFLD) that is not NASH. NAFLD is accepted as a broad term that includes NASH, and current guidance documents use the term nonalcoholic fatty liver (NAFL) to describe NAFLD that is not NASH.[4, 5] The term NAFL will be used here for this purpose as well; however, it is important to note that NAFL is not totally benign and can lead to fibrosis and increased liver-related mortality, albeit at a significantly lower rate compared to patients with NASH.[6-8] A summary of definitions for the purposes of clinical trials is shown in **Table 2**.

Another issue is the language used to describe the severity of NASH as it correlates with the development of hard outcomes. From a clinical perspective, NASH with advanced fibrosis (often defined as stage 3 or 4 fibrosis) is a more severe condition than NASH with little or no fibrosis. However, from a pathophysiological standpoint and for the purposes of a rigorous definition to be used in the context of clinical trials, steatohepatitis has a spectrum of severity independent of the amount of fibrosis **(Figure 1)**. Pathologists grade biopsy features of NAFLD severity based on the degree of steatosis, ballooning and inflammation and separately assign a stage to the amount of fibrosis. A disease activity score based on the degree of ballooning and parenchymal inflammation has also been developed and is discussed below. The separation of NAFLD severity from fibrosis severity is currently used in regulatory language and aligns with the name steatohepatitis which denotes fat accumulation with inflammation and cellular injury. Any discussion of NAFLD or NASH severity should provide clarity regarding the term severity. A summary of definitions recommended for use in the context of treatment trials is shown in **Table 3**.

The current regulatory state for assessing improvement in clinical trials

U.S. and European regulatory authorities recognize the health burden of NASH and the need to develop pharmacotherapy for prevention and reversal of this disease with the ultimate goal of preventing premature death and other consequences of cirrhosis.[1] Resolution of NASH without worsening of fibrosis or improvement in liver fibrosis without worsening of NASH or combined

resolution of NASH and improvement in liver fibrosis are currently accepted as meaningful surrogate endpoints for accelerated (US) or conditional (EU) approval in Phase 3 trials (Table 3). Resolution of NASH is defined as the disappearance of hepatocyte ballooning (reduction to grade 0) and the resolution or minimal persistence of lobular inflammation (grade 0 or 1). Improvement of fibrosis is defined as improvement by at least one stage using the Brunt criteria as modified by Kleiner and Brunt.[9] The clinical significance of changes in these composite endpoints still needs to be demonstrated through long term outcome studies.

Based on these surrogate endpoints, confirmatory registration trials have been able to move forward using an accelerated pathway. In the U.S., the Accelerated Approval Pathway outlined in 21 CFR 314.510 and 601.41, Subparts H and E, allows for approval of medications that are likely to have a positive impact on how patients feel, function, or survive based on surrogate endpoints that have biological plausibility and supporting data as long as post-marketing data are collected to demonstrate improved survival, decreased liver-related events or other measurable clinical benefits.[1] The EMA has taken a quite similar approach, accepting a conditional approval pathway (according to Commission Regulation (EC) No 507/2006), based on the co-primary evaluation of these two surrogate endpoints.

NASH severity and fibrosis severity: two separate concepts

Current histological assessment of NASH (**Supplemental Material**) uses categorical variables to describe the magnitude of the features of steatohepatitis and fibrosis severity although these changes occur on a continuum. The use of numerical descriptors for fibrosis severity may be misinterpreted to mean amount of fibrosis is represented by these numbers (i.e. stage 2 is twice as much fibrosis as stage 1 and half as much as stage 4) and this linear relationship is assumed whenever an arithmetic mean of fibrosis stages is calculated to characterize a group of patients. However it is plausible that the absolute amount of fibrosis in a liver biopsy may increase exponentially (rather than linearly) with each numerical stage. If liver-related mortality is proportional to the actual amount of fibrosis in a liver, then a recent meta-analysis showing an exponential increase in mortality by fibrosis stage supports this notion.[10-12] Additional shortcomings of histological assessment including issues related to sampling variability are discussed further in the Supplemental Material.

Available data on longitudinal changes in liver biopsies suggest that patients with NAFLD can progress from NAFL to NASH or regress from NASH to NAFL.[7] A minority of patients with severe steatohepatitis appear to be relatively resistant to developing fibrosis, while some patients with only NAFL may develop fibrosis.[6-8] This latter phenotype has been termed steatofibrosis,[13] meaning fibrosis in the presence of steatosis but absence of steatohepatitis. It is unclear if some of these less common cases represent sampling variability of the liver biopsy with missed ballooning or lobular inflammation or regression of these features due to natural disease fluctuation. Nonetheless, since

the majority of NAFL patients do not develop NASH and do not experience liver-related outcomes, pharmacologic treatments have been focused on patients with NASH. Most patients with NAFL have similar outcomes to matched patient populations without NAFL, with the majority of adverse outcomes in this population being cardiovascular. However, NASH, especially with associated fibrosis, significantly increases progression to liver-related hard outcomes, including cirrhosis with its complications and liver-related mortality.[6, 10, 11, 14] Although recent studies have emphasized fibrosis as the major determinant of progression to hard endpoints,[10, 11] NASH is the major driver of fibrogenesis. Thus NASH may not be identified as a risk factor for hard outcomes in multivariable analyses since it is not an independent variable but rather related collinearly to the accumulation of fibrosis.[15] This was recently demonstrated in a large cohort study where 35% vs 94% of patients with stage 0 vs 4 fibrosis respectively had NASH[11] and in a study showing worse outcomes in those with NASH compared to NAFL.[16] Thus both NASH and fibrosis are appropriate therapeutic targets (**Figure 2**).

To be accepted in the context of clinical trials, measures of NASH severity need to correlate with the risk of progression to hard outcomes. Lobular inflammation and ballooning are two measures of NASH severity and in the phase 2B elafibranor study, improvement in lobular inflammation and ballooning correlated with improvement in fibrosis, while worsening of inflammation and ballooning correlated with worsening of fibrosis.[17] Using the NAFLD activity score (NAS) which is the sum of the steatosis, inflammation and ballooning scores,[9] natural history data from a NASH-CRN cohort with serial liver biopsies also showed a higher degree of NAFLD severity (defined as NAS>5) is associated with fibrosis progression, whereas patients with a baseline biopsy NAS<5 experience, on average, less fibrosis progression on a subsequent biopsy.[18]

Measures of NASH severity and changes in NASH severity

Describing the severity of NAFLD and NASH is limited by the paucity of outcomes data that are correlated with specific histological, laboratory, and other measures of NASH severity such as the degree of ballooning or inflammation, ALT elevation, and other serum and imaging biomarkers. In the context of treatment trials, pharmacotherapy could also potentially improve one or more of the histologic features of steatohepatitis without changing the diagnostic category. This may have occurred in the FLINT trial of obeticholic acid in which the improvement in NAS was statistically significant (treatment 45% vs placebo 21%, p 0.0002) but the proportion of patients with NASH resolution did not reach significance (treatment 22% vs placebo 13%, p = 0.08).[19] Similarly, improvement in fibrosis on a biopsy with initial portal and dense perisinusoidal fibrosis to persistent portal fibrosis but with trivial perisinusoidal fibrosis would both be classified as stage 2 fibrosis and thus fibrosis would be assessed as unchanged in the current categorical scoring method. The magnitude of change needed in more granular measurements that predict improvement in hard outcomes will need to be defined as more precise assessments are developed. Additional data would be needed to justify the use of changes in single components of the NAS for clinical trial endpoints for trials to support a marketing application.

Assessment of Changes in Steatosis

Of the diagnostic features of NASH, steatosis is the most reliably quantified variable on a continuous scale using readily available magnetic resonance imaging techniques. The most accepted and validated imaging-based biomarker for the detection and quantification of liver fat content is proton-density-fat-fraction (PDFF).[20] Furthermore, a number of blood test panels have been developed to identify the presence of steatosis (**Supplemental Material**) and have been validated mostly in cross-sectional studies. In a phase 2 trial of elafibranor, the change in steatosis as graded by histologic assessment was associated with changes in SteatoTest and Fatty Liver Index.[21] Several studies support that the decrease of steatosis (assessed by MRI-PDFF) occurs in parallel with other improvements in histopathology[22] but the clinical relevance of steatosis change with treatment is still unclear. Longitudinal data has not consistently demonstrated a link between the amount of steatosis and risk of progression to hard outcomes.[6, 16, 23]

Change in steatosis is currently not used as a confirmatory phase 3 trial endpoint. However, results from phase 2 trials of compounds with pleiotropic activities have demonstrated NASH resolution only with concomitant improvement in steatosis. Thus, in the context of phase 2a trials where evidence of potential benefit is sought without histologic confirmation, assessment of liver fat content can have a role as an endpoint, typically in conjunction with additional markers reflective of reduced liver injury such as reductions in serum ALT.

Conventional ultrasound imaging is easily accessible and relatively inexpensive but has limitations, including high observer variability, poor sensitivity for steatosis less than 30%, and limited ability to accurately quantify steatosis.[24, 25] The quantifiable loss of ultrasound signal as it penetrates the liver due to steatosis can be measured with the controlled attenuation parameter (CAP) during vibration controlled transient elastography (VCTE). CAP was found to be significantly and independently associated with steatosis and obesity but not inflammation or ballooning.[26] Although CAP is accurate for diagnosing the presence of steatosis, it is not adequate for determining different grades of steatosis;[27] as a result, it may not be adequately sensitive to dynamic changes in steatosis in the context of clinical trials.

Many clinical trials are using standard MRI to quantify steatosis with an algorithm to calculate the PDFF.[25] Although MR spectroscopy (MRS) also measures liver fat accurately,[24] its use in clinical trials is limited by the need for additional equipment and analytical algorithms. Changes in MRS and PDFF assessment of steatosis were closely correlated with each other in the colesevelam treatment trial,[28] and a subsequent cross-sectional study demonstrated greater accuracy with PDFF compared to MRS.[29]

MRI-PDFF is also more accurate than CAP for quantifying steatosis.[30, 31] The precision and reliability of PDFF quantification has been demonstrated with an inter-examination standard deviation less than 0.5%, suggesting that a numerical change of 2% (e.g., decrease in PDFF from 25% to 23%) may be real, although establishing the pathophysiological significance of such small incremental changes will require further validation.[32] In the phase 2 trial of obeticholic acid, an absolute 6% change in PDFF correlated with change in histologic steatosis grade after treatment.[33] Similarly, in a pediatric treatment trial, changes in PDFF correlated with histologic changes.[34] A treatment trial with ezetimibe showed that histologic responders had a significant decrease in PDFF[35] and a secondary analysis of this data showed that an absolute reduction in PDFF by 4.1% or a relative reduction in PDFF by 29.3%(+/-29% [SD]) correlated with a ≥ 2 point reduction in NAS although the relatively small sample size (N = 10) and variability in PDFF among patients highlight the need for further validation of this threshold value.[22] Some validation may be found in the phase 2 trial of selonsertib in which patients with improved steatosis by ≥1 grade had a relative reduction of ≥30% in PDFF.[36] Additionally, in the trial of colesevelam, a statistically significant change in liver fat by MRI-PDFF was found that was not detected using the categorical grading of steatosis on liver biopsies.[28] Further analysis of these data found that patients with an absolute reduction ≥1% in MRI-PDFF had parallel reductions in body weight and serum aminotransferase levels but did not show a significant change in histologic grading of steatosis,[37] providing evidence that small changes found by more precise measures than categorical grading may have clinical significance. Thus, MRI-PDFF may be more sensitive than histologic grading for assessing changes in steatosis in a treatment group, although the magnitude of change that correlates with meaningful outcomes needs further validation.

Assessment of Changes in Steatohepatitis

Liver biopsy remains necessary for the diagnosis of steatohepatitis. A biopsy identifies not only the characteristic features of hepatocyte ballooning, Mallory Denk bodies, inflammation and fibrosis in the setting of steatosis, but also their characteristics (e.g., type of inflammatory cells) and location. Multiple studies have linked serum and imaging biomarkers with the presence of steatohepatitis and large ongoing studies of metabolomic, lipidomic, proteomic and microRNA species in the blood are underway to find new biomarkers.

While serum ALT and AST elevations may prompt the initial investigation for steatohepatitis, aminotransferase levels have suboptimal diagnostic utility (sensitivity 42%, specificity 80% using ALT >30U/L as a cutoff[38]) in diagnosing NASH. Reductions in aminotransferase levels have correlated with histological improvement in NASH clinical trials. An analysis of the PIVENS and TONIC trials demonstrated an odds ratio of 1.3 for histological improvement or resolution of NASH for every 10 U/L reduction in ALT.[39] However, threshold levels of ALT and AST that correlate with resolution of NASH have not been established and the degree of reduction in ALT and AST levels in clinical trials that correlates with prevention of progression to cirrhosis or liver-related events is unknown. Until such data become available from clinical trials, the use of reductions in aminotransferase levels as an endpoint is limited to proof-of-concept studies.

Ultrasound, CT, and MRI lack the necessary sensitivity to differentiate between NAFL and NASH. Magnetic resonance elastography (MRE) is used to determine liver stiffness, and the results are influenced by the presence of both fibrosis and steatohepatitis. Thus, in the context of treatment trials, interpretation of MRE changes must take into account the effects of both fibrosis and inflammation.

Additional investigative MRI techniques include the use of gadoxetic acid which shows relatively lower enhancement in steatohepatitis, specifically lobular inflammation and ballooning, but not with steatosis. Mmultiparametric MRI (MMRI, e.g, Perspectum® Multiscan). MMRI is another exploratory test based on iron-corrected T1 MR imaging[40] that is analyzed to provide a continuous Liver Inflammation and Fibrosis (LIF) score that is proposed to correlate with the severity of steatohepatitis.[41] A small study suggested the LIF score might predict progression to hard outcomes as none of the 56 patients with a LIF score <2 developed liver-related outcomes in up to 40 months of follow up in one study.[42] The correlation between changes in MMRI and changes in histology with treatment of NASH is currently under investigation, and further data are needed to determine the relative contributions of inflammation and fibrosis to the LIF score.

Multiple exploratory biomarkers to identify changes in NASH severity and resolution and many are being evaluated in ongoing phase 2 and 3 trials (**Supplemental Material)** and the role of these in assessing treatment response may be better understood after these trials are completed.

Measures of changes in fibrosis

Since the degree of liver fibrosis in NAFLD is the feature most strongly associated with hard outcomes,[6, 10, 11, 14] measuring changes in fibrosis is essential in assessing the benefits of NASH therapeutics. Whether the improvements demonstrated in treatment trials will correlate with reductions in hard outcomes will be determined in the continued assessment of patients in these trials.

A large number of diagnostic tests and algorithms have been developed to estimate the degree of liver fibrosis in cross-sectional studies.[43] Serum markers and panels have been developed based on direct measures of collagen synthesis or degradation and measures of altered liver function. Scoring systems that have been studied or validated in patients with NAFLD include the APRI, ELF, BAAT, BARD, BARDI, NAFLD fibrosis, and FIB-4 scores (**Supplemental Table 3**). The utility of these tests and algorithms to assess changes in response to therapeutics will need to be further validated because the current data are limited. A natural history study of 118 patients in Italy found changes in fibrosis on liver biopsy correlated with changes in APRI scores (includes AST) but not with FIB4 or NAFLD fibrosis scores (both incorporate ALT and AST).[44] In a recent early phase clinical trial, the ELF test

(includes neither AST nor ALT) demonstrated a significant absolute and relative change with therapeutic intervention.[45] Longitudinal data from the phase 2B elafibranor trial showed a significant improvement in the NAFLD fibrosis score and Fibrotest after one year of treatment.[21] In general, the fibrosis scores perform best at the high ends of their respective scales, and the AUROC test applied to judge "how good" the test is only shows how well the test performs in distinguishing stages of fibrosis. The relationships between continuous scores and total collagen or other continuous measures of fibrosis may be forthcoming from analyses of the ongoing studies. Reductions in N-terminal type III collagen pro-peptide (Pro-C3) levels were found in studies of an FGF-19 analogue and pegylated FGF-21, but how these changes correlated with histology was not reported.[46, 47] Also, many of the tests and algorithms that have been used to assess liver fibrosis in cross-sectional studies include parameters that also reflect ongoing injury such as serum ALT and AST levels. Since aminotransferase levels are influenced by changes in necroinflammatory activity, changes in fibrosis scores that include ALT and AST could be misinterpreted as showing fibrosis changes if the pharmacotherapy reduces necroinflammation.[21, 44, 45]

Imaging techniques are increasingly used to assess liver fibrosis.[24] While standard ultrasound, CT, and MRI can reliably diagnose late stages of cirrhosis based on a nodular appearing liver or signs of portal hypertension, these imaging modalities are less reliable in detecting earlier stages of fibrosis and even early cirrhosis. VCTE measures shear wave velocity to estimate liver stiffness. However, liver stiffness measurements may be overestimated in nonalcoholic fatty liver disease due to steatosis and increased distance between the probe and liver. However, this can be partly corrected by using the XL probe.[26, 48] Acoustic radiation force impulse imaging and supersonic shear wave elastography have similar diagnostic capabilities with relatively poor performance for identifying mild fibrosis but good accuracy in advanced fibrosis.[49]

MRE, unlike the aforementioned modalities, evaluates the entire liver and is more accurate than VCTE for identifying the presence of the spectrum of fibrosis in patients with NASH,[30] but its limited availability and higher cost may be impediments to widespread use. MRE correlates with histologic assessment in numerous prospective studies,[29, 30] but the cut-offs for each stage of fibrosis are not clearly defined. In the phase 2 selonsertib trial, patients with improvement by ≥1 fibrosis stage had a trend towards reduction in liver stiffness measured by MRE, but no change was detectable by VCTE measurements. These results suggest MRE may have greater sensitivity to detect changes in fibrosis in response to treatment, although the small sample size and lack of a placebo group limit interpretation.[36]

Reversal of cirrhosis

Reversal of existing cirrhosis to lesser degrees of fibrosis or complete resolution of excess fibrosis is an ambitious goal. The possibility that cirrhosis can be reversed has been shown in patients with viral hepatitis and other forms of chronic liver disease when the underlying cause of liver injury is

eliminated. Data in patients with NASH cirrhosis before bariatric surgery also suggests that cirrhosis resolution can occur in NASH, but in what proportion of patients remains to be established. Detecting early reductions in portal hypertension or improvements in liver function that might predict subsequent reversal of cirrhosis with continued treatment has been challenging. Some studies have used measurement of the hepatic-portal venous pressure gradient (HVPG), but this measurement is technically challenging, invasive and expensive for incorporation into large clinical trials. Current trials in patients with cirrhosis capture outcomes related to cirrhosis such hepatic encephalopathy, ascites and variceal hemorrhage as well as the laboratory components of the Child-Pugh-Turcotte (CPT) score. How changes in these clinical parameters align with treatment compared to placebo remains to be established by current and future trials.

Changes in functional testing

Functional testing is a novel approach to assess the severity of liver disease and changes in actual liver function in the context of treatment trials, especially in patients with cirrhosis. In general, functional testing has been challenging because the multiple functions of the liver (e.g. protein synthesis, drug disposal, bile secretion) may not change in parallel with liver disease progression. Additionally, as portal hypertension progresses, the development of collateral blood flow may alter hepatic drug metabolism without directly reflecting the metabolic or synthetic capacity of the liver. These issues have been addressed in the HepQuant® test by using two different stable isotope-labeled bile acids to simultaneously measure clearance from portal and systemic circulation as well as portal-systemic shunting.[50] Preliminary data indicate this assessment of liver impairment correlates with hard endpoints but more extensive validation is needed, particularly correlation with histology at early time points and clinical outcomes. Another test is the methacetin breath test that reflects just one aspect of metabolic capacity of the liver. This is being evaluated in several ongoing clinical trials for NASH and longitudinal data that correlates this assessment with histologic changes may be forthcoming. Current studies in cirrhotic patients typically include changes in the INR, total bilirubin and MELD score as secondary endpoints. How the HepQuant, methacetin breath test and other tests of specific metabolic functions of the liver compare to these commonly used measures remains to be determined.

Summary and recommendations

Describing meaningful changes in NASH severity and NASH-related fibrosis in response to treatment presents linguistic and operational challenges. Disease definitions were outlined previously[3] but describing changes in the context of treatment trials adds complexity. Outlined here is an overview of clinical, blood-based, imaging, functional, and biopsy measures identified in cross-sectional studies and how some of these have been followed longitudinally in clinical trials and correlated with histological changes. As the field of NASH therapeutics evolves, continued acquisition of samples and strategic analysis of these samples may lead to the validation of biomarkers as measures of

treatment response that may be validated as reliable predictors (surrogate markers) of hard outcomes. These measures were summarized previously.[2] Continued progress will rely on the use of consistent definitions of disease status and treatment responses which will allow direct comparisons of different biomarkers across clinical trials. The ultimate goal is to identify biomarkers for which changes with treatment predict a subsequent reduction in the risk of developing hard outcomes such as the complications of cirrhosis, liver cancer, liver transplantation or death.

Figure legends

Figure1. Graphical visualization of changes in NAFLD and fibrosis severity. In this visualization of the major histological changes in NAFLD and NASH, the severity of NAFLD and its subset NASH are shown on the horizontal axis while the severity of fibrosis is shown on the vertical axis. The two axes are labeled with the currently used descriptors and categorical degrees of severity. The impact of hypothetical therapies is shown with biopsy 1 before therapy and biopsy 2 at a subsequent time point after therapy. Panel A shows the ultimate therapeutic goal with resolution of both NAFLD and fibrosis from biopsy 1 to biopsy 2. Whether this is achievable is unknown. Panel B shows the hypothetical response to an effective antifibrotic drug with improved fibrosis but no impact on the severity of NAFLD. Panel C shows the effect of a purely anti-NASH drug, in which NASH resolves and a measurable secondary beneficial reduction in fibrosis occurs due to the decreased stimulus for fibrogenesis. Panel D shows the effect of a hypothetical anti-NASH drug with the undesirable effect of worsening fibrosis. Whether this could occur is unknown but current composite endpoints for anti-NASH drugs include the caveat that fibrosis should not worsen to avoid developing such agents. A goal of the development of non-invasive markers for NAFLD severity and fibrosis severity is to describe both with continuous variables with precision and accuracy with respect to the probability of hard outcomes.

Figure 2. The pathway to hard outcomes in patients with NASH, from healthy to cirrhosis, hepatocellular carcinoma, transplantation or death. Blood tests, imaging and liver biopsies can identify changes during the progression of NASH and improvement in some of these may ultimately be validated to reflect reduction in the progression to hard endpoints of liver related complications (bleeding varices, ascites, hepatic encephalopathy), cirrhosis, death, liver transplantation, and hepatocellular carcinoma. Because hepatocellular injury is the driver of fibrogenesis, a reasonable approach to assessing response to therapy of drugs that target mechanisms of injury is to assess markers of that injury such as ALT and biopsy features of steatohepatitis. Drugs that mechanistically target fibrosis are unlikely to directly influence steatohepatitis and thus measures of change in fibrosis will be essential for assessing their efficacy. Drugs that target steatohepatitis may indirectly improve fibrosis by decreasing the stimulus for fibrogenesis, but the primary means of assessing their efficacy will include measuring improvement in steatohepatitis. No worsening of fibrosis is a reasonable endpoint for drugs that address steatohepatitis because trials of such drugs may be too short in duration to achieve measurable improvement in fibrosis.

Appendix A: Case Definitions Working Group Participants

Manal Abdelmalek, Duke University School of Medicine

Quentin Anstee, Newcastle University Medical School

Rajarshi Banerjee, Perspectum Diagnostics Ltd.

Mustafa Bashir, Duke University Medical Center

Pierre Bedossa, University of Paris Diderot

Mark Berner-Hansen, Novo Nordisk A/S

Manu Chakravarthy, Axcella Health

Jean Chan, Conatus Pharmaceuticals, Inc.

Edgar Charles, Bristol-Myers Squibb

Amanda Cheung, Stanford University

Lara Dimick-Santos, U.S. Food and Drug Administration

Judith Ertle, Boehringer Ingelheim Pharma GmbH & Co. KG

Sven Francque, University Hospital Antwerp

Scott Friedman, Icahn School of Medicine at Mount Sinai

Goran Gannedahl, AstraZeneca R&D

Katherine Greene, Forum for Collaborative Research

Michael Hambleton, DiaPharma Group Inc.

Stephen Harrison, Pinnacle Clinical Research

Dean Hum, GENFIT SA

Joanne Imperial, Conatus Pharmaceuticals, Inc.

David Kleiner, National Cancer Institute

Diana Julie Leeming, Nordic Bioscience

Rohit Loomba, University of California, San Diego

Sophie Jeannin Megnien, Summit Clinical Research

Ruby Mehta, U.S. Food and Drug Administration

Veronica Miller, Forum for Collaborative Research

Nikolai Naoumov, Novartis Pharmaceuticals

Stephanie Omokaro, U.S. Food and Drug Administration

Melissa Palmer, Shire Pharmaceuticals

Dan Peres, Immuron Limited

Margaret Powell, TARGET PharmaSolutions

Vlad Ratziu, Hôpital Pitié Salpêtrière

Arie Regev, Eli Lilly & Company

Mary Rinella, Northwestern University

Glenn Rosen, Bristol-Myers Squibb

Arun Sanyal, Virginia Commonwealth University

Elmer Schabel, Bundesinstitut für Arzneimittel und Medizinprodukte

Corinna Schölch, Boehringer Ingelheim Pharma GmbH & Co. KG

Jeffrey Schwimmer, University of California, San Diego

David Shapiro, Intercept Pharmaceuticals, Inc.

Reshma Shringarpure, Intercept Pharmaceuticals

Dean Tai, Histoindex Pte. Ltd.

Brent Neuschwander-Tetri, Saint Louis University School of Medicine

Richard Torstenson, Novo Nordisk A/S Denmark

Chinweike Ukomadu, Novartis Pharmaceuticals

Vincent Wong, The Chinese University of Hong Kong

Teresa Wright, Veterans Administration Medical Center

References

1) Sanyal AJ, Friedman SL, McCullough AJ, Dimick-Santos L. Challenges and opportunities in drug and biomarker development for nonalcoholic steatohepatitis: findings and recommendations from an American Association for the Study of Liver Diseases-U.S. Food and Drug Administration Joint Workshop. Hepatology 2015;61:1392-1405.

2) Patel YA, Imperial JC, Muir AJ, Anstee QM, DeBrota D, Dimick-Santos L, et al. Baseline parameters in clinical trials for nonalcoholic steatohepatitis: Recommendations from the Liver Forum. Gastroenterology 2017;153:621-625.

3) Siddiqui MS, Harrison SA, Abdelmalek MF, Anstee QM, Bedossa P, Castera L, et al. Case definitions for inclusion and analysis of endpoints in clinical trials for nonalcoholic steatohepatitis through the lens of regulatory science. Hepatology 2018;67:2001-2012.

4) Chalasani N, Younossi Z, 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;67:328-357.

5) 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;64:1388-1402.

6) Angulo P, Kleiner DE, Dam-Larsen S, Adams LA, Bjornsson ES, Charatcharoenwitthaya P, et al. Liver fibrosis, but no other histologic features, is associated with long-term outcomes of patients with nonalcoholic fatty liver disease. Gastroenterology 2015;149:389-397.

7) McPherson S, Hardy T, Henderson E, Burt AD, Day CP, Anstee QM. Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: Implications for prognosis and clinical management. J Hepatol 2015;62:1148-1155.

8) 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;13:643-654.

9) 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;41:1313-1321.

10) Dulai PS, Singh S, Patel J, Soni M, Prokop LJ, Younossi Z, et al. Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. Hepatology 2017;65:1557-1565.

11) Hagström H, Nasr P, Ekstedt M, Hammar U, Stål 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;67:1265-1273.

12) Pavlides M, Birks J, Fryer E, Delaney D, Sarania N, Banerjee R, et al. Interobserver variability in histologic evaluation of liver fibrosis using categorical and quantitative scores. Am J Clin Pathol 2017;147:364-369.

13) Bedossa P, Flip Pathology Consortium. 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;60:565-575.

14) Ekstedt M, Hagström H, Nasr P, Fredrikson M, Stål P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. Hepatology 2015;61:1547-1554.

15) Ratziu V. Back to Byzance: *Querelles byzantines* over NASH and fibrosis. J Hepatol 2017;67:1134-1136.

16) Younossi ZM, Stepanova M, Rafiq N, Makhlouf H, Younoszai Z, Agrawal R, et al. Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality. Hepatology 2011;53:1874-1882.

17) Ratziu V, Francque SM, Harrison S, Anstee QM, Bedossa P, Brozek J, et al. Improvement in NASH histological activity highly correlates with fibrosis regression (abstract). Hepatology 2016;64:1140A.

18) Kleiner DE, Brunt EM, Belt PH, Wilson L, Guy CD, Yeh MM, et al. Diagnostic pattern and disease activity are related to disease progression and regression in nonalcoholic fatty liver disease (abstract). Hepatology 2016;65:19A.

19) 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;385:956-965.

20) Caussy C, Alquiraish MH, Nguyen P, Hernandez C, Cepin S, Fortney LE, et al. Optimal threshold of controlled attenuation parameter with MRI-PDFF as the gold standard for the detection of hepatic steatosis. Hepatology 2018;67:1348-1359.

21) Ratziu V, Harrison SA, Francque S, Bedossa P, Lehert P, Serfaty L, et al. Elafibranor, an agonist of the peroxisome proliferator-activated receptor-a and -d, induces resolution of nonalcoholic steatohepatitis without fibrosis worsening. Gastroenterology 2016;150:1147-1159.

22) Patel J, Bettencourt R, Cui J, Salotti J, Hooker J, Bhatt A, et al. Association of noninvasive quantitative decline in liver fat content on MRI with histologic response in nonalcoholic steatohepatitis. Therapeutic Advances in Gastroenterology 2016;9:692-701.

23) Ajmera V, Park CC, Caussy C, Singh S, Hernandez C, Bettencourt R, et al. Magnetic resonance imaging proton density fat fraction associates with progression of fibrosis in patients with nonalcoholic fatty liver disease. Gastroenterology 2018;155:307-310.

24) Hannah WN, Jr., Harrison SA. Noninvasive imaging methods to determine severity of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. Hepatology 2016;64:2234-2243.

25) Kramer H, Pickhardt PJ, Kliewer MA, Hernando D, Chen GH, Zagzebski JA, et al. Accuracy of liver fat quantification with advanced CT, MRI, and ultrasound techniques: prospective comparison with MR spectroscopy. AJR Am J Roentgenol 2017;208:92-100.

26) Petta S, Wong VW, Camma C, Hiriart JB, Wong GL, Marra F, et al. Improved noninvasive prediction of liver fibrosis by liver stiffness measurement in patients with nonalcoholic fatty liver disease accounting for controlled attenuation parameter values. Hepatology 2017;65:1145-1155.

27) Siddiqui MS, Vuppalanchi R, Van Natta ML, Hallinan E, Kowdley KV, Abdelmalek M, et al. Vibration-controlled transient elastography to assess fibrosis and steatosis in patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2019;17:156-163.

28) Le T-A, Chen J, Changchien C, Peterson MR, Kono Y, Patton H, et al. Effect of colesevelam on liver fat quantified by magnetic resonance in nonalcoholic steatohepatitis: A randomized controlled trial. Hepatology 2012;56:922-932.

29) Imajo K, Kessoku T, Honda Y, Tomeno W, Ogawa Y, Mawatari H, et al. Magnetic resonance imaging more accurately classifies steatosis and fibrosis in patients with nonalcoholic fatty liver disease than transient elastography. Gastroenterology 2016;150:626-637 e627.

30) Park CC, Nguyen P, Hernandez C, Bettencourt R, Ramirez K, Fortney L, et al. Magnetic resonance elastography vs transient elastography in detection of fibrosis and noninvasive measurement of steatosis in patients with biopsy-proven nonalcoholic fatty liver disease. Gastroenterology 2017;152:598-607.

31) Tang A, Tan J, Sun M, Hamilton G, Bydder M, Wolfson T, et al. Nonalcoholic fatty liver disease: MR imaging of liver proton density fat fraction to assess hepatic steatosis. Radiology 2013;267:422-431.

32) Tyagi A, Yeganeh O, Levin Y, Hooker JC, Hamilton GC, Wolfson T, et al. Intra- and inter-examination repeatability of magnetic resonance spectroscopy, magnitude-based MRI, and complex-based MRI for estimation of hepatic proton density fat fraction in overweight and obese children and adults. Abdom Imaging 2015;40:3070-3077.

33) Middleton MS, Heba ER, Hooker CA, Bashir MR, Fowler KJ, Sandrasegaran K, et al. Agreement between magnetic resonance imaging proton density fat fraction measurements and pathologist-assigned steatosis grades of liver biopsies from adults with nonalcoholic steatohepatitis. Gastroenterology 2017;153:753-761.

34) Middleton MS, Van Natta ML, Heba ER, Alazraki A, Trout AT, Masand P, et al. Diagnostic accuracy of magnetic resonance imaging hepatic proton density fat fraction in pediatric nonalcoholic fatty liver disease. Hepatology 2018;67:858-872.

35) Loomba R, Sirlin CB, Ang B, Bettencourt R, Jain R, Salotti J, et al. Ezetimibe for the treatment of nonalcoholic steatohepatitis: assessment by novel magnetic resonance imaging and magnetic resonance elastography in a randomized trial (MOZART trial). Hepatology 2015;61:1239-1250.

36) Loomba R, Lawitz E, Mantry PS, Jayakumar S, Caldwell SH, Arnold H, et al. The ASK1 inhibitor selonsertib in patients with nonalcoholic steatohepatitis: a randomized, phase 2 trial. Hepatology 2018;67:549-559.

37) Noureddin M, Lam J, Peterson MR, Middleton M, Hamilton G, Le TA, et al. Utility of magnetic resonance imaging versus histology for quantifying changes in liver fat in nonalcoholic fatty liver disease trials. Hepatology 2013;58:1930-1940.

38) Kunde SS, Lazenby AJ, Clements RH, Abrams GA. Spectrum of NAFLD and diagnostic implications of the proposed new normal range for serum ALT in obese women. Hepatology 2005;42:650-656.

39) Vuppalanchi R, Jain AK, Deppe R, Yates K, Comerford M, Masuoka HC, et al. Relationship between changes in serum levels of keratin 18 and changes in liver histology in children and adults with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol 2014;12:2121-2130.

40) Banerjee R, Pavlides M, Tunnicliffe EM, Piechnik SK, Sarania N, Philips R, et al. Multiparametric magnetic resonance for the non-invasive diagnosis of liver disease. J Hepatol 2014;60:69-77.

41) Pavlides M, Banerjee R, Tunnicliffe EM, Kelly C, Collier J, Wang LM, et al. Multi-parametric magnetic resonance imaging for the assessment of non-alcoholic fatty liver disease severity. Liver Int 2016;doi: 10.1111/liv.13284.

42) Pavlides M, Banerjee R, Sellwood J, Kelly CJ, Robson MD, Booth JC, et al. Multiparametric magnetic resonance imaging predicts clinical outcomes in patients with chronic liver disease. J Hepatol 2016;64:308-315.

43) Kaswala DH, Lai M, Afdhal NH. Fibrosis assessment in nonalcoholic fatty liver disease (NAFLD) in 2016. Dig Dis Sci 2016;61:1356-1364.

44) Pelusi S, Petta S, Rosso C, Borroni V, Fracanzani AL, Dongiovanni P, et al. Renin-angiotensin system inhibitors, type 2 diabetes and fibrosis progression: An observational study in patients with nonalcoholic fatty liver disease. PLoS ONE 2016;11:e0163069.

45) Harrison SA, Abdelmalek MF, Trotter JF, Paredes AH, Arnold HL, Kugelmas M, et al. NGM282, a novel variant of FGF19, significantly reduces hepatic steatosis and key biomarkers of NASH: results of a Phase 2, multicenter, randomized, double-blinded, placebo controlled trial in biopsy-confirmed NASH patients. J Hepatol 2017;66:S92-S93.

46) Harrison SA, Rinella ME, Abdelmalek MF, Trotter JF, Paredes AH, Arnold HL, et al. NGM282 for treatment of non-alcoholic steatohepatitis: a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. Lancet 2018;391:1174-1185.

47) Sanyal A, Charles ED, Neuschwander-Tetri BA, Loomba R, Harrison SA, Abdelmalek MF, et al. Pegbelfermin (BMS-986036), a PEGylated fibroblast growth factor 21 analogue, in patients with non-alcoholic steatohepatitis: a randomised, double-blind, placebo-controlled, phase 2a trial. The Lancet 2018:http://dx.doi.org/10.1016/S0140-6736(1018)31785-31789.

48) Vuppalanchi R, Siddiqui MS, Van Natta ML, Hallinan E, Brandman D, Kowdley K, et al. Performance characteristics of vibration-controlled transient elastography for evaluation of non-alcoholic fatty liver disease. Hepatology 2018;67:134-144.

49) Cassinotto C, Boursier J, de Ledinghen V, Lebigot J, Lapuyade B, Cales P, et al. Liver stiffness in nonalcoholic fatty liver disease: A comparison of supersonic shear imaging, FibroScan, and ARFI with liver biopsy. Hepatology 2016;63:1817-1827.

50) Helmke S, Colmenero J, Everson GT. Noninvasive assessment of liver function. Curr Opin Gastroenterol 2015;31:199-208.

Table 1. Major current phase 2b and 3 drug trials with pending results in the United States and Europe

Drug/Trial Name	Phase	Primary endpoint(s)	Long-term outcomes assessment	Enrollment Start date
Obeticholic acid (REGENERATE, FXR ligand) NCT02548351	3	≥ 1 stage fibrosis improvement with no worsening of NASH, or NASH resolution with no worsening of fibrosis at 18 months	All cause death, MELD ≥ 15, liver transplant, HCC, ascites requiring treatment, hospitalization for variceal bleeding, HE, or SBP at 7 years	Sept 2015
Elafibranor (RESOLVE-IT, PPARα/δ ligand) NCT02704403	3	NASH resolution with no worsening of fibrosis at 72 weeks	All-cause mortality, cirrhosis, liver-related outcomes at 4 years	Mar 2016
Selonsertib in stage 4 fibrosis (STELLAR-4, ASK-1 inhibitor) NCT03053063	3	≥ 1 stage fibrosis improvement with no worsening of steatohepatitis at 48 weeks	Time to first clinical event at 240 weeks	Jan 2017
Selonsertib in stage 3 fibrosis (STELLAR-3, ASK-1 inhibitor) NCT03053050	3	≥ 1 stage fibrosis improvement with no worsening of steatohepatitis at 48 weeks	Time to first clinical event at 240 weeks	Feb 2017
Cenicriviroc (AURORA, CCR2/5 inhibitor) NCT03028740	3	≥ 1 stage fibrosis improvement with no worsening of NASH at 12 months	All-cause mortality, cirrhosis, liver-related clinical outcomes at 5 years	Apr 2017
Emricasan (ENCORE-NF, pan-caspase inhibitor) NCT02686762	2b	≥ 1 stage fibrosis improvement with no worsening of NASH at 72 weeks	None	Jan 2016
MSDC 0602K (EMMINENCE, mTOT inhibitor) NCT02784444	2b	≥ 2 point reduction in NAS with no worsening of fibrosis at 12 months	None	July 2016
Tropifexor (FLIGHT-FXR, FXR ligand)	2b	AEs, ALT, AST, PDFF at 12 weeks;	None	Aug 2016

NCT02855164		NAS and fibrosis changes at 48 weeks		
Semaglutide (GLP-1 receptor agonist) NCT02970942	2b	NASH resolution with no worsening of fibrosis at 72 weeks	None	Nov 2016
Lanifibranor (NATIVE, pan-PPAR ligand) NCT03008070	2b	≥2 point reduction in SAF at 24 weeks	AEs at 24 weeks	Feb 2017
Emricasan in decompensated cirrhosis (ENCORE-LF, pan-caspase inhibitor) NCT03205345	2b	Event-free survival at 48-120 weeks	Change in MELD and Child-Pugh scores, new decompensation event, liver transplantation, all-cause and liver-specific mortality, quality of life at 120 weeks	June 2017
Obeticholic acid in cirrhosis (REVERSE, FXR ligand) NCT03439254	2b	≥ 1 stage fibrosis improvement with no worsening of NASH at 12 months	NoneNone	Aug 2017
Selonsertib, GS-0976, GS-9674 in stage 3-4 fibrosis (ATLAS, ASK-1 inhibitor, ACC inhibitor, FXR ligand) NCT03449446	2b	≥ 1 stage fibrosis improvement with no worsening of NASH at 48 weeks	None	Mar 2018
Seladelpar (PPARδ ligand) NCT03551522	2b	Relative change in MRI-PDFF at 12 weeks	AEs at 52 weeks	Apr 2018
Tropifexor + cenicriviroc (TANDEM, FXR ligand, CCX2/5 inhibitor) NCT03517540	2b	≥ 1 stage fibrosis improvement, NASH resolution at 48 weeks	AEs at 48 weeks	Sept 2018

NASH: nonalcoholic steatohepatitis; NAS: NAFLD activity score; MELD: model for end-stage liver disease; HCC: hepatocellular carcinoma; HE: hepatic encephalopathy; SBP: spontaneous bacterial peritonitis; SAF: steatosis, activity (inflammation + ballooning), fibrosis score; PPAR: peroxisomal proliferator active receptor; FXR: farnesoid-X receptor; ASK-1: apoptosis signaling-regulated kinase-1; CCR2/5: chemokine receptor 2 and 5; mTOT: mitochondrial target of thiazolidinediones; ACC: acetyl-CoA carboxylase; GLP-1: glucagon-like peptide-1, AE: adverse event

Table 2: Disease Definitions and Common Methods of Assessment

	NAFLD[a]	NAFL	NASH	NASH cirrhosis
Defining features	Steatosis identified by any means	Steatosis with insufficient ballooning or inflammation to diagnose NASH	Steatosis with ballooning and inflammation in a characteristic pattern	Cirrhosis with current or past clinical or histological evidence of steatosis or steatohepatitis
Liver Biopsy Steatosis[b] Inflammation Ballooning Fibrosis[c]	 + any degree any degree any degree	 + none to minimal none to minimal uncommon	 +[d] + + any degree	 +/-[e] +/-[e] +/-[e] substantial, disrupts normal architecture
Serum AST and ALT Platelets INR	 ↑/- - -	 ↑/- - -	 ↑/- -[f] -	 ↑/- ↓/- ↑/-
Imaging	US: increased hepatic echogenicity indicates NAFLD CT: decreased hepatic attenuation indicates NAFLD MRI: increased signal intensity on T1-weight images, signal loss from in-phase to opposed-phase indicates NAFLD MRI-PDFF: quantifies steatosis MRE: assesses fibrosis VCTE, SWE, ARFI: identifies advanced fibrosis (stage 3 and 4) No defining features on current imaging modalities to differentiate between NAFL and NASH			Advanced cirrhosis: • Liver surface nodularity • Findings of portal hypertension: ascites, splenomegaly, collateral vessels/varices

NAFLD: nonalcoholic fatty liver disease; NAFL: nonalcoholic fatty liver; NASH: nonalcoholic steatohepatitis; AST: aspartate aminotransferase; ALT: alanine aminotransferase; INR: prothrombin time internationalized ratio; US: ultrasound; CT: computerized tomography scan; MRI: magnetic resonance imaging; MRI-PDFF: MRI-estimated proton density fat fraction; MRE: magnetic resonance elastography; VCTE: vibration controlled transient elastography; SWE: shear wave elastography; ARFI: acoustic radiation force impulse imaging

[a] NAFLD encompasses both NAFL and NASH; lack of alcohol abuse as a cause of steatosis is confirmed clinically

[b] histologic steatosis: >5% of hepatocytes with lipid droplets

[c] fibrosis: accumulation of excess extracellular matrix proteins (primarily collagens), typically pericellular zone 3 fibrosis in adults in the early stages

[d] steatosis may be < 5% in some cases of NASH with ballooning and inflammation sufficient to meet criteria for NASH

[e] features of steatosis, inflammation, and ballooning may no longer be present in some cases of "burnt out" NASH cirrhosis

[f] thrombocytopenia is a surrogate marker for portal hypertension that typically develops with advanced fibrosis or cirrhosis but in rare cases may be present in NASH with earlier stages of fibrosis

Table 3. Proposed terminology to describe changes in the context of treatment trials.

	Definition	Comments
Resolution of NAFLD	• Reduction of steatosis to grade 0[a], or • Reduction of PDFF to < 5%	If inflammation and ballooning persist with even trivial steatosis, histologic criteria for NASH may still be met
Improved NAFLD	• ≥1 grade reduction of steatosis[a], or • ≥10% absolute reduction of PDFF[a], or • ≥30% relative reduction of PDFF (see text)	Independent of changes in inflammation, ballooning or fibrosis
Improved NASH	• ≥1 point reduction in ballooning[b], and • ≥2 point reduction in NAS[d]	Independent of changes in fibrosis Acceptable as an endpoint in early phase trials but not to support a marketing application
Resolution of NASH (same as "complete resolution" of NASH)	• Disappearance of hepatocyte ballooning[b] (grade 0), and • Absent or minimal lobular inflammation[c] (grade 0 or 1)	Independent of changes in steatosis or fibrosis
Worsened NASH	• ≥1 point increase in ballooning[b], or • ≥1 point increase in inflammation[c]	Independent of changes in steatosis or fibrosis Characterizing the changes in inflammation (cell type and distribution) with treatment requires further analysis
Not worsened NASH	• No increase in ballooning, inflammation or steatosis	Independent of changes in fibrosis
Improved fibrosis	• ≥1 stage reduction in fibrosis[e] • Not yet defined for continuous variables such as liver stiffness by elastography or collagen proportional area by image analysis of collagen staining	Independent of changes in steatohepatitis
Worsened fibrosis	• ≥1 stage increase in stage of fibrosis[e] • Not yet defined for continuous variables such as liver stiffness by	Independent of changes in steatohepatitis

	elastography or collagen proportional area by image analysis of collagen staining	

PDFF: proton density fat fraction; HVPG: hepatic venous pressure gradient; NAS: NAFLD activity score

[a] Steatosis is graded by % of hepatocytes involved: grade 0 is <5%, grade 1 is 5-33%, grade 2 is >33-66%, grade 3 is >66%[(9)]; criteria for meaningful change in steatosis assessed by all measures is rapidly evolving

[b] Ballooning is graded by number of hepatocytes with changes: grade 0 is none, grade 1 is few cells, grade 2 is many cells/prominent ballooning[(9)]

[c] Lobular inflammation is graded by number of inflammatory foci in 200x field: grade 0 is no foci, grade 1 is <2 foci, grade 2 is 2-4 foci, grade 3 is >4 foci[(9)]

[d] NAS is the sum of the graded scores for steatosis, ballooning, and lobular inflammation[(9)]

[e] Fibrosis is staged by patterns of fibrosis: stage 1 is perisinusoidal or periportal fibrosis, stage 2 is perisinusoidal and portal/periportal fibrosis, stage 3 is bridging fibrosis, stage 4 is cirrhosis[(9)]

[a, b, c, d, e] Grading and staging determined by an expert liver pathologist assessing the overall pattern of injury