Case Definitions for Inclusion and Analysis of Endpoints in Clinical Trials for Nonalcoholic Steatohepatitis Through the Lens of Regulatory Science

Mohammad Shadab Siddiqui,1* Stephen A. Harrison,2* Manal F. Abdelmalek,3 Quentin M. Anstee,4 Pierre Bedossa,5 Laurent Castera,6 Lara Dimick-Santos,7 Scott L. Friedman,8 Katherine Greene,14 David E. Kleiner,9 Sophie Megnien,10 Brent A. Neuschwander-Tetri,11 Vlad Ratziu,12 Elmer Schabel,13 Veronica Miller,14 and Arun J. Sanyal1; on behalf of the Liver Forum Case Definitions Working Group

Nonalcoholic steatohepatitis (NASH) is an important cause of liver-related morbidity and mortality. There are no approved therapies, and the results of clinical trials have been difficult to compare due to inconsistent definitions of relevant disease parameters in patients with NASH. The natural course of the disease has not been rigorously characterized, particularly with respect to the contributions of underlying obesity, type 2 diabetes, and other comorbidities and the treatments provided for these comorbidities. Efforts to perform analyses of pooled data are limited by heterogeneous case definitions used across studies to define disease states. There remains a major unmet need in the field to develop standardized definitions for populations for interventional trials. Such definitions are expected to impact how endpoints for clinical trials are constructed. The Liver Forum is a multistakeholder effort including US and European regulatory agencies, academic investigators, professional and patient representative organizations, and industry to catalyze therapeutic development for NASH by developing potential solutions to barriers to development. The Case Definitions Working Group was established by The Liver Forum to evaluate the validity of case definitions for populations to be included in clinical trials for NASH from a regulatory science perspective. Based on such analyses, specific recommendations are provided noting the strengths and weaknesses of the case definitions along with knowledge gaps that require additional study. (HEPATOLOGY 2018;67:2001-2012)

Nonalcoholic fatty liver disease (NAFLD) affects one third of adults in Western nations.1) NAFLD may manifest histologically as a fatty liver (FL) or nonalcoholic steatohepatitis (NASH). FL progresses slowly compared to NASH, which is more likely to lead to progressive fibrosis and cirrhosis.2,3) NASH cirrhosis is rapidly increasing as an indication for transplantation,4) and while pharmacologic therapy is not yet available, NASH is increasingly being targeted by drug development efforts. Regulatory pathways for developing diagnostic tests and therapeutic interventions for NASH require granular understanding of disease course and associated clinical outcomes. Despite a plethora of literature, the variable case definitions, methods for identification of subjects, and assessment of outcomes limit the ability to anchor diagnostics and therapeutic development on a robust model of the disease, impeding assessment of diagnostics and benefit of therapeutic interventions.

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; AUROC, area under the receiver operating characteristic curve; CAP, controlled attenuation parameter; FDA, Food and Drug Administration; FL, fatty liver; HVPG, hepatic venous pressure gradient; LSM, liver stiffness measurement; MRE, magnetic resonance elastography; MRI-PDFF, magnetic resonance imaging–estimated proton density fat fraction; NAFLD, nonalcoholic fatty liver disease; NAS, NAFLD activity score; NASH, nonalcoholic steatohepatitis; NASH CRN, NASH Clinical Research Network; SAF, Steatosis–Activity–Fibrosis; VCTE, vibration-controlled transient elastography.

Received February 15, 2017; accepted October 13, 2017.
Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.29607/suppinfo.
The Liver Forum was established in 2014 following a workshop sponsored by the US Food and Drug Administration (FDA) and the American Association for the Study of Liver Diseases (AASLD) and provides a mechanism for dialogue between regulatory agencies and other stakeholders in a neutral, collaborative, and non-binding manner to accelerate development efforts for NAFLD. The Liver Forum model involves collaboration between the FDA, the European Medicines Agency, academic investigators, industry stakeholders, and patient representatives. A key priority identified was the need for case definitions that meet regulatory standards and are uniformly applicable across clinical trials, which is essential for comparing the utility of various therapies and demonstrating overall benefit of therapeutic interventions. Standardized definitions will facilitate integration of new diagnostic tools into the overall drug development plan for NASH and will allow future pooled analyses to robustly model the disease, taking into account effects of comorbidities and their treatments.

The Case Definitions Working Group analyzed existing literature, generated the “best” definitions from a regulatory science perspective, and identified gaps in the field. This article represents the working group’s output, which has been presented to and critiqued by the overall Liver Forum membership. Recommendations put forth in this article are for the purposes of regulatory science rather than clinical practice guidance. These definitions and...
recommendations are nonbinding and do not represent official positions of the FDA or the European Medicines Agency.

Methods

The working group was divided into subgroups to evaluate the case definitions currently used for the following NAFLD phenotypes: (1) FL, (2) indeterminate NASH, (3) NASH without fibrosis, (4) NASH with early fibrosis, (5) NASH with bridging fibrosis, (6) compensated cirrhosis, and (7) decompensated cirrhosis.

Following a standardized discussion format, members described the characteristic clinical phenotype, liver histology, and means of noninvasive assessment. The data were synthesized to summarize whether definitions were objective, quantifiable, analyzable using quantitative approaches, sensitive to change, and logistically feasible to operationalize in the context of multi-center clinical trials. The relationship of definitions to clinical outcomes was assessed from the literature. The findings of the subgroups were discussed within the larger working group and presented to Liver Forum members for feedback. Consensus regarding optimal definitions was not imposed but allowed to develop through discussion. Areas requiring additional data were identified.

We present the salient points raised by the working group, which were reviewed within the FDA and the European Medicines Agency for alignment with regulatory requirements. Our intent is to identify populations for pooled outcomes analyses, development of patient-reported outcomes instruments, and diagnostics. We include limited references due to space constraints, with a list of additional references reviewed provided in the Supporting Information.

Results

The core requirement defining NAFLD is documentation of hepatic steatosis that is not related to excessive alcohol consumption. For the purpose of this report, we assume that excessive alcohol consumption has been excluded along with secondary causes of hepatic steatosis (Supporting Information).

The subpopulations were collapsed into three broad groups: (1) non-NASH, (2) NASH, and (3) NASH cirrhosis. Each comprises several subcategories with regard to clinical phenotype, histology, and noninvasive assessment.

The field of NAFLD is anchored by histological assessment, and the National Institute of Diabetes and Digestive and Kidney Diseases’ NASH Clinical Research Network (CRN) system of grading disease activity and staging fibrosis separately is the most widely used and validated. The Steatosis–Activity–Fibrosis (SAF) reports but does not add the three major histological features (steatosis, activity, and fibrosis) into one formula, permitting greater granularity; however, SAF has not been used in phase 2B or 3 clinical trials. The Goodman classification combines disease activity and fibrosis, with greater weight given to advanced fibrosis. It is the least validated system and was included to determine if it adds utility from a regulatory perspective to more established systems.

ARTICLE INFORMATION:

From the 1Virginia Commonwealth University, Richmond, VA; 2Pinnacle Clinical Research, San Antonio, TX; 3Division of Gastroenterology & Hepatology, Duke University School of Medicine, Durham, NC; 4Institute of Cellular Medicine, Newcastle University, Newcastle Upon Tyne, United Kingdom; 5Department of Pathology, Physiology and Imaging, University of Paris Diderot, Clichy, France; 6Department of Hepatology, Hopital Beaujon, Clichy, France; 7Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Springs, MD; 8Division of Liver Disease, Icahn School of Medicine at Mount Sinai, New York, NY; 9National Cancer Institute, Center for Cancer Research, Bethesda, MD; 10Genfit Corp, Cambridge, MA; 11Division of Gastroenterology and Hepatology, Saint Louis University School of Medicine, St. Louis, MO; 12Department of Hepatology and Gastroenterology, Hopital Pitie Salette et Universite Pierre et Marie Curie, Paris, France; 13Gastroenterology & Hepatology Unit, Bundesinstitut fur Arzneimittel und Medizinprodukte, Bonn, Germany; 14Forum for Collaborative Research, University of California Berkeley, Washington, DC.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO:

Veronica Miller, Ph.D.
Forum for Collaborative Research
1608 Rhode Island Ave, Suite 212
Washington DC 20036
E-mail: veronicam@berkeley.edu
Tel: +1-202-974-6290

2003
Non-NASH

Non-NASH includes the following histological subtypes: FL, indeterminate NASH, and steatofibrosis. Although described as separate entities for the sake of regulatory classification, clinically they may represent overlapping points on the histological spectrum. Compared to NASH, FL has a low risk of liver-related outcomes\(^{(2,3,9)}\). In indeterminate NASH, inflammation and hepatocyte injury are present but atypical of definite steatohepatitis\(^{(6)}\) and steatofibrosis represents non-NASH with fibrosis\(^{(7)}\). The natural history of these phenotypes is not well characterized, and they are included to distinguish them from NASH. If future studies indicate that the natural histories of these are similar to NASH, they may be reclassified as “at-risk” disease states.

HISTOLOGY

All histological systems define non-NASH by the presence of predominantly macrovesicular steatosis in \(\geq 5\%\) of hepatocytes\(^{(6-8,11)}\). There is relatively good intraobserver and interobserver concordance with steatosis assessment (Table 1)\(^{(6,10)}\). While microvesicular steatosis can occur in non-NASH, the presence of predominantly microvesicular steatosis should raise concern for an alternate etiology.

Steatosis severity is categorized ordinally by all histological systems (Supporting Table S1) and relates linearly to techniques measuring hepatic triglyceride content\(^{(6-8,11)}\). However, steatosis represents a continuous process, and ordinal classification introduces assessment errors at categorical boundaries. Despite this, the intraobserver and interobserver concordance \(\kappa\) statistic for assessment of steatosis severity is 0.79 and 0.83, respectively\(^{(6)}\).

Mild lobular inflammation can coexist with steatosis and does not appear to impact progression to cirrhosis\(^{(6,12)}\). While data quality is not optimal, steatosis with grade 1 inflammation is included in the current spectrum of non-NASH.

Indeterminate NASH, a subclassification only recognized by the NASH CRN system, is characterized by grade \(\geq 1\) steatosis with mild lobular inflammation and none to rare ballooning\(^{(6)}\). It can be associated with varying amounts of fibrosis and may represent more advanced disease than FL. Hepatic steatosis with varying fibrosis stages short of cirrhosis in the absence of inflammation and ballooning has been observed, but no data link it to clinical outcomes\(^{(6,7)}\). This entity is classified as steatofibrosis to allow for recognition and classification. Better natural history data are needed on these phenotypes before large-scale efficacy trials are performed.

CLINICAL PHENOTYPE

No clinical phenotype permits identification of subjects with non-NASH with high specificity. Many subjects have one or more of the features of metabolic syndrome, and polycystic ovarian syndrome may be present in women\(^{(13,14)}\). Although common, overt insulin resistance is not a diagnostic requisite\(^{(14,15)}\).

NONINVASIVE DIAGNOSIS

Aggregate scores from routine bioclinical data have been investigated for their accuracy in predicting steatosis (Table 2). The most promising include the FL index, the hepatic steatosis index, the NAFLD-liver fat score, the visceral adiposity index, the lipid

### TABLE 2. Accuracy of Noninvasive Biomarkers in Diagnosing and Quantifying Hepatic Steatosis

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Models</th>
<th>CK-18</th>
<th>Ultrasound</th>
<th>CAP</th>
<th>MRI/MRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Subjective</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Quantifiable (separation of steatosis grade)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Interobserver reliability</td>
<td>N/A</td>
<td>N/A</td>
<td>Moderate</td>
<td>Moderate</td>
<td>High</td>
</tr>
<tr>
<td>Sensitive to change</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Yes</td>
</tr>
<tr>
<td>Used in clinical trials</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2. Accuracy of Noninvasive Biomarkers in Diagnosing and Quantifying Hepatic Steatosis

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Subjective</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Quantifiable (separation of steatosis grade)</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Interobserver reliability</td>
<td>N/A</td>
<td>N/A</td>
<td>Moderate</td>
<td>Moderate</td>
<td>High</td>
</tr>
<tr>
<td>Sensitive to change</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Yes</td>
</tr>
<tr>
<td>Used in clinical trials</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Abbreviations: CK-18, cytokeratin 18; MRS, magnetic resonance spectroscopy; N/A, not analyzed.
accumulation product, and the triglyceride × glucose index. The sensitivity and specificity of these for detecting >5% and >33% steatosis are summarized in Supporting Table S2. While not sufficiently accurate for use as inclusion criteria, they may be used to enrich populations screened. No circulating biomarkers, including serum aminotransferases, are sufficiently accurate as stand-alone measures of hepatic steatosis for clinical trials.

IMAGING MODALITIES

Ultrasonography relies on increased hepatic echogenicity and blurring of deep vessels and is a sensitive but relatively nonspecific way to identify steatosis (Supporting Table S2); however, sensitivity drops when used to detect 5%-20% steatosis. The hepatorenal index measures the mean ratio of echo intensities of the liver and renal cortex and is highly sensitive and specific in detecting >5% hepatic steatosis (Table 2). While unsuitable as an entry criterion for enrollment in advanced-phase NASH clinical trials, it may be useful in identifying patients for proof-of-concept trials.

Vibration-controlled transient elastography (VCTE) assesses steatosis by reporting the loss of ultrasound signal through the liver, which is reported as the controlled attenuation parameter (CAP). In a meta-analysis, CAP had sensitivity and specificity of 78% and 79%, respectively, for detecting ≥S1 steatosis. A CAP > 236 dB/m has sensitivity and specificity of 82% and 91%, respectively, for detecting steatosis. However, CAP is less robust in discriminating between steatosis grades, with an area under the receiver operating characteristic curve (AUROC) of 0.73 and 0.70 for distinguishing S3 versus S0-2 and S2-3 versus S0-1, respectively (Table 2). CAP can be used as an enrichment tool along with elastography to identify subjects for (1) biopsy for advanced-phase trials, (2) proof-of-concept trials, or (3) in trials in subjects with metabolic syndrome where NAFLD as a covariate needs to be documented.

Magnetic resonance spectroscopy and magnetic resonance imaging–estimated proton density fat fraction (MRI-PDFF) quantify steatosis (Supporting Table S2). Although magnetic resonance spectroscopy is highly accurate, it only measures fat in small regions of interest, while MRI-PDFF allows mapping of the entire liver. Both have been validated against hepatic triglyceride content and histological assessment of steatosis and are sensitive to change. MRI-PDFF is reproducible across scanners of varying field strength and outperformed CAP in steatosis detection. MRI-PDFF is a validated modality that can be used in early-phase studies to detect and assess change in steatosis (Table 2).

In summary, histological assessment represents the current standard for detecting hepatic steatosis. Non-NASH is distinguished from NASH by having steatosis along with no or minimal disease activity. MRI-PDFF is an accurate and sensitive method for steatosis assessment and quantification and useful in early-phase studies evaluating drug efficacy on steatosis. FL, indeterminate steatohepatitis, and steatofibrosis are histological phenotypes whose natural histories are not well defined. These populations may be studied where specific questions regarding these phenotypes are posed or to confirm if responses to treatment are similar to those with NASH.

NASH

The broad category of NASH includes definite NASH, NASH with early fibrosis (stages 1a, 1b, 1c, 2; Supporting Table S1), and NASH with bridging fibrosis. NASH with cirrhosis is considered separately because the clinical, laboratory, histological, and imaging phenotype along with goals of treatment and study design are different from those in patients with precirrhosis NASH.

HISTOLOGY

The NASH CRN system defines NASH by a pattern of injury composed of steatosis, lobular inflammation, and ballooning degeneration. The SAF system requires a grade of 1 or higher for steatosis, lobular inflammation, and ballooning. Neither NAFLD activity score (NAS) nor SAF activity score is influenced by portal inflammation or fibrosis. The NAS is related to the probability of NASH; however, definite NASH cannot be inferred from the NAS alone. Histological diagnosis of NASH has been linked to risk of disease progression and is sensitive to change.
In the Goodman classification, NASH is diagnosed by hepatic steatosis accompanied by either (1) centrilobular inflammation and/or Mallory–Denk bodies or (2) centrilobular pericellular/perisinusoidal fibrosis or bridging fibrosis. This definition was a better predictor of liver-related mortality than the NASH CRN system, consistent with recognition that fibrosis is a primary driver of liver-related mortality. However, whether improvement in NASH disease activity related to effective therapy can be assessed by composite scores including noninvasive measures of fibrosis remains unknown. Additional data linking different diagnostic criteria to clinically meaningful outcomes are needed to validate entry and response criteria, as generally acceptable surrogates for regulatory purposes.

DISEASE ACTIVITY

Disease activity refers to histological findings reflective of activation of disease pathways driving the disease toward cirrhosis. The NAS grades steatosis (0–3), lobular inflammation (0–2), and ballooning (0–2). It has been validated in terms of intraobserver and interobserver variability and sensitivity to change (Table 3) but has not yet been shown to relate to clinical outcomes. The activity score component of the SAF is the sum of lobular inflammation and ballooning but does not include steatosis. It has not been tested against outcomes in prospective, adequately powered, long-term trials. The two systems are comparable in terms of observer variation if expert histopathologists are used. The Goodman classification has not been externally validated.

STAGING

The NASH CRN system categorizes fibrosis progressively from early fibrosis (stages 1a, 1b, 1c, 2; Supporting Table S1) to bridging (stage 3) and cirrhosis (stage 4). The NASH CRN fibrosis staging system has been validated with respect to intraobserver and interobserver concordance and sensitivity to change. No data indicate the NASH CRN subclassifications of stage 1 fibrosis into 1a, 1b, and 1c are linked to risk of fibrosis progression or clinical outcomes. In the Goodman classification, fibrosis is subdivided into portal fibrosis, bridging fibrosis, or cirrhosis with additional subcategories (Supporting Table S1). This is not yet externally validated. Similarly, quantitative assessment of fibrosis content by morphometry needs additional data in NASH. For purposes of clinical trials, precirrhosis stages of NASH are defined by the presence of steatohepatitis with fibrosis stage ≤3.

CLINICAL PHENOTYPE

The clinical phenotype of NASH is similar to non-NASH, and there are no clinical features that differentiate between non-NASH and NASH. The likelihood of NASH and advanced fibrosis increases with age, body mass index, presence of type 2 diabetes mellitus, and metabolic syndrome. There are no data relating the duration of these conditions with the likelihood of definite NASH.

NONINVASIVE METHODS

Liver biopsy is subject to sampling variability; however, there are no noninvasive tools to diagnose NASH with enough accuracy to serve as stand-alone entry criteria for advanced-phase trials to avoid using histology-based endpoints (Supporting Table S3). For instance, the sensitivity and specificity of cytokeratin-18 fragments for detecting NASH are 58% and 68%, respectively, making this an unacceptable diagnostic tool from a regulatory perspective. Most clinical trials focus on subjects with NASH with some degree of fibrosis due to logistical challenges of demonstrating clinically meaningful benefit in those without fibrosis, who have a very low rate of progression to cirrhosis and liver-related outcomes. Moreover, if improvement in fibrosis is sought, then some degree of fibrosis is needed at trial entry.

The elements identifying the population of interest are NASH and fibrosis stage. As noted, no noninvasive laboratory biomarkers can reliably identify the presence of NASH. Noninvasive fibrosis models including aspartate aminotransferase (AST)-to-platelet ratio index, FIB4, FibroTest, FibroMeter, NAFLD fibrosis score, and enhanced liver fibrosis score all have a high negative predictive value (>80%) for excluding fibrosis stage ≥3 (Supporting Table S4). However, the reliability of these methods is greatly decreased for differentiating various fibrosis stages and between bridging fibrosis and cirrhosis.

The AUROCs for distinguishing F2 versus higher fibrosis stages using VCTE are 0.79–0.87 using liver stiffness measurement (LSM) cutoffs of 6.7–7.7 kPa. The diagnostic accuracy of VCTE to distinguish F3 from lesser fibrosis stages is 0.77–0.98 using cutoff values of 8.0–10.4 kPa. The main drawback
of VCTE is its high failure rate in obese patients using the M probe, which can be circumvented by using the XL probe.\(^{(22,29)}\) Similarly, magnetic resonance elastography (MRE) had an AUROC of 0.82 and 0.87 for detecting any fibrosis and advanced fibrosis, respectively.\(^{(22)}\) Using previously reported cutoffs, VCTE and MRE can be used to exclude the presence of cirrhosis in \(>95\%\) of cases.\(^{(21,22)}\) Three studies have evaluated the accuracy of VCTE and MRE in head-to-head comparison.\(^{(21,22,29)}\) MRE outperformed VCTE for detection of any fibrosis (AUROC 0.82 versus 0.67, \(P = 0.012\)) but had similar accuracy for detecting other fibrosis stages.\(^{(22)}\) In patients with body mass index \(\geq 35\, \text{kg/m}^2\), MRE had a lower failure rate than VCTE (5% versus 19%) despite use of the XL probe.\(^{(29)}\) Accuracy of MRE and VCTE was similar in patients with successful examinations; however, MRE outperformed VCTE in detection of moderate and advanced fibrosis when both reliable (interquartile range/median LSM value \(\leq 30\%\) or interquartile range/median LSM \(> 30\%\) if LSM value \(< 7.1\, \text{kPa}\)) and unreliable examinations were included.

There are limited data comparing the accuracy of VCTE and MRE, and use of one over the other cannot be recommended.\(^{(21,22)}\) Both VCTE and MRE have been shown to detect cirrhosis with an AUROC >0.9; however, sample size in MRE studies is small, and more data are needed.\(^{(21,22,29,38)}\)

Several studies have examined combining noninvasive markers of steatosis with markers of cellular injury or fibrosis as a surrogate marker of NASH. This approach is based on the concept that fibrosis develops primarily in those with steatohepatitis rather than steatosis, although recent identification of fibrosis development in those with steatosis alone challenges this assumption. A study combining CAP score \(>250\, \text{dB/second},\) LSM \(> 7\, \text{kPa},\) and alanine aminotransferase (ALT) \(> 60\, \text{IU/L}\) identified NASH with an AUROC of 0.81.\(^{(40)}\) In a preliminary report of screening data from a phase 2B trial, the presence of two or more features of metabolic syndrome and a FIB4 \(> 1.1\) identified NASH with fibrosis, with a positive predictive value of 88%.\(^{(41)}\) The potential for using two FIB4 cutoffs (to exclude those without fibrosis and to exclude cirrhosis) along with imaging documenting the presence of steatosis to identify definite NASH with precirrhosis stages of fibrosis needs further exploration but may be useful for population enrichment.

In summary, for phase 2B and 3 trials, liver biopsy is required for the diagnosis of NASH. Activity and fibrosis stage should be defined by NASH CRN criteria. For phase 1 and 2A trials of short duration where risks and costs of liver biopsy are not justifiable, the following criteria can be used to define the population of interest, realizing that the positive predictive value is 0.70–0.80: (1) two or more features of metabolic syndrome, (2) evidence of steatosis by MRI-PDFF or CAP, and (3) evidence of \(> F0\) fibrosis by either VCTE (\(> 7.0\, \text{kPa}\)) or MRE (\(> 2.88\, \text{kPa}\)). Magnetic resonance–based methods are highly accurate and validated for the diagnosis of steatosis or precirrhosis stages of fibrosis. VCTE may alternatively be used when centers without MRE capability are involved or in exploratory studies. This approach is most applicable when the agents’ safety profile can be assumed to be excellent based on preclinical or early-phase studies or when a drug with extensive safety data is being repurposed for NASH.

**NASH Cirrhosis**

NASH with cirrhosis and cryptogenic cirrhosis due to NASH were both considered here. Following development of cirrhosis, subjects can remain stable and asymptomatic before complications of cirrhosis appear. It is important to consider how the diagnosis was established and whether the patient meets criteria for compensated or decompensated cirrhosis with NASH as the etiology of the cirrhosis.

**HISTOLOGY**

Histology represents the current standard for diagnosing cirrhosis, although certain clinical features can establish the diagnosis in the absence of histology. Histological diagnosis of cirrhosis requires widespread architectural disruption and annular fibrosis surrounding hepatocyte nodules.\(^{(6,42)}\) According to NASH CRN criteria, cirrhosis is described as present or absent, while in the Goodman classification cirrhosis is graded as absent, incomplete, or established.\(^{(6-8)}\)

**CLINICAL PHENOTYPE**

Cirrhosis may be suspected in compensated subjects from laboratory tests such as decreased platelet counts and/or an AST/ALT ratio > 1.\(^{(43)}\) However, an AST/ALT ratio > 2 should raise the possibility of alternate etiologies.\(^{(44)}\) A clinical diagnosis can only be made if there is accompanying portal hypertension, with the presence of splenomegaly, varices, or other signs of decompensation.
NONINVASIVE METHODS

AST/ALT ratio, FIB-4, BARD score, NAFLD fibrosis score, and AST-to-platelet ratio index are useful in identifying patients unlikely to have advanced fibrosis (Supporting Table S4) but less useful in predicting the presence of cirrhosis. Age-adjusted thresholds for FIB-4 and NAFLD fibrosis score further improve the specificity of these tests. There are limited data using these models to distinguish between bridging fibrosis and cirrhosis. FibroTest and the enhanced liver fibrosis score have also shown high specificity in ruling out advanced fibrosis, but utility in differentiating bridging fibrosis and cirrhosis remains uncertain.13

Cross-sectional imaging can provide convincing evidence of cirrhosis when a nodular liver contour or evidence of portal hypertension is identified; however, absence of these does not exclude cirrhosis. Pooled analysis of VCTE and MRE indicates that these can detect cirrhosis with a sensitivity and specificity between 63%-100% and 66%-92%, respectively.21,22,29 The threshold for distinguishing cirrhosis from lower stages in those with NASH is 14.0 kPa using VCTE and >4.52 kPa using MRE, but thresholds vary based on the study (Supporting Table S5).21 VCTE can predict the development of portal hypertension–related complications; however, total subjects with cirrhosis in these studies are low. In the absence of elastography, the appearance of nodular liver on cross-sectional imaging had positive predictive value and negative predictive values of 74% and 88%, respectively.45

Evidence of portal hypertension in the form of collaterals, identified by imaging, endoscopy, or splenomegaly, without portal or splenic vein thrombosis can be used to diagnose cirrhosis in patients with chronic liver disease.48 Similarly, an increase in hepatic venous pressure gradient (HVPG), especially > 10 mm Hg, is diagnostic of sinusoidal portal hypertension, which may be caused by both acute and chronic liver disease. There are multiple caveats for the performance of high-quality HVPG measures, and when done correctly, the results are relatively reproducible.48 There are limited data on the ability of HVPG alone to diagnose cirrhosis of any etiology, including NASH. There is a similar paucity of rigorously performed studies on day-to-day variability of HVPG, and even the coefficient of variation of the measure itself when performed multiple times in succession. Available literature suggests that the coefficient of variation is low if the procedure is executed properly.49

In subjects with ascites, a serum to ascites albumin gradient of >1.1g/dL is diagnostic of sinusoidal portal hypertension–related ascites, but this can be due to cirrhosis or passive congestion.50

COMPENSATED VERSUS DECOMPENSATED CIRRHOSIS

The criteria for decompensated cirrhosis remain the same as those for compensated cirrhosis except that biopsy risks increase with development of decompensation, and most clinical texts recommend against percutaneous liver biopsies in those with ascites or coagulopathy.51 The presence of decompensation is identified by ascites, encephalopathy, variceal hemorrhage, or liver function failure characterized by elevated bilirubin level, international normalized ratio >1.4, and decreased albumin. There are multiple caveats that must be considered, such as increase in total bilirubin due to biliary obstruction or Gilbert syndrome, elevation of international normalized ratio due to malnutrition, or hypoalbuminemia due to nephrotic syndrome. These must be excluded if certainty is needed to show that changes in these indices reflect liver dysfunction. The risk of mortality is related to the Model for End-Stage Liver Disease (MELD) score, particularly as it rises over 1452, and A MELD score >14 may be considered a manifestation of decompensated cirrhosis.

DIAGNOSIS

The current standard for identifying NASH as the etiology of cirrhosis is histological evidence of NASH with cirrhosis or a prior liver biopsy demonstrating NASH and subsequent clinical or histological evidence of progression to cirrhosis. With development of cirrhosis, classical histological parameters may decrease or even disappear completely. However, on careful examination, up to 33% of patients with cryptogenic cirrhosis may have NASH.53 In true cryptogenic cirrhosis or where a recent biopsy is not available, a prior biopsy showing NASH along with an absence of other common causes of cirrhosis may be used to infer NASH as the cause of cirrhosis. If recent liver histology is unavailable and cirrhosis is inferred from an absence of alternate causes of cirrhosis, presence of NASH on a prior biopsy, or evidence of steatosis on recent or prior imaging along with multiple risk factors for NASH such as overweight–obesity, type 2 diabetes mellitus, hypertension, or
dyslipidemia. In the absence of risk factors or non-invasive evidence of steatosis, diagnosis of NASH as the etiology of cirrhosis is uncertain. Dyslipidemia and hypertension can improve with progression to cirrhosis.

In summary, histology remains the current standard for diagnosis of NASH cirrhosis. Presence of cirrhosis may be inferred from elastography with increased liver stiffness along with imaging or endoscopic evidence of collaterals, nodular liver, or splenomegaly without portal or splenic vein thrombosis. In the absence of radiographic or endoscopic evidence of cirrhosis, a diagnosis of cirrhosis can be inferred if patients have two or more of the following: (1) HVPG > 6 mm Hg, (2) AST/ALT ratio > 1 but < 2 and ALT < 10 x the upper limit of normal, or (3) platelet count < 150,000/mm³. Decompensated cirrhosis is defined by the presence of ascites, encephalopathy, variceal hemorrhage, Model for End-Stage Liver Disease score > 14, or an otherwise unexplained abnormality of ≥2 of bilirubin, international normalized ratio, or albumin. If biopsy is not available, the diagnosis of NASH cirrhosis may be inferred from the following after a negative laboratory workup for common causes of chronic liver disease (Supporting Table S5): (1) current documentation of hepatic steatosis with noninvasive profile of cirrhosis, (2) prior biopsy showing NASH with evidence of cirrhosis, (3) prior imaging demonstrating hepatic steatosis with evidence of cirrhosis, or (4) two or more risk factors for NASH with evidence of cirrhosis. The diagnosis of cirrhosis based on prior imaging or presence of risk factors may be suitable for early-phase clinical trials, though the level of evidence is not strong enough to be allowed for phase 2B or 3 trials.

**Discussion**

This document summarizes the current consensus or lack thereof regarding key elements required to diagnose phenotypes of NAFLD for clinical trials (Table 4). The choice of specific recommendations reflects their relationship to disease development and progression, sensitivity, specificity, reproducibility, and sensitivity to change. There are limited data to link such operational definitions to clinically meaningful outcomes; such data will be required to establish the foundation for development of surrogate endpoints for clinical trials that can be translated into clinical practice. The working group will revisit this topic and update recommendations as new data on existing and/or new biomarkers, for example, genetic polymorphisms, become available. Finally, a tabular format (Supporting Table S6) representing the population being studied in a clinical trial is presented to allow across-trial comparison of study populations.

**Acknowledgment:** The authors are grateful for support from the Case Definitions Working Group (Appendix A), Myrna Cozen, Lauren Smith, and Malene Cobourne.

**Appendix**

**WORKING GROUP MEMBERS**

Manal Abdelmalek, M.D., M.P.H., Duke University School of Medicine

Quentin M. Anstee, Ph.D., F.R.C.P., Newcastle University Medical School
REFERENCES


5) 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—

Author names in bold designate shared co-first authorship.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep.29607/suppinfo.