Non-alcoholic fatty liver disease: the pathologist’s perspective

Wei-Qiang Leow1,*, Anthony Wing-Hung Chan2,*, Paulo Giovanni L. Mendoza3, Regina Lo4, Kihan Yap5, Haeryoung Kim6

1Department of Anatomical Pathology, Singapore General Hospital, Singapore, Singapore; 2Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Shatin, Hong Kong; 3Pathology and Laboratory Department, Cardinal Santos Medical Center, San Juan, Philippines; 4Department of Pathology and State Key Laboratory of Liver Research (HKU), The University of Hong Kong, Hong Kong, China; 5Yong Loo Lin School of Medicine, National University of Singapore, Singapore; 6Department of Pathology, Seoul National University Hospital, Seoul National University College of Medicine, Seoul, Republic of Korea

*Wei-Qiang Leow and Anthony Wing-Hung Chan equally contributed to this article as first authors.

Corresponding author: Haeryoung Kim, Department of Pathology, Seoul National University Hospital, Seoul National University College of Medicine, 103 Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea

Tel: +82-2-740-8322, Fax: +82-2-765-5600, E-mail: haeryoung.kim@snu.ac.kr

Funding support: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2022R1A2C2010348).
Abstract

Non-alcoholic fatty liver disease (NAFLD) is a spectrum of diseases characterized by fatty accumulation in hepatocytes, ranging from steatosis, non-alcoholic steatohepatitis, to cirrhosis. While histopathological evaluation of liver biopsies plays a central role in the diagnosis of NAFLD, limitations such as the problem of interobserver variability still exist and active research is underway to improve the diagnostic utility of liver biopsies. In this article, we provide a comprehensive overview of the histopathological features of NAFLD, the current grading and staging systems, and discuss the present and future roles of liver biopsies in the diagnosis and prognostication of NAFLD.

Key words: Non-alcoholic fatty liver disease, liver biopsy, diagnosis, prognosis, histology

1. Introduction

Non-alcoholic fatty liver disease (NAFLD) is a condition in which there is fatty infiltration in the liver in the absence of secondary causes, including significant alcohol consumption. The morphological spectrum of non-alcoholic fatty liver disease (NAFLD) encompasses “simple” steatosis, non-alcoholic steatohepatitis (NASH), and cirrhosis. Histological evaluation by liver biopsy plays an important role in the diagnosis of NAFLD and NASH, and in excluding the possibility of other diseases. Another role of the liver biopsy is prognostication, as the histological parameters may potentially provide important information for identifying groups of NAFLD patients at risk for developing cirrhosis, liver failure and hepatocellular carcinoma (HCC). Grading and staging systems, such as the NAFLD Activity Score (NAS) and Steatosis-Activity-Fibrosis (SAF) scores, are currently widely used to assess disease severity and prognosis, and also to evaluate response to treatment in both the practical setting and clinical trial setting. In this review, we summarize the histopathological features of NAFLD, the grading and staging systems, and the recent advances in ancillary tool development for the accurate diagnosis and prognostic prediction of NAFLD.
2. Definition and diagnostic criteria

Steatosis, or fatty change, is the accumulation of fat droplets in the hepatocyte cytoplasm, and can be classified as macrovesicular or microvesicular based on the size of the lipid droplets (described in more detail in the subsequent section). NAFLD is defined as the presence of steatosis in \( \geq 5\% \) of hepatocytes, in the absence of significant alcohol use or other causes of steatosis, including viral hepatitis or drug/toxin-induced liver injury\(^1\)-\(^3\). NASH is characterized by the presence of active injury, in the form of hepatocellular ballooning degeneration and lobular inflammation (mostly lymphocytic with some neutrophils), in addition to varying degrees of steatosis. Although there are slight differences in the definitions in various practice guidelines, the presence of hepatocellular ballooning is regarded as an important factor for the diagnosis of NASH; in fact, it is considered the *sine qua non* of steatohepatitis for practical purposes, and its presence differentiates NASH from simple steatosis\(^1\)-\(^3\). Fibrosis is typically located in zone 3 with a perivenular and perisinusoidal pattern, and this feature is helpful in corroborating the diagnosis of NASH. Mallory-Denk body formation, apoptotic hepatocytes (acidophilic bodies), and lipogranulomas are other histological features of NASH. NASH-cirrhosis is defined as cirrhosis associated with current or previous histological evidence of NAFL or NASH\(^2\)-\(^3\).

**Steatosis**

The typical steatosis in NAFLD is of the macrovesicular pattern\(^4\). Macrovesicular steatosis is classically characterized by a large lipid droplet occupying the cytoplasm of a hepatocyte, pushing its nucleus to the periphery\(^5\) (Figure 1). It is also increasingly being recognized that the lipid droplets may vary in size as the triglycerides accumulate in the hepatocytes over time, and thus a range of lipid droplet sizes may occur. As such, the terms large, medium and small droplet steatosis have been used to describe this variance in lipid droplet sizes, and it is understood that these findings fall under the macrovesicular pattern of hepatic steatosis.

Of relatively more importance is the distinction of small droplet steatosis from microvesicular pattern
of hepatic steatosis. Microvesicular steatosis is characterized by the cytoplasm of hepatocytes being filled with numerous tiny lipid droplets and the presence of a central nucleus. While small droplet steatosis may morphologically mimic microvesicular steatosis, typical NAFLD will only show patches of small droplet steatosis accompanied by other areas of large and medium droplet steatosis (Figure 1). For most pathologists, the terminology of microvesicular steatosis is more often preferred for instances whereby this pattern diffusely involves the liver. This is also important clinically as the differential diagnoses for microvesicular steatosis is distinctly different from macrovesicular steatosis (Table 1).

The steatosis of NAFLD typically begins in the perivenular region (zone 3), and is graded semiquantitatively as “mild”, “moderate”, or “severe”, when 5-33%, 33-66%, or more than 66% of the hepatocytes are affected, respectively. A zone 1-predominant distribution of steatosis is rare in adults (~1%), while more commonly found in children and teenagers (~12%). An azonal distribution is more likely to be associated with ballooning degeneration, Mallory-Denk bodies and advanced fibrosis.

**Hepatocellular ballooning**

Hepatocyte ballooning is characterized as an enlarged hepatocyte with rarefied pale cytoplasm, usually with the presence of a large, hyperchromatic nucleus and a prominent nucleolus, indicating the presence of hepatocellular injury. The cytoplasmic changes reflect injury to the cytoskeleton of these hepatocytes, with loss of intact keratin 8 and 18 and increased detection of keratin fragments. As the cytoskeleton injury progresses, the increased clumping of these keratin fragments contributes to Mallory-Denk body formation.

It is worthwhile noting that in chronic cholestatic conditions, hepatocytes may also suffer from similar cytoskeleton injury resulting in morphological changes similar to ballooning. This is classically described as “feathery degeneration”. One may easily make the distinction by observing the adjacent steatotic or cholestatic changes, in order to decide which term to use. Mimics of ballooned hepatocytes include hydropic change of hepatocytes and microvesicular steatosis.
Ballooned hepatocytes exist in an “undead” state where they are unable to undergo apoptosis while releasing factors such as sonic hedgehog (Shh) to aid with tissue repair and healing. These ballooned hepatocytes were found to lack caspase 9 - a protease critical for apoptosis12.

Ballooned hepatocytes are also associated with activation of the stress kinase c-Jun N-terminal kinase, which upregulates the hedgehog signaling pathway in the absence of apoptosis12-14. Prolonged hepatocyte lipotoxicity leads to persistent activation of the pathway. This is further exacerbated by the downregulation of protective enzymes such as HSP27, a protein with antioxidant properties that responds to cellular stress15.

In NAFLD, the activity of the hedgehog signaling pathway correlates with the severity of liver damage and fibrosis16. Analysis of a representative subset of subjects enrolled in the PIVENS clinical trial also found that response to treatment corresponds to a greater decrease in Shh-producing hepatocytes17. Increased Shh also is also associated with an increased risk of primary liver cancers, via the upregulation of cyclin B1 and CDK1 mitotic proteins, as well as the induction of the epithelial-mesenchymal transition in malignant cells16.

**Mallory-Denk bodies**

Mallory-Denk bodies (MDB), also known as Mallory hyaline in the past, are cytoplasmic aggregates that could be identified in some cases of steatohepatitis. MDBs appear as aggregates of hepatocytic keratins, K8 and K18, as well as ubiquitin and p62 in the cytoplasm18-21. The aggregates could be highlighted by immunohistochemical staining. Of note, MDB is not a specific histological feature for NAFLD, and is also observed in various inflammatory diseases, including alcoholic hepatitis and primary biliary cholangitis, and hepatocellular carcinoma22.

**Lobular necroinflammation**
Inflammatory infiltrates in the hepatic lobules are commonly seen in steatohepatitis. The number of inflammatory cells may vary but are usually more accentuated in zone 3, in contrast to the portal/peribronchial distribution as seen in viral hepatitis. Mononuclear cells are the major constituent cells; some polymorphonuclear leukocytes and histiocytes are also present (Figure 3). Microgranulomas, which represent macrophages engulfing lipid droplets, may be observed. Apoptosis of hepatocytes (acidophilic bodies) may be present, in accordance to the severity of inflammation. Lobular inflammation may become less conspicuous in the cirrhotic stage of the disease.

Other histological findings

Enlarged mitochondria, or megamitochondria, are detectable under light microscopy as eosinophilic inclusions in the cytoplasm. It has been proposed that megamitochondria signify the presence of cell injury or an adaptation process secondary to lipid peroxidation. Glycogenated nuclei – clear intranuclear inclusions of hepatocytes – are associated with diabetes mellitus, and are more readily observed in NAFLD compared with alcoholic liver disease.

Although the typical NASH histology is characterized by a lobular distribution of inflammation, there is often also a mild degree of portal mononuclear infiltration. In fact, portal inflammation that is moderate (but patchy) can be seen in the setting of severe NASH, NASH in the pediatric population or young adults, and also in the setting of disease resolution post-treatment. However, when there is a significant amount of portal inflammation (diffuse, moderate/severe) that is disproportionate to the degree of lobular inflammation, one should consider the possibility of a concurrent disease, including chronic viral hepatitis and autoimmune hepatitis (AIH). The differential diagnosis of NAFLD/NASH is discussed in more detail in a subsequent section.

Fibrosis

Hepatic fibrosis is caused by the excessive production, deposition, and net accumulation of extracellular
matrix by activated hepatic stellate cells and other myofibroblasts. In line with the preferential and initial deposition of steatosis in zone 3 of the hepatic lobule, the subsequent hepatocellular injury via the presence and accumulation of these lipotoxic lipids culminate in fibrosis commencing in the perivenular and zone 3 regions.

The characteristic histologic pattern of fibrosis in NASH is the zone 3 pericellular and/or perisinusoidal pattern (often described as a “chicken-wire pattern”), resulting from the deposition of collagen and other extracellular matrix fibers around the hepatocytes (Figure 4). In advanced disease, the fibrosis extends to involve the portal and periportal (zone 1) regions, with subsequent central-portal bridging fibrosis and eventually cirrhosis.

In contrast, pediatric cases of NASH are more commonly associated with periportal fibrosis and the absence of perisinusoidal fibrosis. This is due to the preferential and initial deposition of fat in the zone 1 region. As a result, the subsequent downstream hepatocellular injury and fibrosis is centered on zone 1 rather than zone 3.

3. Ancillary tests

1) Connective tissue stains

A good quality connective tissue stain is essential to identify hepatic fibrosis and especially crucial in detecting early-stage fibrosis of NAFLD. Connective tissue stains widely used in liver pathology include trichrome, Sirius red and Gordon-Sweets reticulin stains.

Trichrome stain is the connective tissue stain of choice for the assessment of fibrosis in most laboratories because of its wide availability. However, a good trichrome stain requires proper optimization to avoid overstaining or understaining, which may lead to misinterpretation of the degree of fibrosis. Although both trichrome and Sirius red stains are employed in computer-assisted morphometric quantitation of liver fibrosis, Sirius red stain is shown to be superior to trichrome stain because of its highly detailed and contrasted staining of collagen fibers and high sensitivity in identifying early perivenular and
pericellular fibrosis. Nevertheless, both trichrome and Sirius red stains are equivalently good for routine daily practice. The choice between these two stains largely depends on personal preference and reagent availability. Gordon-Sweets reticulin stain primarily highlights type III collagen, and therefore it is used to assess hepatocyte cord thickness, reticulin framework integrity, and nodular architecture. Although it can also evaluate fibrosis by highlighting type I collagen (the predominant collagen in hepatic fibrosis), it is less sensitive for the detection of early perivenular fibrosis. Of note, reticulin loss may be focally present in areas of steatosis, which may lead to the erroneous interpretation of a well-differentiated hepatocellular neoplasm, especially when the tissue is sampled with the clinical impression of a “hepatic nodule”.

2) Immunohistochemical stains

Cytokeratin 8/18 (CK8/18) is normally distributed in the cytoplasm with a strong intensity. In hepatocytes with ballooning, expression is loss or diminished in majority of the cytoplasm, and immunoreactivity is only retained in the Mallory-Denk bodies. Immunohistochemical staining for p62 and ubiquitin also highlights Mallory-Denk bodies. p62 is an autophagy substrate and a biomarker for the activity of autophagy, while ubiquitin is involved degradation of proteins. Expression of sonic hedgehog (Shh) is identified in the hepatocytes of NAFLD. It was reported that hepatic Shh expression was associated with the degree of liver injury by histological evaluation and by circulatory biochemical profile.

4. Grading and staging systems

Grading and staging are histological markers of activity (severity of active necroinflammation) and chronicity (degree of fibrosis) of chronic liver disease, respectively. Scoring systems of grading and staging are utilized in chronic viral hepatitis to semiquantitatively evaluate disease severity and monitor disease progression. They are useful in clinical management guideline development, pathology report
standardization and histology assessment for clinical trials. Nevertheless, scoring systems for chronic viral hepatitis cannot be simply applied in NAFLD because they do not account for steatosis and ballooning degeneration, which are crucial in assessing disease activity in NAFLD. Additionally, they also do not consider perivenular and perisinusoidal fibrosis, which is the distinctive fibrosis pattern in NAFLD. Hence, the development of scoring systems designed for NAFLD is necessary to fill the gap.

In 1999, the first scoring system for NAFLD was developed by Brunt et al. It was derived from a cohort of 51 patients with NAFLD undergoing liver biopsy. The disease activity grade (0-3) was assigned according to a constellation of histological features composed of steatosis, lobular and portal inflammation, and ballooning degeneration. The fibrosis stage (0-4) was based on the fibrosis pattern of adult NAFLD from perivenular and pericellular fibrosis (stage 1), periportal fibrosis (stage 2), bridging fibrosis (stage 3) and cirrhosis (stage 4).

In 2005, the Non-Alcoholic Steatohepatitis Clinical Research Network (NASH-CRN) proposed the NASH-CRN scoring system, also known as the Kleiner scoring system, based on a cohort of 50 NAFLD patients (32 adults and 18 children). In this system, the disease activity grade (NAS) is the unweighted sum of semiquantitative scores (0-8) for steatosis (0-3), ballooning degeneration (0-2), and lobular inflammation (0-3) (Table 2). The fibrosis stage (0-4) is similar to the Brunt fibrosis stage; however, the early fibrosis stage (stage 1) was refined and stratified into 1a (delicate pericellular fibrosis visualized by connective tissue stain only), 1b (dense pericellular fibrosis visualized by hematoxylin-eosin section) and 1c (portal/periportal fibrosis only). Stage 1c was added to represent the characteristic early fibrosis pattern among pediatric NAFLD patients. The NAS was demonstrated to be associated with the histological diagnosis of steatohepatitis: over 85% of patients with NAS ≥5 were diagnosed as steatohepatitis, whereas 99% of patients with NAS 0-2 were categorized as not diagnostic of steatohepatitis. The NAS of 4 or more is used as one of the inclusion criteria in various clinical trials of NASH patients. One should note that the primary objective of the NAS is to evaluate the overall histological changes. It has been repeatedly emphasized that the NAS should not be regarded as a numerical diagnostic criterion that substitutes the histological diagnosis of steatohepatitis.
In 2012, Bedossa et al. established a diagnostic algorithm and a scoring system from a cohort of 679 obese patients undergoing bariatric surgery. The Fatty Liver Inhibition of Progression (FLIP) algorithm classified a biopsy into either steatosis (without NASH) or NASH by semiquantification of steatosis, ballooning degeneration, and lobular inflammation. This algorithm improved the interobserver agreement in differentiating between steatosis and NASH [from moderate (kappa 0.54) to substantial (kappa 0.66)] among expert liver pathologists. Such an improvement was significantly more substantial among general pathologists [from fair (kappa 0.35) to substantial (kappa 0.61)]. The SAF score was the combination of semiquantitative scores of steatosis (S0-S3), activity [A0-A4; ballooning degeneration (0-2) and lobular inflammation (0-2)] and fibrosis (F0-F4) (Table 3). Although the NAFLD-CRN and SAF scoring systems are apparently similar, direct inter-translation between these two systems is not feasible. It is noteworthy that there are several considerable differences. First, steatosis is not integrated into the activity score of the SAF compared to the NAS because the prognostication of steatosis in long-term outcomes and fibrosis progression remains controversial.

Second, the grading scheme for hepatocellular ballooning differs in the two systems – the NAFLD-CRN system assesses the quantity, while the SAF system evaluates the morphology of the ballooned cells (Tables 1, 2). Third, the NAFLD-CRN system grades lobular inflammation from 0 to 3 (0, none; 1: <2 foci/200x field; 2: 2-4 foci/200x field; 3: >4 foci/200x field), while the SAF system only grade lobular inflammation from 0 to 2 (0, none; 1: 1-2 foci/200x field; 3: >2 foci/200x field). Last but not least, both NAFLD-CRN and SAF systems have been externally validated by other groups but only the NAFLD-CRN system is currently widely used for clinical trials.

Histological features in NAFLD apart from ballooning degeneration and lobular inflammation are also shown to have prognostic significance. Portal inflammation and Mallory-Denk bodies are two histological parameters that have been consistently demonstrated to be associated with adverse clinical outcomes and fibrosis. A more comprehensive but more complicated scoring system, the expanded NAS, has been proposed recently to provide a more accurate evaluation of the histological activity of NAFLD by incorporating portal inflammation and Mallory-Denk bodies. The clinical significance and applicability of the expanded NAS require further studies.
Any scoring system is inevitably subject to have intraobserver and interobserver variabilities. While the agreement in the evaluation of steatosis and fibrosis has been demonstrated to be substantial to almost perfect among different pathologists (kappa 0.79-0.80 and 0.54-0.84, respectively) and for the same pathologist (kappa 0.82-0.85 and 0.73-0.85, respectively), the agreement in the grading of ballooning degeneration and lobular inflammation is only fair to substantial among different pathologists (kappa 0.20-0.69 and 0.35-0.60, respectively) and for the same pathologist (kappa 0.66-0.72 and 0.60-0.70, respectively). Computer-assisted image analysis may provide a more reliable way to minimize intraobserver and interobserver variabilities in the future.

5. Pediatric NAFLD

In the pediatric population, about half of NASH cases demonstrate the features of “type 2” NASH, characterized by moderate-to-severe steatosis with a panacinar distribution, portal inflammation, and portal fibrosis. Hepatocyte ballooning and Mallory-Denk bodies are less frequently seen compared to adults. This pattern is not restricted to children; “type 2” NASH has also been described in a subset of young adults.

6. Looking at NAFLD under the microscope: applications in unique settings and differential diagnoses

1) Identifying ballooned hepatocytes

As the presence of hepatocyte ballooning is the key to the histopathological diagnosis of NASH, it is of paramount importance that this is identified with confidence by pathologists. Although ballooned hepatocytes demonstrate the characteristic appearance as described earlier, pathologists not infrequently encounter situations in which the hepatocyte in question demonstrates equivocal changes that fall short of a “classic” balloon cell (Figure 3). Some of these “equivocal” balloon cells would belong to the “grade 1” ballooning of the SAF score, proposed by Bedossa et al., while others could represent other
changes with similar morphology, such as hydropic change of hepatocytes and microvesicular steatosis. In order to increase the accuracy of balloon cell identification, ancillary immunohistochemical stains such as CK8/18, ubiquitin, or Shh could be used. In addition, artificial intelligence (AI)-based technologies may have a role in the future.

2) Steatosis and steatohepatitis of other etiologies

Steatosis or steatohepatitis occurs in a variety of other settings, such as alcoholic liver disease, metabolic disorders (e.g., Wilson disease), chronic viral hepatitis, and drug/toxin-induced liver injury. Steatosis or steatohepatitis associated with alcoholic liver disease often demonstrate histological features that overlap with those of NAFL or NASH, respectively. Although alcoholic liver disease also commonly presents with macrovesicular steatosis in the perivenular zone, the general histological picture of steatohepatitis is more pronounced in ALD compared to NASH, with more abundant ballooned hepatocytes, Mallory-Denk bodies, acidophil bodies, lipogranulomas, and neutrophilic infiltration. Neutrophils may predominate in alcohol-related steatohepatitis, sometimes forming aggregates around ballooned hepatocytes (“neutrophilic satellitosis”). Alcoholic foamy degeneration and sclerosing hyaline necrosis are not features of NAFLD. The presence of cholestasis may help in the differential between alcoholic steatohepatitis and NASH, as it is not a typical histological feature of the latter. The pattern of fibrosis is similar to that of NASH, with the zone 3-predominant perisinusoidal fibrosis that eventually progresses to bridging fibrosis and cirrhosis. Most importantly, the key distinguishing feature is the patient's history of alcohol consumption, and therefore clinicopathological correlation is necessary.

Among the different viral hepatitides, steatosis has been described to be a common histological feature of chronic hepatitis C. However, the degree of steatosis in chronic hepatitis C alone should be at most mild, and in the presence of moderate or severe steatosis in patients with chronic hepatitis C, a co-existing cause of fatty liver should be investigated. Drug/toxin-induced liver injury may present as steatosis or even steatohepatitis (“drug-induced steatohepatitis, DISH”); examples of offending drugs
include glucocorticoids, tamoxifen, irinotecan and amiodarone. As the histological features are most often similar to that of NAFL or NASH, the clinical information is the most important key to the diagnosis.

3) NAFLD with serum autoantibody positivity

Coexistence of AIH with NASH is not a rare occurrence; in such cases, there is a significant amount of portal lymphoplasmacytic infiltration and interface hepatitis in addition to the histological features of NASH. Correlation with the clinical findings, including elevated serum immunoglobulin G levels and positive autoantibodies, is important when contemplating the possibility of a combined AIH, as portal mononuclear cell infiltration with focal mild interface hepatitis may be encountered in NASH. Moreover, serum autoantibody positivity has been identified in up to 34% of NAFLD patients in the absence of AIH, and no significant differences in the histology of NAFLD have been found according to serum autoantibody status.

4) NAFLD in the post-liver transplantation setting

NAFLD may occur as a recurrent disease or de novo disease in the post-liver transplantation setting. In a study over a 10-year-period that analyzed 11 cases of recurrent disease and 80 de novo NAFLD in post-liver transplant patients, a higher prevalence of diabetes mellitus was observed in recurrent NAFLD. Severe fibrosis and steatohepatitis were more readily observed in recurrent NAFLD versus de novo NAFLD. Interestingly, serial biopsies have demonstrated resolution of steatosis in 22.5% patients with de novo NAFLD but in none of the patients with recurrent NAFLD.

5) Association of NAFLD with steatohepatitic HCC

Steatohepatitic HCC is associated with metabolic syndrome, a key driver of NAFLD. This HCC variant
shows features resembling steatohepatitis within the tumor itself, including macrovesicular steatosis, balloon cells, intratumoral inflammation and intratumoral pericellular fibrosis\textsuperscript{67,68}.

Salomao et al. demonstrated that their cohort that patients with steatohepatitic HCCs had significantly higher number of metabolic syndrome risk factors (2.44 vs 1.48, \(p=0.01\)) and higher percentage of patients with at least 3 metabolic syndrome components (50\% vs 22.5\%, \(p=0.02\))\textsuperscript{69}. However, this intuitive association has been challenged in another study by Yeh et al. that evaluated 12 steatohepatitic HCCs arising in patients without metabolic syndrome. In this cohort, a subset of tumor showed loss of 9q12-q31-1 via genomic microarray analysis\textsuperscript{70}.

7. State-of-art and future trends

1) Role of digital pathology and AI

Due to the limitations of current methods to assess NAFLD and liver fibrosis, there is considerable interest in the use of AI to improve these systems for risk stratification, diagnosis, monitoring, and prognostication of NAFLD in patients\textsuperscript{71}. AI can be integrated in AI-based digital pathology systems to assess NAFLD. Digital pathology is defined as the process of utilizing whole slide scanners for digitizing of histopathology slides, producing images that allow for quantitative analyses\textsuperscript{72}. When combined with AI, these systems have the potential to diagnose and prognosticate NAFLD via automated processes\textsuperscript{73}.

Taylor-Weiner et al. developed a machine learning-based approach for the assessment of liver histology in NAFLD\textsuperscript{74}. For the assessment of the diagnostic features of NAFLD, the model’s predictions were significantly correlated with the consensus NAS grades of pathologists’ assessments – steatosis: \(\rho=0.66\), lobular inflammation: \(\rho=0.54\), hepatocellular ballooning: \(\rho=0.62\). For the assessment of fibrosis, the model’s predictions were also significantly correlated with the consensus staging of pathologists, with a weighted Cohen’s kappa of 0.801 and 0.817 for the NASH CRN and the Ishak classifications.
respectively. This level of agreement is within the range of agreement between individual pathologists and the consensus staging by pathologists.

Machine learning models also enabled the identification and quantification of novel and complex parameters that are usually difficult to evaluate with conventional methods. The study identified the steatosis to hepatocellular ballooning ratio to be a significant parameter of NAFLD progression, where subjects with more hepatocellular ballooning and less steatosis at baseline were significantly more likely to experience a clinical event. The study also proposed the DELTA Liver Fibrosis Score – a machine learning-derived metric used to measure changes in the intra-sample distribution of fibrosis associated with disease progression or therapy. When a stringent DELTA Liver Fibrosis Score threshold was applied comparing images pre- and post-treatment, significant differences could be found in samples that previously did not demonstrate any significant difference using conventional pathologist staging methods. Therefore, the DELTA Liver Fibrosis Score could be a more sensitive method for assessing histological response to treatment, potentially being a useful tool in NAFLD clinical trials.

Forlano et al. developed an automated image analysis-based system to quantify steatosis, ballooning, inflammation, and fibrosis from the histological images of NAFLD patients. There was excellent concordance between manual annotations of histopathologists and the automated measurements, with an intraclass correlation coefficient of 0.95-0.99 for the four parameters measured. The fully automated model was described to be straightforward to install, not requiring specialized equipment, only requiring modest computational effort, and being able to produce results within 2 minutes.

2) Second-Harmonic Generation (SHG) microscopy

Second-Harmonic Generation (SHG) microscopy and Two-Photon Excited Fluorescence (TPEF) Microscopy are both imaging techniques under the umbrella of Non-Linear Optimal (NLO) microscopy techniques, which were described to produce images of good spatial resolution, depth of penetration,
and excitation capability. Both SHG and TPEF imaging can be performed regardless of the means of sample preparation – where both frozen and formalin-fixed paraffin-embedded tissues can be used without staining.

In the liver, TPEF microscopy enables the visualization of the liver background and lobular organization, while SHG microscopy characterizes the morphology of collagen (Figure 5). Combined SHG/TPEF microscopy can localize and quantify fibrillar collagen in 2D and 3D, enabling the automated quantification of fibrosis. These features tackle known limitations of traditional histological scores with semiquantitative grading systems such as inter- and intraobserver variation.

Other than NAFLD, SHG/TPEF microscopy have been initially used to quantify fibrosis in other liver conditions, especially chronic hepatitis B. Developed by Xu et al, qFibrosis, a combined index based on 87 parameters, was first validated with core biopsies of chronic hepatitis B patients. qFibrosis was found to be able to reliably replicate the Metavir fibrosis staging by histopathologists, and was more sensitive in differentiating fibrosis stages compared to Collagen Proportionate Area (CPA). qFibrosis was also described to have decreased sensitivity to sampling error, and can aid in the correction of intra- and interobserver bias. For chronic hepatitis B patient post-antiviral therapy, qFibrosis was not only able to detect the changes observable by histopathologists, but could also detect and characterize subtle changes in fibrosis, potentially being more sensitive in evaluating changes in fibrosis.

Following the successes of SHG/TPEF microscopy in chronic hepatitis B, several models in the same vein have been developed for NAFLD.

Quantifiable fibrosis-related parameter (q-FP)

Established by Wang et al, the q-FP model was the first established SHG based model that quantified fibrosis-related parameters in NAFLD. The q-FPs included the geometric and textural features of collagen fibers, and the number of collagen fibers. The collagen fibers at defined regions such as the general liver section, perisinusoidal space, vessels, and vessel bridges were measured and characterized.
Seventy of the q-FPs had inter- and intraobserver concordance ≥0.8 and were strongly related to the NAS fibrosis staging. Sixteen of these q-FPs with the strongest concordance were included in a principal component analysis model, differentiating any stage of fibrosis versus no fibrosis, and cirrhosis versus earlier fibrosis stages with an AUC of 0.88 and 0.93 respectively. Four q-FPs – number of collagen strands, strand length, strand eccentricity, and strand solidity – were found to also be independently associated with fibrosis stages. These 4 q-FPs could model fibrosis along a continuous linear scale using desirability functions, with the obtained measurements being significantly correlated with actual fibrosis stage.

**SHG B-index**

Chang et al. developed a SHG-based model, the SHG B-index, to scan and analyze the SHG properties of collagen in unstained liver tissue specimens of NAFLD patients, and is able to grade the severity of liver fibrosis. A total of 14 parameters that correlated strongly with the Brunt fibrosis staging classification were selected.

The SHG B-index had a high correlation with Brunt fibrosis staging, with an excellent ability to differentiate advanced fibrosis from no or mild fibrosis. However, between Brunt stages 0-2, the SHG B-index had a poorer discriminatory ability. The SHG B-index was also able to identify different fibrosis stages, with AUROCs of 0.853-0.985 for the prediction of mild fibrosis, significant fibrosis, bridging fibrosis, and cirrhosis.

The study also utilized Youden’s index to derive optimal SHG B-index cut-off values to identify specific Brunt fibrosis groups. The cut-off value for advanced fibrosis had an overall diagnostic accuracy of 98.5% for prediction of the presence of bridging fibrosis, with a positive predictive value of 96.6% and a negative predictive value of 92.6%. This suggests that the SHG B-index has high accuracy for the discrimination of advanced fibrosis compared to milder stages of fibrosis. This is clinically important as bridging fibrosis is a clinically important feature that is associated with poor prognosis in NAFLD.
patients.

**qFibrosis/qFIBS**

Liu et al. modified features of qFibrosis to compare the features of collagen and fibrosis in pediatric and adult NAFLD\textsuperscript{85}. The study found that there was more baseline collagen in livers of adult NAFLD, and a predominance of portal fibrosis in pediatric NAFLD compared to centrilobular fibrosis in adult NAFLD. qFibrosis was also able to detect subtle differences not apparent in histology, such as wider central vein lumens in pediatric NAFLD, possibly indicating the presence of increased portal-central vascular shunting. The same group expanded SHG/TPEF microscopy further to produce qFIBS, an algorithm that provides an automated quantitative assessment of histological features pertinent to NASH. qFIBS quantifies the four key histopathological features of NAFLD – fibrosis (qFibrosis), inflammation (qInflammation), hepatocyte ballooning (qBallooning), and steatosis (qSteatosis), with the goal of predicting the severity of NAFLD. Each parameter in qFIBS correlated well their corresponding histological counterparts, and could distinguish between different grades of the histological feature with an AUC between 0.813-0.939. qFIBS was also validated in both adult and pediatric NAFLD liver biopsy samples\textsuperscript{86}.

Leow et al. refined the qFibrosis algorithm further, including 26 new periportal parameters to produce an algorithm with a better discriminatory ability for F1 and F2 fibrosis according to the NAS. These new parameters are able to better compensate for limitations of previous AI-based SHG algorithms, where they are less discerning in discriminating between early stages of fibrosis. Having a better ability to discriminate between early fibrosis stages can play an important role in clinical trials – increasing the accuracy of patient enrollment, while more accurately monitoring treatment responses\textsuperscript{87}.

Therefore, it can be seen that AI has great potential and could have a large role to play in multiple aspects of NAFLD.
3) The role of liver biopsy in clinical trials

Despite the large amount of resources invested into NAFLD clinical trials, no drug has been specifically approved for the treatment of NAFLD yet\(^{88,89}\). While the complex and multifactorial pathophysiology of NAFLD provides numerous potential targets for intervention, this complexity also hampers the ability to define clear, measurable, and objective clinical endpoints in clinical trials\(^90\).

Liver biopsies are still considered as the gold standard for the diagnosis and evaluation of NAFLD. The quality of the obtained sample can be affected by the method of procurement, location, type, and dimensions of the liver biopsy\(^91\). For the same sample, the intra- and interobserver variability of histopathologist evaluation could also affect the reported results. The limitations of the procurement and interpretation of liver biopsies could affect the enrollment of participants into clinical trials, as well as incorrectly assess the histological treatment responses in serial liver biopsies. In addition, the presence of co-morbidities such as type 2 diabetes, metabolic syndrome, and cardiovascular diseases, along with the lack of uniformity of confounders such as alcohol, diet, and physical activity also complicates the interpretation of NAFLD clinical trial results\(^92,93\).

Aside from key clinical endpoints such as liver-related mortality, liver transplantation, liver decompensation, and hepatocellular carcinoma (HCC), histological changes in serial liver biopsies have also been used as the main surrogate endpoints in clinical trials, especially for NAFLD patients without cirrhosis. Currently, meaningful endpoints that indicate an improvement in NAFLD include a reduction of the NAS \(\geq 2\) with \(\geq 1\)-point reduction in either lobular inflammation or hepatocellular ballooning without worsening of fibrosis, resolution of NAFLD without worsening of fibrosis, and the improvement in liver fibrosis without the worsening of NAFLD\(^92,94\). An improvement of fibrosis is defined as an improvement by at least 1 fibrosis stage using the Brunt criteria.

Other proposed surrogate endpoints include the use of non-invasive imaging and biochemical modalities, but these modalities are not validated for and have limited use in late-phase clinical trials. Magnetic Resonance Imaging-Proton Density Fat Fraction (MRI-PDFF) is a validated technique used in early-phase clinical trials to assess the extent of steatosis in each segment of the liver, and can detect
small changes in steatosis better than histopathologist interpretation of liver biopsies. Liver stiffness can also be determined using elastography-based methods such as VCTE, MRE, and SWE, but have not been validated to be used as surrogate endpoints in clinical trials92,95. Numerous serum biomarkers and algorithms have been investigated to prognosticate the severity of NAFLD. Acute-phase proteins, cytokines, and markers of oxidative stress and apoptosis have been evaluated in NAFLD patients but were found to have limited utility. Previously mentioned algorithms such as the NAFLD Fibrosis Score and FIB-4 have also been considered for use in clinical trials90. However, these algorithms only showed a modest ability to predict fibrosis, as well as lacking conclusive data on how these measures change in response to disease progression, thus not being suitable surrogate endpoints for clinical trials95.

Unfortunately, there are also no clear endpoints for NAFLD clinical trials in the pediatric age group. This is contributed and complicated by the presence of knowledge gaps in pediatric NAFLD, as well as the numerous added limitations involved with conducting research in pediatric patients92.

8. Conclusion

Despite the remarkable advances in non-invasive biomarker development during the recent years, liver biopsy evaluation still has important roles in the setting of NAFLD diagnosis, such as confirmation or exclusion of the diagnosis, distinction of NASH from simple steatosis, assessment of disease severity and stage, and other histological alterations96. In fact, currently, only liver biopsy can provide simultaneous information on steatosis, inflammation, hepatocellular injury, fibrosis and concurrent liver disease. In addition, liver biopsy is essential in clinical trials, for confirming the presence of NASH, assessing and semiquantitating individual features and evaluating the effects of the therapeutic intervention. To overcome the current limitations of liver biopsy, such as the problem of inter/intraobserver variability, new diagnostic tools are being developed – with the recent burst of research on AI-based pathology tools and the increasing implementation of digital pathology into routine diagnostic practice, it will probably not be long before these new technologies will make their way into routine clinical care.
Figure 1. Steatosis. A combination of large and small droplet macrovesicular steatosis is seen in this example of non-alcoholic steatohepatitis. In large droplet macrovesicular steatosis, the fat droplet occupies more than half of the hepatocyte cytoplasm and pushes the nucleus to the edge of the cell (black arrowheads). Smaller droplets are also seen. A small patch of microvesicular steatosis is noted on the right (black star), characterized by innumerable tiny fat droplets in the hepatocyte cytoplasm. A few ballooned hepatocytes are also noted (white arrowheads) (H/E, original magnification x200)
Figure 2. The many faces of ballooned hepatocytes. (A) A cluster of classical ballooned cells. (B) Occasionally the cytoplasmic keratins aggregate to form tighter and more eosinophilic clumps, also known as Mallory-Denk bodies (black arrowhead). (C) A lonely non-classical ballooned cell (black arrow) which is similar in size to the adjacent non-ballooned hepatocytes.
Figure 3. Lobular necroinflammation. Foci of lobular spotty necrosis are seen in this example of non-alcoholic steatohepatitis (yellow circles). The inflammatory cell infiltrates are mainly composed of mononuclear cells (H/E, original magnification x200).
Figure 4. Stages of fibrosis and the utility of histochemical stains in accentuating the histological appearance. (A1 and A2) The classical ‘chicken-wire’ appearance of pericellular fibrosis is accentuated with a Sirius Red histochemical stain, which reveals the collagen fibers in red. (B1 and B2) Stage 2 fibrosis is the co-presence of pericellular fibrosis and also portal fibrosis. Masson Trichrome stain shows the blue pericellular collagen fibers on the left and the portal fibrosis on the right. (C1 and C2) The fibrosis extends across the hepatic lobules and forms bridging fibrosis. (D1 and D2) The presence of hepatocytic nodules heralds cirrhosis, with Masson Trichrome stain confirming the broad fibrous bands. [Original magnification x200 (A1, A2), x40 (B1-D2); H/E (A1, B1, C1, D1; Sirius red (A2); Masson Trichrome (B2, C2, D2)]
Figure 5. An example of a case of non-alcoholic fatty liver disease (NAFLD)-cirrhosis seen by second harmonic generation/two-photon excitation fluorescence (SHG/TPEF) (SHG/TPEF microscopy, scanning power).
References


50. Bedossa P, Consortium FP. Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease. Hepatology 2014;60:565-575.


Ting Soon GS, Wei A. Liver biopsy in the quantitative assessment of liver fibrosis in nonalcoholic fatty liver disease. Indian J Pathol Microbiol 2021;64:S104-S111.


96. Kim HY. Recent advances in nonalcoholic fatty liver disease metabolomics. Clin Mol Hepatol 2021;27:553-559.
Table 1. Differential diagnoses for macrovesicular and microvesicular steatosis

<table>
<thead>
<tr>
<th>Differentials for macrovesicular steatosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alcoholic liver disease</td>
</tr>
<tr>
<td>Non-alcoholic fatty liver disease</td>
</tr>
<tr>
<td>Other metabolic conditions, such as diabetes mellitus, growth hormone deficiency and hyperthyroidism</td>
</tr>
<tr>
<td>Genetic diseases, such as cystic fibrosis, PFIC1 mutations and Wilson disease</td>
</tr>
<tr>
<td>Malnutrition and related causes, including inflammatory diseases affecting the small bowel and gastrointestinal surgery</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Differentials for microvesicular steatosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute fatty liver of pregnancy</td>
</tr>
<tr>
<td>Alcoholic foamy degeneration</td>
</tr>
<tr>
<td>Genetic mitochondrial disease</td>
</tr>
<tr>
<td>Other genetic diseases, such as ornithine transcarbamylase (OTC) deficiency, fatty acid oxidation disorders, and Wolman disease/cholesterol ester storage disease</td>
</tr>
<tr>
<td>Infections, including human herpes virus 8 and toxin of bacillus cereus</td>
</tr>
<tr>
<td>Toxins, including arsenic toxicity and industrial solvents</td>
</tr>
<tr>
<td>Medication effect, including linezolid, Reye syndrome, amiodarone, nucleoside analog reverse transcriptase inhibitors used in HIV treatment, valproate, high-dose tetracycline</td>
</tr>
</tbody>
</table>
Table 2. NAFLD Activity Score (NAS) and fibrosis stage by NASH-CRN.7

<table>
<thead>
<tr>
<th>NAFLD Activity Score</th>
<th>Steatosis</th>
<th>Lobular inflammation</th>
<th>Ballooning degeneration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>&lt;5%</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>1</td>
<td>5-33%</td>
<td>&lt;2 foci/20x field</td>
<td>Few</td>
</tr>
<tr>
<td>2</td>
<td>&gt;33-66%</td>
<td>2-4 foci/20x field</td>
<td>Many</td>
</tr>
<tr>
<td>3</td>
<td>&gt;60%</td>
<td>&gt;4 foci/20x field</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fibrosis Score</th>
<th>Histological findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Mild pericellular fibrosis (only seen on connective tissue stain)</td>
</tr>
<tr>
<td>1b</td>
<td>Moderate pericellular fibrosis (readily seen on H&amp;E)</td>
</tr>
<tr>
<td>1c</td>
<td>Portal/periportal fibrosis without pericellular fibrosis</td>
</tr>
<tr>
<td>2</td>
<td>Pericellular and portal/periportal fibrosis</td>
</tr>
<tr>
<td>3</td>
<td>Bridging fibrosis</td>
</tr>
<tr>
<td>4</td>
<td>Cirrhosis</td>
</tr>
</tbody>
</table>
Table 3. SAF (steatosis, activity, fibrosis) score and FLIP algorithm.

<table>
<thead>
<tr>
<th>SAF (steatosis, activity, fibrosis) Score</th>
<th>Steatosis</th>
<th>Steatosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>S0</td>
<td>&lt;5%</td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>5-33%</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>&gt;33-66%</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>&gt;60%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity</th>
<th>Score</th>
<th>Lobular inflammation (LI)</th>
<th>Ballooning degeneration (BD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A0-4 (LI+ BD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>1</td>
<td>≤2 foci/20x field</td>
<td>Hepatocytes with a rounded shape and pale cytoplasm usually reticulated. Size is quite similar to that of normal hepatocytes</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>&gt;2 foci/20x field</td>
<td>Hepatocytes with a rounded shape and pale cytoplasm usually reticulated. Some cells are twice the size of normal hepatocytes</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fibrosis</th>
<th>Histological findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1a</td>
<td>Mild pericellular fibrosis (only seen on connective tissue stain)</td>
</tr>
<tr>
<td>F1b</td>
<td>Moderate pericellular fibrosis (readily seen on H&amp;E)</td>
</tr>
<tr>
<td>F1c</td>
<td>Portal/periportal fibrosis without pericellular fibrosis</td>
</tr>
<tr>
<td>F2</td>
<td>Pericellular and portal/periportal fibrosis</td>
</tr>
<tr>
<td>F3</td>
<td>Bridging fibrosis</td>
</tr>
<tr>
<td>F4</td>
<td>Cirrhosis</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FLIP Algorithm</th>
<th>Steatosis</th>
<th>Ballooning degeneration</th>
<th>Lobular inflammation</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1, 2 or 3</td>
<td>0</td>
<td>0, 1 or 2</td>
<td>NAFLD</td>
</tr>
<tr>
<td></td>
<td>1, 2 or 3</td>
<td>1 or 2</td>
<td>0</td>
<td>NAFLD</td>
</tr>
<tr>
<td></td>
<td>1, 2 or 3</td>
<td>1 or 2</td>
<td>1 or 2</td>
<td>NASH</td>
</tr>
</tbody>
</table>