LES JOURNÉES SCIENTIFIQUES
DE L'INSTITUT SERVIER

HETEROGENEITY IN DIABETES AND BETA CELLS

20e COLLOQUE 2021

Nonalcoholic steatohepatitis: a pathologist's point of view

Valérie PARADIS
Université Paris Cité, AP-HP Nord, Hôpital Beaujon, Department of Pathology, Clichy, France

L'INSTITUT SERVIER

Springer Healthcare Communications

Abstract

As well as being a key risk factor for the development of nonalcoholic fatty liver disease (NAFLD), type 2 diabetes (T2D) also appears to accelerate NAFLD progression. Liver biopsy is the reference standard for the diagnosis of the severe form of NAFLD, nonalcoholic steatohepatitis (NASH), and there are two key, non-interchangeable, standardized histologic scoring systems used to evaluate disease features. Although fibrosis staging is not included as a criterion for the diagnosis of NASH, it is a key prognostic indicator, and (using biopsy samples), the agreement between pathologists concerning fibrosis staging is very good. Liver biopsy is an imperfect reference standard, with issues including its invasive nature, sampling errors, and intra- and inter-observer variability. Thus, numerous researchers are focussing on developing improved assessment tools. Studies have shown the utility of digital pathology and machine learning models in providing improved objectivity, sensitivity, and accuracy in assessing and quantifying morphologic features, and in classifying patients. Furthermore, our ongoing study in patients with T2D (QUID NASH) is investigating the use of non-invasive tools for the assessment of NAFLD, and initial results are encouraging.

Key words: digital pathology, histology, machine learning, NASH, type 2 diabetes

Introduction

Type 2 diabetes (T2D) is a key risk factor for the development of nonalcoholic fatty liver disease (NAFLD), although the association between the two diseases is bidirectional.[1] Amongst those with T2D, the estimated global prevalence of NAFLD is >55%,[1] which is approximately double the prevalence in the general population.[2] Furthermore, progression to severe liver disease appears to be accelerated in patients with T2D.[1] Globally, in patients with T2D, the prevalence of nonalcoholic steatohepatitis (NASH) is estimated to be close to 40%, and in biopsied patients with both NALFD and T2D, the estimated prevalence of advanced fibrosis is 17%.[1]

Pathologically, NAFLD can be characterized by four main hallmarks. First, a diagnosis requires the presence of an accumulation of hepatic fat (steatosis) in more than 5% of hepatocytes.[3] Second, NAFLD is a complex disease with heterogeneous pathologic patterns, ranging from simple steatosis to steatohepatitis. Third, there is also heterogeneity in the progression of the disease, and only some patients will progress beyond simple steatosis and develop advanced fibrosis and cirrhosis. Finally, at the time of writing, liver biopsy is still required for the diagnosis of NASH, as clinical, biochemical, or imaging measurements cannot currently distinguish NASH from steatosis. Although liver biopsy is considered the reference standard for the diagnosis of NASH, it is imperfect,[4] and several researchers are investigating possible improvements or alternatives to this standard.[5] This article discusses, from a pathologist's viewpoint, the current systems for diagnosing and assessing NAFLD and NASH, and also summarizes recent developments in this field.

Liver biopsy: the reference standard

In patients with suspected NAFLD, a liver biopsy can serve several purposes. A diagnosis of NASH requires three key histologic features: steatosis, lobular inflammation, and hepatocellular ballooning.[2] In addition to confirming the diagnosis, a biopsy can help assess the severity of NAFLD, in terms of inflammatory activity (disease grade based on inflammation and hepatocellular damage) and fibrosis stage. Biopsy results highlight the above-mentioned heterogeneity of the disease, which is not obvious from currently-available noninvasive tests, and we now know that NAFLD is a continuous morphologic spectrum with many intermediates.[2] Standardized histologic scoring systems also offer a common language between pathologists and clinicians, enable improved diagnostic accuracy and inter-reader agreement, aid in evaluating treatment responses, and support patient inclusion in clinical trials. Liver biopsy may also identify, or exclude, some potential comorbidity risk factors.[6]

Histologic scoring systems

Two key NAFLD histologic scoring systems have been developed and validated, and both provide standardized evaluation of disease features that cannot be achieved by any available noninvasive tests. The two tests differ markedly and are not interchangeable. The NAFLD activity score (NAS) is a composite scoring system based on the unweighted sum of semiquantitative analyses of steatosis (0–3), lobular inflammation (0–3), and hepatocellular ballooning (0–2).[7] NAS was designed as a continuous scale to compare disease activity in clinical trials and should not be used for diagnosis.[2,8] In the seminal study of NAS, a score of ≥5 generally correlated with a NASH diagnosis; a score of 0–2 was associated with a diagnosis of 'not NASH,' whereas a score of 3–4 was inconclusive.[7] NAS does not assess fibrosis; however, the fibrosis stage is generally reported alongside the NAS score. The second system, the Steatosis Activity Fibrosis (SAF) score, can be used for both grading/staging of disease severity and diagnosis.[9,10] Rather than a sum of scores, SAF gives separate semiquantitative scores for steatosis (0–3), activity (0–4), and fibrosis (0–4), with the activity score based on the unweighted sum of the scores for two features, ballooning (0–2) and lobular inflammation (0–2).[10] In the first study of this system, the activity score correlated strongly with a NASH diagnosis; 92% of patients with an activity score of ≥2 had NASH, and all patients with an activity score of <2 did not have NASH. The two scoring systems also differ in the way hepatocyte ballooning is evaluated.[8] With the NAS, the intensity of the ballooning is evaluated; the ballooning score would be 1 if only a few ballooned cells were observed, and 2 if there were many ballooned cells. When using the SAF system, the ballooning score is related to the size of the hepatocyte; the presence of ballooned cells a similar size to normal hepatocytes would be graded as 1, while an enlarged ballooned cell (≥2x the size of a normal hepatocyte) would be graded as 2.[9]

Few studies have compared both scoring systems in large cohorts. However, one study performed by a group in Delhi used both NAS and SAF systems to retrospectively evaluate liver biopsy samples from 1000 patients with NAFLD.[8] For NAS, a score of >5 was classified as NASH, 3–4 as borderline NASH, and <3 as non-NASH. For SAF, NASH was indicated by a score of ≥1 for each of the following three categories: steatosis, hepatocellular ballooning, and lobular inflammation. As per the definition of NASH, the fibrosis stage was not used for diagnosis. Using NAS, the diagnosis was definite NASH for 618 patients, borderline/possible NASH

for 310, and not NASH for 72, while using SAF, 883 were diagnosed as definite NASH, and 117 patients as not NASH. Thus, of the 310 patients diagnosed as borderline/possible NASH using NAS, 273 (88.06%) were diagnosed as definite NASH using SAF. Furthermore, of those with NAS scores of 3, 79.4% were diagnosed as NASH using the SAF score, compared with 94.4% with a NAS score of 4. As mentioned above, the NAS system was not designed as a diagnostic tool, and these findings highlight the importance of providing individual grades for each histologic feature of NAFLD, to enable better interpretability for clinicians and pathologists, along with easier evaluation of treatment responses.

NASH diagnosis: should additional morphologic features be included?

Based on the current definition, the diagnosis of NASH via liver biopsy is relatively straightforward, requiring the presence of the three above-mentioned histologic features. However, it is possible that this definition has been oversimplified, with a consequential loss of diagnostic accuracy. Older papers outlining pathologic criteria for the diagnosis of NASH consider additional features, such as perisinusoidal fibrosis and portal inflammation;[11-13] perhaps these additional features should be included in a new scoring system?

The rationale for including perisinusoidal fibrosis is that NAFLD is a dynamic process, in which disease activity can lead to fibrosis. Perisinusoidal fibrosis is an early marker of aggressive disease, which typically starts in the centrilobular area of the liver and is triggered by lobular inflammation and, more importantly, by hepatocellular ballooning. Furthermore, the between-pathologist reproducibility of perisinusoidal fibrosis scoring is far greater than that for lobular inflammation and ballooning. Therefore, taking the presence of perisinusoidal fibrosis into account may improve the accuracy of NASH diagnoses.

Another important histologic feature, briefly discussed by Dominique Valla in another article in this series, is portal inflammation. In fact, portal inflammation was originally proposed to be included in the NAS system; however, due to the low number of patients included in the validation study (N=50), this feature did not reach significance.[7]

Fibrosis staging: the most relevant histologic endpoint

Although fibrosis is not included as a diagnostic criterion, it is observed in the majority of patients with NASH,[14] and, as outlined by Dominique Valla, it is a key prognostic indicator for NAFLD.[15] In the context of NAFLD, fibrosis staging is very specific, and takes into account sinusoidal fibrosis occurring within the lobule and portal fibrosis.[10] The stages range from 0 (no fibrosis) to 4 (cirrhosis). As mentioned above, in contrast to grading of lobular inflammation or ballooning, the agreement between pathologists concerning fibrosis staging is very good. However, the semiquantitative nature of this system means that the heterogeneity of fibrosis within stages is not completely captured.[16] The magnitude of fibrosis in stage 3 is the best example of this heterogeneity, with biopsies showing any number of fibrous septa (between one and many) without nodulation.

Liver biopsy: an imperfect reference standard

Liver biopsy is associated with several challenging issues that make it an imperfect reference standard. It is an invasive procedure, with an accompanying risk of complications,[14] and is not suitable for screening large numbers of patients with T2D and NAFLD.[15] There is the possibility of sampling errors; even when of adequate size, a biopsy sample may not be representative of the overall histology of a patient's liver.[16] Furthermore, although reduced by standardizing histologic scoring systems, intra- and inter-observer variability remains an issue, the degree of which can depend upon the level of a pathologist's expertise, as well as the type of morphologic feature being assessed. As an example, for lobular inflammation, the consistency of assessment is only fair to good between pathologists (kappa 0.33–0.45) and even for different readings by the same pathologist (kappa 0.37–0.60).[14]

NAFLD assessment: the future for pathologists

Digital pathology

Improved, objective, and reliable quantitative NAFLD assessment tools are required, and numerous researchers are working towards this goal.[5] One such improvement is digital pathology, in which a stained tissue section is converted to a digital image via a scanning device. Thus, pathologists use a computer rather than a microscope to evaluate tissue sections, allowing deeper image analysis. Dedicated software allows quantitative and qualitative analysis of various morphologic features, resulting in improved objectivity, sensitivity, and accuracy. For example, algorithms have been developed to quantify the amount of steatosis,[17] and the area or number of inflammatory cells.[18] It is also possible to obtain automated quantification of fibrosis using conventional staining, such as Sirius red, or more sophisticated tools, such as dual photon imaging microscopy.[19] In addition, we can use artificial intelligence-based models that will help us to recognize, from routine stained slides, morphologic features that are not accessible to the human eye, and to apply

computer-aided deep learning methods based on convolutional neuronal networks (CNNs).[20]

Digital pathology: experience in NAFLD

Several papers outlining the use of digital pathology in NAFLD assessment have been published recently. One such study used data from more than 200 patients with biopsy-proven NAFLD.[21] Expert pathologists manually annotated histologic features from slides obtained from the first 100 patients (the derivation cohort), and these slides were then scanned and used to facilitate training of a machine learning algorithm that was able to identify these annotated features. Biopsy data from the remaining patients (the validation cohort) were used to validate the algorithm. In the derivation cohort, there was excellent concordance between the manual annotation from the pathologist and the automatic measurement. However, there was only a moderate correlation between the semi-quantitative analyses performed by the pathologist and the quantitative analysis performed by the software. Importantly, the results showed an overestimation of steatosis by the pathologist, and confirmed that as the fibrosis stage increases, the amount of fibrosis tissue increases in an exponential, not linear, fashion.

Classification accuracy rates of >90% were achieved in a study that used biopsy data from 79 patients with NAFLD to implement and test a topological data analysis method of assessing hepatocellular ballooning to classify patients as having NASH or NAFLD.[22] Another study, which used more than 5000 liver biopsy samples obtained in 3 clinical trials involving more than 3000 patients with advanced fibrosis and NASH, developed and validated deep CNNs to quantify histologic features of NASH.[23] One important finding from this study was the heterogeneity of fibrosis, both within a single patient's biopsy and between different patients with the same stage of fibrosis. For example, within a single slide from one patient with advanced fibrosis, in some areas the staging will be F0, while in other areas it will be F3 or F4. This study also identified morphologic parameters that are predictive of outcome, including the ratio of steatosis to hepatocellular ballooning, and the level of portal inflammation.

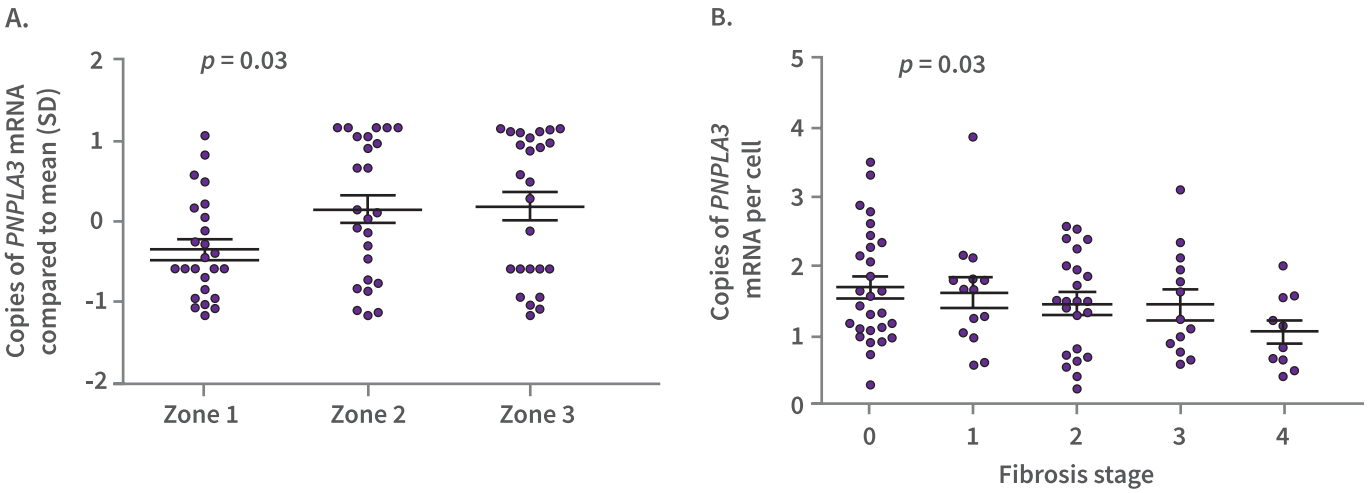
Another study used an automated tool to quantify the amount and assess the architectural patterns of fibrosis.[24] As pathologists, we emphasize that assessing the location and architectural organization of the fibrosis is even more important than evaluating the severity and extent of fibrosis. Although this study used biopsies from only 18 patients, they covered the entire spectrum of NAFLD and fibrosis severity, and two pathologists provided almost 1000 annotations to develop, train, and test machine learning models. They showed good to excellent correlation between the automatic quantification of fibrosis (as measured by collagen proportionate area [CPA]) and the staging by the pathologists. However, they found a significant overlap in CPA across the different fibrosis stages; that is, a patient with F2 may have a similar CPA to one with F4, meaning that CPA alone should not be used to guide patient management. Importantly, the automated tool was also able to distinguish different patterns of fibrosis.

As a final example of the use of digital pathology, a study using samples from 87 patients with NAFLD (≈80% with NASH) aimed to evaluate the contribution of gene transcription to disease severity.[25] RNA-scope results showed that the majority of patatin-like phospholipase domain-containing protein 3 (PNPLA3) mRNA was in hepatocytes, while collagen 1α (COL1α) mRNA was predominately seen in the portal area, as expected, within the myofibroblast; approximately half of the myofibroblasts also expressed PNPLA3. They also showed that PNPLA3 mRNA was less abundant in the hepatocytes located in zone one, and that the level of PNPLA3 transcription decreased as the severity of fibrosis increased (**Figure 1**).

Non-invasive diagnosis of NASH in T2D

Although the digital pathology studies discussed above represent significant advances in the diagnosis and classification of NAFLD, they rely on the use of liver biopsies. Considerable research is also being invested into minimally or non-invasive tests for NAFLD assessment. As an example, our ongoing Quantitative Imaging in Diabetes-NASH (QUID-NASH) research study includes a large number of patients with T2D who require a liver biopsy as part of their care.[26] The ultimate goal of this study is to develop a 'virtual liver biopsy' to assess the main histologic features of NASH, and interim results are encouraging.[4] As part of this ongoing study, we are performing semiquantitative analyses of histologic features and developing quantitative algorithms for steatosis, lobular inflammation, portal inflammation, and fibrosis discrimination (perisinusoidal vs. portal). We are also investigating ductular reactions within the portal tract. We will compare the standard semi-quantitative analyses with non-invasive approaches such as magnetic resonance imaging and blood tests, and also evaluate the correlation of histologic features with clinical biologic parameters and NAFLD disease severity in these patients with T2D.

Figure 1: *Patatin-like phospholipase domain-containing protein 3 (PNPLA3) transcription: A) in zone 1, 2, and 3 hepatocytes compared with mean (standard deviation [SD]) mRNA/cell in all zones, and B) per cell across different fibrosis stages. Reprinted from JHEP Rep, 1(3), Sandhu B, Perez Matos MC, Tran S, Zhong A, Csizmadia E, Kim M, Herman MA, Nasser I, Lai M, Jiang ZG, Quantitative digital pathology reveals association of cell-specific PNPLA3 transcription with NAFLD disease activity, 199-202, Copyright 2019, with permission from Elsevier.*



Conclusion

Although liver biopsy is the reference standard for the diagnosis of NASH, it is an imperfect standard. For biopsy analysis, there are two key histologic scoring systems available. They are not interchangeable, and they each have specific issues. Some of these issues may be resolved through the use of digital pathology which, as an example, allows objective and accurate quantification of the elementary morphologic features, thus enabling improved patient stratification and monitoring of response to therapy. Furthermore, using a machine learning approach will allow the development of computer-assisted classification of morphologic patterns, and also produce single-cell analyses that may provide powerful mechanistic insights. In addition, our study investigating non-invasive tools for the assessment of NAFLD in T2D is underway, and initial results are encouraging.

Medical writing assistance was provided by Toni Dando, on behalf of Springer Healthcare Communications, and funded by L'Institut Servier.

References

1. Younossi ZM, Golabi P, de Avila L, et al. The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: A systematic review and meta-analysis. *J Hepatol.* 2019;71(4):793-801.
2. Chalasani N, Younossi Z, Lavine JE, 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(1):328-357.
3. EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *J Hepatol.* 2016;64(6):1388-1402.
4. Poynard T, Paradis V, Mullaert J, et al. Prospective external validation of a new non-invasive test for the diagnosis of non-alcoholic steatohepatitis in patients with type 2 diabetes. *Aliment Pharmacol Ther.* 2021;54(7):952-966.
5. Zhou YJ, Wong VW, Zheng MH. Consensus scoring systems for nonalcoholic fatty liver disease: an unmet clinical need. *Hepatobiliary Surg Nutr.* 2021;10(3):388-390.
6. Tannapfel A, Dienes HP, Lohse AW. The indications for liver biopsy. Dtsch Arztebl Int. 2012;109(27-28):477-483.
7. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology.* 2005;41(6):1313-1321.
8. Rastogi A, Shasthry SM, Agarwal A, et al. Non-alcoholic fatty liver disease - histological scoring systems: a large cohort single-center, evaluation study. *Apmis.* 2017;125(11):962-973.
9. Bedossa P. 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(2):565-575.
10. Bedossa P, Poitou C, Veyrie N, et al. Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. *Hepatology.* 2012;56(5):1751-1759.
11. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol.* 1999;94(9):2467-2474.
12. Brunt EM, Kleiner DE, Wilson LA, et al. Portal chronic inflammation in nonalcoholic fatty liver disease (NAFLD): a histologic marker of advanced NAFLD-Clinicopathologic correlations from the nonalcoholic steatohepatitis clinical research network. *Hepatology.* 2009;49(3):809-820.
13. Younossi ZM, Stepanova M, Rafiq N, et al. Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality. *Hepatology.* 2011;53(6):1874-1882.
14. Wang L, Shuangni Y, Chan APoN-AFLDIJDD, 2:1. Pathology of Non-Alcoholic Fatty Liver Disease. *Int J Dig Dis.* 2016;2(1):1-7.
15. Geier A, Tiniakos D, Denk H, Trauner M. From the origin of NASH to the future of metabolic fatty liver disease. *Gut.* 2021;70(8):1570-1579.
16. Soon G, Wee A. Updates in the quantitative assessment of liver fibrosis for nonalcoholic fatty liver disease: Histological perspective. *Clin Mol Hepatol.* 2021;27(1):44-57.
17. Munsterman ID, van Erp M, Weijers G, et al. A Novel Automatic Digital Algorithm that Accurately Quantifies Steatosis in NAFLD on Histopathological Whole-Slide Images. *Cytometry B Clin Cytom.* 2019;96(6):521-528.
18. Marti-Aguado D, Rodríguez-Ortega A, Mestre-Alagarda C, et al. Digital pathology: accurate technique for quantitative assessment of histological features in metabolic-associated fatty liver disease. *Aliment Pharmacol Ther.* 2021;53(1):160-171.
19. Xu S, Wang Y, Tai DCS, et al. qFibrosis: a fully-quantitative innovative method incorporating histological features to facilitate accurate fibrosis scoring in animal model and chronic hepatitis B patients. *J Hepatol.* 2014;61(2):260-269.
20. Calderaro J, Kather JN. Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers. *Gut.* 2021;70(6):1183-1193.
21. Forlano R, Mullish BH, Giannakeas N, et al. High-Throughput, Machine Learning-Based Quantification of Steatosis, Inflammation, Ballooning, and Fibrosis in Biopsies From Patients With Nonalcoholic Fatty Liver Disease. *Clin Gastroenterol Hepatol.* 2020;18(9):2081-2090.e9.
22. Teramoto T, Shinohara T, Takiyama A. Computer-aided classification of hepatocellular ballooning in liver biopsies from patients with NASH using persistent homology. *Comput Methods Programs Biomed.* 2020;195:105614.
23. Taylor-Weiner A, Pokkalla H, Han L, et al. A Machine Learning Approach Enables Quantitative Measurement of Liver Histology and Disease Monitoring in NASH. *Hepatology.* 2021;74(1):133-147.
24. Gawrieh S, Sethunath D, Cummings OW, et al. Automated quantification and architectural pattern detection of hepatic fibrosis in NAFLD. *Ann Diagn Pathol.* 2020;47:151518.
25. Sandhu B, Perez Matos MC, Tran S, et al. Quantitative digital pathology reveals association of cell-specific PNPLA3 transcription with NAFLD disease activity. *JHEP Rep.* 2019;1(3):199-202.
26. The QUID-NASH research program. https://rhu-quidnash.com/. Published 2022. Accessed 11 February, 2022.