



Published in final edited form as:

J Clin Gastroenterol. 2011 January ; 45(1): 55–58. doi:10.1097/MCG.0b013e3181dd1348.

Generalizability of the NASH CRN Histological Scoring System for Nonalcoholic Fatty Liver Disease

Ravi Juluri, MD¹, Raj Vuppalanchi, MD¹, John Olson, MD², Aynur Ünalp, MD⁴, Mark L. Van Natta, MHS⁴, Oscar W. Cummings, MD³, James Tonascia, PhD⁴, and Naga Chalasani, MD¹

¹ Department of Medicine; Indiana University School of Medicine, Indianapolis, Indiana ²Witham Health Services, Lebanon, Indiana ³ Department of Pathology; Indiana University School of Medicine, Indianapolis, Indiana ⁴ Center for Clinical Trials, The Johns Hopkins University, Bloomberg School of Public Health, Baltimore, Maryland

Abstract

Goals and Background—The recently developed histological scoring system for NAFLD by the NASH Clinical Research Network (NASH CRN) is becoming increasingly popular. However, its generalizability to a community setting has not been evaluated. We conducted a study to compare a community general pathologist to an expert hepatopathologist in assessing NAFLD using the NASH CRN scoring system.

Study—Forty eight consecutive patients with suspected NAFLD underwent liver biopsy. Histological features of interest such as steatosis, lobular inflammation, ballooning degeneration, fibrosis, NAFLD Activity Score (NAS) and the presence of NASH were scored in a blinded fashion by the two pathologists on two separate occasions 3 months apart.

Results—The mean (\pm SD) length of the liver biopsy samples was 25 ± 5 mm. Inter-observer agreement (kappa) between two pathologists was 0.62(0.45-0.80) for steatosis, 0.44(0.23-0.65) for lobular inflammation, 0.25(0.11-0.38) for ballooning, 0.40 for NAS(0.28-0.52) and 0.35 (0.19-0.52) for fibrosis. The two pathologists diagnosed “definite NASH” in a similar proportion of patients (56% vs. 57%), but their inter-observer agreement was only 0.46 (0.24-0.67) as they both diagnosed different levels of NASH (borderline vs. definite) in different subjects. Intra-observer agreement was generally comparable for steatosis, lobular inflammation, NAS and diagnosis of NASH, but not for fibrosis.

Conclusions—Clinically important differences exist between community general pathologist and expert hepatopathologist in assessing NAFLD using the NASH CRN scoring system. More studies are needed to investigate its suitability for community-based clinical practice.

Keywords

Sampling error; Liver biopsy; Non-alcoholic steatohepatitis

Author for Correspondence Raj Vuppalanchi, MD Indiana University School of Medicine Regenstrief Health Center, Room 4100 1050 Wishard Blvd. Indianapolis, IN 46202-2859 Tel (317) 278-1652 Fax (317) 278-1949 rvuppala@iupui.edu.

Financial disclosures: Authors have no relevant conflicts of interest to declare.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

Non alcoholic fatty liver disease (NAFLD) is one of the most common forms of chronic liver disease affecting nearly a third of adults in the United States.¹⁻³ It is the hepatic component of metabolic syndrome and is associated with deposition of triglycerides in the hepatocytes.⁴ The clinical spectrum of NAFLD ranges between hepatic steatosis that is generally considered benign to non-alcoholic steatohepatitis (NASH) that can progress to cirrhosis, liver failure or hepatocellular cancer.⁵⁻⁷ While the diagnosis of NAFLD can be made based on clinical and imaging criteria, steatohepatitis (NASH) is a histological diagnosis and requires liver biopsy.

Since the original description by Ludwig et al., in 1980,⁸ many histological features that characterize NASH and several scoring schema have been published.⁹⁻¹² The recently published histological scoring scheme of the NASH Clinical Research Network (NASH CRN) is increasing in popularity by both clinical and research communities.¹³ This scoring system was developed by the Pathology Subcommittee of the NASH CRN that comprised of nine expert hepato-pathologists. In principle, this scoring system comprises of NAFLD Activity Score (NAS), fibrosis stage and identification of NASH by pattern recognition (gestalt).¹³ The NAS can range from 0 to 8 and is calculated by the sum of scores of steatosis (0-3), lobular inflammation (0-3) and hepatocyte ballooning (0-2). In patients with NAFLD, NAS score of ≥ 5 strongly correlated with a diagnosis of “definite NASH” whereas $\text{NAS} \leq 3$ correlated with a diagnosis of “not NASH”.¹³ As the NASH CRN Scoring System was developed and validated by a group of expert academic hepatopathologists, it is unclear if it can be generalized to community-based general pathologists. Furthermore, the relationship between different NAS and the diagnosis of NASH by pattern recognition for a community pathologist is not known. Therefore, we conducted a study to examine if the NASH CRN Scoring System is generalizable to community-based general pathologists. Our study objectives were (a) to assess the performance of a community-based general pathologist in the interpretation of histological features of NAFLD as compared to an expert hepatopathologist and (b) to assess the relationship between NAS and the diagnosis of “definite NASH” for community-based general pathologist as compared to expert hepato-pathologist.

Materials and Methods

This study consisted of forty eight consecutive patients with suspected NAFLD who had undergone percutaneous liver biopsy for clinical purposes at Indiana University Hospital from July 2004 to June 2006. These patients were extensively evaluated by blood tests and imaging studies to exclude competing etiologies. None of the subjects had ≥ 7 drinks per week of alcohol on average over the preceding 5 year period. All liver biopsies were performed by radiologists percutaneously under ultrasound guidance using an 18 gauge automated biopsy gun (Bard® Monopty®). Liver biopsy samples are sent to the histopathology lab in a single formalin bottle for standard tissue fixation. The liver biopsy slides from each patient were stained with H&E and Masson trichrome stain under a standard protocol. This study has been reviewed and approved by the local institutional review board. These liver biopsies were used for another study that examined the relationship between biopsy length and histological yield and the results are published elsewhere.¹⁴

A community-based general pathologist (J.O) and an expert hepato-pathologist (O.C) independently examined all liver biopsy slides, 3 months apart on two separate occasions in a blinded fashion to score steatosis, lobular inflammation, hepatocellular ballooning, fibrosis and calculated NAS using the published NASH CRN criteria.¹³ In addition, using pattern

recognition, both pathologists assessed for steatohepatitis in each set of slides and categorized them as “definite”, “borderline” or “not”. The community-based general pathologist has been in practice for nearly two decades and serves as the lead pathologist for one of the local community hospitals and interprets a wide variety of human pathological specimens. The expert hepato-pathologist is a member of the Pathology Subcommittee that developed the NASH CRN Scoring System.¹³ To mimic real-life practice, we offered no training session for the community general pathologist but he familiarized himself with the NASH CRN scoring system as published in the paper. For remainder of the manuscript, community-based general pathologist will be referred to as community pathologist and expert hepatopathologist as expert pathologist.

Data analysis

Steatosis, lobular inflammation, hepatocyte balloon degeneration, fibrosis, NAS and the presence of NASH by pattern recognition were systematically assessed according to the published NASH CRN Scoring System. These outcome measures were compared between community pathologist and expert pathologist. Coefficient of concordance (Kappa statistic) was utilized to assess the intra- and inter-observer agreement in the interpretation of histological features. A Kappa value of 0.2 – 0.39 was considered as “fair”, 0.4 - 0.59 as “moderate”, 0.6 – 0.79 as “substantial” and ≥ 0.8 as “perfect” agreement.¹⁵ P-values were nominal and were derived from ordered logistic regression with robust variance estimation to account for within patient correlation. The test of whether the relationship of diagnosis of NASH on NAS varied by type of pathologist was assessed using ordered logistic regression on NASH with an interaction term for NAS by pathologist. Statistical analyses used both Stata 9.0 (StataCorp, Stata Statistical Software: Release 9. College Station, TX: StataCorp LP, 2005) and SAS 8.0 (SAS Institute Inc., SAS/STAT User's Guide, Version 8, Cary NC: SAS Institute Inc., 1999).

Results

Liver biopsy samples from 48 patients with NAFLD were included in this study. All patients were Caucasian (50% women) and their mean age (S.D.) was 46.2 ± 9.7 years and mean BMI was 33.2 ± 6.2 kg/m². The biochemical tests revealed normal total bilirubin, alkaline phosphatase, albumin and coagulation parameters but AST and ALT were elevated at values 72 ± 52 U/l and 81 ± 47 IU/L, respectively. The mean \pm S.D. length of liver biopsy sample was 25 ± 5 mm.

Compared to expert pathologist, general pathologist's mean (s.e.) grade for steatosis, (1.77 ± 0.13 vs. 1.38 ± 0.13 , $p=0.0002$), ballooning (1.38 ± 0.07 vs. 0.70 ± 0.11 , $p<0.0001$), fibrosis (2.13 ± 0.14 vs. 1.46 ± 0.17 , $p<0.0001$) and NAS (4.71 ± 0.21 vs. 3.57 ± 0.23 , $p<0.0001$) were significantly higher. There was no difference between two pathologists' in the grade for lobular inflammation (Table 1). The inter-observer agreement between general and expert pathologists was “substantial” [0.62 (0.45-0.80)] for steatosis, “moderate” [0.44 (0.23-0.65)] for lobular inflammation, “fair” for both ballooning [0.25 (0.11-0.38)] and fibrosis [0.35 (0.19-0.52)] (Table 1). Compared to expert pathologist, general pathologist classified a significantly lower proportion of liver biopsies as “no NASH” (6% vs. 17%, $p=0.01$) but a comparable number of biopsies as “borderline” (38% vs. 26%, $p=0.2$) and “definite NASH” (56% vs. 57%, $p=0.9$) (Table 1). In their first reading out of 47 available pairs, the general and expert pathologists agreed there were 3, 6, and 20 cases of no, borderline, and definite NASH, respectively. Disagreement occurred in which 5, 6, and 7 cases of no, borderline, and definite NASH, respectively, diagnosed by the expert were called as borderline, definite, and borderline NASH, respectively, by the general pathologist. Note, there were no cases where the general pathologist called no NASH and the expert pathologist called borderline or definite NASH; similarly, there were no cases where the

expert pathologist called no NASH and the general pathologist called definite NASH.” Although they diagnosed “definite” and “borderline” NASH in comparable proportion of liver biopsies, but their inter-observer agreement was only moderate ($\kappa=0.46$) because they diagnosed different levels of NASH (borderline vs. definite) in different liver biopsies.

The intra-observer agreements for general and expert pathologists are shown in Table 2. In general, their intra-observer agreements were comparable for steatosis, lobular inflammation, NAS and diagnosis of NASH but not for ballooning or fibrosis. For both ballooning and fibrosis, community pathologist had lower intra-observer agreement than the hepatopathologist (Table 2).

The relationship between NAS and the diagnosis of NASH are shown in Figures 1 and 2. For biopsies read by the general pathologist, biopsies with NAS = 2 (there were no cases of NAS = 1), NAS = 3 or 4 and NAS \geq 5 were primarily diagnosed as “not NASH”, borderline NASH, and “definite NASH”, respectively (Figure 1). However, the relationship between diagnosis of NASH and NAS was more complex for the expert pathologist (Figure 2). Only NAS = 1 and NAS > 5 completely correlated with “not NASH” and “definite NASH”, respectively. However, there was no statistical evidence that the relationship between NASH diagnosis and NAS varied by type of pathologist ($p=0.53$ by ordered logistic regression). When receiver operating curves were constructed for the outcome of definite NASH (regressed on NAS as the explanatory variable), the area under the curve was 0.89 (95% CI: 0.80-0.98) for the expert pathologist and 0.99 (95%CI=0.97, 1.00) for the general pathologist.

Discussion

Since its publication, the NASH CRN histological scoring system has become widely popular among the clinical investigators for systematically assessing liver histology in patients with NAFLD.¹⁶⁻¹⁹ Although not necessarily intended for use in clinical practice, it has been our anecdotal experience that many community pathologists have started to apply the NASH CRN scoring system in their clinical practice. As the NASH CRN scoring system was developed by a group of academic pathologists with extensive experience in liver pathology, it is not known if it can be generalized to community-based general pathologists. This prompted us to conduct this study in which we systematically compared and found important clinical differences between community-based general pathologist and expert hepatopathologist in assessing liver histology in patients with NAFLD according to the NASH CRN scoring system.

Our study makes several important observations. First, compared to an expert pathologist, community-based general pathologist assigned higher scores for almost all histological variables which translated into significantly higher NAS and also diagnosed more patients with borderline/definite NASH. If we were to assume expert pathologist as the gold standard, then our study general pathologist has misdiagnosed NASH in 10% of our study patients. Second, although both pathologists diagnosed definite NASH in almost equal proportion of study patients, inter-observer agreement was only moderate because they diagnosed definite NASH in different patients. This finding may not have clinical implications currently because there are no approved treatments for NASH to be administered in community clinical practice. However, if there were ever to be a proven medical treatment then this discrepancy may lead to unnecessary treatment for some patients. Third, compared to expert pathologist, community general pathologist had generally comparable intra-observer agreement for steatosis, inflammation, NAS and NASH diagnosis but much lower consistency in assessing ballooning and fibrosis. Although consistency in assessing NAS and diagnosing NASH is reassuring, significantly lower

consistency in staging fibrosis is somewhat surprising because it is generally thought to be less influenced by subjectivity.²⁰

Although it has been stated NASH should be histologically diagnosed only by pattern recognition rather than based on a numerical score,¹³ the relationship we found between NAS and NASH diagnosis is quite compelling. We believe that more work needs to be done to explore if NAS or its modification including fibrosis stage could be used to diagnose NASH.

One limitation of our study is that we evaluated only one community-based general pathologist and thus it is unknown if our findings are valid reflection of community pathologists at large. We have considered having more than one community-based general pathologist participate in our study, but we believed that repeating our study at a different geographic location by a different group of investigators might be more informative.

In conclusion, we found clinically important differences between community-based general pathologist and expert hepatopathologist in assessing NAFLD histology according to the NASH CRN scoring system. As there is compelling relationship between NAS and the diagnosis of NASH, more work needs to be done to explore the utility of NAS or its modification with fibrosis in diagnosing NASH in clinical practice.

Acknowledgments

Supported by in part by NIH grant K24 DK69290 (NC).

References

1. Clark JM, Brancati FL, Diehl AM. The prevalence and etiology of elevated aminotransferase levels in the United States. *Am J Gastroenterol* 2003;98:960–7. [PubMed: 12809815]
2. Browning JD, Szczepaniak LS, Dobbins R, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *Hepatology* 2004;40:1387–95. [PubMed: 15565570]
3. Liangpunsakul S, Chalasani N. Unexplained elevations in alanine aminotransferase in individuals with the metabolic syndrome: results from the third National Health and Nutrition Survey (NHANES III). *Am J Med Sci* 2005;329:111–6. [PubMed: 15767815]
4. Marchesini G, Bugianesi E, Forlani G, et al. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. *Hepatology* 2003;37:917–23. [PubMed: 12668987]
5. Adams LA, Lymp JF, St Sauver J, et al. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. *Gastroenterology* 2005;129:113–21. [PubMed: 16012941]
6. Fassio E, Alvarez E, Dominguez N, et al. Natural history of nonalcoholic steatohepatitis: a longitudinal study of repeat liver biopsies. *Hepatology* 2004;40:820–6. [PubMed: 15382171]
7. Harrison SA, Neuschwander-Tetri BA. Nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *Clin Liver Dis* 2004;8:861–79. ix. [PubMed: 15464659]
8. Ludwig J, Viggiano TR, McGill DB, et al. Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clin Proc* 1980;55:434–8. [PubMed: 7382552]
9. Brunt EM, Janney CG, Di Bisceglie AM, et al. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol* 1999;94:2467–74. [PubMed: 10484010]
10. Brunt EM, Neuschwander-Tetri BA, Oliver D, et al. Nonalcoholic steatohepatitis: histologic features and clinical correlations with 30 blinded biopsy specimens. *Hum Pathol* 2004;35:1070–82. [PubMed: 15343508]
11. Matteoni CA, Younossi ZM, Gramlich T, et al. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. *Gastroenterology* 1999;116:1413–9. [PubMed: 10348825]
12. Younossi ZM, Gramlich T, Liu YC, et al. Nonalcoholic fatty liver disease: assessment of variability in pathologic interpretations. *Mod Pathol* 1998;11:560–5. [PubMed: 9647594]

13. Kleiner DE, Brunt EM, Van Natta M, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005;41:1313–21. [PubMed: 15915461]
14. Vuppalanchi R, Unalp A, Natta ML, et al. Effects of Liver Biopsy Sample Length and Number of Readings on Histologic Yield for Nonalcoholic Fatty Liver Disease. *Clin Gastroenterol Hepatol.* 2008
15. Ratziu V, Charlotte F, Heurtier A, et al. Sampling variability of liver biopsy in nonalcoholic fatty liver disease. *Gastroenterology* 2005;128:1898–906. [PubMed: 15940625]
16. Guruprasad PA, James AT, Philip VK, et al. Randomized, placebo controlled trial of pioglitazone in non-diabetic subjects with nonalcoholic steatohepatitis (NASH). *Gastroenterology.* 2008
17. Belfort R, Harrison SA, Brown K, et al. A placebo-controlled trial of pioglitazone in subjects with nonalcoholic steatohepatitis. *N Engl J Med* 2006;355:2297–307. [PubMed: 17135584]
18. Ratziu V, Giral P, Jacqueminet S, et al. Rosiglitazone for nonalcoholic steatohepatitis: one-year results of the randomized placebo-controlled Fatty Liver Improvement with Rosiglitazone Therapy (FLIRT) Trial. *Gastroenterology* 2008;135:100–10. [PubMed: 18503774]
19. Harrison SA, Fecht W, Brunt EM, et al. Orlistat for overweight subjects with nonalcoholic steatohepatitis: A randomized, prospective trial. *Hepatology* 2009;49:80–6. [PubMed: 19053049]
20. Wilson S, Chalasani N. Noninvasive markers of advanced histology in nonalcoholic fatty liver disease: are we there yet? *Gastroenterology* 2007;133:1377–8. discussion 1378-9. [PubMed: 17919509]

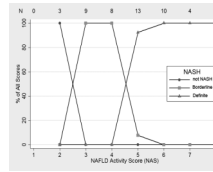


Figure 1. Relationship between NAFLD Activity Score and the Diagnosis of NASH by Pattern Recognition by the Community General Pathologist

For each activity score represented on the X axis, the percentage of observations with a particular histological diagnosis is shown on the Y axis. The total number of observations for each activity score is shown across the top of the graph.

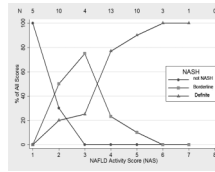


Figure 2. Relationship between NAFLD Activity Score and the Diagnosis of NASH by Pattern Recognition by the Expert Hepatopathologist

For each activity score represented on the X axis, the percentage of observations with a particular histological diagnosis is shown on Y axis. The total number of observations for each activity score is shown across the top of the graph.

Table 1

Inter-observer agreement between community-based general pathologist and expert hepatopathologist for assessing NAFLD histology according to the NASH CRN scoring system.

Histological features	Histology Score[¶]		Inter-observer Agreement Kappa (95% CI)[*]
	Community General Pathologist	Expert Hepatopathologist	
Steatosis (0-3)	1.77±0.13	1.38±0.13	0.62 (0.45-0.80)
Lobular Inflammation (0-3)	1.56±0.10	1.47±0.09	0.44 (0.23-0.65)
Ballooning (0-2)	1.38±0.07	0.70±0.11	0.25 (0.11-0.38)
Fibrosis (stage 0-4)	2.13±0.14	1.46 ±0.17	0.35 (0.19-0.52)
NAFLD Activity Score(0-8)	4.71±0.21	3.57±0.23	0.40 (0.28-0.52)
Diagnosis of NASH			
No	6%	17%	
Borderline	38%	26%	
Definite	56%	57%	0.46 (0.24-0.67)

* P-value < 0.001 for all kappa values.

¶ Values represent mean ± s.e. unless indicated otherwise.

Table 2

Intra-observer agreement for community general pathologist and expert hepatopathologist for assessing NAFLD histology according to the NASH CRN scoring system.[¶]

Histological features	Community General Pathologists			Expert Hepatopathologist		
	1 st reading	2 nd reading	Kappa (95% CI)*	1 st reading	2 nd reading	Kappa (95% CI)*
Steatosis (0-3)	1.77±0.13	1.85±0.12	0.71 (0.55-0.87)	1.38±0.13	1.48±0.13	0.77 (0.62-0.92)
Lobular Inflammation (0-3)	1.56±0.10	1.65±0.10	0.39 (0.17-0.61)	1.47±0.09	1.62±0.08	0.51 (0.27-0.75)
Ballooning (0-2)	1.38±0.07	1.65±0.07	0.43 (0.23-0.64)	0.70±0.11	0.91±0.11	0.68 (0.51-0.84)
Fibrosis (stage 0-4)	2.13±0.14	2.46±0.13	0.48 (0.28-0.68)	1.46±0.17	1.54±0.18	0.80 (0.69-0.90)
NAFLD Activity Score (0-8)	4.71±0.21	4.98±0.21	0.58 (0.46-0.71)	3.57±0.23	3.98±0.22	0.61 (0.49-0.74)
Presence of NASH						
% No	6	4		17	15	
% Borderline	38	38	0.80 (0.65-0.95)	26	23	0.81 (0.65-0.96)
% Definite	56	58		57	62	

[¶] Values indicate mean ± s.e. unless indicated otherwise.

* P-value < 0.001 for all kappa values.