

# Design and Validation of a Histological Scoring System for Nonalcoholic Fatty Liver Disease

David E. Kleiner,<sup>1</sup> Elizabeth M. Brunt,<sup>2</sup> Mark Van Natta,<sup>3</sup> Cynthia Behling,<sup>4</sup> Melissa J. Contos,<sup>5</sup> Oscar W. Cummings,<sup>6</sup> Linda D. Ferrell,<sup>7</sup> Yao-Chang Liu,<sup>8</sup> Michael S. Torbenson,<sup>9</sup> Aynur Unalp-Arida,<sup>3</sup> Matthew Yeh,<sup>10</sup> Arthur J. McCullough,<sup>11</sup> and Arun J. Sanyal<sup>12</sup> for the Nonalcoholic Steatohepatitis Clinical Research Network<sup>13</sup>

Nonalcoholic fatty liver disease (NAFLD) is characterized by hepatic steatosis in the absence of a history of significant alcohol use or other known liver disease. Nonalcoholic steatohepatitis (NASH) is the progressive form of NAFLD. The Pathology Committee of the NASH Clinical Research Network designed and validated a histological feature scoring system that addresses the full spectrum of lesions of NAFLD and proposed a NAFLD activity score (NAS) for use in clinical trials. The scoring system comprised 14 histological features, 4 of which were evaluated semi-quantitatively: steatosis (0-3), lobular inflammation (0-2), hepatocellular ballooning (0-2), and fibrosis (0-4). Another nine features were recorded as present or absent. An anonymized study set of 50 cases (32 from adult hepatology services, 18 from pediatric hepatology services) was assembled, coded, and circulated. For the validation study, agreement on scoring and a diagnostic categorization ("NASH," "borderline," or "not NASH") were evaluated by using weighted kappa statistics. Inter-rater agreement on adult cases was: 0.84 for fibrosis, 0.79 for steatosis, 0.56 for injury, and 0.45 for lobular inflammation. Agreement on diagnostic category was 0.61. Using multiple logistic regression, five features were independently associated with the diagnosis of NASH in adult biopsies: steatosis ( $P = .009$ ), hepatocellular ballooning ( $P = .0001$ ), lobular inflammation ( $P = .0001$ ), fibrosis ( $P = .0001$ ), and the absence of lipogranulomas ( $P = .001$ ). The proposed NAS is the unweighted sum of steatosis, lobular inflammation, and hepatocellular ballooning scores. **In conclusion**, we present a strong scoring system and NAS for NAFLD and NASH with reasonable inter-rater reproducibility that should be useful for studies of both adults and children with any degree of NAFLD. NAS of  $\geq 5$  correlated with a diagnosis of NASH, and biopsies with scores of less than 3 were diagnosed as "not NASH." (HEPATOLOGY 2005;41:1313-1321.)

---

Abbreviations: NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; NIDDK, National Institute of Diabetes & Digestive & Kidney Diseases; NIH, National Institutes of Health; H&E, hematoxylin and eosin; NAS, NAFLD activity score.

From the <sup>1</sup>Laboratory of Pathology, National Cancer Institute, Bethesda, MD; <sup>2</sup>Saint Louis University Liver Center, Department of Pathology, Saint Louis University School of Medicine, St. Louis, MO; <sup>3</sup>Johns Hopkins Center for Clinical Trials, Baltimore, MD; <sup>4</sup>Department of Pathology, University of California San Diego, San Diego, CA; <sup>5</sup>Department of Pathology, Virginia Commonwealth University, Richmond, VA; <sup>6</sup>Department of Pathology, University Hospital, Indianapolis, IN; <sup>7</sup>Department of Pathology, University of California San Francisco, San Francisco, CA; <sup>8</sup>Department of Pathology, and <sup>11</sup>Division of Gastroenterology, MetroHealth Medical Center, Cleveland, OH; <sup>9</sup>Department of Pathology, Johns Hopkins Hospital, Baltimore, MD; <sup>10</sup>Department of Pathology, University of Washington Medical Center, Seattle, WA; and <sup>12</sup>Department of Internal Medicine, Division of Internal Medicine, Division of Gastroenterology, Virginia Commonwealth University Medical Center, Richmond, VA. <sup>13</sup>A list of members of the Nonalcoholic Steatohepatitis Clinical Research Network is located in the Appendix.

Received November 12, 2004; accepted February 25, 2005.

Supported by the National Institute of Diabetes & Digestive & Kidney Diseases (grants 1U01DK061718, 1U01DK061728, 1U01DK061731, 1U01DK061732, 1U01DK061734, 1U01DK061737, 1U01DK061738, 5U01DK061730, 7U01DK061713, and N01HR761019) and the National Institute of Child Health and Human Development.

Address reprint requests to: David E. Kleiner, Laboratory of Pathology, National Cancer Institute, Bldg 10, Room 2N212, MSC 1516, Bethesda, MD 20892. E-mail: KleinerD@mail.nih.gov; fax: 301-480-9488.

Copyright © 2005 by the American Association for the Study of Liver Diseases.

Published online in Wiley InterScience (www.interscience.wiley.com).

DOI 10.1002/hep.20701

Potential conflict of interest: Nothing to report.

**N**onalcoholic fatty liver disease (NAFLD) is increasingly recognized as the hepatic manifestation of insulin resistance and the systemic complex known as metabolic syndrome.<sup>1-4</sup> Prevalence estimates of NAFLD have used a variety of laboratory and imaging assessments and suggest that NAFLD may be the most common form of chronic liver disease in adults in the United States, Australia, Asia, and Europe, paralleling the “epidemic” of obesity in developed countries.<sup>5-9</sup>

Ludwig et al.<sup>10</sup> are credited with solidifying the nomenclature and pathological findings of nonalcoholic steatohepatitis (NASH) in a seminal manuscript published in 1980. Now recognized as a progressive form of fatty liver disease, NASH has been documented to have the potential to progress to cirrhosis and hepatocellular carcinoma.<sup>3,5</sup> NASH may be a leading cause of “cryptogenic cirrhosis” in which etiologically specific clinical laboratory or pathological features can no longer be identified.<sup>11</sup> NAFLD is also gaining recognition as a significant form of liver disease in pediatric populations, which in some patients may progress to cirrhosis in adulthood.<sup>12,13</sup>

Whereas laboratory test abnormalities and radiographic findings may be suggestive of NAFLD, histological evaluation remains the only means of accurately assessing the degree of steatosis, the distinct necroinflammatory lesions and fibrosis of NASH, and distinguishing NASH from “simple” steatosis, or steatosis with inflammation.<sup>14</sup> Matteoni et al.<sup>15</sup> showed that cirrhosis developed in 21% to 28% of patients whose index biopsies had shown the combination of lesions of steatosis, inflammation, ballooning, and Mallory’s hyaline or fibrosis, whereas only 4% of patients with simple steatosis and none of the patients with steatosis and inflammation alone had evidence of cirrhosis during the 10 years of follow-up.

A system for semiquantitative evaluation for the unique lesions recognized for NASH proposed by Brunt et al.<sup>16</sup> in 1999 was developed to parallel the concepts and terminology used in chronic hepatitis for semiquantitative evaluation, commonly referred to as “grading” and “staging.”<sup>17</sup> The proposed system was based on the concept that the histological diagnosis of NASH rests on a constellation of features rather than any individual feature. However, it was developed for NASH and was not developed to encompass the entire spectrum of NAFLD as defined by Matteoni et al.<sup>15</sup> A different semiquantitative feature-based scoring system for NAFLD was used in a recently published treatment trial from Promrat et al.<sup>18</sup> Neither of these systems was designed to evaluate pediatric NAFLD, which may show different histological features than adult NASH.<sup>12</sup>

Beginning in 2002, the National Institute of Diabetes & Digestive & Kidney Diseases (NIDDK) sponsored the development of a multicenter cooperative Clinical Research Network for NASH.<sup>19</sup> Among the goals of the network were: (1) to form a database for long-term natural history observations of patients with NAFLD; (2) to collect clinical samples for metabolic, immunological, molecular, and genetic studies of NAFLD focusing on defining the etiology of this disease; and (3) to evaluate promising therapies for NASH in both adult and pediatric populations. One of the charges to the NASH Clinical Research Network was to develop and validate a system of histological evaluation that would encompass the spectrum of NAFLD, that could be applied to pediatric NAFLD, and that would allow for assessment of changes with therapy. This report describes the multi-center effort to accomplish this task.

## Materials and Methods

The Pathology Subcommittee, composed of pathologists from each Clinical Research Network Clinical Center (total of eight), a pathologist from the National Institutes of Health (NIH), two principal investigators from Clinical Centers, the NIDDK project scientist, and the principal investigator from the Data Coordinating Center, met to devise a scoring system and to discuss plans for case review. A study set was assembled from cases contributed by each center; the cases provided were drawn from locally evaluated patients who underwent biopsies to “rule out” NAFLD or NASH. Cases were specifically included to cover the range of possible diagnoses: (1) diagnostic of steatohepatitis, (2) borderline or possible steatohepatitis, and (3) not diagnostic of steatohepatitis. Each case had hematoxylin and eosin (H&E) and Masson trichrome stains submitted. The study set cases were stripped of all clinical information except for the designation as “adult” or “pediatric,” and the pathologists were blinded to even this piece of information. The Institute Review Boards at each Clinical Center, the Data Coordinating Center, and the Office of Human Subjects Research of the NIH approved the research plan.

**Scoring System Definition.** The Committee reviewed the study set cases as a group and then proposed an evaluation method for each of the recognized features of NAFLD. The Committee agreed that only H&E and Masson’s trichrome stains should be necessary to perform the evaluation. The histological features were grouped into five broad categories: steatosis, inflammation, hepatocellular injury, fibrosis, and miscellaneous features. The system of evaluation is detailed in Table 1, and examples of ballooning, microvesicular steatosis, and megamitochondria are shown in Figs. 1 and 2.

**Table 1. NASH Clinical Research Network Scoring System Definitions and Scores in Study Set**

Item	Definition	Score/Code	% Responses in Category for Study Set Cases	
			Adult (n = 576)	Pediatric (n = 162)
Steatosis				
Grade	Low- to medium-power evaluation of parenchymal involvement by steatosis			
	<5%	0	10%	3%
	5%-33%	1	34%	29%
	>33%-66%	2	31%	31%
	>66%	3	26%	36%
Location	Predominant distribution pattern			
	Zone 3	0	31%	14%
	Zone 1	1	1%	12%
	Azonal	2	37%	22%
	Panacinar	3	31%	52%
Microvesicular steatosis*	Contiguous patches			
	Not present	0	90%	96%
	Present	1	10%	4%
Fibrosis				
Stage	None	0	40%	29%
	Perisinusoidal or periportal	1		
	Mild, zone 3, perisinusoidal	1A	15%	4%
	Moderate, zone 3, perisinusoidal	1B	6%	5%
	Portal/periportal	1C	6%	27%
	Perisinusoidal and portal/periportal	2	12%	10%
	Bridging fibrosis	3	15%	17%
	Cirrhosis	4	6%	9%
Inflammation				
Lobular inflammation	Overall assessment of all inflammatory foci			
	No foci	0	14%	15%
	<2 foci per 200× field	1	53%	60%
	2-4 foci per 200× field	2	27%	22%
	>4 foci per 200× field	3	6%	3%
Microgranulomas	Small aggregates of macrophages			
	Absent	0	57%	53%
	Present	1	43%	47%
Large lipogranulomas	Usually in portal areas or adjacent to central veins			
	Absent	0	86%	99%
	Present	1	14%	1%
Portal inflammation	Assessed from low magnification			
	None to minimal	0	67%	67%
	Greater than minimal	1	33%	33%
Liver cell injury				
Ballooning*	None	0	33%	49%
	Few balloon cells	1	38%	36%
	Many cells/prominent ballooning	2	29%	15%
Acidophil bodies	None to rare†	0	89%	90%
	Many	1	11%	10%
Pigmented macrophages	None to rare†	0	87%	82%
	Many	1	13%	18%
Megamitochondria*	None to rare†	0	86%	95%
	Many	1	14%	5%
Other findings				
Mallory's hyaline	Visible on routine stains			
	None to rare†	0	80%	94%
	Many	1	20%	6%
Glycogenated nuclei	Contiguous patches			
	None to rare†	0	57%	71%
	Many	1	43%	29%
Diagnostic classification‡				
	Not steatohepatitis	0	31%	32%
	Possible/borderline	1	26%	33%
	Definite steatohepatitis	2	43%	35%

\*Ballooning classification: few indicates rare but definite ballooned hepatocytes as well as case that are diagnostically borderline; examples are shown in Fig. 1. Examples of patches of microvesicular steatosis and megamitochondria are shown in Fig. 2.

†The "None to rare" category is meant to alleviate the need for time-consuming searches for rare examples or deliberation over diagnostically borderline changes. If the feature is identified after a reasonable search, it should be coded as "many."

‡Diagnostic classification was not available on 2 sets of adult biopsy observations, reducing the total of such observations to 512.



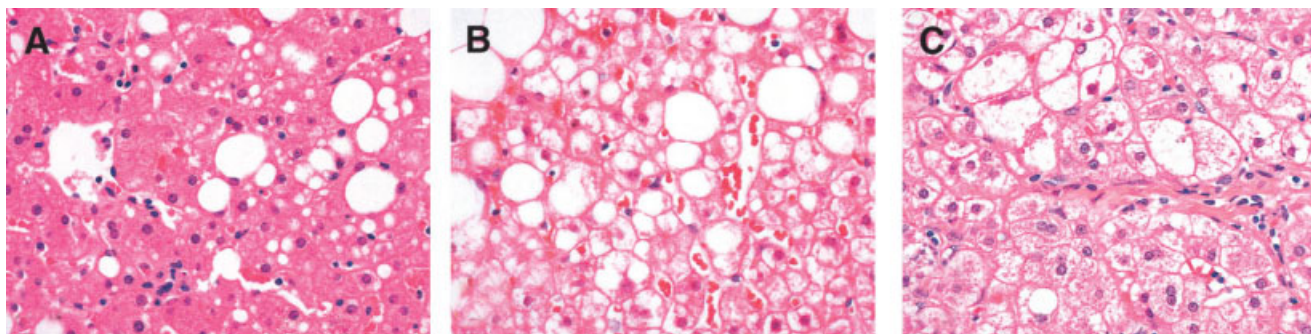


Fig. 1. Scoring of ballooning injury. (A) There is mild steatosis but no ballooning degeneration. All pathologists scored this case as “0” for ballooning injury. (B) In no study set case was there absolute concordance among the nine pathologists for a ballooning score of 1. During the second round of reviews, this case was scored as 1+ ballooning injury by 8 of the 9 pathologists. This field shows several scattered balloon cells that are not much larger than the surrounding steatotic hepatocytes but with the same cytoplasmic characteristics as more obvious balloon cells, such as those seen in panel C. (C) This field is taken from a case scored as 2+ ballooning injury by all pathologists. There is a contiguous patch of hepatocytes showing prominent ballooning injury, sharply contrasted against the more normal hepatocytes in the field. (All photomicrographs: Hematoxylin and eosin; original magnification  $\times 600$ ).

For the validation study, the pathologists were additionally asked to provide an overall diagnosis for each case as “NASH,” “borderline,” or “not NASH.” The biopsy specimens were made anonymous and randomized by an NIH employee not involved in the study. The adult cases were reviewed by each pathologist at his or her home

institutions twice and at separate times; the pediatric cases were circulated once.

**Statistical Analyses.** Pediatric and adult cases were analyzed separately. Weighted kappa scores were used to measure the degree of inter-rater and intra-rater agreement between and within multiple pathologists. Weighted kappa scores were estimated by using intra-class correlation coefficients derived from components of variance models.<sup>20</sup> Chi-square tests were used to compare univariate associations of histological features with diagnosis of steatohepatitis. Multivariate associations with diagnosis of steatohepatitis were assessed by using multiple logistic regression models using generalized estimating equations with robust variance estimation and exchangeable correlation to account for correlations due to multiple readings within and between raters.<sup>21</sup> Adjusted percentage distributions for each histological feature, holding other features constant and retaining the univariate marginal distributions, were calculated from the logistic regression model coefficients for comparison with the unadjusted univariate distributions of features. All statistical analyses were carried out by the Data Coordinating Center (M. Van N.) using SAS 8.0 (SAS Institute Inc., Cary, NC) and Stata 7.0 (Stata Corp., College Station, TX).

## Results

A total of 50 anonymous liver biopsy specimens formed the validation study set. The study set was chosen to sample the range of pathological conditions seen in pediatric and adult NAFLD. The 32 cases from adults were each read twice by the nine pathologists, for a total of 576 sets of observations; the 18 pediatric cases were each read once, for a total of 162 sets of observations. Each

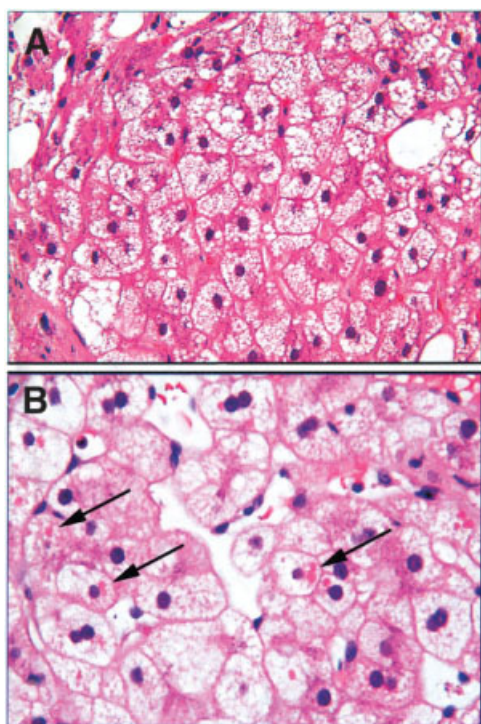


Fig. 2. Microvesicular steatosis and megamitochondria. (A) Contiguous patch of microvesicular steatosis from case of steatohepatitis. This case would be scored positively for this feature. (B) Photomicrograph of cells with easily identified megamitochondria (arrows). This case would be scored positively for this feature, although it is not necessary to see so many positive cells in a single field. (All photomicrographs: Hematoxylin and eosin; original magnification  $\times 400$ ).

**Table 2. Inter- and Intra-rater Variability**

Item	Agreement (Kappa Score)		
	Intra-rater	Interrater	
	Adult Cases (32 cases, 9 raters)	Adult Cases (32 cases, 9 raters)	Pediatric Cases (18 cases, 9 raters)
Steatosis, grade	0.83	0.79	0.64
Steatosis, location	0.39	0.31	0.39
Microvesicular steatosis	0.37	0.34	0.02
Fibrosis	0.85	0.84	0.62
Lobular inflammation	0.60	0.45	0.28
Microgranulomas	0.40	0.18	0.15
Lipogranulomas	0.38	0.26	0.00
Portal inflammation	0.55	0.45	0.42
Ballooning	0.66	0.56	0.22
Acidophil bodies	0.28	0.19	0.27
Pigmented macrophages	0.38	0.15	0.06
Megamitochondria	0.28	0.16	−0.03
Mallory's hyaline	0.64	0.58	0.69
Glycogenated nuclei	0.66	0.58	0.32
Diagnostic classification	0.66	0.61	0.32

NOTE. All values represent weighted kappa statistics where appropriate.

possible score was used at least once during the course of the validation. The distribution of recorded scores is shown in Table 1.

Analysis of agreement on features in the adult cases showed reasonable agreement on the scores for the major scoring categories of steatosis grade, fibrosis, ballooning injury, and Mallory's hyaline, with weighted kappa values for the adult study set of 0.5 and above (Table 2). Agreement was not as strong for the inflammatory changes or for steatosis location; both lobular and portal inflammation had interrater kappa values of 0.45. The interrater agreement on the diagnosis of steatohepatitis was 0.61. There was absolute agreement on the diagnosis between the nine pathologists on six cases and agreement among eight of the nine on a further five cases. The intra-rater agreement was higher in all categories than the interrater agreement.

Feature scoring agreement on the pediatric cases in the study set was not as robust as in the adult study set, with lower weighted kappa scores in all categories except steatosis (0.64). Some of the results were statistically no better than chance (microvesicular steatosis, pigmented macrophages, lipogranulomas and megamitochondria, and glycogenated nuclei); this may be caused in part by the low level of agreement on the presence or absence of the feature as seen in the adult cases and in part by the low frequency of observation of one of the two scores. Many of the pediatric cases appeared to have more zone 1 steatosis, more "periportal only" fibrosis, less ballooning, and rare Mallory's hyaline. In particular, there was disagreement between pathologists on the diagnosis of ste-

atohepatitis when cases had fibrosis and steatosis but little or no ballooning or lobular inflammation.

Although it is accepted that steatohepatitis is a pattern of injury composed of several features, it has been difficult to define exact diagnostic criteria that all pathologists agree on to precisely distinguish NAFLD cases with steatohepatitis from those with only steatosis and inflammation.<sup>14</sup> Data generated by the NASH Network study on how each pathologist categorized each case along with the assigned feature scores allowed statistical examination of which individual features were useful in discriminating definite steatohepatitis from the other two categories. Both crude and adjusted analyses were performed to address these questions; the results are shown in Table 3. Several features showed significant association with the diagnosis of steatohepatitis on both analyses and for both adults and children, including lobular inflammation, ballooning degeneration, and fibrosis. In the adult cases, the degree of steatosis showed a trend toward significance in the crude analysis but was clearly associated with the diagnosis of NASH in the adjusted analysis. In children, the degree of steatosis above the 5% cutoff was not associated with the diagnosis of NASH in either analysis.

Portal inflammation, location of steatosis, Mallory's hyaline, and megamitochondria were associated with steatohepatitis on crude analysis but, when analyzed in the adjusted model, were not significantly associated. In this series, 81% of cases with Mallory's hyaline were also given a ballooning score of 2; this may be why this feature was not independently associated with steatohepatitis.

The proportion of adult cases categorized as steatohepatitis for the most significant features is shown in Fig. 3. Although no single feature or score was absolutely associated with a diagnosis of steatohepatitis, the severity of some individual findings did show association. Pathologists had somewhat varied criteria for steatohepatitis and tended to emphasize or require some features more than others. Even so, the mere presence of a combination of features was not sufficient for a diagnosis of steatohepatitis. For example, only 68% of adult cases with steatosis, ballooning, and lobular inflammation (a common set of minimal criteria<sup>14</sup>) were diagnosed as NASH. Among the individual pathologists, this percentage varied from 56% to 79%. Adding fibrosis increased this fraction to 82%, but at the cost of increasing the fraction of cases that failed to meet these criteria yet still were diagnosed as steatohepatitis. These results emphasize how difficult it is to reduce the histopathological diagnosis of steatohepatitis to the presence or absence of specific features, independent of the variability that is part of any observation.

Based on both the agreement data and the multiple regression analysis, we would propose a NAFLD Activity

**Table 3. Logistic Regression Analysis of Features of NAFLD With Respect to the Diagnosis of NASH**

Histologic Feature of NAFLD	Category	Percent With Diagnosis of NASH			
		Adult (n = 512 Readings From 32 Biopsies)		Pediatric (n = 162 Readings From 18 Biopsies)	
		No Adjustment	Adjusted*	No Adjustment	Adjusted*
Overall diagnosis of NASH (outcome)		43.1%	43.1%	34.6%	34.6%
Steatosis					
Grade	<5%	12.5%	5.8%	†	†
	5%-33%	35.2%	37.3%	21.3%	35.2%
	>33%-66%	52.8%	46.2%	45.1%	32.9%
	>66%	51.8%	59.2%	39.0%	35.5%
	P value‡	.08	.009	.11	.98
Location	Zone 3	45.7%	49.2%	27.3%	16.7%
	Zone 1	25.0%	20.8%	5.0%	18.5%
	Azonal	31.6%	36.6%	38.9%	28.5%
	Panacinar	54.6%	45.6%	41.7%	45.7%
	P value	.01	.61	.17	.46
Microvesicular steatosis	Absent	43.2%	44.8%	35.3%	35.3%
	Present	41.9%	24.6%	16.7%	15.5%
	P value	.60	.28	.44	.39
Fibrosis stage	None	10.1%	16.3%	14.9%	13.2%
	Mild, zone 3	53.3%	61.4%	66.7%	71.6%
	Moderate, zone 3	93.3%	97.7%	87.5%	99.1%
	Portal/periportal	11.1%	9.0%	13.6%	16.7%
	Zone 3 & periportal	76.6%	74.2%	43.8%	48.9%
	Bridging	83.6%	64.4%	70.4%	72.6%
	Cirrhosis	54.6%	42.1%	42.9%	20.1%
	P value	<.0001	<.0001	<.0001	.0001
Inflammation					
Lobular inflammation	No foci	1.2%	1.1%	12.0%	4.7%
	<2 foci	36.0%	38.0%	36.1%	32.5%
	2-4 foci	76.6%	73.3%	42.9%	53.2%
	>4 foci	87.5%	82.1%	60.0%	93.7%
	P value	<.0001	<.0001	.02	.01
Microgranulomas	Absent	42.6%	36.1%	39.5%	38.2%
	Present	43.9%	51.8%	29.0%	30.4%
	P value	.70	.09	.10	.62
Large lipogranulomas	Absent	42.1%	46.5%	33.8%	NC§
	Present	50.8%	18.3%	100.0%	NC§
	P value	.35	.001	NC§	NC§
Portal inflammation	None to minimal	36.5%	43.1%	40.4%	40.6%
	>Minimal	58.0%	43.2%	22.6%	22.2%
	P value	.05	.99	.04	.47
Hepatocellular injury					
Ballooning degeneration	None	7.7%	9.0%	6.2%	3.4%
	Few	44.8%	41.6%	53.4%	54.0%
	Many	81.1%	83.9%	83.3%	91.7%
	P value	<.0001	<.0001	<.0001	<.0001
Acidophil bodies	None to rare	40.3%	40.7%	30.3%	32.1%
	Many	68.6%	65.1%	70.6%	56.1%
	P value	.12	.11	.07	.56
Pigmented macrophages	None to rare	42.7%	45.4%	36.8%	40.1%
	Many	46.4%	24.4%	24.1%	9.0%
	P value	.66	.08	.12	.19
Megamitochondria	None to rare	40.7%	41.2%	31.8%	34.9%
	Many	58.8%	55.5%	87.5%	28.7%
	P value	.02	.11	.01	.87
Other findings					
Mallory bodies	None to rare	34.5%	40.3%	30.9%	31.1%
	Many	81.7%	56.1%	90.0%	87.2%
	P value	.04	.15	.06	.16
Glycogen nuclei	None to rare	41.4%	45.5%	35.6%	25.5%
	Many	45.5%	39.8%	31.9%	56.7%
	P value	.69	.59	.32	.11

\*Adjusted percent with NASH diagnosis calculated from a logistic regression model for correlated data with NASH diagnosis as the outcome and indicator variables for each histological feature. Marginal totals for adjusted percentages were fixed to equal the unadjusted marginal totals, and odds ratios from the adjusted percentages equal those from the logistic regression model.

†Combined 5 observations in the <5% category with 5%-33% category because there were no events in the <5% category.

‡P values were derived from comparisons of percentages with diagnosis of NASH across categories of each histological feature and were calculated from chi-square tests (unadjusted percentages) or from Wald's tests from the logistic regression model (adjusted percentages).

§Not calculable because 2 of 2 observations in large lipogranulomas-present category had the event.



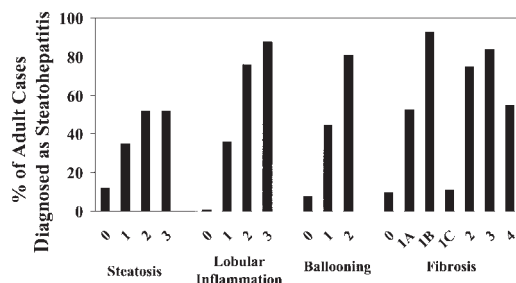


Fig. 3. Fraction of adult cases of NAFLD diagnosed as steatohepatitis. For each subscore of steatosis grade, lobular inflammation, ballooning injury, and fibrosis, the histogram shows the fraction of times the observer made the diagnosis of definite steatohepatitis. For example, of all of the cases in which a 2+ score for ballooning injury was made, 80% were diagnosed as definite steatohepatitis. Similarly, only 11% of the time that a fibrosis score of 0 was made did the observer diagnose the case as definite steatohepatitis.

Score (NAS), which specifically includes only features of active injury that are potentially reversible in the short term. The score is defined as the unweighted sum of the scores for steatosis (0-3), lobular inflammation (0-3), and ballooning (0-2); thus ranging from 0 to 8. Fibrosis, which is both less reversible and generally thought to be a result of disease activity, is not included as a component of the activity score. The separation of fibrosis from other features of activity is an accepted paradigm for staging and grading for both NASH<sup>16</sup> and chronic hepatitis.<sup>17</sup>

The relationship of the NAS to the diagnosis of steatohepatitis is shown in Fig. 4. Cases with NAS of 0 to 2 were largely considered not diagnostic of steatohepatitis; on the other hand, most cases with scores of  $\geq 5$  were diagnosed as steatohepatitis. Cases with activity scores of 3 and 4 were divided almost evenly between the 3 diagnostic categories. A similar analysis of pediatric case scores showed nearly identical results (data not shown).

## Discussion

The aims of this study were to devise and validate a feature-based semiquantitative scoring system to be used in clinical trials and natural history studies of NAFLD. The feature scoring system described used the recognized lesions of NAFLD and NASH<sup>10,14,15</sup> and identified a core group of histological features for evaluation. This system was based on and further refined the grading proposal of Brunt et al.<sup>16</sup> In particular, the fibrosis staging system was subdivided, and the scoring of steatosis was altered. Fibrosis scores for stage 1 were extended to include a distinction between delicate (1A) and dense (1B) perisinusoidal fibrosis, and to detect portal-only fibrosis, without perisinusoidal fibrosis (stage 1C). Minimal steatosis (score 0) (under 5%) was separated from mild (5%-33%) steatosis (score 1) to avoid giving weight to this feature when very

little steatosis is present. A minimum of 5% steatosis was used for the operational minimal definition of histological NAFLD in biopsy specimens from both adults and children. Evaluation of ballooning was limited to three categories (none, few, and many) after discussion among the Committee pathologists as to what might be the most reproducible cutoff points. Lobular inflammation was assessed semiquantitatively on a scale that is the same as the method of Brunt et al.<sup>16</sup> The proposed system differed from the previous methods in that the histologically distinct lesions of NAFLD were assessed and summed to provide an NAFLD Activity Score (NAS), in contrast with methods that depended more on an aggregate assessment to grade severity<sup>16</sup> or to make disease categories.<sup>22</sup> None of the features was weighted in this analysis; future studies may identify elements that indeed deserve greater weight.

With respect to these major disease features, the current system had fair to good interobserver reproducibility in both adult and pediatric cases and was equivalent to previously reported interobserver reproducibility studies in fatty liver disease.<sup>22,23</sup> Each of the major features (steatosis grade, lobular inflammation, ballooning, and fibrosis) showed independent correlation with a diagnosis of steatohepatitis. Based on this observation and the reproducibility studies, we defined a NAS for evaluating histological changes after therapeutic intervention trials. It is important to note that the primary purpose of the NAS is to assess overall histological change; it is not intended that numeric values replace the pathologist's diagnostic determination of steatohepatitis. The NAS has also not been studied as a measure of the rapidity of disease progression, nor should it be taken as an absolute severity scale.

With respect to other features assessed in this scoring system, some were more reproducible than others. The less reproducible features were considered an optional part of the scoring system, and for the purposes of the

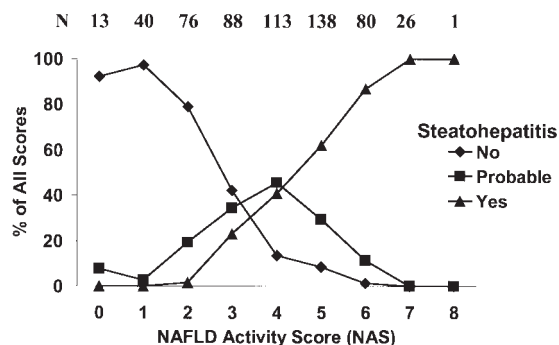


Fig. 4. Fraction of adult cases with activity scores by diagnosis. For each activity score, the fraction of observations with a particular diagnostic categorization is shown. The total number of observations for each activity score is shown across the top of the graph.

NASH Clinical Research Network, will only be evaluated in group review sessions, where differences in observations between pathologists can be immediately addressed.

One of the findings of this study was the lesser degree of interobserver agreement in scoring the features of the pediatric cases of NAFLD submitted for review. As has been previously reported in pediatric NAFLD,<sup>12,13,24,25</sup> pediatric study cases included a significant proportion with only periportal fibrosis. This feature is considered unusual in adults, although it has been reported.<sup>26</sup> Pediatric biopsy specimens also showed less lobular inflammation, less ballooning, and only rare Mallory's hyaline when compared with adult cases. The only pediatric case that all 9 pathologists agreed was diagnostic of steatohepatitis resembled "typical" adult steatohepatitis, with prominent ballooning, moderate lobular inflammation, and Mallory's hyaline. We postulate that the increased variability of scores was attributable to different patterns of disease in pediatric NAFLD, and we will continue to study this in the cases accrued by the NASH Clinical Research Network. Nevertheless, the scoring system that was developed covers the range of features in pediatric NAFLD.

One concern for any new scoring system is how it applies in actual clinical trials. Although the data are not presented here, this method for scoring and the activity score (NAS) were used in a blinded re-review of biopsy specimens from two recently published therapeutic trials.<sup>18,27</sup> The conclusions that could be drawn regarding histological changes were essentially the same as those in the original publications. These results were useful in protocol design for the clinical trials that have been developed by the NASH Network.

In summary, we have designed and validated a semi-quantitative scoring system that is useful for assessing the range of histological features of NAFLD. The system is simple and requires only routine histochemical stains (H&E and Masson trichrome stains), so that the system can be used by practicing pathologists. Of course, other special stains for additional analyses may be used at the discretion of each pathologist. The method proposed showed reasonable interrater agreement among experienced hepatopathologists similar to other studies of variability in fatty liver disease.<sup>22,23</sup> Multiple regression analysis of the scores with respect to the diagnosis of NASH confirmed previous observations that the diagnosis of steatohepatitis is not dependent on a single histological feature, but rather involves assessment of multiple independent features. As a reflection of this fact, the NAS was able to discriminate between NASH and non-NASH fatty liver disease in this patient population. To make it easy for other pathologists to make use of this system, it is

our plan to show examples of the lesions on the NASH Clinical Research Network web site.

### Appendix: Members of the Nonalcoholic Steatohepatitis Clinical Research Network

*Case Western Reserve University, Cleveland, OH:* Yao-Chang Liu, M.D.; Arthur J. McCullough, M.D. (Principal Investigator); *Duke University Medical Center, Durham, NC:* Anna Mae Diehl, M.D. (Principal Investigator); Marcia Gottfried, M.D.; Michael S. Torbenson, M.D.; *Indiana University School of Medicine, Indianapolis, IN:* Naga Chalasani, M.D. (Principal Investigator); Oscar W. Cummings, M.D.; *Johns Hopkins University Center for Clinical Trials (Data Coordinating Center), Baltimore, MD:* Pat Belt, B.S.; Aynur Ünalp-Arida, M.D., Ph.D.; Mark Van Natta, M.H.S.; James Tonascia, PhD (Principal Investigator) *National Cancer Institute (NCI), Bethesda, MD:* David E. Kleiner, M.D., Ph.D. (Co-Lead Pathologist); *National Institute of Child Health and Human Development (NICHD), Bethesda, MD:* Gilman D. Grave, MD (Project Scientist) *National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), Bethesda, MD:* Jay Hoofnagle, M.D. (Project Scientist); Patricia R. Robuck, Ph.D. (Project Scientist); *St Louis University Hospital, St. Louis, MO:* Elizabeth M. Brunt, M.D. (Co-Lead Pathologist); Brent A. Tetri, M.D. (Principal Investigator); *University of California San Diego, San Diego, CA:* Cynthia Behling, M.D.; Joel E. Lavine, M.D., Ph.D. (Principal Investigator); *University of California San Francisco, San Francisco, CA:* Nathan M. Bass, M.D., Ph.D. (Principal Investigator); Linda D. Ferrell, M.D.; *University of Washington, Seattle, WA:* Kris V. Kowdley, M.D. (Principal Investigator); Matthew Yeh, M.D., Ph.D. *Virginia Commonwealth University, Richmond, VA:* Melissa J. Contos, M.D.; Arun J. Sanyal, M.D. (Principal Investigator)

### References

- Marchesini G, Brizi M, Bianchi G, Tomassetti S, Bugianesi E, Lenzi M, et al. Nonalcoholic fatty liver disease: a feature of the metabolic syndrome. *Diabetes* 2001;50:1844-1850.
- Day C, Saksena S. Non-alcoholic steatohepatitis: Definitions and pathogenesis. *J Gastroenterol Hepatol* 2002;17(Suppl 3):S377-S384.
- Sanyal AJ. AGA technical review on nonalcoholic fatty liver disease. *Gastroenterology* 2002;123:1705-1725.
- Marchesini G, Bugianesi E, Forlani G, Cerrelli F, Lenzi M, Manini R, et al. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. *HEPATOLOGY* 2003;37:917-923.
- McCullough AJ. Update on nonalcoholic fatty liver disease. *J Clin Gastroenterol* 2002;34:255-262.
- Mulhall BP, Ong JP, Younossi ZM. Non-alcoholic fatty liver disease: an overview. *J Gastroenterol Hepatol* 2002;17:1136-1143.
- Clark JM, Brancati FL, Diehl AM. The prevalence and etiology of elevated aminotransferase levels in the United States. *Am J Gastroenterol* 2003;98:960-967.
- Farrell GC. Non-alcoholic steatohepatitis: what is it, and why is it important in the Asia-Pacific region? *J Gastroenterol Hepatol* 2003;18:124-138.
- Ruhl CE, Everhart JE. Relation of elevated serum alanine aminotransferase activity with iron and antioxidant levels in the United States. *Gastroenterology* 2003;124:1821-1829.
- Ludwig J, Viggiano TR, McGill DB, Oh BJ. Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clin Proc* 1980;55:434-438.
- Caldwell SH, Oelsner DH, Iezzoni JC, Hespenheide EE, Battle EH, Driscoll CJ. Cryptogenic cirrhosis: clinical characterization and risk factors for underlying disease. *HEPATOLOGY* 1999;29:664-669.
- Schwimmer JB, Behling C, Newberry R, McGreal N, Rauch J, Deutsch R, et al. The histological features of pediatric nonalcoholic fatty liver disease (NAFLD) [Abstract]. *HEPATOLOGY* 2002;36(Suppl):412A.



13. Roberts EA. Steatohepatitis in children. *Best Pract Res Clin Gastroenterol* 2002;16:749-765.
14. Brunt EM. Nonalcoholic steatohepatitis. *Semin Liver Dis* 2004;24:3-20.
15. Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, McCullough AJ. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. *Gastroenterology* 1999;116:1413-1419.
16. 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;94:2467-2474.
17. Brunt EM. Grading and staging the histopathological lesions of chronic hepatitis: the Knodell histology activity index and beyond. *HEPATOLOGY* 2000;31:241-246.
18. Promrat K, Lutchman G, Uwaifo GI, Freedman RJ, Soza A, Heller T, et al. A pilot study of pioglitazone treatment for nonalcoholic steatohepatitis. *HEPATOLOGY* 2004;39:188-196.
19. Nonalcoholic steatohepatitis clinical research network. *HEPATOLOGY* 2003;35:244.
20. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979;86:420-428.
21. Diggle PJ, Liang KY, Zeger SL. *Analysis of Longitudinal Data*. Oxford, England: Oxford University Press, 1995.
22. Younossi ZM, Gramlich T, Liu YC, Matteoni C, Petrelli M, Goldblum J, et al. Nonalcoholic fatty liver disease: assessment of variability in pathologic interpretations. *Mod Pathol* 1998;11:560-565.
23. Bedossa P, Poynard T, Naveau S, Martin ED, Agostini H, Chaput JC. Observer variation in assessment of liver biopsies of alcoholic patients. *Alcohol Clin Exp Res* 1988;12:173-178.
24. Baldrige AD, Perezatayde AR, Graemecook F, Higgins L, Lavine JE. Idiopathic steatohepatitis in childhood: a multicenter retrospective study. *J Pediatr* 1995;127:700-704.
25. Rashid M, Roberts EA. Nonalcoholic steatohepatitis in children. *J Pediatric Gastroenterol Nutr* 2000;30:48-53.
26. Ratziu V, Giral P, Charlotte F, Bruckert E, Thibault V, Theodorou I, et al. Liver fibrosis in overweight patients. *Gastroenterology* 2000;118:1117-1123.
27. Neuschwander-Tetri BA, Brunt EM, Wehmeier KR, Oliver D, Bacon BR. Improved nonalcoholic steatohepatitis after 48 weeks of treatment with the PPAR-gamma ligand rosiglitazone. *HEPATOLOGY* 2003;38:1008-1017.