

Original Article

A Modification of the Brunt System for Scoring Liver Histology of Patients with Non-Alcoholic Fatty Liver Disease

Shahin Merat MD*, Farzaneh Khadem-Sameni MD**, Mehdi Nouraei MD****, Mohammad H Derakhshan MD*†, Seyed Mohammad Tavangar MD**, Sara Mossaffa MD*, Reza Malekzadeh MD*, Masoud Sotoudeh MD****

Background: Nonalcoholic steatohepatitis is a common cause of chronic liver disease. It is important to have a uniform and validated method for scoring liver histology in these patients. Therefore, we propose a modification of the Brunt system by scoring the four histologic features separately and reporting their sums as the nonalcoholic steatohepatitis activity index.

Methods: A nonalcoholic steatohepatitis activity index was defined which scored the grade of disease activity between 0 and 12 according to four histologic features: steatosis, hepatocyte ballooning, portal inflammation, and lobular inflammation. Fibrosis was separately scored between 0 and 4. A total of 60 liver specimens collected from nonalcoholic steatohepatitis patients were scored by two pathologists at the time of biopsy and after three months. Liver enzyme levels were also correlated with the histologic score.

Results: The intra-rater agreement (weighted kappa) for various variables of the scoring system was between 0.59 and 0.80 for one pathologist and 0.78 to 0.95 for the other. The repeatability of the nonalcoholic steatohepatitis activity index was excellent with only 4% of the repeated measures out of the acceptable range of agreement. In addition, liver enzyme levels strongly correlated with the nonalcoholic steatohepatitis activity index.

Conclusion: This system provides a fine graduation of liver histologic damage in nonalcoholic steatohepatitis and is well suited for use in clinical trials or natural history studies.

Archives of Iranian Medicine, Volume 13, Number 1, 2010: 38 – 44.

Keywords: Fatty liver • observer variation • pathology

Introduction

Non-alcoholic fatty liver disease (NAFLD) is an entity histologically similar to alcoholic liver disease but observed in patients without excessive alcohol intake.¹ NAFLD includes a wide spectrum of liver

diseases ranging from simple steatosis to nonalcoholic steatohepatitis (NASH), and to end-stage cirrhosis.² After widespread vaccination against hepatitis B, acute liver diseases resulting from hepatitis C, alcohol abuse and NAFLD, itself, have emerged as major causes of chronic liver disease. Unfortunately these three conditions appear to be increasing in frequency. This is especially true for NAFLD because of its relation to obesity, which has reached epidemic proportions.^{3,4}

Clearly, treatment is required for NAFLD patients who have advanced liver pathology. Many studies have used serum levels of liver enzymes to document the effectiveness of treatment.^{5,6} Others have considered the histopathological changes as a clue for the effectiveness of treatment.⁷⁻⁹ However, the histological criteria are subjective if not interpreted in the context of a valid and reliable

Authors' affiliation: *Digestive Disease Research Center, Shariati Hospital, Tehran University of Medical Sciences, **Department of Pathology, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran.

***Department of Internal Medicine and Center for Sickle Cell Disease, Howard University, Washington, DC

†Section of Gastroenterology, Medical Faculty, University of Glasgow, Glasgow, UK

•Corresponding author and reprints: Masoud Sotoudeh MD, Department of Pathology, Shariati Hospital, N. Kargar St. Tehran, 14117, Iran

Tel: +98-218-801-2992, E-mail: sotoudeh@ams.ac.ir

Accepted for publication: 16 December 2009

scoring system.

It is well known that the majority of cases of NAFLD never lead to significant chronic liver damage and one challenge is to identify those who are at risk for progression by non-invasive methods.¹⁰ Serial liver biopsies are used to study the natural course of NAFLD,^{11,12} but again, a validated scoring system is required.

The Hepatitis Activity Index (HAI) and METAVIR systems are widely used to evaluate the effectiveness of various treatments and to guide therapy, especially in patients with viral hepatitis.^{13,14} There has been no general agreement on a similar system for NASH. The key diagnostic histological features are quite different between viral hepatitis and NAFLD.^{15,16} For example, inflammation from viral hepatitis is mostly portal and periportal in contrast to the predominantly parenchymal and pericentral inflammation that is seen in NAFLD. Additionally, there is a peculiar pericentral and perisinusoidal fibrosis in NAFLD which is not usually present in chronic viral liver diseases.¹⁶ Furthermore, the degree of steatosis should certainly be included in any scoring system for NAFLD. Given these differences, it is unlikely that a scoring system developed for any other liver disease would be suitable for NAFLD.

In response to this need, Brunt et al. have studied several histologic features before finally proposing and validating a system to scoring the degree of disease activity in NAFLD (grade) by using the four major and consistent criteria of steatosis, hepatocyte ballooning, lobular, and portal inflammation.¹⁷ They classify NAFLD as mild, moderate, or severe according to the degree of these four variables. Grading lesions can occasionally be confusing since the severities of the four variables used are not always in agreement.

In 2005, Kleiner et al. proposed a system for scoring histologic findings in patients with NAFLD. They studied 14 histologic features and finally proposed a NAFLD activity score (NAS) which included three of these features: steatosis, lobular inflammation, and ballooning.¹⁸ Although NAS was originally validated to differentiate "NASH" from "not NASH", it has been frequently used as a scoring system for NAFLD. Although easy to use, NAS does not include portal inflammation which has been recently shown to be important in determining the activity of NASH.¹⁹

Therefore, we propose a modification of the Brunt system by scoring the four histologic

features separately and reporting their sums as the NASH activity index (NAI).

Materials and Methods

Liver specimens were collected from patients diagnosed as NASH from three referral centers in Tehran: Shariati Hospital, Emam Hospital, and Mehr Genral Hospital. The diagnosis of NASH was made by clinical, laboratory, and pathologic features as confirmed by two gastroenterologists. Patients were labeled as NASH if they had any degree of steatosis in association with inflammation and hepatocellular injury. Patients taking any medications during the three months prior to biopsy and those with an alcohol intake of over 40 g per week were excluded. Serum AST and ALT levels obtained within two weeks prior to biopsy were also recorded for each patient.

Several step sections of the biopsy material as stained by H&E, Masson's trichrome and reticulin methods were reviewed by two pathologists who agreed on terms and definitions. The same pathologists reviewed and scored the slides again after three months. Both pathologists were blinded to the clinical data and the previous baseline readings. The original variables defined by Brunt et al. were each scored between 0 and 3 according to Table 1. The sums of these scores were reported as NAI. The stage (fibrosis) was scored exactly as defined by Brunt et al. (Table 1, Figure 1).¹⁷

Statistical analysis

The agreement within each rater (intra-rater agreement) for the five individual histologic variables was calculated using weighted kappa statistics (weights= 1, 0.9, 0.8, 0.7). Subsequently, a chi-square test was applied to test the five kappa values between the two raters in order to determine the overall reliability of the raters. A summary score for each variable was calculated for each rater. Agreement between the two raters (inter-rater agreement) for each of the five variables as well as NAI was calculated by the Bland-Altman method. A graph of the difference between scores versus the mean scores for both raters was developed for each variable and the number of observations beyond the mean \pm 1.96SD was calculated.²⁰ Intra-cluster correlation (ICC) for each item was calculated to show repeatability of the measurements across different ratings. ICC was calculated from an ANOVA table with four fixed groups (two ratings for each of the two raters). ICC

Table 1. Proposed scoring system for steatohepatitis

Variable	Score	Description
Steatosis	0	None
	1	Up to 33% of acini, mainly macrovesicular
	2	34–66% of acini, commonly mixed steatosis
	3	Over 66% of acini (panacinar), commonly mixed steatosis
Hepatocyte ballooning	0	None
	1	Occasional in zone III
	2	Obvious in zone III
	3	Marked, predominantly in zone III
Lobular inflammation	0	None
	1	Scattered neutrophils, occasional mononuclear cells, 1 or 2 foci per 20x objective.
	2	Neutrophils associated with ballooned hepatocytes, mild chronic inflammation, 3 or 4 foci per 20x objective
	3	Acute and chronic inflammation, neutrophils may concentrate in zone III, over 4 foci per 20x objective
Portal inflammation	0	None
	1	Mild, some portal areas
	2	Mild to moderate, most portal areas
	3	Moderate to severe, most portal areas
Stage	0	No fibrosis
	1	Zone III perivenular, perisinusoidal (pericellular) fibrosis
	2	Stage 1 changes + periportal fibrosis
	3	Bridging fibrosis
	4	Cirrhosis

showed the degree of total variance attributable to between-subjects variation. Spearman's rho was used to evaluate the correlation between liver enzymes and histologic variables.

Ethics

The study was approved by the Institutional Review Board and the Ethics Committee of the Digestive Disease Research Center at Tehran University of Medical Sciences.

Results

A total of 60 liver biopsy samples were collected. Patients' characteristics are listed in

Table 2. All samples contained seven or greater portal spaces and were considered adequate for

Table 2. Characteristics of patients

Total number	60
Male/Female	35/25
Age (yrs, mean±SD, range)	38±12, 19–71
AST (IU/L, mean±SD, range)	70±41, 17–200
ALT (IU/L, mean±SD, range)	97±61, 22–289
Weight (kg, mean±SD, range)	82±13, 60–115
Height (cm, mean±SD, range)	168±9, 150–184
BMI (kg/m ² , mean±SD, range)	29±5, 21–48

histological evaluation. The lengths of the biopsies ranged from 9 to 26 mm with a mean of 21 mm. The degree of intra-rater agreement, (an indicator of repeatability) is given in Table 3. The most consistent repeatability was seen for portal

Table 3. Intra-rater agreements for raters scoring 60 liver biopsy samples

	Weighted Kappa (SE)	
	Rater 1	Rater 2
Steatosis	0.80 (0.09)	0.86 (0.09)
Hepatocyte ballooning	0.59 (0.09)	0.78 (0.10)
Portal inflammation	0.84 (0.09)	0.95 (0.09)
Lobular inflammation	0.60 (0.09)	0.89 (0.10)
Stage	0.80 (0.09)	0.86 (0.09)

inflammation followed by stage, steatosis and the least for hepatocyte ballooning. There was no statistically significant difference between the two raters for the degree of intra-rater agreement.

For each variable, the difference between the two raters was charted against their mean score. Results beyond the mean±1.96 SD were considered unacceptable. The most consistent interpretation was observed for NAI with only 4% of the readings that were out of the acceptable range of agreement (Figure 2). The greatest discrepancy was observed in portal inflammation (14%; Table 4).

The ICC for each variable is given in Table 5. ICC indicates the general correlation between all four ratings (two ratings by each rater). Excellent correlation is observed for steatosis, stage, portal inflammation, and NAI, whereas the correlations for lobular inflammation and hepatocyte ballooning are less reliable. As in the intra-rater agreement, the poorest correlation is seen for hepatocyte ballooning.

Lobular and portal inflammation were moderately but significantly correlated ($r = 0.5$,

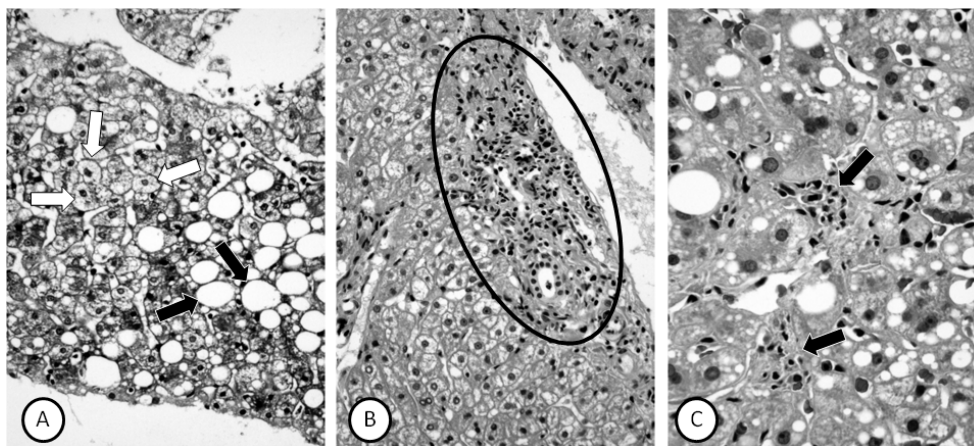


Figure 1. Example of liver histology scoring. **A)** Steatosis (black arrows), almost 50% of acinus display steatosis, score=2. Hepatocyte ballooning (white arrows), obvious in zone III, score=2. **B)** Portal inflammation, some portal areas, score=1. **C)** Lobular inflammation (arrows), neutrophils associated with ballooned hepatocytes, mild chronic inflammation, 3 or 4 foci per 20x objective, score=2. NASH Activity Index=7(2+2+1+2)

$P<0.001$). Steatosis slightly correlated with lobular inflammation ($r=0.34$, $P=0.002$) and the stage was fairly correlated with ballooning and portal inflammation ($r=0.3$, $P=0.01$ for both).

Table 4. Discrepancy between the raters scoring 60 liver biopsy samples and the percent of unacceptable agreements

	Range of difference	Mean difference (SD)	Unacceptable observation
Steatosis	-1.0 to 0.5	-0.05 (0.24)	7%
Ballooning	-2.0 to 2.0	0.27 (0.87)	11%
Portal inflammation	-1.0 to 1.0	0.04 (0.42)	14%
Lobular inflammation	-1.0 to 1.0	0.05 (0.44)	11%
NAI*	-3.0 to 2.0	-0.2 (0.93)	4%
Stage	-2.0 to 1	-0.07 (0.51)	5%

*NAI=NASH activity index

The levels of AST and ALT significantly correlated with steatosis, ballooning, lobular inflammation, stage, and NAI; but not with portal inflammation (Table 6). The ratio of AST to ALT did not correlate with any of the histologic variables.

Table 5. Intra-cluster correlation for the raters for twice scoring 60 liver biopsy samples

	ICC (95%CI)
Steatosis	0.93 (0.90 – 0.96)
Hepatocyte ballooning	0.68 (0.57 – 0.78)
Portal inflammation	0.85 (0.78 – 0.90)
Lobular inflammation	0.75 (0.86 – 0.83)
NAI*	0.86 (0.81 – 0.91)
Stage	0.82 (0.74 – 0.88)

*NAI=NASH activity index

Discussion

The prevalence of NAFLD is increasing, partly as a result of its relation to obesity which has reached epidemic proportions in many countries.^{3,4} NASH is rapidly evolving as the major cause of chronic liver disease and recent evidence shows that it can be potentially as dangerous as hepatitis C.^{21,22} Some authors report up to 20% progression to cirrhosis,²¹ much higher than what is reported for hepatitis C.²³ Unfortunately, effective medical treatment against this disease is lacking and a large number of clinical trials should be performed before an effective treatment can be established.²⁴ In addition, although there are many studies on the natural course of NAFLD,²¹ still more are required. Such studies will need an endpoint in order to evaluate the effectiveness of treatment or disease progression. Liver enzyme levels are an obvious and easily available endpoint used in many previous studies.⁶ Unfortunately, AST or ALT levels cannot be solely relied upon to indicate histologic changes and thus their use as an endpoint is limited. Analysis of liver histology is probably the best endpoint and is used in many newer studies.^{7,8,25} Unfortunately such studies lack uniformity in reporting histologic changes, occasionally using non-validated arbitrary methods.

A good scoring system should be simple, repeatable, and include the key pathologic features of NASH. Most importantly, the score should be predictive of the clinical severity of the disease and its prognosis. Unfortunately, due to the slow progression of NASH, it is very difficult to find the

Table 6. Correlation between liver enzyme levels and histologic variables used in scoring NASH, Spearman's rho (*P* value)

	Steatosis	Ballooning	Lobular inflammation	Portal inflammation	NAI*	Stage
AST	0.37 (0.004)	0.37 (0.004)	0.33 (0.01)	0.25 (0.055)	0.54 (<0.001)	0.37 (0.003)
ALT	0.32 (0.012)	0.35 (0.007)	0.34 (0.008)	0.18 (0.17)	0.48 (<0.001)	0.26 (0.049)

*NAI: NASH activity index

predictive value of any scoring system without studying large numbers of patients for extended periods of time.

A landmark system has been proposed by Brunt et al. in 1999.¹⁷ In this system, the severity of NASH is classified as mild, moderate, or severe. Although widely used, the small number of levels in this system and the slowness of changes in liver histology make this system less sensitive in detecting potential histologic changes when used as an endpoint in clinical trials or cohort studies. In our proposed system, each case can be assigned a number ranging from 0 to 12 which provides a finer gradation of the spectrum of histologic damage; thus it becomes more probable to show changes during the relatively short period of a clinical study.

Other than Brunt's system, the NAS, as proposed by Kleiner et al. is also frequently used.¹⁸ NAS incorporates the severity of steatosis, hepatocyte ballooning, and lobular inflammation

but not portal inflammation. It has been originally designed to confirm the diagnosis of "NASH" versus "not NASH" and not meant to be an indicator of the severity of histologic damage, although frequently used as such. NAS is validated against the over-all diagnosis which is made by a pathologist reviewing the histology in the absence of clinical data. NAS has not been designed nor validated for indicating the severity of disease. Unlike the original Brunt system, Kleiner et al. have not included portal inflammation in their system. This variable might not be important in differentiating "NASH" from "not NASH", but it has recently been shown to be an important indicator of severity and thus should be included in the scoring of NASH.¹⁹ We believe the inclusion of portal inflammation in our proposed system is an important advantage.

We have validated our system by showing its inter- and intra-rater agreement and its correlation with liver enzyme levels. Among individual histologic features, we observed the best intra- and inter-rater agreement in steatosis and stage (Figure 2, Tables 3 and 4). This acceptable agreement has also been reported by other researchers studying intra- and inter-rater variability.^{18,26} Interestingly, our proposed NAI had the best inter-rater agreement, which was better than any of the individual histologic features (Table 4). This indicates that NAI is well suited for use in research and even clinical practice.

At the other end of the spectrum, hepatocyte ballooning had the worst repeatability figures in our study. This is in contrast with reports by Kleiner et al. and Younossi et al. where lobular and portal inflammation were less repeatable than hepatocyte ballooning.^{18,26} Our raters are highly skilled and experienced in scoring lobular and portal inflammation, as it is routinely performed for viral hepatitis. In contrast, scoring hepatocyte ballooning is a relatively new concept. Less experience on scoring this variable might be responsible for the poor repeatability we observed.

Serum AST and ALT levels significantly correlated with NAI but not with portal inflammation (Table 6). This correlation indicates

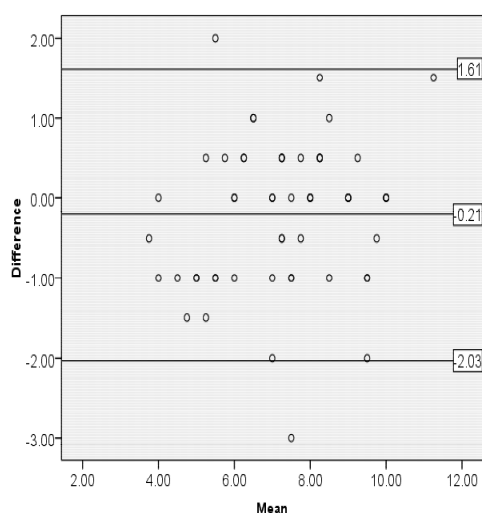


Figure 2. Inter-rater agreement in scoring 60 liver samples with nonalcoholic steatohepatitis using NASH Activity Index. The middle line indicates the mean difference between raters. The other two lines indicate $\text{mean} \pm 1.96 \text{ SD}$. Differences between the raters lying outside these lines are considered unacceptable

the validity of our system and its relation to clinical variables. The lack of significant correlation between portal inflammation and liver enzymes is consistent with previous reports.^{19,27}

The value of a scoring system is best demonstrated during clinical studies. We have used our proposed system in a small clinical trial evaluating the effectiveness of probucol as an antioxidant in NASH. We compared the liver histology of eight patients before and after one year of treatment with probucol and observed that NAI decreased from 7.4 to 5.6 ($P=0.03$). This improvement in histology was paralleled by a very significant improvement in AST and ALT levels.⁷

We believe that NAI is a reasonable option for scoring liver histologies of adult patients with NAFLD and NASH. It is both helpful and easy to use in the clinical as well as research settings, however it should be noted that NAI has only been validated for adult patients. In order to apply this system to pediatric patients, a similar pediatric study would be required.

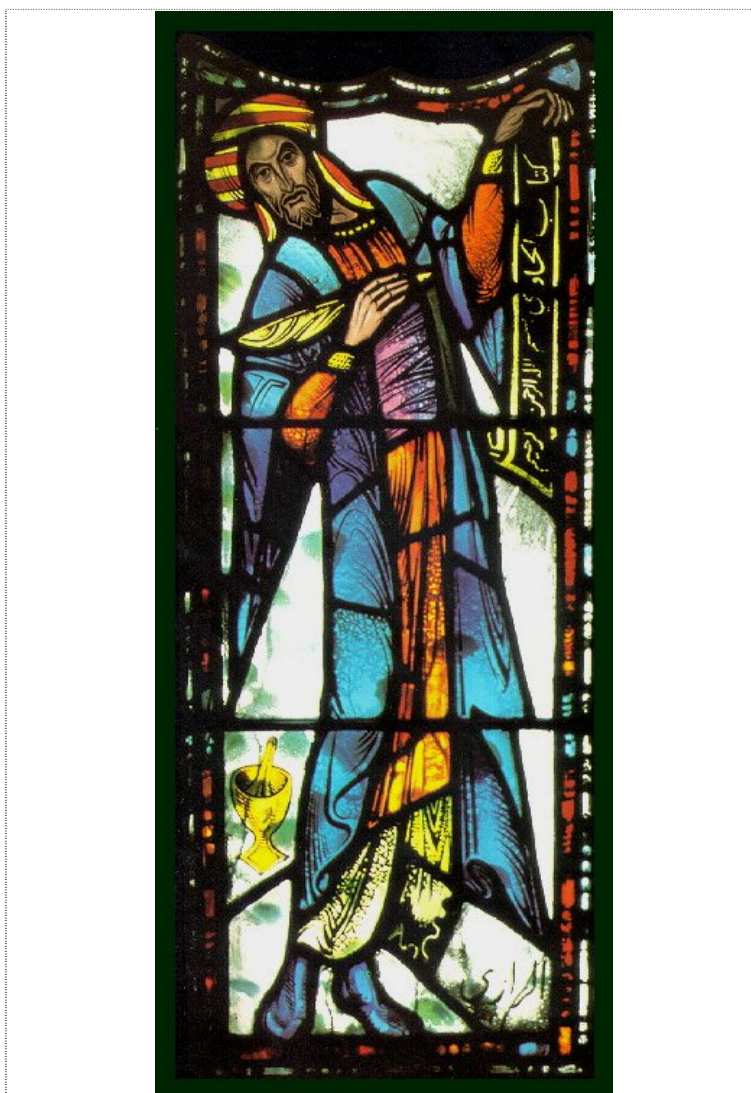
Acknowledgment

This study was supported by a grant from Tehran University of Medical Sciences.

References

- 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.
- 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.
- Bedogni G, Miglioli L, Masutti F, Tiribelli C, Marchesini G, Bellentani S. Prevalence of and risk factors for nonalcoholic fatty liver disease: the Dionysos nutrition and liver study. *Hepatology.* 2005; **42**: 44 - 52.
- Malekzadeh R, Mohamadnejad M, Merat S, Pourshams A, Etemadi A. Obesity pandemic: an Iranian perspective. *Arch Iranian Med.* 2005; **8**: 1 – 7.
- Marchesini G, Brizi M, Bianchi G, Tomassetti S, Zoli M, Melchionda N. Metformin in non-alcoholic steatohepatitis. *Lancet.* 2001; **358**: 893 – 894.
- Merat S, Malekzadeh R, Sohrabi MR, Hormazdi M, Naserimoghadam S, Mikaeli J, et al. Probuco in the treatment of nonalcoholic steatohepatitis: an open-labeled study. *J Clin Gastroenterol.* 2003; **36**: 266 – 268.
- Merat S, Aduli M, Kazemi R, Sotoudeh M, Sedighi N, Sohrabi M, et al. Liver histology changes in nonalcoholic steatohepatitis after one year of treatment with probucol. *Dig Dis Sci.* 2008; **53**: 2246 – 2250.
- Belfort R, Harrison SA, Brown K, Darland C, Finch J, Hardies J, et al. A placebo-controlled trial of pioglitazone in subjects with nonalcoholic steatohepatitis. *N Engl J Med.* 2006; **355**: 2297 – 2307.
- Bugianesi E, Gentilcore E, Manini R, Natale S, Vanni E, Villanova N, et al. A randomized controlled trial of metformin versus vitamin E or prescriptive diet in nonalcoholic fatty liver disease. *Am J Gastroenterol.* 2005; **100**: 1082 – 1090.
- Harrison SA, Oliver D, Arnold HL, Gogia S, Neuschwander-Tetri BA. Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. *Gut.* 2008; **57**: 1441–1447.
- Adams LA, Sanderson S, Lindor KD, Angulo P. The histological course of nonalcoholic fatty liver disease: a longitudinal study of 103 patients with sequential liver biopsies. *J Hepatol.* 2005; **42**: 132 - 138.
- Harrison SA, Torgerson S, Hayashi PH. The natural history of nonalcoholic fatty liver disease: a clinical histopathological study. *Am J Gastroenterol.* 2003; **98**: 2042 – 2047.
- Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR cooperative study group. *Hepatology.* 1996; **24**: 289 – 293.
- Ishak K, Baptista A, Bianchi L, Callea F, De Groote J, Gudat F, et al. Histological grading and staging of chronic hepatitis. *J Hepatol.* 1995; **22**: 696 - 699.
- Bondini S, Kleiner DE, Goodman ZD, Gramlich T, Younossi ZM. Pathologic assessment of non-alcoholic fatty liver disease. *Clin Liver Dis.* 2007; **11**: 17 – 23, vii.
- Hubscher SG. Histological assessment of non-alcoholic fatty liver disease. *Histopathology.* 2006; **49**: 450 – 465.
- 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.
- Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology.* 2005; **41**: 1313 – 1321.
- Brunt EM, Kleiner DE, Wilson LA, Unalp A, Behling CE, Lavine JE, 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**: 809 – 820.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; **1**: 307 – 310.
- Ong JP, Younossi ZM. Epidemiology and natural history of NAFLD and NASH. *Clin Liver Dis.* 2007; **11**: 1 – 16, vii.
- Yatsuji S, Hashimoto E, Tobarai M, Taniai M, Tokushige K, Shiratori K. Clinical features and outcomes of cirrhosis due to non-alcoholic steatohepatitis compared with cirrhosis caused by chronic hepatitis C. *J Gastroenterol Hepatol.* 2009; **24**: 248 – 254.
- Ascione A, Tartaglione T, Di Costanzo GG. Natural history of chronic hepatitis C virus infection. *Dig Liver Dis.* 2007; **39** (suppl 1): S4 – S7.
- Torres DM, Harrison SA. Diagnosis and therapy of nonalcoholic steatohepatitis. *Gastroenterology.* 2008; **134**: 1682 –1698.
- Harrison SA, Fecht W, Brunt EM, Neuschwander-Tetri

- BA. Orlistat for overweight subjects with nonalcoholic steatohepatitis: a randomized, prospective trial. *Hepatology*. 2009; **49**: 80 – 86.
- 26 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.
- 27 Loria P, Lonardo A, Leonardi F, Fontana C, Carulli L, Verrone AM, et al. Non-organ specific autoantibodies in nonalcoholic fatty liver disease: prevalence and correlates. *Dig Dis Sci*. 2003; **48**: 2173 – 2181.



This stained glass window of Razi (Rhazes), the Great Iranian Physician is found in Cambridge University's Medical School in London, England.