



# Interobserver Agreement on Pathologic Features of Liver Biopsy Tissue in Patients with Nonalcoholic Fatty Liver Disease

Eun Sun Jung<sup>1,2</sup> · Kyoungbun Lee<sup>1,3</sup>  
Eunsil Yu<sup>1,4</sup> · Yun Kyung Kang<sup>1,5</sup>  
Mee-Yon Cho<sup>1,6</sup> · Joon Mee Kim<sup>1,7</sup>  
Woo Sung Moon<sup>1,8</sup> · Jin Sook Jeong<sup>1,9</sup>  
Cheol Keun Park<sup>1,10</sup> · Jae-Bok Park<sup>1,11</sup>  
Dae Young Kang<sup>1,12</sup> · Jin Hee Sohn<sup>1,13</sup>  
So-Young Jin<sup>1,14</sup>

<sup>1</sup>Gastrointestinal Pathology Study Group of Korean Society of Pathologist; <sup>2</sup>Department of Pathology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul; <sup>3</sup>Department of Pathology, Seoul National University College of Medicine, Seoul; <sup>4</sup>Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul; <sup>5</sup>Department of Pathology, Inje University Seoul Paik Hospital, Seoul; <sup>6</sup>Department of Pathology, Yonsei University Wonju College of Medicine, Wonju; <sup>7</sup>Department of Pathology, Inha University Hospital, Incheon; <sup>8</sup>Department of Pathology, Chonbuk National University Medical School, Jeonju; <sup>9</sup>Department of Pathology, Dong-A University College of Medicine, Busan; <sup>10</sup>Department of Pathology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul; <sup>11</sup>Department of Pathology, Daegu Catholic University College of Medicine, Daegu; <sup>12</sup>Department of Pathology, Chungnam National University Hospital, Chungnam National University School of Medicine, Daejeon; <sup>13</sup>Department of Pathology, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul; <sup>14</sup>Department of Pathology, Soon Chun Hyang University Seoul Hospital, Seoul, Korea

**Received:** January 11, 2016  
**Revised:** February 25, 2016  
**Accepted:** March 1, 2016

## Corresponding Author

So-Young Jin, MD  
Department of Pathology, Soon Chun Hyang University Seoul Hospital, 59 Daesagwan-ro, Yongsan-gu, Seoul 04401, Korea  
Tel: +82-2-709-9424  
Fax: +82-2-709-9441  
E-mail: jin0924@schmc.ac.kr

**Background:** The histomorphologic criteria for the pathological features of liver tissue from patients with non-alcoholic fatty liver disease (NAFLD) remain subjective, causing confusion among pathologists and clinicians. In this report, we studied interobserver agreement of NAFLD pathologic features and analyzed causes of disagreement. **Methods:** Thirty-one cases of clinicopathologically diagnosed NAFLD from 10 hospitals were selected. One hematoxylin and eosin and one Masson's trichrome-stained virtual slide from each case were blindly reviewed with regard to 12 histological parameters by 13 pathologists in a gastrointestinal study group of the Korean Society of Pathologists. After the first review, we analyzed the causes of disagreement and defined detailed morphological criteria. The glass slides from each case were reviewed a second time after a consensus meeting. The degree of interobserver agreement was determined by multi-rater kappa statistics. **Results:** Kappa values of the first review ranged from 0.0091–0.7618. Acidophilic bodies ( $k=0.7618$ ) and portal inflammation ( $k=0.5914$ ) showed high levels of agreement, whereas microgranuloma ( $k=0.0984$ ) and microvesicular fatty change ( $k=0.0091$ ) showed low levels of agreement. After the second review, the kappa values of the four major pathological features increased from 0.3830 to 0.5638 for steatosis grade, from 0.1398 to 0.2815 for lobular inflammation, from 0.1923 to 0.3362 for ballooning degeneration, and from 0.3303 to 0.4664 for fibrosis. **Conclusions:** More detailed histomorphological criteria must be defined for correct diagnosis and high interobserver agreement of NAFLD.

**Key Words:** Non-alcoholic fatty liver disease; Pathologic features; Interobserver agreement

Nonalcoholic fatty liver disease (NAFLD) is a clinicopathological spectrum characterized by hepatic steatosis without history of significant alcohol use or other known liver disease. The commonly associated factors include obesity, insulin resistance syndrome, and hyperlipidemia. Other associations include jejunio-ileal bypass/gastroplasty surgery for morbid obesity, parenteral nutrition, forms of malnutrition, bacterial contamination of the small bowel, certain inherited metabolic disorders, and a wide range of drugs and environmental toxins.<sup>1</sup> The natural history and histological spectrum of NAFLD range from stable simple steatosis to progressive or advanced disease, such as steatohepatitis, cirrhosis, and even hepatocellular carcinoma.<sup>2-13</sup> A clinicopathological correlation is needed to diagnose NAFLD, as clinical assessment, laboratory tests, and imaging techniques alone provide limited pertinent information. Liver biopsy is the only reliable diagnostic tool to evaluate a patient with suspected NAFLD. Degree of steatosis, liver injury, and fibrosis associated with NAFLD can be identified and differentiated from simple steatosis and nonalcoholic steatohepatitis (NASH), which is the progressive form of NAFLD.<sup>14,15</sup> A NAFLD histopathological grading system was proposed by Brunt *et al.* in 1999.<sup>16</sup> They proposed to classify each of steatosis, inflammation, and ballooning degeneration into three grades and fibrosis into four stages. Since then, Kleiner *et al.*<sup>17</sup> of the NASH Clinical Research Network have proposed a more relevant classification. Other than steatosis, inflammation, and ballooning degeneration, they proposed the classification of fibrosis around the central area into three more detailed grades. Also, they scored each finding and suggested that a score greater than five indicates NASH. However, until now, classifications for NAFLD or NASH have differed depending on the researcher, and no NAFLD histopathological grading system has been standardized, causing confusion for clinicians and pathologists. Furthermore, the histomorphological criteria for NAFLD pathologic features in liver tissue remain subjective. Thus, in this study, we determined the interobserver agreement of NAFLD pathologic features between pathologists and analyzed the causes of the disagreement in order to define the histopathological features in more detail as the basis for a grading system.

## MATERIALS AND METHODS

Thirty-one patients with clinically and pathologically diagnosed NAFLD from 10 hospitals (Daegu Catholic University Medical Center, Dong-A University Hospital, Samsung Medical Center, Seoul National University Hospital, Inje University Seoul

Paik Hospital, Seoul St. Mary's Hospital, Soon Chun Hyang University Seoul Hospital, Wonju Severance Christian Hospital, Inha University Hospital, Chungnam National University Hospital) were selected. Selection criteria were clinical NAFLD (non-alcoholic, serologically negative for viral and autoimmune markers, abnormal levels of liver enzymes such as aspartate aminotransferase and alanine aminotransferase) and age  $\geq 19$  years. Cirrhosis cases were included, and cases of drug and toxic injury conditions were excluded. Fifty-one liver biopsies from 10 hospitals were collected. Among them, 31 biopsies ( $\geq 1.5$  cm in length and  $\geq 16$  G needle size) were selected. One hematoxylin and eosin (H&E)- and one Masson's trichrome-stained slide were selected from each of the 31 cases. The biopsy specimens were anonymized and randomized by a researcher not involved in the study. All selected slides were scanned by a virtual slide scanning system (3DHistotech Ltd., Budapest, Hungary) at Asan Medical Center in Seoul.

The following 12 NAFLD pathologic features were selected: steatosis grade, steatosis location, microvesicular steatosis, fibrosis stage, lobular inflammation, microgranuloma, large lipogranuloma, portal inflammation, ballooning degeneration, acidophilic bodies, Mallory's hyaline, and glycogenated nuclei. Each parameter was reviewed and scored using the detailed scoring criteria shown in Table 1.

One H&E- and one Masson's trichrome-stained virtual slide from each case were reviewed for the 12 parameters. Reviews were performed blindly by 13 pathologists from a gastrointestinal study group of the Korean Society of Pathologists. The degree of interobserver agreement for the first review was analyzed by multi-rater Kappa statistics.

The results were shared with all 13 pathologists, and a consensus meeting was held after the first review to analyze the reasons for disagreement and to define the morphologic criteria in more detail.

After the consensus meeting, a second review of the 12 pathological parameters was performed using glass slides from each of the 31 cases. The degree of interobserver agreement after the second review was analyzed by multi-rater Kappa statistics and compared with the results of the first review.

The Institutional Review Board of Seoul St. Mary's Hospital approved this study (KIRB-00562\_5-001).

## RESULTS

Kappa values of interobserver agreement for the first review ranged from 0.0091 to 0.7618 (Table 2). The order of agreement,

**Table 1.** Pathological parameters and detailed scoring criteria of non-alcoholic fatty liver disease

	Pathologic parameter	Criteria for scoring
1	Steatosis grade	1: <5%, 2: 5–33%, 3: 34–66%, 4: >66%
2	Steatosis location	1: zone 1, 2: zone 3, 3: azonal, 4: panacinar
3	Microvesicular steatosis	0: absent, 1: present
4	Fibrosis	0: none 1A: mild, zone 3, perisinusoidal 1B: moderate, zone 3, perisinusoidal 1C: portal/periportal 2: perisinusoidal and portal/periportal 3: bridging fibrosis 4: cirrhosis
5	Lobular inflammation	0: no foci 1: <2 foci per 200 × field 2: 2–4 foci per 200 × field 3: >4 foci per 200 × field
6	Microgranuloma	0: absent, 1: present
7	Large lipogranuloma	0: absent, 1: present
8	Portal inflammation	0: none to minimal, 1: greater than minimal
9	Ballooning	0: absent, 1: few, 2: many/prominent
10	Acidophilic bodies	0: none to rare, 1: many
11	Mallory's hyaline	0: none to rare, 1: many
12	Glycogenated nuclei	0: none to rare, 1: many

according to the kappa value, was acidophilic bodies ( $k = 0.7618$ ), portal inflammation ( $k = 0.5914$ ), large lipogranuloma ( $k = 0.4822$ ), Mallory's hyaline ( $k = 0.4603$ ), steatosis grade ( $k = 0.3830$ ), steatosis location ( $k = 0.3388$ ), fibrosis ( $k = 0.3303$ ), glycogenated nuclei ( $0.3218$ ), ballooning degeneration ( $k = 0.1923$ ), lobular inflammation ( $0.1398$ ), microgranuloma ( $0.0984$ ), and microvesicular fatty change ( $0.0091$ ). The kappa values of the four major pathologic features (steatosis grade, portal inflammation, ballooning degeneration, and fibrosis) were measured as  $0.3829$ ,  $0.5913$ ,  $0.1923$ , and  $0.3303$ , respectively. In particular, ballooning degeneration ( $k = 0.1923$ ), which is an important feature for diagnosis of NASH, showed a low level of agreement.

Kappa values of interobserver agreement for the second review ranged from  $0.1199$  to  $0.7386$  (Table 2). The order of kappa values for interobserver agreement after the second review were portal inflammation ( $k = 0.7386$ ), acidophilic bodies ( $k = 0.6493$ ), steatosis grade ( $k = 0.5638$ ), Mallory's hyaline ( $k = 0.5236$ ), large lipogranuloma ( $k = 0.5004$ ), fibrosis ( $k = 0.4664$ ), steatosis location ( $k = 0.4502$ ), glycogenated nuclei ( $k = 0.3846$ ), ballooning degeneration ( $k = 0.3362$ ), microvesicular fatty change ( $k = 0.2916$ ), lobular inflammation ( $k = 0.2815$ ), and microgranuloma ( $k = 0.1199$ ). The kappa values of interobserver agreement increased for all parameters except acidophilic bodies. Microvesicular steatosis demonstrated the largest improvement ( $k = 0.0091$  to  $0.2916$ ), and microgranuloma the

**Table 2.** Comparison of kappa values after the first and second reviews of non-alcoholic fatty liver disease pathological features

	Pathologic parameters	First review	Second review
1	Steatosis grade	0.3830	0.5638
2	Steatosis location	0.3388	0.4502
3	Microvesicular steatosis	0.0091	0.2916
4	Fibrosis	0.3303	0.4664
5	Lobular inflammation	0.1398	0.2815
6	Microgranuloma	0.0984	0.1199
7	Large lipogranuloma	0.4822	0.5004
8	Portal inflammation	0.5914	0.7386
9	Ballooning	0.1923	0.3362
10	Acidophilic bodies	0.7618	0.6493
11	Mallory's hyaline	0.4603	0.5236
12	Glycogenated nuclei	0.3218	0.3846

smallest ( $k = 0.0984$  to  $0.1199$ ). All kappa values of the four major pathological features increased as follows: steatosis grade from  $k = 0.3830$  to  $0.5638$ , portal inflammation from  $k = 0.5914$  to  $0.7386$ , ballooning degeneration from  $k = 0.1923$  to  $0.3362$ , and fibrosis from  $k = 0.3303$  to  $0.4664$ .

## DISCUSSION

Since the first description of NASH by Ludwig *et al.* in 1980,<sup>18</sup> several NAFLD scoring schemes have been proposed.<sup>16,17,19,20</sup> Among them, the NAFLD activity score (NAS) proposed by Kleiner *et al.*<sup>17</sup> is the most well-known and popular system. Their



proposed NAS system is based on agreement data and a multiple regression analysis of the<sup>14</sup> histological features of steatosis grade, steatosis location, microvesicular steatosis, fibrosis, lobular inflammation, microgranuloma, large lipogranuloma, portal inflammation, ballooning degeneration, acidophilic bodies, pigmented macrophages, megamitochondria, Mallory's hyaline, and glycogenated nuclei. The NAS is defined as the unweighted sum of the scores for steatosis (0–3), lobular inflammation (0–3), and ballooning (0–2) and ranges from 0 to 8. Fibrosis was not included as an NAS component. Kleiner *et al.*<sup>17</sup> reported that the interobserver agreement values for the four major features were 0.79 for steatosis grade, 0.45 for lobular inflammation, 0.56 for ballooning degeneration, and 0.84 for fibrosis. The agreement for other histologic features ranged from  $k = 0.15$  to 0.58. However, the histomorphological features of some parameters remain ambiguous, contributing to low interobserver agreement. We studied interobserver agreement among 13 pathologists for each of the 12 well-known parameters and analyzed the reasons for disagreement. At the first circulation of slides, we reviewed each case without consensus to identify current discrepancies in diagnostic criteria. The kappa values in this review ranged widely from 0.0091 to 0.7618. Kappa values of the four

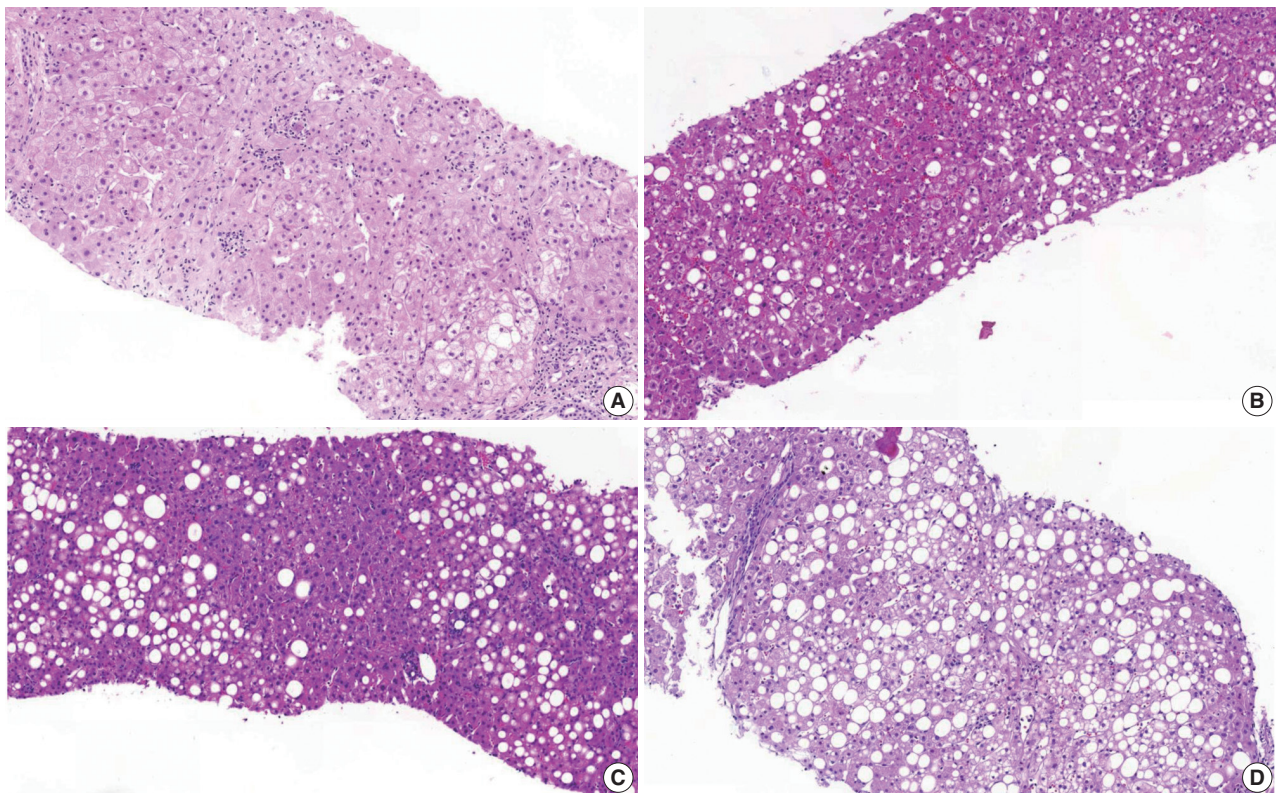
major pathological features at the first review were 0.3830 for steatosis grade, 0.1398 for lobular inflammation, 0.1923 for ballooning degeneration, and 0.3303 for fibrosis, lower than those of Kleiner *et al.*<sup>17</sup>

After the first review, we discussed several points of debate surrounding the definition of each parameter. As a result, we identified several details regarding steatosis grade, lobular inflammation, ballooning degeneration, fibrosis, Mallory's body, and microvesicular fatty change and recommend the following:

(1) Steatosis grade: steatosis grade should be determined by fat volume rather than the number of fatty hepatocytes at  $100\times$  optical magnification (Fig. 1).

(2) Lobular inflammation: lobular inflammation should be graded under  $200\times$  magnification throughout the entire biopsy field, and the mean, not the maximum number in the most active field, should be determined. Spotty necrosis of hepatocytes, lymphocyte aggregations, and acidophilic bodies should be included, whereas lipogranuloma resulting from fat phagocytosis should be excluded as in the histomorphologic criteria for chronic hepatitis grading (Fig. 2).

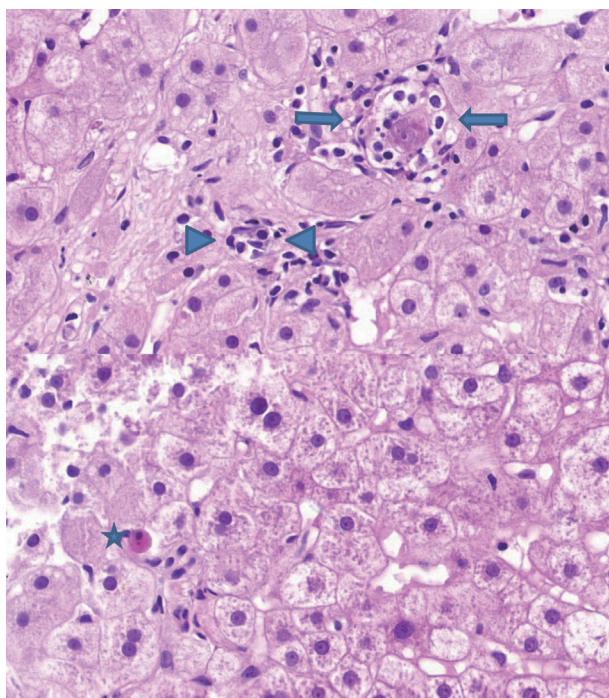
(3) Ballooning degeneration: the histomorphologic criteria of ballooning degeneration are enlarged round cells with loss of



**Fig. 1.** Grading of steatosis. It is determined by fat volume rather than number of fatty hepatocytes. (A) <5%. (B) 5%–33%. (C) 34%–66%. (D) > 66%.



polygonal features and cytoplasm showing heterogeneous granular features (Fig. 3). Hydropic swelling and microvesicular fatty changes should be carefully distinguished from ballooning degeneration. In cases of hydropic swelling, the hepatocyte has large, swollen, and homogenously granular cytoplasm with



**Fig. 2.** Lobular inflammation. A focus spotty necrosis of hepatocytes (arrows), lymphocyte aggregation (arrowheads), and acidophilic bodies (asterisk) can be seen.

well-preserved polygonal features. In particular, microvesicular fatty changes can also be enlarged and can be confused with ballooning degeneration. Microvesicular fatty changes show centrally located nuclei indented by a small fat droplet with a lipoblast-like feature (Fig. 4). When only one or two ballooned cells can be seen throughout the entire field, the term “few” can be applied.

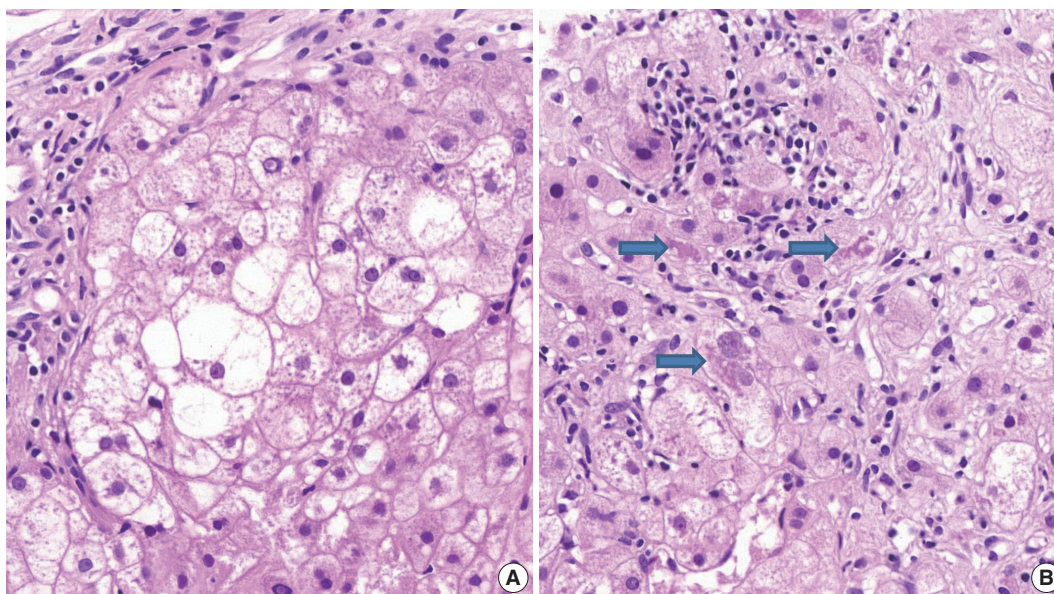
(4) Fibrosis: mild fibrosis should be carefully distinguished from the normal framework around the central area. Only obvious fibrosis with pericellular collagen deposition should be considered as existence of fibrosis (Fig. 5).

(5) Mallory’s body: Mallory’s body should be defined as a definite eosinophilic lump in the cytoplasm (Fig. 3B).

(6) Microvesicular fatty changes: these are defined as lipoblast-like features showing centrally located nucleus and numerous intracytoplasmic micro-fat vacuoles inducing nuclear indentation. Microvesicular fatty changes should be carefully differentiated from cells having small- or medium-sized fat vacuoles without cytoplasmic enlargement and nuclear indentation, as mentioned by Yeh and Brunt (Fig. 6).<sup>21</sup>

All kappa values increased in the second review based on the above criteria. In particular, kappa values for the four major parameters increased from 0.3830 to 0.5638 for steatosis grade, from 0.1398 to 0.2815 for lobular inflammation, from 0.1923 to 0.3362 for ballooning degeneration, and from 0.3303 to 0.4664 for fibrosis.

This increased agreement likely resulted from the consensus

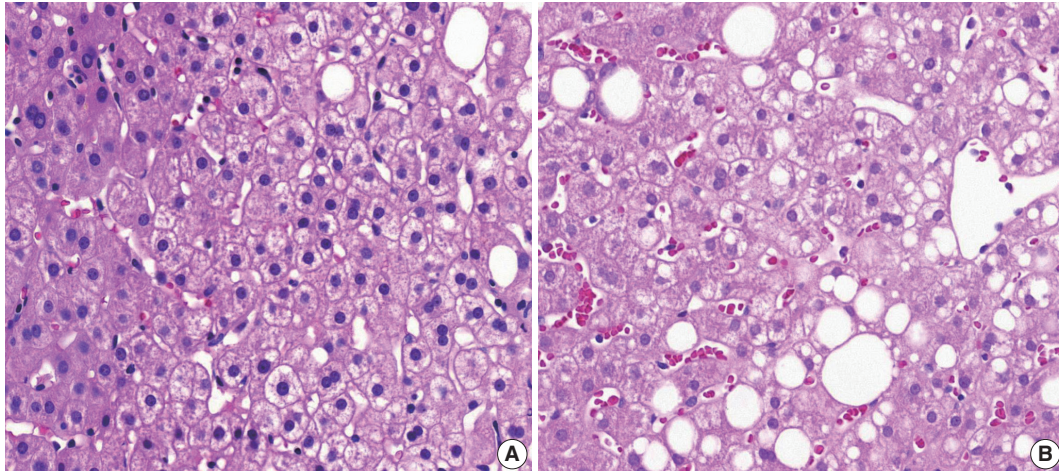


**Fig. 3.** Ballooning degeneration. (A) Group of enlarged round cells with loss of polygonal features and heterogeneous granular cytoplasm are noted. (B) Ballooning degeneration with cytoplasmic Mallory’s bodies (arrows) are found.

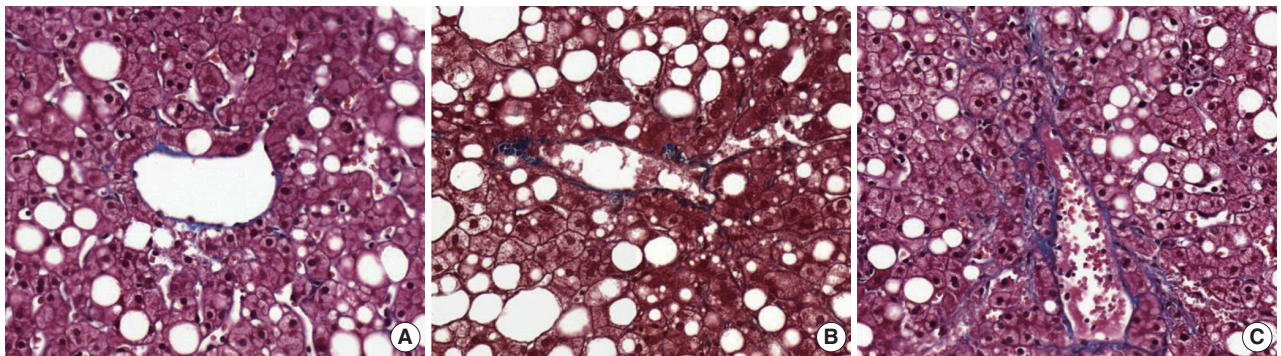


meeting and the determination of more detailed histomorphologic criteria. However, the method used in the second review is more familiar to most pathologists and may have also contrib-

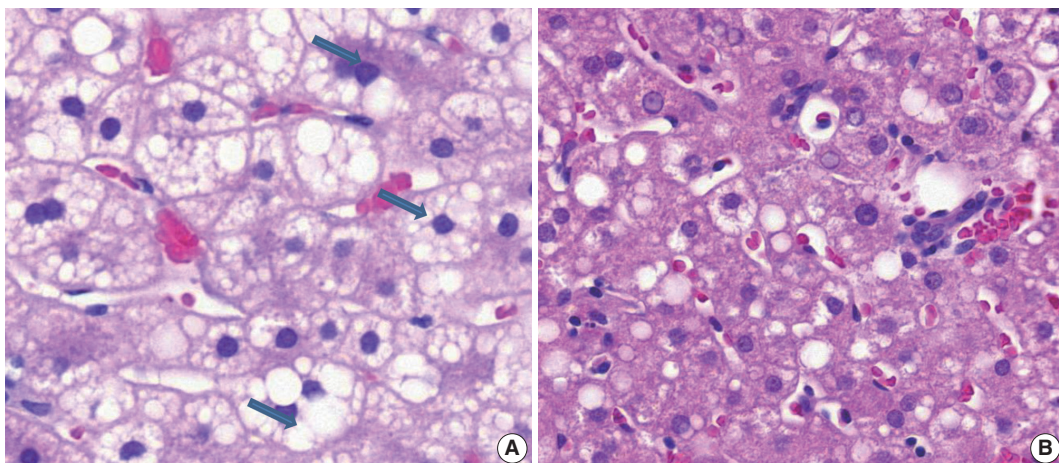
uted to increased agreement. Despite the differences in review method (virtual vs glass), there is no doubt that the exact histomorphologic criteria of NAFLD remain ambiguous and con-



**Fig. 4.** Mimickers of ballooning degeneration. (A, B) Hydropic swelling and microvesicular fatty changes mimicking ballooning degeneration.



**Fig. 5.** Fibrosis around the central area (A), normal framework (B), normal range framework without pericellular collagen deposition (C), and mild perisinusoidal fibrosis with definite pericellular collagen deposition.



**Fig. 6.** Microvesicular steatosis. (A) Microvesicular fatty changes including lipoblast-like features showing numerous intracytoplasmic micro-fat vacuoles with nuclear indentation (arrows). (B) Excluding small- or medium-sized fat vacuoles without cytoplasmic enlargement and nuclear indentation should not be confused with microvesicular steatosis.

tribute to low interobserver agreement between pathologists.

Therefore, our more detailed suggestions for NAFLD histomorphologic criteria—including steatosis grade, lobular inflammation, ballooning degeneration, and fibrosis, as mentioned above—will increase the accuracy of diagnosis and grading of NAFLD and improve interobserver agreement. Through this work and recommendations, we expect that a more exact basis for research of NAFLD and development of a new grading and scoring system will follow.

### Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

### Acknowledgments

This study was supported by the Academic Research Fund 2012 from the Korean Society of Pathologists. We appreciate all members of the Gastrointestinal Pathology Study Group of the Korean Society of Pathologists, particularly Kyung Bun Lee for statistical analysis and Eunsil Yu for scanning the virtual slides.

## REFERENCES

1. Brunt EM, Tiniakos DG. Pathology of steatohepatitis. *Best Pract Res Clin Gastroenterol* 2002; 16: 691-707.
2. Lee RG. Nonalcoholic steatohepatitis: a study of 49 patients. *Hum Pathol* 1989; 20: 594-8.
3. Powell EE, Cooksley WG, Hanson R, Searle J, Halliday JW, Powell LW. The natural history of nonalcoholic steatohepatitis: a follow-up study of forty-two patients for up to 21 years. *Hepatology* 1990; 11: 74-80.
4. Teli MR, James OF, Burt AD, Bennett MK, Day CP. The natural history of nonalcoholic fatty liver: a follow-up study. *Hepatology* 1995; 22: 1714-9.
5. Ong JP, Younossi ZM. Epidemiology and natural history of NAFLD and NASH. *Clin Liver Dis* 2007; 11: 1-16.
6. Dunn W, Xu R, Wingard DL, *et al.* Suspected nonalcoholic fatty liver disease and mortality risk in a population-based cohort study. *Am J Gastroenterol* 2008; 103: 2263-71.
7. Ong JP, Pitts A, Younossi ZM. Increased overall mortality and liver-related mortality in non-alcoholic fatty liver disease. *J Hepatol* 2008; 49: 608-12.
8. Carter-Kent C, Yerian LM, Brunt EM, *et al.* Nonalcoholic steatohepatitis in children: a multicenter clinicopathological study. *Hepatology* 2009; 50: 1113-20.
9. Musso G, Gambino R, Cassader M, Pagano G. A meta-analysis of randomized trials for the treatment of nonalcoholic fatty liver disease. *Hepatology* 2010; 52: 79-104.
10. Starley BQ, Calcagno CJ, Harrison SA. Nonalcoholic fatty liver disease and hepatocellular carcinoma: a weighty connection. *Hepatology* 2010; 51: 1820-32.
11. Abrams GA, Kunde SS, Lazenby AJ, Clements RH. Portal fibrosis and hepatic steatosis in morbidly obese subjects: a spectrum of nonalcoholic fatty liver disease. *Hepatology* 2004; 40: 475-83.
12. 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: 132-8.
13. Xanthakos S, Miles L, Bucuvalas J, Daniels S, Garcia V, Inge T. Histologic spectrum of nonalcoholic fatty liver disease in morbidly obese adolescents. *Clin Gastroenterol Hepatol* 2006; 4: 226-32.
14. Brunt EM. Pathology of nonalcoholic fatty liver disease. *Nat Rev Gastroenterol Hepatol* 2010; 7: 195-203.
15. Kleiner DE, Brunt EM. Nonalcoholic fatty liver disease: pathologic patterns and biopsy evaluation in clinical research. *Semin Liver Dis* 2012; 32: 3-13.
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: 2467-74.
17. 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.
18. Ludwig J, Viggiano TR, McGill DB, Oh BJ. Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clin Proc* 1980; 55: 434-8.
19. 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: 1413-9.
20. Promrat K, Lutchman G, Uwaifo GI, *et al.* A pilot study of pioglitazone treatment for nonalcoholic steatohepatitis. *Hepatology* 2004; 39: 188-96.
21. Yeh MM, Brunt EM. Pathology of nonalcoholic fatty liver disease. *Am J Clin Pathol* 2007; 128: 837-47.