

Original Article

Sonographic assessment of fatty liver: intraobserver and interobserver variability

Mustafa Cengiz¹, Senem Sentürk², Bulent Cetin³, Aylin Hasanefendioğlu Bayrak⁴, Senem Uysal Bilek⁴

¹Department of Gastroenterology, Dr. A.Y. Ankara Oncology Education and Research Hospital, Ankara, Turkey;

²Department of Radiology, Göztepe Training and Research Hospital, Istanbul, Turkey; ³Department of Oncology, Gazi University Faculty of Medicine, Ankara, Turkey; ⁴Department of Radiology, Dicle University Faculty of Medicine, Diyarbakır, Turkey

Received September 24, 2014; Accepted November 25, 2014; Epub December 15, 2014; Published December 30, 2014

Abstract: Non-alcoholic fatty liver disease (NAFLD) is a common liver disease worldwide and ultrasonography is widely used in the diagnosis and the follow-up we purposed to assess intraobserver and interobserver variability in the sonographic evaluation of the existence and steatosis grades of NAFLD. Alanine aminotransferase (ALT) and aspartate aminotransferase (AST) levels and AST to ALT (AST/ALT) ratio were compared between the grades of hepatosteatosi. Hepatic ultrasonography (US) examinations consisted of 5-10 static images of 113 successive adult patients, whose records were in the picture archiving and communication system (PACS) of our hospital were retrospectively evaluated by two experienced radiologists. Hepatic images were graded into 4 groups; as normal, mild, moderate or severe hepatic steatosis. Evaluation of hepatic steatosis of the same set of images was repeated after one month under the same conditions. Interobserver and intraobserver agreement was assessed by using kappa (κ) statistics. In each group, the percentage of individuals with high ALT and/or AST, or AST/ALT ratio over 1 was calculated. The intraobserver agreement was 51%, fair kappa ($\kappa=0.356$) for observer 1; and 68%, moderate ($\kappa=0.591$) for observer 2. The interobserver agreements in the initial and second readings were 39% and 40%, fair ($\kappa=0.208$) and ($\kappa=0.225$), respectively. Elevations of ALT and/or AST levels were similar between groups depending on the degree of hepatosteatosi among the patients. Visual assessment of NAFLD by ultrasonography has substantial interobserver variability, and reproducibility of results is limited. More objective imaging modalities are needed to evaluate the degree of hepatosteatosi.

Keywords: Hepatosteatosi, interobserver-intraobserver variability, NAFLD, ultrasonography

Introduction

Hepatic steatosis is a result of the deposition of triglycerides in the hepatocytes and divided into two subgroups, alcoholic and nonalcoholic fatty liver diseases (NAFLD). NAFLD is a widespread form of hepatic disorder with a prevalence between 10% and 24% in normal population and nearly 74% in obese population [1-5]. NAFLD is a group of hepatic disorders, ranges from simple hepatic steatosis characterized with no inflammation except macrovesicular or microvesicular steatosis and non-alcoholic steatohepatitis (NASH) an inflammatory reaction with balloonic degeneration, inflammation with or without fibrosis [6, 7]. NASH is described as the coexistence of fat accumulation, inflammation, and this ultimately may result in liver cirrhosis

as well as hepatocellular carcinoma [8-10]. As hepatocyte injury caused by fat accumulation may yield to steatohepatitis, fibrosis and cirrhosis, it is important to diagnose and follow up NAFLD [11].

Currently, the only definite way to diagnose NAFLD is liver biopsy, a costly, invasive and morbidity associated diagnostic process [12]. Ultrasonography (US), computed tomography (CT) and magnetic resonance imaging (MRI) may be used to detect NAFLD noninvasively. However, inflammation and early period of fibrosis in the NAFLD can not be detected by conventional imaging modalities [11]. US, a simple, non-invasive, easily applicable and safely repeatable imaging modality, is widely used in the diagnosis and the follow-up of NAFLD The most suitable

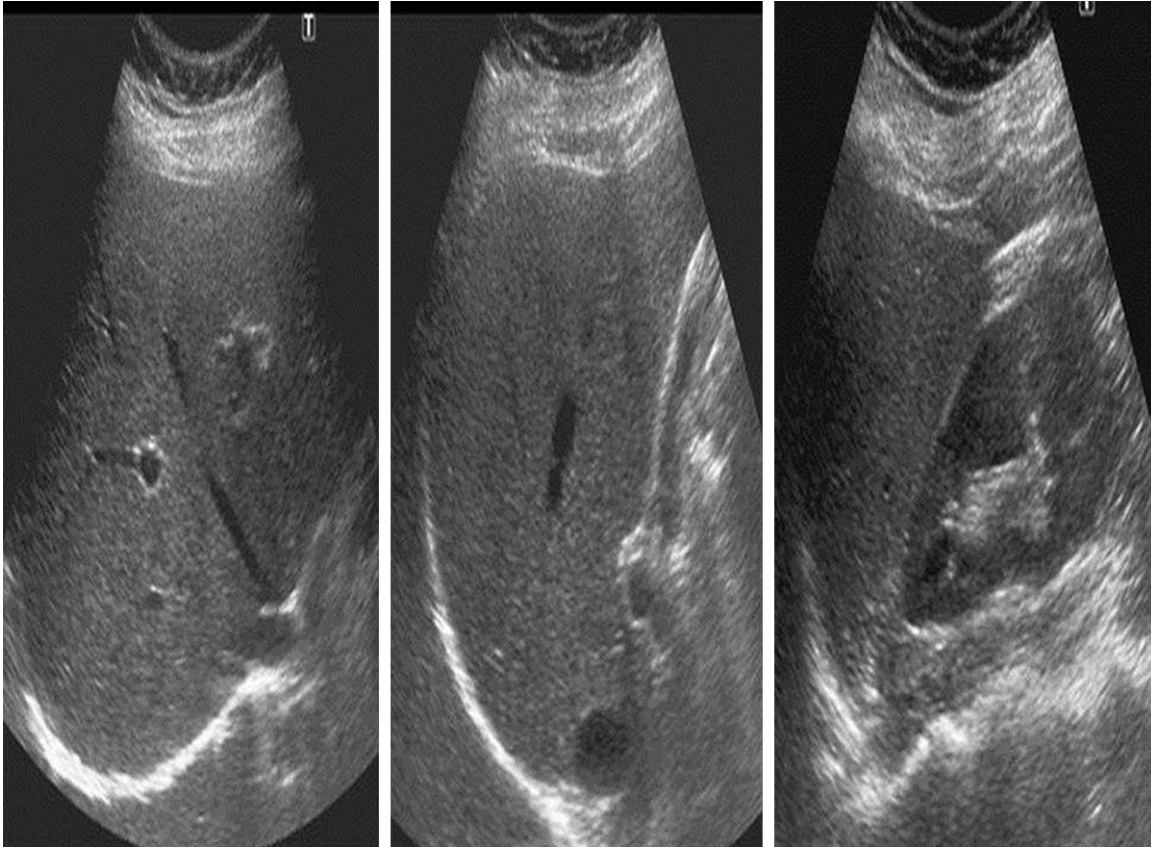


Figure 1. US images of a 36-year-old female patient. Observer 1 graded the liver as normal on the first evaluation; and as mildly fatty on the second evaluation. Observer 2 graded liver as mildly fatty on both evaluations.

non-invasive method in detecting hepatic steatosis is US with the sensitivity 60-94% and specificity 66-97%, but sensitivity and positive predictive values are low in mild steatosis [13]. It is actually an operator dependent and the assessment of fatty liver depends primarily on the subjective evaluation of hepatic echogenicity and posterior attenuation of the ultrasound beam [14, 15]. The dependency of the diagnosis on the subjective judgments of operators is also problematic [16].

Mild or moderate increase in serum aminotransferase levels is one of the most typical and frequent laboratory presentation of patients with NAFLD. The actual aspartate aminotransferase (AST) to alanine aminotransferase (ALT) ratio (AST/ALT) is generally lower than 1, however the ratio rises as the liver fibrosis increases [17, 18].

We aimed to investigate the interobserver and intraobserver variability of the sonographic evaluation of NAFLD in routine clinical practice and

evaluate whether there was a correlation between the degree of hepatic steatosis and ALT and/or AST elevation or AST/ALT ratio.

Materials and methods

Patients and ultrasonography evaluation

A retrospective review of two hundred abdominal sonograms obtained during 2010-2011 and stored in the picture archiving and communication system (PACS) had been preliminarily carried out by one investigator. One hundred and twenty nine consecutive adult patients referred for possible NAFLD who met the following criteria were selected for this study; the evaluation was theoretically sufficient when integrated each transverse and longitudinal views from the liver with 5-10 images, absolutely no central hepatic lesion and both right kidney and liver were seen within a minimum of one of all the images. Patients with heterogeneous liver structure or ascites were excluded from the study. Clinical and laboratory findings

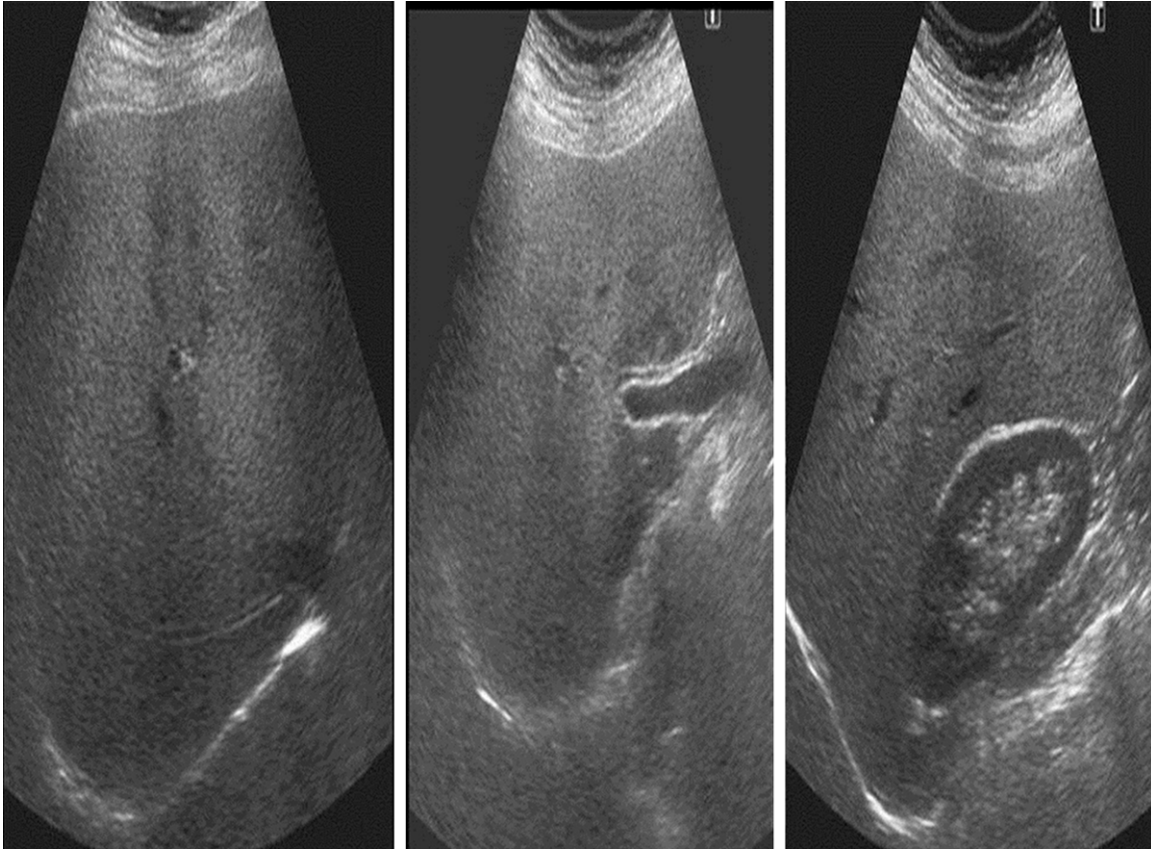


Figure 2. US images of a 55-year-old female patient. Observer 1 graded the liver as moderately fatty; whereas observer 2 graded the liver as severely fatty on both observations.

of the patients were acquired from the hospital records. Alcohol consumption of more than 20 gr/day, the presence of any liver diseases that could affect the fatty liver were considered as exclusion criteria. Sixteen patients depending on the history of alcohol consumption more than 20 gr/day and/or the presence of any liver disorders were excluded from the study.

All US studies were performed by the same radiologists and imaging unit (Nemio XG; Toshiba, San Jose, CA, USA) with a 3.5- to 5-MHz convex probe. Five to ten US images of 113 patients free from alcoholic liver disease and no extra liver related disorders were assessed by 2 skilled radiologists. One radiologist had 8 years of experience and the other had 9 years of experience in abdominal US. The images were examined by the experts on the same monitor under the exact same lighting conditions. The experts were blinded to the clinical and laboratory data of patients and unaware of the previously reported, written interpretation and the other observer's assessments.

The observers rated each situation as normal liver, mild, moderate or severe fatty liver. The liver was assessed as normal when the consistency was homogeneous, displayed fine level echoes, minimally hyperechoic or even isoechoic in contrast to regular renal cortex. Mild steatosis was evaluated as the minor increase in liver echogenicity. In moderate steatosis, there were visual images associated with intrahepatic vessels, the slightly damaged diaphragm and the existence of increased liver organ echogenicity. Severe steatosis was named as the marked increase in hepatic echogenicity, poor penetration of posterior segment from the right lobe of the liver, poor or any visual images from the hepatic vessels and diaphragm [19]. After an interval of one month, the observers reassessed the presence and severity of steatosis in the same 113 cases under the same conditions and were blinded to the initial evaluations. Finally, the degree of steatosis in the liver was re-evaluated as a consensus between the two observers and the patients were placed in 4 groups. In each group the percentage of pati-

Ultrasonography and NAFLD

Table 1. Intraobserver agreement rates

	Intraobserver agreement (kappa)	Intraobserver agreement (percentage)
Observer 1	$\kappa=0.356$	51%
Observer 2	$\kappa=0.591$	68%

Table 2. Interobserver agreement rates

	Interobserver agreement (kappa)	Interobserver agreement (percentage)
1st evaluation	$\kappa=0.208$	39%
2nd evaluation	$\kappa=0.225$	40%

ents with high levels of ALT, AST and AST/ALT ratio over 1 were calculated and the results were compared.

Statistical analysis

Intraobserver and interobserver variations were investigated by using the Kappa statistics. The interobserver and intraobserver agreement percentages were calculated by dividing the number of occasions of complete agreement by the total number of occasions. Weighted kappa statistics were used to determine the degree of agreement after correction for the agreement expected by chance. The kappa statistic was interpreted as follows: less than 0.00, poor agreement; 0.00-0.20, slight agreement; 0.21-0.40, fair agreement; 0.41-0.60, moderate agreement; 0.61-0.80, substantial agreement; and 0.81-1.00, almost perfect agreement. The frequency percentages of patients who had high serum transaminase levels and AST/ALT ratio were compared between groups by using Chi square or Fisher-exact test where appropriate.

Ethics

This research study was appropriate according to the ethical guidelines of the 1975 Declaration of Helsinki, as updated in 2008. The local ethics committee approved the research protocol.

Results

Figures 1 and 2 illustrate the variability between and within observers. Intraobserver agreement for grading the severity of hepatic steatosis between the first and second evaluations was fair (kappa (κ)=0.356) for the first observer, whereas it was moderate (κ =0.591) for the second observer. The first observer agreed with

herself in 58 (51%) of the 113 cases, the second in 77 (68%) of all the cases (κ =0.51).

Interobserver agreement for the evaluation of the occurrence or lack of fatty liver was 39% (κ =0.208) in the first evaluation and 40% (κ =0.225) in the second. There was fair agreement between the observers at both evaluations. The observer agreed in 44 (40%) of the 113 cases at the first evaluation, and 45 (40%) of the cases at the second evaluation. Intraobserver and interobserver agreement rates of grading the severity of hepatic steatosis are summarized in **Tables 1 and 2**.

The consensus of both observers graded 27 cases as normal, 29 cases as mild fatty liver, 37 cases as moderate fatty liver, and 20 cases as severe fatty liver. ALT and/or AST elevation was detected in 3 (11%) of patients with a normal US evaluation. ALT and/or AST was elevated in 8 (27%) in patients with mild fatty liver, 5 (13%) of patients with moderate fatty liver, and 2 (10%) of patients with severe fatty liver. There was no association between the degree of steatosis and elevation of ALT and/or AST. Thirteen (48%) of cases with normal liver evaluation had an AST/ALT ratio more than 1. AST/ALT ratio was also more than 1 in 10 (34%) of individuals with mild fatty liver, 11 (30%) with moderate, and 6 (30%) with severe hepatic steatosis. There was no association between the degree of steatosis and elevated (> 1) AST/ALT ratio. The results are demonstrated in **Table 3**.

Discussion

In this particular study, we investigated the interobserver and intraobserver variability of the sonographic evaluation of NAFLD in clinical practice and evaluate whether there was a correlation between the degree of hepatic steatosis and ALT and/or AST elevation or ALT/AST ratio. We showed that the radiologists sometimes may differ substantially in their evaluation of grading the fatty liver, and observed that elevation of ALT and/or AST, or AST/ALT ratio (> 1) were not correlated to the degree of the steatosis in the liver. Visual evaluation of NAFLD by US had important interobserver variability, and reproducibility of results was limited. Hepatic steatosis, a clinical problem, is one of the most common hepatic disorders. NAFLD is one of the components of metabolic syndrome, and has a

Ultrasonography and NAFLD

Table 3. Number of individuals with elevated ALT and/or AST and with an elevated (> 1) AST/ALT ratio in each group

	Individuals with elevated ALT and/or AST (percentage)	Individuals with elevated (> 1) AST/ALT ratio (percentage)
Normal	3/27 (11%)	13/27 (48%)
Mild hepatic steatosis	8/29 (27%)	10/29 (34%)
Moderate hepatic steatosis	5/37 (13%)	11/37 (30%)
Severe hepatic steatosis	2/29 (10%)	6/20 (30%)

probability of enhancing to cirrhosis as well as liver failure [11]. Numerous individuals with hepatic steatosis have no signs or symptoms of liver disorders in the course of examination, and steatosis is frequently found incidentally on cross sectional images. In some individuals, hepatic steatosis may be the reason behind hepatomegaly in addition to increased liver enzyme levels, prompting a specific hepatobiliary US imaging.

US is considered the most popular used imaging technique within the evaluation of abdominal disorders generally and hepatic illnesses particularly. It is operator-dependent and also diagnosis of hepatic steatosis is reliant primarily on the opinion based evaluation regarding liver echogenicity. Once the liver possesses an echogenic appearance, it may be construed to be fatty. Liver echogenicity typically equals or even slightly surpasses renal cortical echogenicity, however, this examination is dependent upon the visual notion of the viewer. Moreover, the typical echogenicity of the renal cortex may also change by renal disorders which makes comparability less trustworthy. In this specific investigation, a well-recognized categorization was utilized intended for determining hepatic steatosis as; mild, moderate, and severe degrees of steatosis according to the rise in hepatic echogenicity, diminished visual images of hepatic vessels and the diaphragm, and inadequate transmission of the posterior areas of the liver [19]. By this research we demonstrated that interobserver variability within the sonographic evaluation of hepatic steatosis is actually significant. Among determinants associated with diagnostic precision is actually reproducible, from the evaluation that produces exactly the same or even similar outcomes when repeated. Although a well-organized classification of hepatic steatosis was used, the proportion of agreement among observers was 39% in the first evaluation and 40% in the second. There was

only slight agreement between the two observers at both readings. Intraobserver agreement was better than interobserver agreement. In more than one third of the conditions each expert assessed the severity of steatosis diversely on the second evaluation, and the intraobserver agreement was fair ($\kappa=0.356$) for the first observer, and moderate ($\kappa=0.591$) for the second observer. Interobserver agreement was fair on both readings ($\kappa=0.208$ and $\kappa=0.225$). We showed that both intra- and interobserver reproducibility of the results was low. The usefulness of US for the diagnosis of NAFLD is evaluated, to some extent, because of its simplicity. Quantification of fatty change using US to supplement elastography has also occasionally been reported, and further development of this application is expected [20]. It is impossible to differentiate between NASH and simple steatosis using any imaging methods. At the same time, certain US and CT findings, such as irregularity of the liver surface, blunt margins of the liver, and splenomegaly, suggest the presence of chronic liver diseases, including NASH with advanced fibrosis, and can indicate the need for further attention. It has been reported that the differentiation of NASH and simple hepatic steatosis may be possible using contrast-enhanced US [21].

Saadeh et al. have reported the variability of radiologic interpretations of images of individuals with NAFLD [16]. They found that while the intraobserver agreement of the severity of steatosis was substantial ($\kappa=0.63$), the particular interobserver agreement was fair ($\kappa=0.40$). Strauss et al. have also evaluated inter- and intraobserver variability of the US evaluation in the existence and severity of hepatic steatosis [22]. The particular interobserver and intraobserver agreement proportions for the existence of hepatic steatosis were 72% ($\kappa=0.43$) and 76% ($\kappa=0.54$), respectively. They found moderate intra- and interobserver agreement. Intra-

and interobserver agreement rates in our study was slightly lower than the previous studies. The observers in our study had been working in the different sonography units for many years. This aspect probably have created decreased agreement degrees compared to previous reports. There was no substantial or perfect or intra- or interobserver agreement in the previous studies. The results of previous studies and our study demonstrated that radiologists may vary significantly within their US evaluation of hepatic steatosis.

In a study, the authors evaluated the patients with NAFLD to find new non-invasive modalities of ultrasound attenuation measurements and they stated that hepatorenal-indexes, ultrasound attenuation, and tissue elasticity might be useful in differentiating steatosis and healthy individuals and expressing the differences [23]. The stage of liver fibrosis can now be estimated non-invasively by using several techniques of elastography (including FibroScan and acoustic radiation force impulse imaging (ARFI)) The FibroScan investigation also may enable differentiation between fibrosis and cirrhosis [23], but when morbid obesity exists it is not easy. Sonographically the prevalence rate varied between 17% and 46%, differs according to the population study. The predicted frequency of NAFLD among common population is approximately between 6% and 33% worldwide with a median of 20% but for NASH it is markedly lower and ranges from 3% to 5% [24]. While the ultrasonography scan may be a critical component in the evaluation of patients referred for possible NAFLD/NASH, it commonly does not identify advanced liver disease. US evaluation may not be the final step in the evaluation. It has been concluded that the lack of steatosis on ultrasonography is not convincing especially in advanced liver fibrosis [25].

Younossi *et al.* found inter- and intraobserver variability in the pathological diagnosis of NAFLD and it has also been shown that inter-observer variability existed in histopathological evaluations also [26, 27]. Some measures should be considered to solve this problem. Currently, liver biopsy is crucial in the diagnosis of NASH, but in the future, combining scoring systems and imaging methods may efficiently diagnose NAFLD/NASH. Whether these scoring systems reflect the long term prognosis and carcinogenesis their potential remains to be investigated. An improved scoring system will provide benefit

in detecting NASH and reducing liver disease-related deaths in the future [28]. To date, no clear guidance is given in the literature with regard to defining an indication for liver biopsy in NAFLD. The severity of fatty changes is not correlated with the advancement of fibrosis and it decreases with the progression of fibrosis in NASH. The grade of fatty change obtained from imaging modalities should not be employed as an evaluation criterion for NAFLD severity.

In this study, we also searched individuals with ALT and/or AST elevation in each group, and determined percentage of patients with an AST/ALT ratio over 1. Some patients without fatty liver had elevated ALT and/or AST, and moreover there was no difference in ALT and/or AST elevation rate in patients with different degree of steatosis. Nearly fifty percent of patients without fatty liver had an AST/ALT ratio over 1. There was no difference in percentage of patients with AST/ALT ratio over 1 between patients with different degree of steatosis.

In patients with steatohepatitis and fibrosis, elevated AST/ALT ratio usually accompanies by hepatic enzyme elevation [17]. We observed that most of our patients with elevated AST/ALT ratio (> 1) did not have elevated ALT and/or AST. Even with severe steatosis most of our cases had no elevation of AST and/or ALT. Our findings showed that AST/ALT ratio was not affected by the degree of steatosis. We suggest that in the absence of steatohepatitis and fibrosis in patients with or without fatty liver AST/ALT ratio is nonspecific.

An essential restriction of our study was that the examined images were stationary; the experts were not present through the entire assessments and retrospective design of the study. The liver echogenicity might be affected by the settings of the US unit in real time imaging. A further limitation was the absence of the liver biopsy if we had compared the results of US and liver biopsy it would have better define the results of the study.

In conclusion, we showed that visual evaluation of NAFLD by US has considerable interobserver variability and radiologists sometimes may differ substantially in their evaluation of grading hepatic steatosis and the reproducibility of results was limited. We also observed that ele-

vation of ALT and/or AST, or AST/ALT ratio (> 1) were not correlated with the degree of the hepatic steatosis. Much more objective as well as quantitative approaches are needed for evaluating the hepatic steatosis.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Mustafa Cengiz, Department of Gastroenterology, Dr. A.Y. Ankara Oncology Training and Research Hospital, Yenimahalle, Ankara 06200, Turkey. Tel: +90 (312) 3060909; Fax: +90 (312) 3340352; E-mail: drmustafacen@gmail.com

References

- [1] Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, Grundy SM, Hobbs HH. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *Hepatology* 2004; 40: 1387-95.
- [2] Clark JM. The epidemiology of nonalcoholic fatty liver disease in adults. *J Clin Gastroenterol* 2006; 40: 5-10.
- [3] Lazo M, Clark JM. The epidemiology of nonalcoholic fatty liver disease: a global perspective. *Semin Liver Dis* 2008; 28: 339-50.
- [4] Zhou YJ, Li YY, Nie YQ, Ma JX, Lu LG, Shi SL, Chen MH, Hu PJ. Prevalence of fatty liver disease and its risk factors in the population of South China. *World J Gastroenterol* 2007; 13: 6419-24.
- [5] Hamer OW, Aguirre DA, Casola G, Lavine JE, Woenckhaus M, Sirlin CB. Fatty liver: imaging patterns and pitfalls. *Radiographics* 2006; 26: 1637-53.
- [6] Chalasani N, Younossi Z, Lavine JE, Diehl AM, Brunt EM, Cusi K, Charlton M, Sanyal AJ. The diagnosis and management of non-alcoholic fatty liver disease: practice Guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatology* 2012; 55: 2005-23.
- [7] Cohen JC, Horton JD, Hobbs HH. Human fatty liver disease: old questions and new insights. *Science* 2011; 332: 1519-23.
- [8] Chiang DJ, Pritchard MT, Nagy LE. Obesity, diabetes mellitus, and liver fibrosis. *Am J Physiol Gastrointest Liver Physiol* 2011; 300: 697-702.
- [9] Hebbard L, George J. Animal models of nonalcoholic fatty liver disease. *Nat Rev Gastroenterol Hepatol* 2011; 8: 35-44.
- [10] Schattenberg JM, Schuppan D. Nonalcoholic steatohepatitis: the therapeutic challenge of a global epidemic. *Curr Opin Lipidol* 2011; 22: 479-88.
- [11] Lall CG, Aisen AM, Bansal N, Sandrasegaran K. Nonalcoholic fatty liver disease. *AJR Am J Roentgenol* 2008; 190: 993-1002.
- [12] Rockey DC, Caldwell SH, Goodman ZD, Nelson RC, Smith AD. American Association for the Study of Liver Diseases. Liver biopsy. *Hepatology* 2009; 49: 1017-44.
- [13] Machado MV, Cortez-Pinto H. Non-invasive diagnosis of non-alcoholic fatty liver disease. A critical appraisal. *J Hepatol* 2013; 58: 1007-19.
- [14] Hernaez R, Lazo M, Bonekamp S, Kamel I, Brancati FL, Guallar E, Clark JM. Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: a meta-analysis. *Hepatology* 2011; 54: 1082-90.
- [15] Dasarathy S, Dasarathy J, Khiyami A, Joseph R, Lopez R, McCullough AJ. Validity of real time ultrasound in the diagnosis of hepatic steatosis: a prospective study. *J Hepatol* 2009; 51: 1061-7.
- [16] Saadeh S, Younossi ZM, Remer EM, Gramlich T, Ong JP, Hurley M, Mullen KD, Cooper JN, Sheridan MJ. The utility of radiological imaging in nonalcoholic fatty liver disease. *Gastroenterology* 2002; 123: 745-50.
- [17] Angulo P. GI epidemiology: nonalcoholic fatty liver disease. *Aliment Pharmacol Ther* 2007; 25: 883-9.
- [18] Angulo P. Nonalcoholic fatty liver disease. *N Engl J Med* 2002; 346: 1221-31.
- [19] Rumack CM. Diagnostic Ultrasound. In: Rumack CM, editor. St Louis: Mosby; 1998. pp. 110-112.
- [20] de Ledinghen V, Vergniol J, Foucher J, Merrouche W, Le Bail B. Non-invasive diagnosis of liver steatosis using controlled attenuation parameter (CAP) and transient elastography. *Liver Int* 2012; 32: 911-8.
- [21] Iijima H, Moriyasu F, Tsuchiya K, Suzuki S, Yoshida M, Shimizu M, Sasaki S, Nishiquchi S, Maeyama S. Decrease in accumulation of ultrasound contrast microbubbles in non-alcoholic steatohepatitis. *Hepatol Res* 2007; 37: 722-30.
- [22] Strauss S, Gavish E, Gottlieb P, Katsnelson L. Interobserver and intraobserver variability in the sonographic assessment of fatty liver. *AJR Am J Roentgenol* 2007; 189: 320-323.
- [23] von Volkman HL, Havre RF, Loberg EM, Haaland T, Immervoll H, Haukeland JW, Hausken T, Gilja OH. Quantitative measurement of ultrasound attenuation and Hepato-Renal Index in Non-Alcoholic Fatty Liver Disease. *Med Ultrason* 2013; 15: 16-22.

Ultrasonography and NAFLD

- [24] Vernon G, Baranova A, Younossi ZM. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther* 2011; 34: 274-85.
- [25] Tapper EB, Krajewski K, Lai M, Challies T, Kane R, Afdhal N, Lau D. Simple non-invasive biomarkers of advanced fibrosis in the evaluation of non-alcoholic fatty liver disease. *Gastroenterol Rep* 2014; [Epub ahead of print].
- [26] Younossi ZM, Gramlich T, Liu YC, Matteoni C, Petrelli M, Goldblum J, Rybicki L, McCullough AJ. Nonalcoholic fatty liver disease: assessment of variability in pathologic interpretations. *Mod Pathol* 1998; 11: 560-565.
- [27] Gawrieh S, Knoedler DM, Saeian K, Wallace JR, Komorowski RA. Effects of interventions on intra- and interobserver agreement on interpretation of nonalcoholic fatty liver disease histology. *Ann Diagn Pathol* 2011; 15: 19-24.
- [28] Sumida Y, Nakajima A, Itoh Y. Limitations of liver biopsy and non-invasive diagnostic tests for the diagnosis of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. *World J Gastroenterol* 2014; 20: 475-85.