INTER-OBSERVER AGREEMENT FOR DETECTION AND GRADING OF HEPATIC STEATOSIS - AN ULTRASOUND BASED STUDY

Saba Sohail,[1] Mukhtiar Ahmed Memon,[1] Kausar Jahan Siddiqui[2]

[1] Department of Radiology, Dow Medical College / Civil Hospital (DUHS), Karachi, Pakistan.
[2] Department of Radiology, Civil Hospital, Karachi, Pakistan.

ABSTRACT

OBJECTIVES: To determine the inter observer agreement of ultrasound diagnosis of the severity and grading of steatosis in patients with HCV. **METHODS:** Patients with HCV were evaluated, from March 2008-August 2010 at Radiology department, Dow Medical College and Civil Hospital Karachi, with ultrasound for detection and grading of steatosis and fibrosis using a standardized set of criteria. The same sets of images were reviewed by the same radiologists 5 years later (2015) for determining the agreement in the grading. Kappa (k) statistics were utilized. Present results were compared with that of the pilot study conducted in 2007 on 100 images. **RESULTS:** 452 set of images were reviewed by three radiologists, designated A, B and C. The inter observer agreement was satisfactory to good with k=0.8 for no steatosis, 0.4 for mild steatosis, 0.7 for moderate steatosis and 0.9 for severe steatosis. The inter-observer agreement in the pilot study in 2007 had shown k - values of 0.81 for no steatosis, 0.5 for mild steatosis, 0.8 for moderate steatosis and 0.9 for severe steatosis. **CONCLUSION:** The studied criteria for ultrasound diagnosis and grading of hepatic steatosis had persistently good inter-observer agreement for absence, and moderate to severe grades of steatosis. The agreement was low for mild steatosis.
Key words: Fatty liver, Hepatic steatosis, ultrasound, diagnosis, grading, Inter observer agreement.

Introduction

Hepatic steatosis or fatty liver is increasingly being recognized as a cause of NASH (nonalcoholic steato-hepatitis) which may lead to many complications including hepatocellular carcinoma.[1] The condition can be diagnosed and evaluated by many means including laboratory tests, biopsy, elastography and ultrasound. Ultrasound (U/S) scan is a commonly used modality for the evaluation of hepatic pathologies particularly chronic liver disease of infective or staetotic origin. It is an effective and widely available modality to identify many structural changes in the liver by virtue of visualizing the alteration in hepatic paren-chymal echo pattern and echogenicity. Operator skill remains a practical issue in the general application and accuracy of ultrasound due to the use of subjective non-standardized criteria, single ultrasound parameter utilization, and ultrasound conducted by sonographers. There are variable reports of ultrasound accuracy ranging from low to high reliability.[2] One of the suggested reasons for this variation is operator skill and variation in the sonographic interpretation.[3]

Using objectively structured approach for ultrasound evaluation to identify the stage of steatosis as accurately as possible would help in homogenizing the reports and increasing the reliability. In Pakistani scenario, this can be very cost effective, particularly if a diagnostic algorithm is developed or a scoring system be used for this evaluation on structured basis. A set of such criteria was developed and tested against histopathology results by the author in a small

Correspondence : Dr. Saba Sohail
Department of Radiology,
Dow Medical College / Civil Hospital (DUHS),
Karachi, Pakistan.
Email: drsabasohail@hotmail.com

set of patients.[4] However it remained to be seen as to whether these criteria have good long term agreement among the radiologists familiar with these criteria on a larger number of patients. The present study was conducted to determine the agreement between radiologists regarding the ultrasound diagnosis of presence and grading of hepatic steatosis using the above objectively structured criteria.

Methods

It was an analytical study carried out in the Department of Radiology, Dow Medical College and Civil Hospital Karachi, as a follow up of a research project leading to PhD.[4] Adult patients with hepatitis C were recruited from the Sarwar Zuberi Liver Centre (SZLC), Dow Medical College and Civil Hospital Karachi, and from the medical out patient referrals to the Radiology department. Patients aged under 15 years, alcoholics, auto immune or drug induced hepatitis, deposition disorder, and concomitant HIV were excluded. Those with hepatocellular carcinoma and/or ascites were likewise not considered. Ethical approval was taken from the Ethical Review Board, Dow University of Health Sciences, Karachi vide letter no - PhD-1/ERB-31/DUHS-07.

The primary data collection had started in March 2008 and ended in August 2010. However, in October 2007 a pilot study was conducted to test the inter-observer agreement of the ultrasound criteria used for grading of ultrasound on a 100 patients undergoing abdominal ultrasound without trauma or ascites, using a 3.5 MHz convex probe on gray scale and Doppler spectrum. Two MCPS and one FCPS (the author) radiologists had taken part in that pretesting. Following criteria were used:[4,5]

• No steatosis = normal homogenous pattern of liver with echogenicity lower than adjacent pancreas
• mild steatosis = high echogenicity of liver tissue (equal to or greater than the body of pancreas) with adequately visualized of posterior hemi diaphragm and portal radicles' margins inside liver (Fig. 1)
• moderate steatosis = high echogenicity of liver tissue with obscured posterior hemi diaphragm and portal radicles' margins (Fig. 2)
• severe steatosis= high echogenicity of hepatic parenchyma with markedly compromised visualization
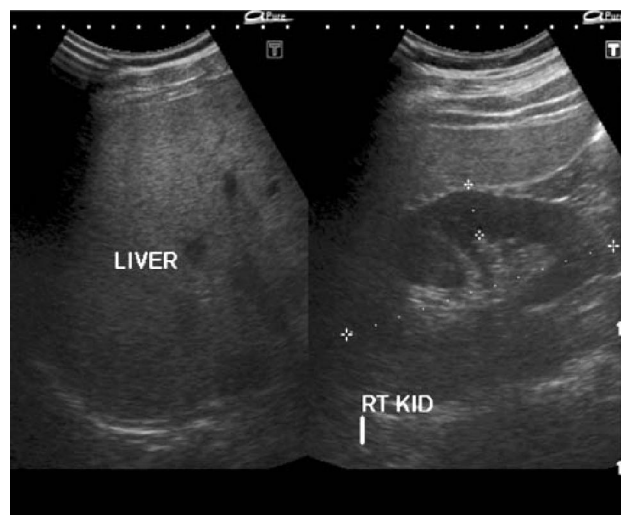

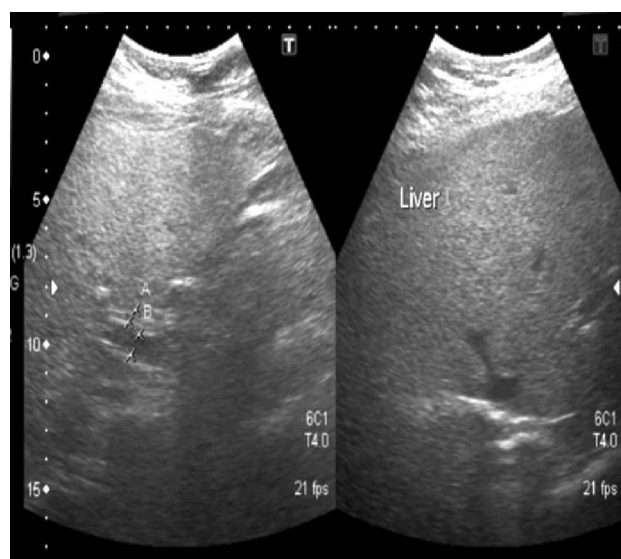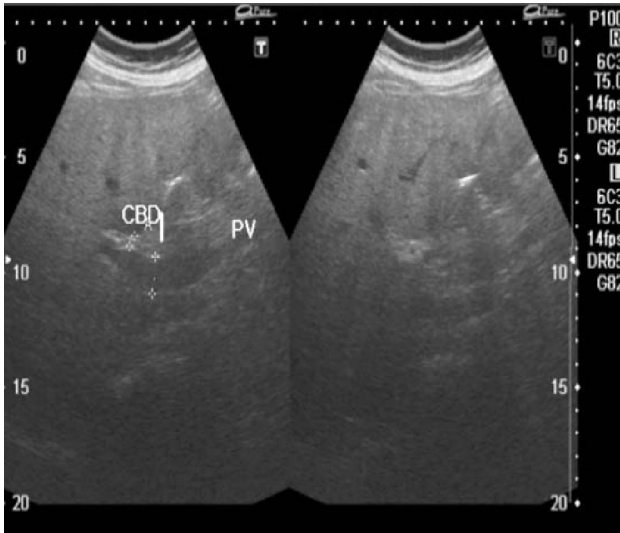**Figure 1:** Ultrasound appearance of mild steatosis

Figure 2: Ultrasound appearance of moderate steatosis

of posterior hemi diaphragm, portal radicles' margins and posterior segments of the right hepatic lobe (Fig. 3).

Later on, 452 patients with hepatitis C were evaluated with ultrasound for diagnosis and grading of hepatic steatosis as part of research leading to PhD, funded by the HEC (project no 1121). Ultrasound findings were recorded and compared with ultrasound. Thermal prints of images clearly showing the hepatic parenchyma, portal radicles, hepatic margins particularly the left lobe, and the adjacent pancreas, were recorded. The same sets of images were reviewed by the same radiologists 5 years later (2015) for determining the agreement in the grading. Kappa

Figure 3: Ultrasound appearance of severe steatosis

(k) statistics were utilized and k values were determined. Present results were compared with that of the pilot study conducted in 2007 on 100 patients.

Results

There were 100 sets of images in the pilot study in 2007-2008, and 452 different sets obtained during 2008-2010, which were reviewed by the same radiologists who had participated in the pilot study. There were three reading radiologists; one being a Faculty member with Membership, Fellowship as well as Doctoral qualifications and the other two being Consultants with Membership of the College of Physicians and Surgeons, Pakistan. All were experienced in ultrasound with 12 years of experience after the first postgraduate qualification for Radiologist A, and 5 years of post-Membership qualification for the other two Radiologists at the time of the pilot study in 2007-2008. All the three radiologists were females.

The inter observer agreement was satisfactory to good with k=0.8 for no steatosis, 0.4 for mild steatosis, 0.7 for moderate steatosis and 0.9 for severe steatosis. The inter-observer agreement in the pilot study in 2007 had shown k - values of 0.81 for no steatosis, 0.5 for mild steatosis, 0.8 for moderate steatosis and 0.9 for severe steatosis between the same radiologists for the respective grades.

Discussion

In current clinical practice, ultrasound is used as the first line imaging modality in the evaluation of hepato-biliary disease being non- invasive and cost effective.[2] The only drawback is that it is operator dependent with significant inter- observer variability.[6] It is therefore necessary to lay down standardized criteria for ultrasound examination so that this variability may be reduced to minimum. This research intended to address to the issue of the lack of literature as well as lack of practical standardized criteria on ultrasound evaluation to overcome the inter- observer variability. The overall accuracy of u/s in diagnosing and grading a fatty liver compared to CT and biopsy is 85-89%.[7-9] A recent meta-analysis stated the sensitivity to range from 55-100% and specificity to range from 126-100%.[10] Despite a large volume of international reports all of which are contrasting, there is little local radiology-oriented research is available on the subject to date with most studies being conducted by clinicians using ultrasound as the primary tool of diagnosing fatty liver.[11,12] In fact, ultrasound has been used to determine the prevalence of fatty liver disease.[13] Specifically in the context of steatosis, the intra-observer and inter-observer interpretation of even the histopathology can vary especially for detecting lobular inflammation and ballooning,[14] which affects the reported severity of steatosis.

Ultrasound is a non-invasive modality with 60-100% sensitivity and specificity particularly for advanced stages, more for steatosis than for fibrosis.[10,15-18] It may serve the need of assessing if a liver biopsy is needed, when practiced with caution, objectivity and attention to detail.

The studied ultrasound criteria were pretested for inter-observer agreement and showed good agreement. In a meta-analysis of 49 ultrasound- based studies for hepatic steatosis, the reliability ranged from 0.54 to 0.92 for intra observer and 0.44 -1.00 for inter observer agreement,[10] confirming our pretesting. The intra-rater reliability was not checked in this study as the images in the second round were not the same set as in the pilot study.

Ultrasound is a highly operator dependent technique. The images used in this study were obtained by an experienced radiologist and were read by two other experienced colleagues of nearly the same technical

caliber with knowledge of the purpose of the study and the standardized criteria using the same grading system but not the findings given by the principal researcher. This was done to check the agreement between them. The lowest agreement was for mild steatosis which was either missed or classified as no-steatosis so early/ mild steatosis is likely to be missed even by experienced operators. This is supported by previous researches.[10,15-18]

Hence ultrasound should better be made a part of diagnostic algorithms for best practices. Ultrasound can very well be the first step in such an algorithm. The standardized criteria used for the purpose of this study can be adopted at centers dedicated to hepatobiliary and/or Gastroenterology practice, at the least.

Conclusion

Using standardized ultrasound criteria leads to good intra - observer agreement between qualified experienced Radiologists.

Acknowledgement:

References

1. Galati G, Dell Unto C, De Vincentis A, Gentilucci UV, Gallo P, Guidi A, et al. Hepatocellular Carcinoma in Alcoholic Liver Disease: Current Management and Recent Advances. Rev Recent Clin Trials. 2016; **11:** 238-52.

2. Charatcharoenwithaya P, Lindor KD. Role of radiologic modalities in the management of non-alcoholic steatohepatitis. Clin Liver Dis 2007; **11:** 37-54.

3. Strauss S, Gavish E, Gottlieb P, Katsnelson L. Interobserver and intraobserver variability in the sonographic diagnosis of fatty liver. AJR Am J Roentgenol 2007; **189:** W 320-3.

4. Sohail S, Aziz S, Mirza T. Ultrasound based evaluation of hepatic steatosis and fibrosis in hepatitis C nonresponders.. J Coll Phy Surg Pak 2013; **23:** 548-52.

5. Scatrige JC, Scott WE, Doriovan PJ. Fatty infiltration of liver: ultrasonographic and computed tomographic correlation. J Ultrasound Med 1984; **3:** 9.

6. Parulekar SG, Bree RL. Liver. In: diagnostic ultrasound- a logical approach. McGahan JP, Goldberg BB (edi). Lippincott - Raven, Philadelphia. 1996: 599-6.

7. Almeida AM, Cotrin HP, Barbosa DBV, de Athay de LGM, Santos AS, Bitencourt GV, et al. Fatty liver disease in severely obese patients: diagnostic value of abdominal ultrasound. World J Gastroenterol 2008; **14:** 1415-8

8. Dick R, Watkinson A. The liver and spleen. In: a textbook of radiology and imaging. Sutton D (edi). 7th ed, 2003; Elsevier Churchill Livingstone. Edinburgh: 737-86.

9. Chen CH, Lin ST, Yong CC, Yeh VH, Kou CL, Nien CK. The accuracy of sonography in predicting steatosis and fibrosis in chronic hepatitis C. Dig Dis Sci 2008; **53:** 1699- 706.

10. Hernaez R, Lazo M, Bonekamp S, Kamel I, Brancati FL, Dualler E, et al. Diagnostic accuracy and reliability of ultrasonography for the detection of fatty liver: a meta -analysis. Hepatology 2011; **54:** 1082-90.

11. Taseer IH, Hussain L, Safdar S, Mirbahar AM, Ahmad I. Frequency of nonalcoholic fatty liver disease (NAFLD) and its biochemical derangements in type-2 diabetic patients. Pak J Med Sci 2009; **25:** 817-20.

12. Ansari A, Bilal N. Frequency of non- alcoholic steatohepatitis (NASH) in type 2 Diabetes mellitus. Ann Pak Inst Med Sci 2005; **1:** 227-9.

13. Jimba S, Nakagami T, Takahashi M, Wakamatsu T, Hirota Y, Iwamoto Y, et al. Prevalence of non-alcoholic fatty liver disease and its association with impaired glucose metabolism in Japanese adults. Diabet Med 2005; **8:** 1141-5

14. Pynard T, Rtaziu V, Bedossa P. Appropriateness of liver biopsy. Can J Gastroenterol 2000; **14:** 543-8.

15. Liu P, Li P, He W, Zhao L. Liver and spleen volume variations in patients with hepatic fibrosis. World J Gastroenetrol 2009; **15:** 3298-302.

16. Soresi M, Giannitrapani L, Cerello M, Licata A, Montalto G. Noninvasive tools for the diagnosis of liver cirhosis. World J Gastroenterol 2014; **20:** 1831-50.

17. Machado MV, Cortez- Pinto H. Noninvasive diagnosis of non- alcoholic fatty liver disease. A critical appraisal. J Hepatol 2013; **58:** 1007-19.

18. Zelbur- Sagi S, Webb M, Assy N, Blendis L, Yeshua H, Lashno N, et al. Comparison of fatty liver index with noninvasive methods for steatosis detection and quantification. World J Gastroenterol 2013; **19:** 57-64.