Original research

Feasibility and reproducibility of liver surface nodularity quantification for the assessment of liver cirrhosis using CT and MRI

Grace C. Lo[a,b], Cecilia Besa[a,b], Michael J. King[a,b], Martin Kang[a], Ashley Stueck[c], Swan Thung[c], Mathilde Wagner[a], Andrew D. Smith[d], Bachir Taouli[a,b,*]

[a] Translational and Molecular Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[b] Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[c] Department of Pathology, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[d] Department of Radiology, University of Alabama at Birmingham, Birmingham, AL, USA

ARTICLE INFO

ABSTRACT

Purpose: To assess intra-observer, inter-observer and inter-modality (CT vs. MRI) reproducibility of liver surface nodularity (LSN) scores measured with software used for detection of liver fibrosis.

Methods: This IRB-approved retrospective study included patients with both abdominal CT and MRI within 6 months of histopathologic sampling. Two independent observers used post-processing software to quantify LSN scores on axial non-contrast CT (NCT), axial contrast-enhanced CT (CECT), axial T2-weighted (T2W) HASTE, and axial and coronal post-gadoxetic acid T1-weighted (T1W) images obtained during the hepatobiliary phase (HBP). Ten slices were used to acquire the LSN scores. Intra-observer, inter-observer, and inter-modality (CT vs. MRI) reproducibility were assessed with intraclass correlation coefficient (ICC) and coefficients of variability (CV). Accuracy for detection of cirrhosis was evaluated for each technique.

Results: 26 patients (M/F 19/7, mean age 57 years), including 7 with cirrhosis (26.9%), were assessed. Technical failure occurred with NCT (1/23, 4.3%) and T2 HASTE (8/28, 28.6%). Intra-observer reproducibility was excellent for NCT, CECT, axial and coronal T1W HBP [ICC ≥ 0.92, CV ≤ 8%]. Inter-observer reproducibility was also excellent for NCT and CECT (ICC ≥ 0.95, CV ≤ 7.3%) and for coronal T1W HBP (ICC = 0.84, CV = 5.6%). There was fair to moderate agreement between CT and MRI (ICC 0.20–0.44). There were significant differences in mean LSN scores between non-cirrhotic and cirrhotic patients with NCT (2.6 vs. 4.2, p = 0.04) and T1W HBP (3.7 vs. 4.6; p = 0.01) images, with AUCs of 0.81 and 0.82, respectively.

Conclusions: LSN measurement is highly reproducible with NCT and post-contrast T1W HBP on MRI, with different results obtained between CT and MRI.

1. Introduction

Liver biopsy remains the reference standard for assessing liver fibrosis and cirrhosis, though this method is invasive, is associated with morbidity, is subject to sampling bias and reader variability and has poor patient acceptance [1,2]. Several non-invasive methods have been developed to assist in evaluation for liver fibrosis and cirrhosis, including a wide variety of clinical and laboratory assessments. Imaging has also played a prominent role in evaluating liver fibrosis and cirrhosis, with several methods currently employed in clinical practice ranging from visual assessment of liver morphology to elastography [3–6].

The use of liver morphology, specifically by subjectively evaluating liver surface nodularity (LSN), to stage liver fibrosis has been well studied in ultrasound imaging, with varying levels of diagnostic accuracy for predicting liver fibrosis stage reported in the literature [7–9]. The use of subjective LSN to stage liver fibrosis has been also applied to computed tomography (CT) [10,11] and magnetic resonance imaging (MRI) [6] with varying levels of reported diagnostic accuracy.

A recent study introduced a novel noninvasive method of objectively quantifying liver fibrosis and cirrhosis by applying semi-automated LSN software to CT to evaluate liver morphology and predict liver fibrosis or cirrhosis [12]. With the software, the user highlights a region of interest along the contour of the liver on continuous slices. The difference between the detected liver edge and a smooth polynomial line is measured on a pixel-by-pixel basis and then averaged for

each slice. The final LSN score calculation is the mean of 10 average measurements, with a higher LSN score denoting greater surface nodularity. This method has the potential to possibly decrease the need for invasive tissue sampling or additional imaging tests, as it can be retrospectively applied to previously acquired images or be used to monitor liver disease progression over time. Furthermore, LSN software may be able to help predict cirrhosis compensation and death [13]. While this LSN software has been evaluated in non-contrast CT (NCT) and contrast-enhanced CT (CECT), it has not been applied to MRI, which is another potential modality where this quantitative method may prove helpful in aiding in staging of liver fibrosis [12,14]. Measurement reproducibility and repeatability as well as diagnostic accuracy are crucial criteria in evaluating the performance of a quantitative imaging technique [15].

Thus, the objective of our study was to assess intra-observer, inter-observer and inter-modality (CT vs. MRI) reproducibility of LSN scores measured with software.

2. Methods and materials

This was a single center retrospective HIPAA-compliant study. The institutional review board at the Icahn School of Medicine at Mount Sinai approved this study and waived the requirement for informed consent.

2.1. Subjects

A computerized search of our pathology files and medical records from July 2013 to August 2014 was conducted to identify patients who underwent abdominal MRI, CT, and liver biopsy, all within 6 months. Our inclusion criteria were: a) patients over the age of 18 years, b) patients with liver biopsy within 6 months of both CT and MRI, c) liver MRI performed at 1.5T or 3T and abdominal NCT or CECT, and d) availability of pathology reports for histopathology classification of parenchymal liver disease. Exclusion criteria included patients with a non-contrast MRI exam, patients with MRI or CT performed at outside institutions without available scan parameters, patients without concomitant CT and MRI, or a non-diagnostic liver biopsy. Fig. 1 shows the flowchart for the study. The final study population included 26 patients (M/F 19/7, mean age 57 years, range 22-75 years). Of these 26 patients, 7 (26.9%) patients had no liver disease and underwent abdominal imaging for metastatic disease work-up (19.2%), splenomegaly (3.8%), and liver mass surveillance (3.8%). 19/26 (73.1%) patients had chronic liver disease, including 8 (30.8%) with chronic hepatitis C infection (HCV), 5 (19.2%) with hepatitis B infection (HBV), 1 (3.8%) with HBV

Table 1
Characteristics of study population (n = 26).

Parameters	
Sex (M/F)	19/7
Mean age (range)	57 y (22–75 y)
Chronic liver disease	
Yes	18
No	8
Etiology of chronic liver disease	
HBV	5
HCV	8
HBV and HCV Co-infection	1
Other	5
Fibrosis staging	
Non-cirrhotic	19
Cirrhotic	7

and HCV co-infection, 3 (11.5%) with non-alcoholic steatohepatitis (NASH), 1 (3.8%) with glycogen storage disease, and 1 (3.8%) with congenital erythropoeitic porphyria.

The mean interval between imaging and pathology was 64 days (range 0–181 days), and the mean interval between CT and MRI was 107 days (range 13–183 days). Characteristics of the patient population are detailed in Table 1.

2.2. CT acquisition

Abdominal CT images were obtained using routine parameters from a variety of both GE Medical and Siemens multi-detector scanners. The patients had either a single phase NCT, a single portal venous phase CECT, or a multiphase CECT. Typically, contrast enhancement was administered as a weight-based dose (1.5 mL/kg) of nonionic iodinated contrast (Isovue 370, Bracco Imaging, Westfield, New York). For single-phase studies, intravenous contrast was typically injected at 3 mL per second, with imaging through the abdomen acquired around 80 s. For multiphase CT studies, contrast enhancement was injected at a speed of 4 mL per second. A real-time bolus-triggering technique was used with a region of interest placed over the abdominal aorta with a trigger set at 100 Hounsfield Unit. Axial images through the abdomen and pelvis were obtained at arterial phase (12 s), portal venous phase (42 s) and delayed imaging (110 s). Typical scan parameters are summarized in Table 2.
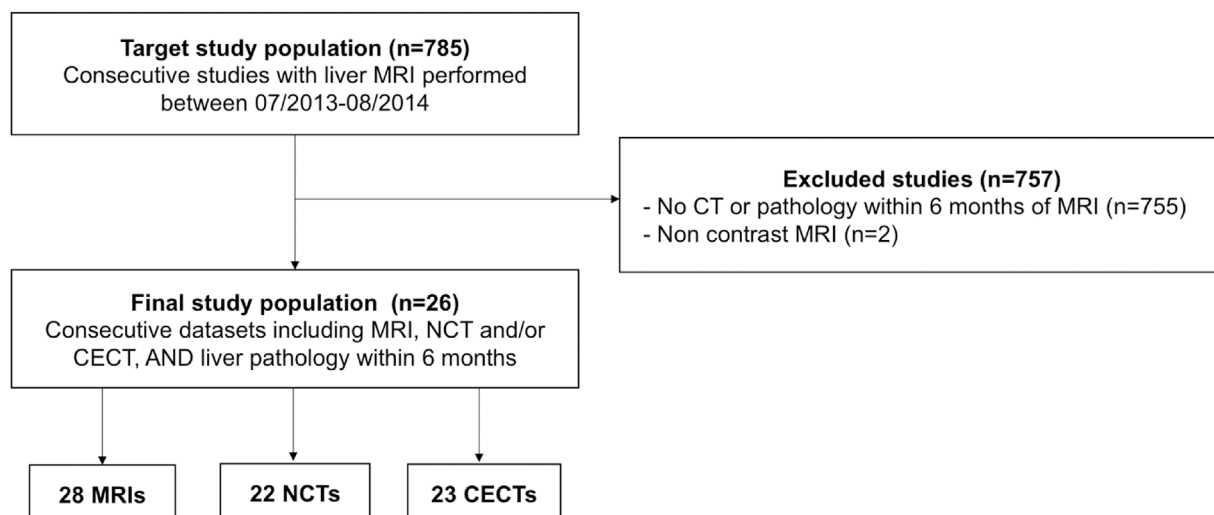

Fig. 1. Flow chart of the study population.

Table 2
CT imaging parameters.

Scan Parameter	Siemens[a] (n = 18)	GE[b] (n = 8)
Kilovolts	100–120	100–120
mAs/mA	Automatic	Automatic
Pitch	0.9–1.6	0.6–1.1
Collimation (mm)	0.6–2.5	0.6–1.15
Field of view	Based on patient size	Based on patient size
Matrix size (pixels)	512 × 512	512 × 512
Section thickness (mm)	2.5–5	2–5
Section interval (mm)	1.25–5	2–5

[a] Siemens scanners included SOMATOM Definition AS+, Sensation Cardiac 64, Definition Flash, Volume Zoom, and Biograph 6.
[b] GE scanners included VCT/Ultra/16/QX/I, Discovery CT750 HD, Lightspeed Ultra, Lightspeed QX/i.

2.3. MRI acquisition

Multichannel systems were used for scanning, using 1.5T (n = 13, Signa HDxt; General Electric Connecticut, USA) or 3T (n = 15, 3T Discovery MR750; General Electric, Connecticut, USA). For all sequences, we used parallel imaging factor 2 with a rectangular field of view of 300 × 400 mm, which was adjusted for patient body size. The MR images were acquired with an 18-channel phased-array receiver coil. Routine liver MRI protocol (Table 3) included non-fat suppressed axial and coronal single-shot fast spin echo T2 (SSFSE), axial fat suppressed fast spin echo (FSE) T2, axial dual-echo chemical shift T1 weighted imaging (T1), axial multi-gradient echo (GRE) T2*, axial diffusion-weighted imaging (DWI), and dynamic contrast-enhanced (CE) T1 acquisition using gadoxetic acid (Eovist; Bayer HealthCare Pharmaceuticals, Berlin, Germany). For gadoxetic acid-enhanced imaging, unenhanced, early and late arterial phases [two sequences back to back in a single breath-hold (BH)], portal venous phase (60 s), transitional phase (180 s), and HBP (at 10 and 20 min) were obtained using a 3D T1 breath-hold fat-suppressed spoiled gradient-recalled echo sequence (VIBE/LAVA) before and after dynamic contrast administration of a fixed dose of 10 mL of gadoxetic acid (mean weight-based dose of 0.025 mmol/kg) injected at a rate of 1.5 mL/sec via a power injector followed by a 20 mL saline flush using a bolus tracking technique. The use of a fixed dose of gadoxetic acid corresponds to our current standard of care clinical practice and is motivated by practicality and costs. Typical scan parameters are summarized in Table 3.

2.4. Liver surface nodularity analysis

Custom semi-automated LSN software was applied by two independent observers (GL, observer 1; MK, observer 2; fourth-year and second-year radiology residents, respectively). The observers were blinded to clinical data. To test intra-observer variability, observer 1, repeated the dataset after a two-week interval while blinded to the

Table 3
MRI acquisition parameters.

	Axial T2 HASTE	Axial T1 HBP	Coronal T1 HBP
TR (ms)	547–1490	3.0–3.5	3.0–3.6
TE (ms)	99–241	1.3–1.5	1.4–1.7
Flip angle (°)	90	12 (1.5T) / 10 (3T)	12 (1.5T) / 10 (3T)
Slice thickness/gap (mm)	5–6 / 1	5 / 0	5 / 0
Matrix	256 × 192 (1.5T)	256 × 160 (1.5T)	256 × 160
	320 × 192 (3T)	320 × 160 (3T)	
Number of averages	1	1	1
Acquisition time	2 breath holds (30 s)	1 breath hold (15 s)	1 breath hold (15 s)

clinical data and prior results. Details of the LSN measurement technique are detailed in a recent paper [12]. In short, the observer highlighted a region of interest along the surface of the liver on continuous slices. The distances between the detected liver edge and a smooth software-generated polynomial line was measured on a pixel-by-pixel basis and then averaged for each slice by the software. The final LSN score calculation was the mean of average distances of ten slices, with a higher LSN score denoting greater surface nodularity. The measurement technique was identical when applied to CT and MRI. On each test date, measurements were made on available axial NCT, axial CECT, axial T2 HASTE, axial T1 HBP, and coronal T1 HBP imaging. On the axial images, measurements were made when achievable along the anterior aspect of the left lobe of the liver if it was bordered by visceral fat or ascites. If the software rejected measurements and measurement attempts had been made on all axial slices, or if the patient was status post left lobe resection, then measurements were made along the anterior and lateral aspects of the right lobe of the liver if it was bordered by visceral fat or ascites. On the coronal images, measurements were made along the lateral aspect of the right lobe, unless the patient was status post right liver resection, in which case measurements were made along the lateral surface of the hypertrophied left lobe. Surfaces of the liver other than those mentioned above, regions where the liver contacts the abdominal wall, natural turns or fissures of the liver, significant image artifact, and the dome of the liver were avoided. Fig. 2 demonstrates examples of LSN measurements.

The images were reviewed again by the same observer after the individual measurements were generated to ensure that no sharp turns were falsely suggested by the imaging software, and any measurements that did not appear to adhere to the liver surface were rejected and did not contribute to the final LSN score. At least ten and at most twelve accepted measurements were made for each image series. A final LSN score was automatically calculated by the software as the mean of the individual slice measurements, with a higher LSN score denoting a higher degree of surface nodularity. Failure was defined as a study for which at least ten measurements could not be made. Processing time and total liver surface sampled was automatically recorded by the software.

2.5. Reference standard

Histopathologic specimens were obtained by transjugular or percutaneous liver biopsy (n = 8, 28.5%), liver resection (n = 13, 46.4%), or liver transplantation (n = 5, 17.9%). The METAVIR semi-quantitative scoring system was used for histopathologic determination of the stage of fibrosis (fibrosis stage: F0–F4) [16]. Fibrosis scores were as follows: F0 (n = 8, 30.7%), F1 (n = 2, 7.6%), F2 (n = 6, 23.1%), F3 (n = 3, 11.5%), and F4 (n = 7, 26.9%). Of the F4 (cirrhotic) patients, 1 (14.3%) patient was Child-Pugh class A, and the remaining 6 (85.7%) were Child-Pugh class B. The median MELD score was 9.7 (range 6.4–13.2).

2.6. Statistical analysis

The coefficient of variation (CV, defined as the standard deviation divided by the mean; expressed as a percentage), intraclass correlation coefficient (ICC), and Bland-Altman limits of agreement were used to assess intra- and inter-observer reproducibility of LSN measurements. ICC agreement was based on Koch and Landis's benchmark for Kappa agreement, with below 0.2 as slight agreement, 0.2–0.4 as fair, 0.41–0.6 as moderate, and 0.61–0.8 as substantial, and 0.81–1 as excellent [17]. A Spearman correlation test was used to assess intermodality (CT and MRI) correlation and inter-sequence correlation within the CT and MRI images. Correlation coefficients were classified with 0–0.19 as very weak, 0.2–0.39 as weak, 0.4–0.59 as moderate, 0.6–0.79 as strong, and 0.8–1.0 as very strong. A Mann-Whitney test was performed to compare the mean distances of liver surface sampled
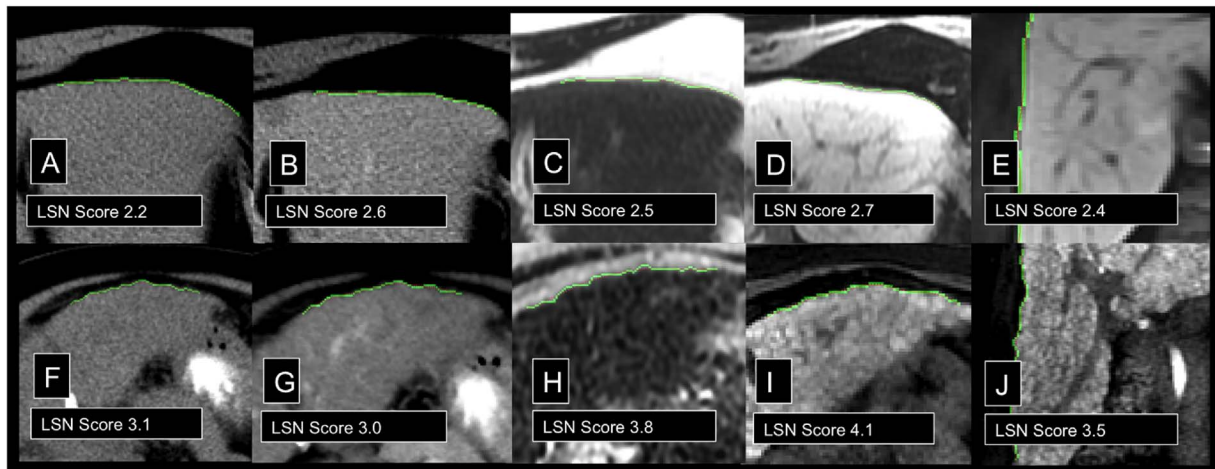
Fig. 2. Top row: 58-year-old patient with non-cirrhotic liver. Bottom row, 45-year-old patient with liver cirrhosis secondary to chronic hepatitis C. Liver surface nodularity drawings (green line) on (A and F) non-contrast CT, (B and G) contrast-enhanced CT, (C and H) axial T2W HASTE, (D and I) axial post-contrast T1W obtained in the hepatobiliary phase (HBP), and (E and J) coronal post-contrast T1W HBP. Higher LSN scores are seen with increased surface nodularity in the cirrhotic liver.

for axial vs. coronal imaging. The diagnostic performance of LSN score for differentiating non-cirrhotic from cirrhotic livers was evaluated by area under the receiver-operating characteristic (AUC) analysis. All data was statistically analyzed using IBM SPSS (IBM Software; New York, USA) and GraphPad Prism (GraphPad Software; California, USA). For all statistical tests, a p value less than 0.05 was considered statistically significant.

3. Results

3.1. LSN measurements

Twenty-six patients with both CT and MRI were analyzed with LSN software, with a total of 22 NCT, 23 CECT, and 28 MRI. For CT measurements, there was one (4.5%) failure for obtaining a LSN score for NCT, and no failures for obtaining a LSN score for CECT. The median time involved in making LSN measurements was 1.3 min (range, 0.5–5.3 min) for NCT and 1.1 min (range 0.6–4.4 min for CECT). The median total distance of liver surface evaluated was 47.8 cm (range 34.7–87.3 cm) for NCT and 68.3 cm (range 38.9–68.3 cm) for CECT.

For MRI measurements, there were 9 (32.1%) failures for obtaining a LSN score for axial T2W HASTE, and no failures for obtaining a LSN score for axial or coronal post-contrast T1W HBP. The median time involved in making LSN measurements was 1.6 min (range, 0.9–8.3 min) for axial T2W HASTE, 1.4 min (range, 0.7–7.4 min) for axial T1W HBP, and 0.5 min (range, 0.4–1.1 min) for coronal T1W HBP. The median of total distance of liver surface evaluated was 56.5 cm (range 34.1–106.4 cm), 77.3 cm (range 40.3–122.9 cm), and 109.4 cm (55.1–176 cm), respectively.

There was a longer mean distance of liver surface evaluated for coronal imaging vs. axial imaging (107.9 cm vs. 64.8 cm, p < 0.001).

3.2. Intra-observer reproducibility

There was excellent agreement using NCT, CECT, axial T1W HBP, and coronal T1W HBP and moderate agreement with axial T2 HASTE sequence (Table 4, Fig. 4).

3.3. Inter-observer reproducibility

There was excellent agreement between observers when using NCT, CECT, and coronal T1W HBP sequence and moderate agreement with axial T2W HASTE and axial T1W HBP sequences (Table 4, Fig. 3).

Table 4
Inter- and intra-observer reproducibility in liver surface nodularity software using non-contrast CT (NCT), contrast-enhanced CT (CECT) and MRI.

	ICC	CV (%)	CR (%)	Mean bias (%)	BALA (%)
Inter-observer					
Axial NCT	0.95	7.3	25.7	3.2	−22.5, 28.9
Axial CECT	0.96	6.2	17.9	6.2	−11.7, 24.1
Axial T2 HASTE	0.48	16.2	56.6	15.7	−41.0, 72.3
Axial T1 HBP	0.57	13.3	43.0	11.4	−31.6, 54.4
Coronal T1 HBP	0.84	5.6	19.5	3.3	−16.2, 22.8
Intra-observer					
Axial NCT	0.98	5.2	18.1	− 0.1	−18.2, 17.9
Axial CECT	0.95	6.9	22.2	3.7	−18.6, 29.9
Axial T2 HASTE	0.44	15.8	62.4	− 1.9	−64.4, 60.5
Axial T1 HBP	0.93	8.0	23.0	6.8	−16.1, 29.8
Coronal T1 HBP	0.92	5.1	18.0	1.8	−16.2, 19.9

ICC: intraclass correlation coefficient; CV: coefficient of variation (%); CR: coefficient of reproducibility (%); BALA: Bland-Altman limits of agreement (%).

3.4. Inter-modality and inter-sequence comparison

There was substantial agreement between NCT and CECT (ICC 0.78), with a moderate positive correlation observed between NCT and CECT (r = 0.58, p < 0.05). There was substantial agreement also between axial T2 HASTE and axial HBP (ICC 0.75) with moderate positive correlation observed (r = 0.5, p < 0.05). There was moderate agreement between axial HBP and coronal HBP (ICC 0.57), though there was no statistically significant correlation observed between these sequences (p = 0.21). There was slight to fair agreement between CT and MRI sequences, as well as between axial T2 HASTE and coronal T1 HBP sequences (ICC 0.20–0.44), with moderate positive correlation between NCT and axial T1 HBP (p = 0.49, < 0.02). No significant correlation was observed between other combinations of CT and MRI sequences (p = 0.41–0.98).

3.5. Preliminary diagnostic accuracy of LSN score for diagnosing cirrhosis

There was a significant difference between F0-3 versus F4 with NCT (2.56 vs 4.21, p = 0.04) and coronal T1W HBP (3.73 vs 4.63; p = 0.01) images, with AUC of 0.81 and 0.82, respectively, for identifying cirrhosis (Table 5). The AUC cutoff for predicting cirrhosis (F4) was 2.86 (sensitivity 83.3%, specificity 76.9%) for NCT and 3.86 (sensitivity 87.5%, specificity 47.8%) for coronal T1W HBP.
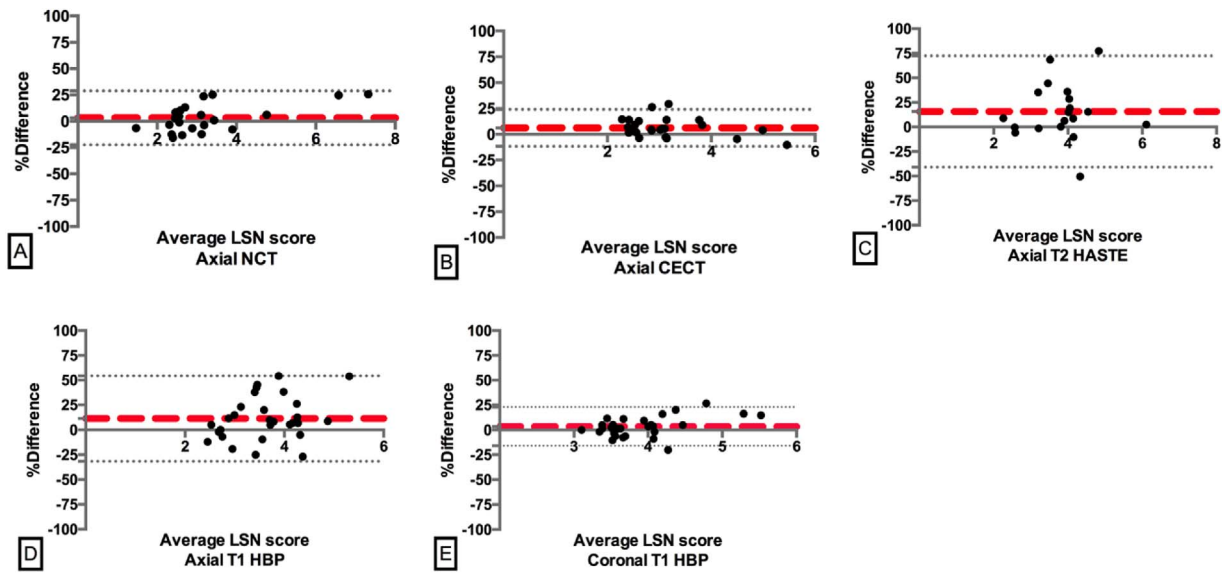
Fig. 3. Bland-Altman plots for inter-observer reproducibility of liver surface nodularity scoring for (A) axial NCT, (B) axial CECT, (C) axial T2W HASTE, (D) axial T1W HBP, and (E) coronal T1W HBP.

4. Discussion

In this study, we compared the precision of a novel semi-automated software for evaluating liver fibrosis stage by measuring LSN with CT vs. MRI. We showed that intra- and inter-observer reproducibility were excellent for NCT, CECT and coronal T1W HBP sequence, as reported by Smith et al. using NCT [12]. In our study, NCT and coronal T1W HBP sequences demonstrated acceptable accuracy in diagnosing liver cirrhosis, though the sample size was small. Both Smith et al. [12] and recently Pickhardt et al. [14] observed excellent diagnostic accuracy of LSN using NCT for predicting cirrhosis (AUC of 0.910 and 0.959, respectively). Our preliminary results also demonstrated a statistically significant, though lower AUC of 0.81. This may be due to our limited sample size (which included only patients with pathology-proven fibrosis staging) and to included populations (Smith et al. included normal subjects and the Pickhardt et al. study focused on patients with chronic end-stage liver disease undergoing transplant evaluation, patients with pre-cirrhotic hepatic fibrosis, and potential kidney donors). Our population included a majority of chronic liver disease patients

Table 5
Receiver-operator curve analysis for diagnosing cirrhosis using liver surface nodularity software.

	AUC	p
Axial NCT	0.81	**0.035**
Axial CECT	0.63	0.33
Axial T2 HASTE	0.63	0.36
Axial T1 HBP	0.5	1
Coronal T1 HBP	0.82	**0.0097**

AUC: area under the receiver-operator curve. Significant p values are in bold.

(73.1%).

Application of the LSN software to MRI has not been previously reported. We found that axial T2W HASTE and axial T1W HBP sequences were neither highly reproducible between readers, nor they were accurate in diagnosing liver cirrhosis. This finding is in contra-distinction to coronal T1W HBP, which was found to be both re-producible and accurate in predicting liver cirrhosis. This discrepancy
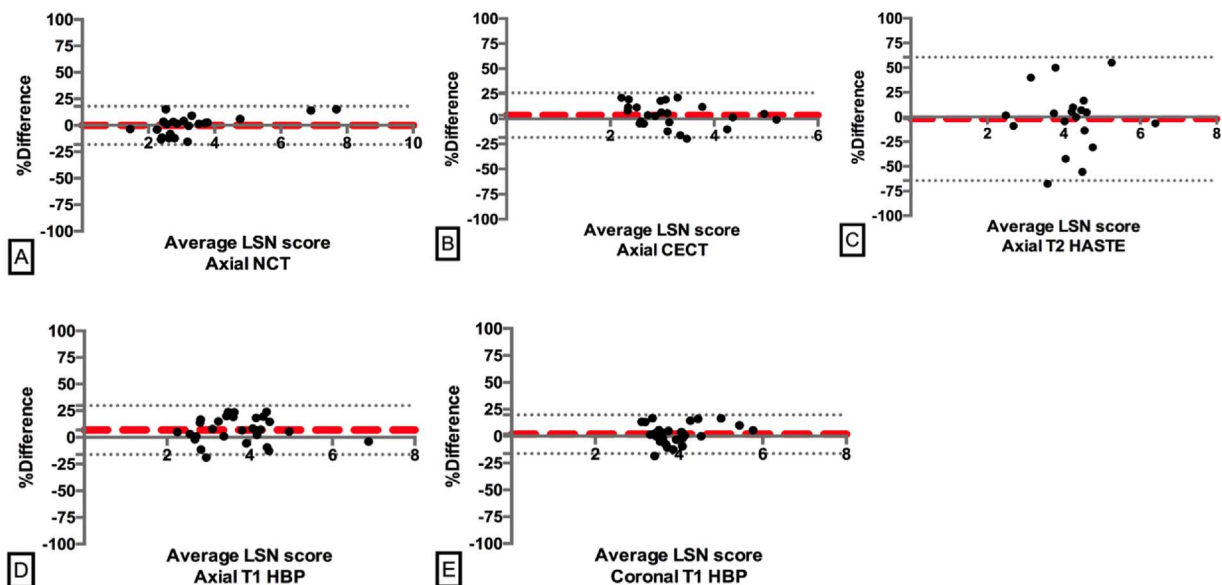

Fig. 4. Bland-Altman plots for intra-observer of liver surface nodularity scoring for (A) axial NCT, (B) axial CECT, (C) axial T2W HASTE, (D) axial T1W HBP, and (E) coronal T1W HBP.

may be explained by the presence of a more well-defined fat plane along the lateral surface of the liver in the right lobe when compared to the anterior surface of the liver on the left lobe, allowing for more reproducible measurements. There may also be more cardiac artifacts affecting measurements on the left lobe. Finally, the total distance sampled for the coronal images was longer than that of the axial images, perhaps providing a more accurate LSN measurement by reducing sampling bias.

LSN software is a promising tool that can be applied to acquired images in less than 5 min. Its applicability on both prospectively and retrospectively available studies gives LSN software the unique ability to both predict liver cirrhosis and to compare a patient's progression over multiple time points, particularly given its high associated repeatability and reproducibility.

There are several limitations to our study. First, the study is retrospective, and the sample size is small, as it reflects our early experience. Second, multiple types of liver disease were grouped together by liver fibrosis staging, some of which may exhibit different patterns or degrees of LSN. Third, we did not assess the accuracy of the software for diagnosing liver fibrosis, given the small sample size. A larger sample size will be necessary to confirm its value in evaluating for liver fibrosis and cirrhosis.

In conclusion, quantitative measurement of LSN is highly reproducible when applied to axial NCT, axial CECT, and coronal T1W HBP. NCT and coronal T1W HBP had acceptable accuracy in diagnosing cirrhosis. This method needs to be compared to other non-invasive techniques such as ultrasound elastography and MR elastography.

Compliance with ethical standards

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent

The Institutional Review Board at our institution approved this single center retrospective study and the need for informed consent was waived.

Conflict of interest

References

[1] D.C. Howlett, et al., Findings of the UK national audit evaluating image-guided or image-assisted liver biopsy. Part II. Minor and major complications and procedure-related mortality, Radiology 266 (1) (2013) 226–235.

[2] F. Piccinino, et al., Complications following percutaneous liver biopsy. A multi-centre retrospective study on 68,276 biopsies, J. Hepatol. 2 (2) (1986) 165–173.

[3] H.A. Dyvorne, et al., Prospective comparison of magnetic resonance imaging to transient elastography and serum markers for liver fibrosis detection, Liver Int. 36 (5) (2016) 659–666.

[4] A. Tang, et al., Ultrasound elastography and MR elastography for assessing liver fibrosis: part 2, diagnostic performance, confounders, and future directions, AJR Am. J. Roentgenol. 205 (1) (2015) 33–40.

[5] S.K. Venkatesh, M. Yin, R.L. Ehman, Magnetic resonance elastography of liver: technique, analysis, and clinical applications, J. Magn. Reson. Imaging 37 (3) (2013) 544–555.

[6] S.K. Venkatesh, et al., Non-invasive detection of liver fibrosis: MR imaging features vs. MR elastography, Abdom. Imaging 40 (4) (2015) 766–775.

[7] V. Simonovský, The diagnosis of cirrhosis by high resolution ultrasound of the liver surface, Br. J. Radiol. 72 (853) (1999) 29–34.

[8] A. Colli, et al., Severe liver fibrosis or cirrhosis: accuracy of US for detection–analysis of 300 cases, Radiology 227 (1) (2003) 89–94.

[9] C.C. Choong, S.K. Venkatesh, E.P. Siew, Accuracy of routine clinical ultrasound for staging of liver fibrosis, J. Clin. Imaging Sci. 2 (2012) 58.

[10] M. Kudo, et al., Diagnostic accuracy of imaging for liver cirrhosis compared to histologically proven liver cirrhosis. A multicenter collaborative study, Intervirology 51 (Suppl. 1) (2008) 17–26.

[11] A. Keedy, et al., Diagnosis of cirrhosis by spiral computed tomography: a case-control study with feature analysis and assessment of interobserver agreement, J. Comput. Assist. Tomogr. 32 (2) (2008) 198–203.

[12] A.D. Smith, et al., Liver surface nodularity quantification from routine CT images as a biomarker for detection and evaluation of cirrhosis, Radiology (2016) 151542.

[13] A.D. Smith, et al., Liver surface nodularity score allows prediction of cirrhosis decompensation and death, Radiology (2016) 160799.

[14] P.J. Pickhardt, et al., Accuracy of liver surface nodularity quantification on MDCT as a noninvasive biomarker for staging hepatic fibrosis, AJR Am. J. Roentgenol. (2016) 1–6.

[15] D.L. Raunig, et al., Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment, Stat. Methods Med. Res. 24 (1) (2015) 27–67.

[16] P. Bedossa, T. Poynard, An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group, Hepatology 24 (2) (1996) 289–293.

[17] J.R. Landis, G.G. Koch, The measurement of observer agreement for categorical data, Biometrics 33 (1) (1977) 159–174.