Regional Analysis of Liver Surface Nodularity in a Single Axial MR Image for Staging Liver Fibrosis

Tae-Hoon Kim, PhD,1 Youe Ree Kim, MD,1,2 Chang-Won Jeong, PhD,1 Hyung Joong Kim, PhD,3 Jin Woong Kim, MD, PhD,4 Young Hwan Lee, MD,1,2 and Kwon-Ha Yoon, MD, PhD1,2*

Background: The assessment of liver surface nodularity (LSN) for staging hepatic fibrosis is restricted in clinical practice because it requires customized software and time-consuming procedures. A simplified method to estimate LSN score may be useful in the clinic.

Purpose: To evaluate the regional analysis of LSN and processing time in a single axial liver MR image for staging liver fibrosis.

Study Type: Retrospective.

Population: A total of 210 subjects, a multicenter study.

Field Strength/Sequence: A 3 T/noncontrast gradient echo T1WI.

Assessment: Subjects were divided into five fibrosis groups (F0 = 29; F1 = 20; F2 = 32; F3 = 50; F4 = 79) based on the METAVIR fibrosis scoring system. The mean LSN (on three slices) and regional LSN (on one slice) measurements, and the processing times, are compared. The regional LSN scores in five regions-of-interests (ROI1-5) were analyzed in a single axial MRI at the level of the hilum by two independent observers.

Statistical Tests: Regional variations in LSN scores were compared using ANOVA with Tukey test. Agreement between the mean and regional LSN measurements was evaluated using Pearson correlation coefficients (r) and Bland–Altman plots. The diagnostic performance of mean and regional LSN scores according to fibrosis stage was evaluated with the AUROC. A P value < 0.05 was considered statistically significant.

Results: Total processing time for a regional LSN measurement (3.6 min) was 75.5% less than that for mean LSN measurement (14.7 min). Mean LSN scores and all five regional LSN scores showed significant differences between fibrosis groups. Among regional LSN scores, ROI5 showed the highest AUROC (0.871 at cut-off 1.12) for discriminating F0-2 vs. F3-4 and the best correlation with mean LSN score (r = 0.800, −0.07 limit of agreement).

Conclusion: Quantitative regional LSN measurement in a single axial MR image reduces processing time. Regional ROI5 LSN score might be useful for clinical decision-making and for distinguishing the difference between early fibrosis (F0-2) and advanced fibrosis (F3-4) in the liver.

Evidence Level: 3

Technical Efficacy: Stage 2

J. MAGN. RESON. IMAGING 2022;56:1781–1791.

Noninvasive evaluation of liver fibrosis is challenging in clinical practice.1 In an initial diagnosis of early stage hepatic fibrosis (compensated liver) or cirrhosis (decompensated liver as end stage), it is difficult to accurately predict hepatic compensation or decompensation using non-invasive methods.2 Therefore, there is an unmet need for...
widely applicable noninvasive methods to diagnose fibrosis and advanced cirrhosis and to predict future risk of hepatic decompensation.

To diagnose fibrosis and cirrhosis of the liver, the importance of assessing liver surface nodularity (LSN) is emerging in the clinical setting.\(^4,5\) The LSN score can be measured from routine liver CT and MR images using post-processing software.\(^3,4-6\) In clinical practice, a quantitative LSN score may be used to diagnose and stage various liver diseases including chronic liver disease (CLD) and non-alcoholic fatty liver disease (NAFLD). Smith et al and Pickhardt et al reported that the quantitative LSN score in CT images progressively increases with higher clinical stages of CLD and/or cirrhosis (end-stage) and is predictive of cirrhosis decompensation and death.\(^4,6\) Recently, Kim et al., who developed a Matlab-based LSN quantification software using multipolynomial curve-fitting analysis,\(^5\) reported that the use of quantitative LSN software can help to differentiate the stage of liver fibrosis in NAFLD (fibrosis grade F0-1 vs. F2-3 based on pathologic NAFLD activity score) and in CLD (fibrosis grades F0-F3 based on serum fibrosis-4 index).\(^3,7\)

The LSN score has high reproducibility on each CT scan,\(^4,6\) and MRI\(^7,8\) study for staging liver fibrosis and end-stage cirrhosis. However, the mean LSN scores reported in MRI studies\(^7,8\) were different from CT studies\(^4,6\) because they used a different software with a different methodology. Moreover, it has restricted use in clinical practice because it requires customized software for quantification and the LSN assessment of the whole liver is a time-consuming procedure. Therefore, the assessment of LSN scores using a simplified method, such as measurement from a single image, may be useful.

Thus, the purpose of this study was to compare the regional LSN scores and processing times derived from a single slice (with mean LSN score derived from the whole liver) for assessing liver fibrosis on a single axial liver MR image and to evaluate the diagnostic performance of LSN score in hepatic fibrosis.

Materials and Methods
The retrospective study was approved by our institutional review board (No. 2018-01-005) and the requirement for written informed consent was waived.

Subject Population
Consecutive patients from January 2010 to June 2020, who were over 20 years of age, who underwent abdominal MRI and who had available pathologic information and serologic tests were retrospectively identified at four medical centers. A total of 210 patients who had undergone 3.0 T liver MRI before or after serum biochemistry and histopathologic investigation were enrolled (n = 85 from Wonkwang University Hospital, n = 82 from Chonbuk National University Hospital, n = 14 from Asan Medical Center and n = 29 from Chonnam National University Hospital). This study excluded patients (n = 9) who had undergone hepatectomy or had a pathologic condition including a large bulging mass that could interfere with measurement of LSN on the anterior or anterolateral surfaces of the liver.\(^5\) Also, we excluded MR images (n = 2) impaired by artifacts that can affect image quality and lead to loss of diagnostic information.\(^9\)

Reference Standard for Liver Fibrosis
All histopathological information was obtained by percutaneous needle biopsy or surgical biopsy. The histological data were analyzed by a pathologist with 9-years’ experience based on the Meta-analysis of Histological Data in Viral Hepatitis (METAVIR) fibrosis scoring system,\(^10,11\) as follows: F0, no fibrosis; F1, portal fibrosis; F2, periportal fibrosis; F3, septal fibrosis; and F4, cirrhosis. The final cohort was divided into five groups (F0-F4) according to hepatic fibrosis stages as follows: F0 (n = 29), F1 (n = 20), F2 (n = 32), F3 (n = 50), and F4 (cirrhosis, n = 79) (Fig. 1). The patients with F0–F2 scores and F3–F4 scores were classified as early fibrosis and advanced fibrosis, respectively.

Acquisition of MR Images
All imaging studies were performed on MR scanners with body matrix coils. Unenhanced fat-suppressed T1-weighted 3D spoiled gradient-echo images in axial plane were acquired with following parameters: TR/TE = 4.2/1.97 msec, field of view = 38 x 38 x 14 cm\(^3\), matrix size = 512 x 512, number of excitations = 2, slice thickness = 2.0 mm, and number of slices = 70.

A total of four 3.0 T MRI scanners were used. Two MR scanners (Achieva [n = 40]; Ingenia CX [n = 45], Philips Healthcare, Best, The Netherlands) were used at Wonkwang University Hospital. At Chonbuk National University Hospital, three MR scanners (Verio [n = 46]; Skyra [n = 13], Siemens healthineers, Erlangen, Germany) (Achieva, Philips Healthcare [n = 23]) were used. At Asan Medical Center, two MR scanners (Skyra, Siemens [n = 13]; Ingenia CX, Philips [n = 1]) were used. At Chonnam National University Hospital, a scanner (Skyra, Siemens [n = 29]) scanner was used.

Preprocessing and Quantification of MR Data for LSN Assessment on Single Slice
The LSN quantification software was developed in Matlab (R2018a; MathWorks, Natick, Massachusetts).\(^3\) The Wonkwang Abdomen and Liver Total Solution (WALTS) is a customized postprocessing program that operates on the Windows platform (ver. XP or higher; Microsoft, Redmond, WA). This was used to measure LSN on the DICOM format images to derive an LSN score with the same procedure as in prior studies.\(^3,7,8\) The main algorithm for assessing the LSN was as follows: bias field correction of images via the level set method, semi-automatic liver boundary detection for drawing the liver reference line, liver segmentation, and LSN measurements with multipolynomial curve-fitting analysis.

LSN Analysis in Clinical Patients
All MR images were reviewed on standard PACS stations with standard window settings. LSN analysis using customized WALTS software was independently performed by two radiologists (abdominal radiologists with 9-years’ [Y.R.K] and 29-years’ [K.H.Y] experience, respectively, in liver imaging) blinded to the radiologic reports of MR datasets. Each radiologist selected three image slices at the level
of the hepatic hilum for quantitative analysis (Fig. 2). Then, bias correction and liver segmentation were performed on the selected slices. LSN is measured on a pixel-by-pixel basis and measures the mean distance for each line (expressed as tenths of a millimeter). The final LSN score (mean LSN score and regional LSN score) is the arithmetic mean of the measurements as following procedures.

To measure mean LSN score as a reference value, the mean LSN score in an individual subject was calculated as an averaged value measured from three slice images at the level of the hilum (vertical lines as shown in Fig. 2a). LSN score on every single slice (Fig. 2b) estimated from five regions-of-interests (ROIs) (Fig. 2d,e).

To measure regional LSN, five ROIs (ROI1-5) on one of three slices were analyzed to measure the regional LSN scores with a fourth-degree polynomial curve-fitting (Fig. 2c). The data points for the selection of ROIs on the segmented liver image are shown in Fig. 2d, e. The datapoints are on the basis of the bottom-left corner origin position (x, y coordinate = 0, 0) in a single axial image (Fig. 2d).

All selected MR images were analyzed using both mean (from three slices) and regional (from one slice) LSN measurements. To determine interobserver agreement, two radiologists were independently measured LSN score on selected images by each method. Details of the LSN measurement technique for clinical patients are described in recent studies.

Statistical Analysis

Differences in mean and regional LSN scores between fibrosis groups (F0–F4) as well as serum biochemistry were analyzed using one-way analysis of variance (ANOVA) with Tukey’s post hoc test using the statistical package for the social sciences program (SPSS ver. 20; Chicago, IL, USA). The coefficient of variance (CV) in each ROI group was calculated to assess the variability of measurements. Pearson correlation between the five regional LSN measurements and the mean LSN scores was performed. Agreement between the LSN measurements of the two methods was evaluated by calculating Pearson correlation coefficients (r) and generating Bland–Altman plots. Pearson correlation and Bland–Altman analysis were performed using MedCalc ver.14.8.1 (Medcalc, Mariakerke, Belgium). Intraclass correlation coefficients (ICC) values were evaluated to assess interobserver agreement in LSN scores and were classified as: poor (<0.40), moderate (0.40 to <0.60), good (0.60 to <0.80), and excellent (0.80–1.00).

To evaluate the diagnostic performance of mean and regional LSN scores according to fibrosis stages, receiver operating characteristics (ROC) curve analysis was performed to calculate the area under the ROC curve (AUROC), sensitivity, specificity, and diagnostic accuracy (DA). Two-sided P values less than 0.05 were considered to indicate statistical significance.
Results

Patient Characteristics
The etiology and serum biochemistry of the enrolled 210 patients (148 men, 62 women; mean age, 56.2 ± 13.2 years) are shown in Table 1. These patients had various liver diseases, including chronic hepatitis B (n = 110), chronic hepatitis C (n = 9), coinfection (n = 4), alcoholic liver disease (n = 18), NAFLD/steatohepatitis (NASH) (n = 18), and autoimmune hepatitis (n = 13). Thirty-eight patients had liver disease of unknown cause. The results of blood chemistry showed significant differences between fibrosis stages in albumin, alanine aminotransferase, aspartate aminotransferase, and platelet count.

LSN Measurements in Different ROIs
Figure 2 shows representative postprocessing procedures (a–e), MR images (f–j), and histopathologic images (k–o). The processing time and LSN scores measured from two observers present an averaged value. The median time for quantifying mean LSN measurement from three image slices was 5.1 minutes (range, 2.1–6.0 minutes). For quantifying a single ROI LSN score, it was 0.4 minutes (range, 0.3–0.6 minutes) (92.2% time reduction). The median time for additional postprocessing (including bias correction, boundary detection, and liver surface line drawing) was 3.2 minutes (range, 2.4–5.4 minutes) per single image. The total processing time was 3.6 minutes (range, 2.7–6.0 minutes) for the single ROI measurement method and 14.7 minutes (range, 9.5–22.2 minutes) for the mean LSN measurement method.

Mean and regional LSN scores in ROIs were significantly different in fibrosis stages (Table 2). The CV value of mean LSN scores was 6.2% (range 5%–8%) and the averaged CV values in each ROI were 15.4% for ROI1, 17.6% for ROI2, 18.4% for ROI3, 13.6% for ROI4, and 13.2% for ROI5 (range 11–24%; average 17.3%). The mean LSN scores in each fibrosis stage were significantly different for all fibrosis stage comparisons. Regional LSN scores were also significantly different for most fibrosis stage comparisons but not for F0 vs. F1 (ROI1 P = 0.493; ROI2 P = 0.082; ROI3 P = 0.959; ROI4 P = 0.996; ROI5 P = 0.516), F2 vs. F3 (ROI1 P = 0.485; ROI2 P = 0.283; ROI3 P = 0.561; ROI4 P = 0.539), and F3 vs. F4 (ROI1 P = 0.072; ROI3 P = 0.144; ROI4 P = 0.112; ROI5 P = 0.205).

Correlations Between Mean and Regional LSN Measurement Methods
Figure 3 shows the Bland–Altman measurement agreement plots between the two measurement methods. Table 3 shows
| Etiology | Total (n = 210) | F0 (n = 29) | F1 (n = 20) | F2 (n = 32) | F3 (n = 50) | F4 (n = 79) | P value*† |
|----------|---------------|-------------|-------------|-------------|-------------|-------------|-----------|
| Age (year) | 56.2 ± 13.2 | 46.8 ± 20.9 | 54.4 ± 16.9 | 57.5 ± 10.9 | 59.7 ± 10.5 | 57.4 ± 9.0 | <0.001bcd† |
| Male: Female | 148:62 | 21:8 | 13:7 | 21:11 | 35:15 | 58:21 | 0.904* |
| Etiology* | | | | | | | |
| Hepatitis B | 110 | 2 | 9 | 14 | 35 | 50 | <0.001* |
| Hepatitis C | 9 | 1 | 0 | 3 | 3 | 2 | |
| Coinfection | 4 | 0 | 0 | 0 | 2 | 2 | |
| Alcohol | 18 | 1 | 1 | 2 | 3 | 11 | |
| NAFLD | 18 | 1 | 6 | 5 | 0 | 6 | |
| Autoimmune | 13 | 0 | 1 | 4 | 4 | 4 | |
| Unknown cause | 38 | 24 | 3 | 4 | 3 | 4 | |
| Serum biochemistry† | | | | | | | |
| Albumin (g/dL) | 4.14 ± 0.03 | 4.35 ± 0.05 | 4.12 ± 0.08 | 4.05 ± 0.06 | 4.27 ± 0.05 | 3.98 ± 0.07 | <0.003dj |
| ALT (IU/L) | 34.6 ± 1.78 | 20.56 ± 1.55 | 47.94 ± 12.54 | 35.73 ± 3.30 | 34.76 ± 3.23 | 36.92 ± 2.26 | <0.011a |
| AST (IU/L) | 41.5 ± 2.65 | 24.26 ± 1.65 | 57.13 ± 21.28 | 44.95 ± 7.69 | 37.31 ± 2.17 | 47.56 ± 2.83 | 0.036† |
| Platelet count (×10³/µL) | 170.9 ± 4.73 | 221.30 ± 17.70 | 175.56 ± 14.48 | 165.04 ± 8.41 | 173.93 ± 6.92 | 143.69 ± 7.12 | <0.001bcd† |

ALP = alkaline phosphatase; ALT = alanine aminotransferase; AST = aspartate aminotransferase; NAFLD = non-alcoholic fatty liver disease.

Etiology data are presented as the number of patients. The value in parenthesis indicates the percentage as the number of patients/total number of patients = 100.

Serum biochemistry data are presented as mean ± SEM. Serum levels in the healthy control group are used as reference ranges.

*The difference among fibrosis groups in etiology data was analyzed by Pearson’s chi-square test.

†The difference among three fibrosis groups was analyzed by one-way ANOVA with Tukey post hoc test as follows: aF0 vs. F1, bF0 vs. F2, cF0 vs. F3, dF0 vs. F4, eF1 vs. F2, fF1 vs. F3, gF1 vs. F4, hF2 vs. F3, iF2 vs. F4, and jF3 vs. F4.
## Table 2. Comparison of LSN Scores in Different Region of Interests According to Fibrosis Stages

| ROI | LSN Score | ROI | LSN Score | ROI | LSN Score | ROI | LSN Score | ROI | LSN Score | Mean LSN |
|-----|-----------|-----|-----------|-----|-----------|-----|-----------|-----|-----------|----------|
| F0  | 0.93 ± 0.14 [15] (1) | ROI1 | 0.93 ± 0.14 [15] (1) | ROI2 | 0.73 ± 0.16 [21] (2) | ROI3 | 0.78 ± 0.14 [18] (4) | ROI4 | 0.80 ± 0.08 [11] (2) | ROI5 | 0.75 ± 0.04 [5] |
| F1  | 1.04 ± 0.12 [12] (1) | ROI1 | 1.04 ± 0.12 [12] (1) | ROI2 | 0.79 ± 0.09 [12] (3) | ROI3 | 0.80 ± 0.10 [13] (4) | ROI4 | 0.89 ± 0.11 [12] (2) | ROI5 | 0.82 ± 0.05 [6] |
| F2  | 1.26 ± 0.17 [14] (1) | ROI1 | 1.26 ± 0.17 [14] (1) | ROI2 | 1.09 ± 0.16 [15] (3) | ROI3 | 1.08 ± 0.19 [18] (4) | ROI4 | 1.14 ± 0.18 [16] (2) | ROI5 | 1.08 ± 0.06 [5] |
| F3  | 1.35 ± 0.27 [20] (1) | ROI1 | 1.35 ± 0.27 [20] (1) | ROI2 | 1.21 ± 0.27 [23] (3) | ROI3 | 1.14 ± 0.22 [19] (4) | ROI4 | 1.26 ± 0.18 [14] (2) | ROI5 | 1.17 ± 0.10 [8] |
| F4  | 1.44 ± 0.22 [16] (1) | ROI1 | 1.44 ± 0.22 [16] (1) | ROI2 | 1.30 ± 0.27 [21] (3) | ROI3 | 1.23 ± 0.22 [18] (4) | ROI4 | 1.33 ± 0.18 [13] (2) | ROI5 | 1.26 ± 0.09 [7] |

*The difference among fibrosis groups within each ROI was analyzed by the one-way ANOVA with Tukey's post hoc test.*

LSN = liver surface nodularity; ROI = region of interest.

Data presented mean ± SD (coefficient of variation [CV]). The values in parenthesis indicated the ranking of LSN score in each fibrosis stage.
the associations between the mean and each ROI\textsubscript{n} LSN score. The systemic bias between mean and ROI\textsubscript{n} LSN scores was $-0.18$ (limits of agreement [LOA] of $-0.55$ and $0.19$) for ROI\textsubscript{1}, $0.26$ (LOA of $-0.01$ and $0.53$) for ROI\textsubscript{2}, $-0.02$ (LOA of $-0.44$ and $0.40$) for ROI\textsubscript{3}, $0.01$ (LOA of $-0.35$ and $0.37$) for ROI\textsubscript{4}, and $-0.07$ (LOA of $-0.37$ and $0.23$) for ROI\textsubscript{5}, respectively. There were significant correlations between mean and regional LSN measurements using the

**FIGURE 3:** Bland–Altman liver surface nodularity (LSN) measurement agreement. Bland–Altman LSN measurement agreement plots between mean LSN and ROI measurement methods: (a) mean LSN vs. ROI\textsubscript{1} LSN measurement; (b) mean LSN vs. ROI\textsubscript{2} LSN measurement; (c) mean LSN vs. ROI\textsubscript{3} LSN measurement; (d) mean LSN vs. ROI\textsubscript{4} LSN measurement; and (e) mean LSN vs. ROI\textsubscript{5} LSN measurement. The mean error is shown as the central dashed line, and the 95% limits of agreement (±1.96 standard deviation [SD]) are shown as the peripheral dashed lines. Note that x- and y-axis scales different from each other.
TABLE 3. Correlation Between Mean LSN Score and Each ROIi, LSN

| ROI  | Pearson’s coefficient (r) | 0.740** | 0.807** | 0.732** | 0.715** | 0.800** |
|------|--------------------------|---------|---------|---------|---------|---------|
| P value* | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

LSN = liver surface nodularity; ROI = region of interest.
The correlation coefficient was analyzed by the Pearson’s test as following significance:*P < 0.05 and **P < 0.01.

Pearson correlation coefficient (r = 0.740 for ROI1, r = 0.807 for ROI2, r = 0.732 for ROI3, r = 0.715 for ROI4, and r = 0.800 for ROI5). Among them, the regional ROI5 LSN score (r = 0.800) with −0.07 LOA had the best correlation with the mean LSN score.

**ROC Analysis for Differential Diagnosis According to Fibrosis Stages**
The AUROC of LSN scores for the differentiation of fibrosis stages and their DA is summarized in Table 4. Mean LSN scores showed excellent diagnostic performance with the highest sensitivities and specificities (AUROC 0.994 [95% confidence interval 0.986–1.00 for F0-F1 vs. F2-F4; 0.946 [0.920–0.973] for F0-F2 vs. F3-F4; 0.884 [0.842–0.926] for F0-F3 vs. F4). The AUROC5 of regional LSN scores in differentiating early fibrosis (F0-F2) from advanced fibrosis (F3-F4) were 0.820 (0.764–0.876) for ROI0 LSN; 0.869 (0.823–0.915) for ROI2 LSN; 0.863 (0.810–0.916) for ROI1 LSN; 0.817 (0.760–0.874) for ROI4 LSN; and 0.871 (0.822–0.920) for ROI5 LSN. ROI5 LSN score showed the highest diagnostic performance for discriminating F0-F2 from F3-F4. DA (% of number of subject) for discriminating between early fibrosis and advanced fibrosis (F0,2 vs. F3,4) was 72.4% (152/210) for ROI1, 78.6% (165/210) for ROI2, 82.9% (174/210) for ROI3, 73.3% (154/210) for ROI4, and 81.9% (172/210) for ROI5. For mean LSN, the DA was 89%.

**Interobserver Variability**
The interobserver agreement of LSN scores from the two readers is shown in Table 5. All ICCs for interobserver agreement were higher than 0.8, indicating good reliability. According to fibrosis stages, the ICCs were 0.844 for LSN measurements in F0 stage, 0.801 for F1, 0.821 for F2, 0.807 for F3 and 0.804 for F4. For ROIs, the ICCs were 0.986 for ROI1, 0.972 for ROI2, 0.989 for ROI3, 0.988 for ROI4, and 0.990 for ROI5.

**Discussion**
This study demonstrated regional LSN analysis on a single axial MR image. The processing time using the WALTS program for a single ROI LSN measurement was reduced by 75.5% compared with the mean LSN measurement. The regional ROI5 LSN score demonstrated the strongest correlation with the mean LSN score and the highest DA in the differentiation between early fibrosis (≤F2) and advanced fibrosis (>F2). In the mean LSN measurement method, LSN scores showed excellent diagnostic performance in the quantification of LSN on routine liver MR scans as a part of a morphologic assessment. In prior studies, Smith et al17 and Pickhardt et al6 reported the excellent DA of mean LSN score using MRI and CT images in predicting fibrosis or cirrhosis (AUROC of 0.910 and 0.959, respectively). Our findings from mean LSN scores are well matched with previous studies.5,6 Our results imply that both mean and regional LSN score from a single MRI image have potential as a diagnostic tool for staging hepatic fibrosis.

The reproducibility, repeatability, and DA are important in evaluating the performance of quantitative analysis.17 This study investigated the variability in LSN measurements, with CV values for mean LSN measurements being <10% and for single ROIi LSN measurements being <20%. Although the variability of the single regional measurement method was approximately double that of the mean measurement method, the processing time for the single ROI measurement method was reduced by 75%. The single ROI measurement method for LSN score might therefore be helpful for clinical application. In this study, the variability in all LSN measurements was reproducible as previous MR studies.7,8 For evaluating hepatic fibrosis, MR elastography (MRE) for assessing liver stiffness provided strong correlation with histopathologic fibrosis stage (r = 0.74) and high diagnostic performances for detection of stages F2–F4, F3–F4, and F4 (AUROCs of 0.87, 0.91, and 0.89, respectively) compared to other methods.18–20 Also, MRE maps showed fair to good accuracy for detection of fibrosis (AUROC range 0.76–0.84).18 MRI-based fibrosis markers negatively correlated with histologic stage: contrast enhancement index (CEI) (r = −0.786); liver-spleen contrast ratio (LSC) (r = −0.718); liver-portal vein contrast ratio (LPC) (r = −0.448); and signal intensity ratio (SIR) (r = −0.617).21 For diagnosis of either significant liver fibrosis (≥F2) or cirrhosis (F4), the CEI provided better diagnostic accuracy (AUROC 0.898 and 0.881) than the aspartate aminotransferase-to-platelet ratio index (APRI) (AUROC 0.699 and 0.715). The CEI showed
TABLE 4. Receiver Operator Curve (ROC) Analysis for Diagnosing Fibrosis Stage Using Regional LSN Scores

| Comparison                | Threshold Value | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) | DA (%) | AUROC   | P value |
|---------------------------|-----------------|-----------------|-----------------|---------|---------|--------|---------|---------|
| Mean LSN score            |                 |                 |                 |         |         |        |         |         |
| $F_{0.1}$ ($n = 49$) vs. $F_{2.4}$ ($n = 161$) | 0.93            | 98.8 (159/161)  | 98.0 (48/49)    | 99.4 (159/160) | 96.0 (48/50) | 98.6 (207/210) | 0.999   | <0.001  |
| $F_{0.2}$ ($n = 81$) vs. $F_{3.4}$ ($n = 129$) | 1.12            | 89.1 (115/129)  | 88.9 (72/81)    | 92.7 (115/124) | 83.7 (72/86) | 89.0 (187/210) | 0.967   | <0.001  |
| $F_{0.3}$ ($n = 131$) vs. $F_{4}$ ($n = 79$)   | 1.18            | 83.5 (66/79)    | 80.2 (105/131)  | 71.7 (66/92)  | 89.0 (105/118) | 81.4 (171/210) | 0.900   | <0.001  |
| ROI1 LSN score            |                 |                 |                 |         |         |        |         |         |
| $F_{0.1}$ ($n = 49$) vs. $F_{2.4}$ ($n = 161$) | 1.13            | 84.5 (136/161)  | 85.7 (42/49)    | 95.1 (136/143) | 62.7 (42/67) | 84.8 (178/210) | 0.942   | <0.001  |
| $F_{0.2}$ ($n = 81$) vs. $F_{3.4}$ ($n = 129$) | 1.24            | 72.1 (93/129)   | 72.8 (59/81)    | 80.9 (93/115) | 62.1 (59/95) | 72.4 (152/210) | 0.837   | <0.001  |
| $F_{0.3}$ ($n = 131$) vs. $F_{4}$ ($n = 79$)   | 1.31            | 69.6 (55/79)    | 68.7 (90/131)   | 57.3 (55/96)  | 78.9 (90/114) | 69.0 (145/210) | 0.774   | <0.001  |
| ROI2 LSN score            |                 |                 |                 |         |         |        |         |         |
| $F_{0.1}$ ($n = 49$) vs. $F_{2.4}$ ($n = 161$) | 0.67            | 90.7 (146/161)  | 89.8 (44/49)    | 96.7 (146/151) | 74.6 (44/59) | 90.5 (190/210) | 0.970   | <0.001  |
| $F_{0.2}$ ($n = 81$) vs. $F_{3.4}$ ($n = 129$) | 0.83            | 78.3 (101/129)  | 79.0 (64/81)    | 85.6 (101/118) | 69.6 (64/92) | 78.6 (165/210) | 0.883   | <0.001  |
| $F_{0.3}$ ($n = 131$) vs. $F_{4}$ ($n = 79$)   | 0.92            | 75.9 (60/79)    | 77.9 (102/131)  | 67.4 (60/89)  | 84.3 (102/121) | 77.1 (162/210) | 0.822   | <0.001  |
| ROI3 LSN score            |                 |                 |                 |         |         |        |         |         |
| $F_{0.1}$ ($n = 49$) vs. $F_{2.4}$ ($n = 161$) | 0.94            | 93.8 (151/161)  | 93.9 (46/49)    | 98.1 (151/154) | 82.1 (46/56) | 93.8 (197/210) | 0.978   | <0.001  |
| $F_{0.2}$ ($n = 81$) vs. $F_{3.4}$ ($n = 129$) | 1.09            | 82.9 (107/129)  | 82.7 (67/81)    | 88.4 (94/116) | 75.3 (67/89) | 82.9 (174/210) | 0.886   | <0.001  |
| $F_{0.3}$ ($n = 131$) vs. $F_{4}$ ($n = 79$)   | 1.16            | 69.6 (55/79)    | 68.7 (90/131)   | 57.3 (55/96)  | 78.9 (90/114) | 69.0 (145/210) | 0.792   | <0.001  |
| ROI4 LSN score            |                 |                 |                 |         |         |        |         |         |
| $F_{0.1}$ ($n = 49$) vs. $F_{2.4}$ ($n = 161$) | 0.93            | 86.3 (139/161)  | 85.7 (42/49)    | 95.2 (139/146) | 65.6 (42/64) | 86.2 (181/210) | 0.935   | <0.001  |
| $F_{0.2}$ ($n = 81$) vs. $F_{3.4}$ ($n = 129$) | 1.06            | 72.9 (94/129)   | 74.1 (60/81)    | 81.7 (94/115) | 63.2 (60/95) | 73.3 (154/210) | 0.834   | <0.001  |
| $F_{0.3}$ ($n = 131$) vs. $F_{4}$ ($n = 79$)   | 1.13            | 69.6 (55/79)    | 70.2 (92/131)   | 58.5 (55/94)  | 79.3 (92/116) | 70.0 (147/210) | 0.763   | <0.001  |
| ROI5 LSN score            |                 |                 |                 |         |         |        |         |         |
| $F_{0.1}$ ($n = 49$) vs. $F_{2.4}$ ($n = 161$) | 1.01            | 94.4 (152/161)  | 93.9 (46/49)    | 98.1 (152/155) | 83.6 (46/55) | 94.3 (198/210) | 0.977   | <0.001  |
| $F_{0.2}$ ($n = 81$) vs. $F_{3.4}$ ($n = 129$) | 1.13            | 82.2 (106/129)  | 81.5 (66/81)    | 87.6 (106/121) | 74.2 (66/89) | 81.9 (172/210) | 0.889   | <0.001  |
| $F_{0.3}$ ($n = 131$) vs. $F_{4}$ ($n = 79$)   | 1.24            | 73.4 (58/79)    | 72.5 (95/131)   | 61.7 (58/94)  | 81.9 (95/116) | 72.9 (153/210) | 0.791   | <0.001  |

Data in parentheses are raw data used to calculate percentages. AUROC = area under the receiver operator curve; DA = diagnostic accuracy = (TP + TN)/(TP + FP + TN + FN); F = fibrosis stages; FN = false negative; FP = false positive; LSN = liver surface nodularity; NPV = negative predictive value; PPV = positive predictive value; TN = true negative; TP = true positive.
Unlike MRE, 18, 21 the WALTS software-based LSN measurement may provide a reproducible way to quantify liver fibrosis using routine MR images in clinical practice. The methods of liver stiffness measurement, such as ultrasound elastography and MRE, also require specific devices and additional data acquisition. 18, 21 The WALTS software-based LSN measurement may provide a reproducible way to quantify fibrosis using routine MR imaging in clinical practice. Several studies suggest that early stage liver fibrosis may be reversible. 25–28 Therefore, an accurate and easy-implemented noninvasive diagnosis method to differentiate early and advanced liver fibrosis is important for therapeutic strategies. In this study, a single ROI5 LSN measurement was useful for rapidly discriminating between early (F0–F2) and advanced fibrosis groups (F3–F4).

### Limitations

First, this study was retrospective and used a nonrandomized design; therefore, enrolled patients had heterogeneous diseases including chronic hepatitis B, chronic hepatitis C, alcoholic liver disease, NALFD/NASH, and autoimmune hepatitis. Our study population predominantly included patients with chronic hepatitis B (54%) and 18% had unknown cause. Moreover, because most cases were hepatitis B viral infection, macronodular liver cirrhosis would be predominant in the liver cirrhosis (F4) group. Therefore, this finding might potentially lead to false-positive or true-negative findings due to the effect of dominant hepatitis B, and thus, the diagnostic performance might be overestimated. To overcome this shortcoming, a large-scale multicenter study in a specific disease such as chronic hepatitis B and NAFLD should be conducted in the future. Second, there was no consideration for the effects of cross-MRI (magnetic field strength and/or different vendors), cross-institutional protocol, or cross-modality reproducibility of LSN score. Thus, further study is necessary to clarify this. Finally, we focused on a regional analysis on a METAVIR classified axial-plane MR image in the LSN measurement. We excluded patients who had undergone hepatectomy or had a pathologic condition that may affect the anterior or anterolateral surface. The axial image plane may be substituted by LSN measurement on another image-plane (coronal or sagittal plane image) and/or another surface.

### Conclusions

This study showed regional differences in LSN scores in the liver. A single ROI measurement to evaluate the LSN score is

---

**TABLE 5. Interobserver Variability in LSN Measurements**

| Fibrosis stage | Reader A | Reader B | P value* | Intrarater Reliability (ICC)† | 95% CI Lower Bound | 95% CI Upper Bound | P value* |
|----------------|----------|----------|----------|-----------------------------|--------------------|--------------------|----------|
| F0 (n = 29)    | 0.75 ± 0.04 [5] | 0.74 ± 0.04 [6] | 0.510 | 0.844 | 0.668 | 0.927 | <0.001 |
| F1 (n = 20)    | 0.83 ± 0.05 [7] | 0.82 ± 0.06 [7] | 0.445 | 0.801 | 0.498 | 0.921 | <0.001 |
| F2 (n = 32)    | 1.08 ± 0.06 [6] | 1.07 ± 0.06 [5] | 0.366 | 0.821 | 0.633 | 0.912 | <0.001 |
| F3 (n = 50)    | 1.17 ± 0.10 [9] | 1.18 ± 0.11 [9] | 0.743 | 0.807 | 0.659 | 0.890 | <0.001 |
| F4 (n = 79)    | 1.26 ± 0.10 [8] | 1.26 ± 0.09 [7] | 0.418 | 0.804 | 0.694 | 0.875 | <0.001 |

ROIs

| ROIs          | Reader A | Reader B | P value* | Intrarater Reliability (ICC)† | 95% CI Lower Bound | 95% CI Upper Bound | P value* |
|---------------|----------|----------|----------|-----------------------------|--------------------|--------------------|----------|
| ROI1 (n = 210) | 1.28 ± 0.29 [23] | 1.28 ± 0.27 [21] | 0.324 | 0.986 | 0.981 | 0.989 | <0.001 |
| ROI2 (n = 210) | 0.84 ± 0.22 [26] | 0.83 ± 0.25 [30] | 0.280 | 0.972 | 0.963 | 0.979 | <0.001 |
| ROI3 (n = 210) | 1.12 ± 0.32 [29] | 1.12 ± 0.31 [28] | 0.628 | 0.989 | 0.986 | 0.992 | <0.001 |
| ROI4 (n = 210) | 1.08 ± 0.26 [24] | 1.09 ± 0.26 [24] | 0.413 | 0.988 | 0.985 | 0.991 | <0.001 |
| ROI5 (n = 210) | 1.17 ± 0.26 [22] | 1.17 ± 0.25 [21] | 0.147 | 0.990 | 0.987 | 0.993 | <0.001 |

ICC = intraclass correlation coefficient; CI = confident interval.
LSN scores of each reader in three groups are presented as means ± SD (mean coefficient of variance, %).
*The LSN score differences of two readers were analyzed by the paired t-test.
†The intrarater reliability between two readers was analyzed by intraclass correlation (ICC) test.
an efficient way of saving processing time in clinical practice. The LSN score in the posterior region of the right lobe (ROIs) may help differentiate fibrosis stages rapidly in clinical practice. Further studies are necessary to validate whether an LSN score is a simple alternative to the mean LSN measurement score for the diagnosis of hepatic fibrosis in various liver diseases and to investigate the external validation of LSN measurement technique.

Acknowledgments
This study was supported by the grants of the National Research Foundation of Korea (NRF) (2021R1A5A8029876, 2020R1I1A1A01073871) and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare (HI18C1216). The authors thank professor Keum Ha Choi in the Department of Pathology for her help in collecting the pathologic data and photographs.

Conflict of Interest
The authors have declared that no conflict of interest exists.

Authors Contributions
T.H.K., Y.R.K., K.H.Y.: substantial contributions to the conception and design; T.H.K., Y.R.K., C.W.J., H.J.K., J.W.K., Y.H.L.: the acquisition, analysis, interpretation of the data; T.H.K., Y.R.K., C.W.J., H.J.K., J.W.K., Y.H.L., K.H.Y.: the drafting of the article or critical revision for important intellectual content; T.H.K., Y.R.K., C.W.J., H.J.K., J.W.K., Y.H.L., K.H.Y.: substantial contributions to the conception and design; T.H.K., Y.R.K., C.W.J., H.J.K., J.W.K., Y.H.L., K.H.Y.: final approval of the version to be published; T.H.K., K.H.Y.: agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.

REFERENCES
1. Gheorghe G, Bungau S, Ceobanu G, et al. The non-invasive assessment of hepatic fibrosis. J Formos Med Assoc 2021;120(2):794-803.
2. Martinez SM, Crespo G, Navasa M, Forns X. Noninvasive assessment of liver fibrosis. Hepatology 2011;53(1):325-335.
3. Kim TH, Kim JE, Ryu JH, Jeong CW. Development of liver surface nodularity quantification program and its clinical application in non-alcoholic fatty liver disease. Sci Rep 2019;9(1):9994.
4. Smith AD, Zand KA, Florez E, et al. Liver surface nodularity score allows prediction of cirrhosis decompensation and death. Radiology 2017;283(3):711-722.
5. Smith AD, Branch CR, Zand K, et al. Liver surface nodularity quantification from routine CT images as a biomarker for detection and evaluation of cirrhosis. Radiology 2016;280(3):771-781.
6. Pickhardt PJ, Malecki K, Kloge J, Lubner MG. Accuracy of liver surface nodularity quantification on MDCT as a noninvasive biomarker for staging hepatic fibrosis. AJR Am J Roentgenol 2016;207(6):1194-1199.
7. Kim T-H, Jeong C-W, Kim JE, et al. Assessment of liver fibrosis stage using integrative analysis of hepatic heterogeneity and nodularity in routine MRI with FIB-4 index as reference standard. J Clin Med 2021;10(8):1697.
8. Kim SW, Kim YR, Choi KH, et al. Staging of liver fibrosis by means of semiautomatic measurement of liver surface nodularity in MRI. AJR Am J Roentgenol 2020;215(3):624-630.
9. Budrys T, Veikutis V, Lukosevicius S, Gleizniene R, Monastyreckiene E, Kulakiene I. Artifacts in magnetic resonance imaging: How can it really affect diagnostic image quality and confuse clinical diagnosis? J Vibroengineering 2018;20(2):1202-1213.
10. Bedossa P, Patel K, Castera L. Histologic and noninvasive estimates of liver fibrosis. Clin Liver Dis 2015;6(1):5-8.
11. Goodman ZD. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. J Hepato 2007;47(4):598-607.
12. Sung P, Lee JM, Joo I, Lee S, Kim TH, Ganeshan B. Evaluation of the impact of iterative reconstruction algorithms on computed tomography texture features of the liver parenchyma using the filtration-histogram method. Korean J Radiol 2019;20(4):558-568.
13. Basso L, Baldi D, Mannelli L, Cavaliere C, Salvatore M, Brancato V. Investigating dual-energy CT post-contrast phases for liver iron quantification: A preliminary study. Dose Response 2021;19(2):15593258211011359.
14. Kim TH, Jun HY, Kim KJ, et al. Hepatic alanine deaminates non-alcoholic Steatohepatitis from simple steatosis in humans and mice. A proton MR spectroscopy study with long Echo time. J Magn Reson Imaging 2017;46(5):1298-1310.
15. Giavarina D. Understanding bBlind Altman analysis. Biochem Med 2015;25(2):141-151.
16. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Clin Exp Med 2016;15(2):155-163.
17. Raunig DL, McShane LM, Pennello G, et al. Quantitative imaging biomarkers: A review of statistical methods for technical performance assessment. Stat Methods Med Res 2015;24(1):27-47.
18. Besa C, Wagner M, Lo G, et al. Detection of liver fibrosis using qualitative and quantitative MR elastography compared to liver surface nodularity measurement, gadoxetic acid uptake, and serum markers. J Magn Reson Imaging 2018;47(6):1552-1561.
19. Liang Y, Li D. Magnetic resonance elastography in staging liver fibrosis in non-alcoholic fatty liver disease. A pooled analysis of the diagnostic accuracy. BMC Gastroenterol 2020;20(1):89.
20. Hsu C, Caussy C, Imaio K, et al. Magnetic resonance vs transient elastography analysis of patients with nonalcoholic fatty liver disease: A systematic review and pooled analysis of individual participants. Clin Gastroenterol Hepatol 2019;17(4):630-7.e8.
21. Jang HU, Min JH, Lee JE, Shin KS, Kim KH, Choi SY. Assessment of liver fibrosis with gadoxetic acid-enhanced MRI: Comparisons with transient elastography, ElastPQ, and serologic fibrosis markers. Abdom Radiol 2019;44(8):2769-2780.
22. Keller S, Aigner A, Zenouzi R, et al. Association of gadolinium-enhanced magnetic resonance imaging with hepatic fibrosis and inflammation in primary sclerosing cholangitis. PLoS One 2018;13(3):e0193929.
23. Petitclerc L, Sebastiani G, Gilbert G, Cloutier G, Tang A. Liver fibrosis: Review of current imaging and MRI quantification techniques. J Magn Reson Imaging 2017;45(5):1276-1295.
24. Verloh N, Utpatel K, Haimerl M, et al. Liver fibrosis and Gd-EOB-DTPA-enhanced MRI: A histopathologic correlation. Sci Rep 2015;5:15408.
25. Hernandez-Gea V. Liver fibrosis: What is reversible and what not? How to assess regression? In: Franchis R, editor. Portal hypertension VI. Switzerland: Springer International Publishing; 2016. p 111-115.
26. Campana L, Iredale JP. Regression of liver fibrosis. Semin Liver Dis 2017;37(1):1-10.
27. Sun M, Kisseleva T. Reversibility of liver fibrosis. Clin Res Hepatol Gastroenterol 2015;39:560-563.
28. Hernandez-Gea V, Friedman SL. Pathogenesis of liver fibrosis. Annu Rev Pathol 2011;6:425-456.