



# CT indices for the diagnosis of hepatic steatosis using non-enhanced CT images: development and validation of diagnostic cut-off values in a large cohort with pathological reference standard

Jieun Byun<sup>1</sup> · Seung Soo Lee<sup>1</sup> · Yu Sub Sung<sup>1</sup> · Youngbin Shin<sup>1</sup> · Jessica Yun<sup>1</sup> · Ho Sung Kim<sup>1</sup> · Eun sil Yu<sup>2</sup> · Sung-Gyu Lee<sup>3</sup> · Moon-gyu Lee<sup>1</sup>

Received: 17 August 2018 / Revised: 26 October 2018 / Accepted: 22 November 2018 / Published online: 19 December 2018

© European Society of Radiology 2018

## Abstract

**Objectives** To compare the performances of CT indices for diagnosing hepatic steatosis (HS) and to determine and validate the CT index cut-off values.

**Methods** Three indices were measured on non-enhanced CT images of 4413 living liver donor candidates (2939 men, 1474 women; mean age, 31.4 years): hepatic attenuation ( $CT_L$ ), hepatic attenuation minus splenic attenuation ( $CT_{L-S}$ ), and hepatic attenuation divided by splenic attenuation ( $CT_{L/S}$ ). The performances of these CT indices in diagnosing HS, relative to pathologic diagnosis, were compared in the development cohort of 3312 subjects by receiver operating characteristic (ROC) analysis. The cut-off values for diagnosing HS > 33% in the development cohort were determined at 95% specificity and 95% sensitivity using bootstrap ROC analysis, and the diagnostic performance of these cut-off values was validated in the test cohort of 1101 subjects.

**Results**  $CT_{L-S}$  showed the highest performance for diagnosing HS  $\geq 5\%$  and HS > 33% (areas under the curve (AUCs) = 0.737 and 0.926, respectively), followed by  $CT_{L/S}$  (AUCs = 0.732 and 0.925, respectively) and  $CT_L$  (AUCs = 0.707 and 0.880, respectively). For CT scans using 120 kVp, the  $CT_{L-S}$  cut-off values for highly specific (i.e., -2.1) and highly sensitive (i.e., 7.6) diagnosis of HS > 33% resulted in a specificity of 96.4% with a sensitivity of 64.0% and a sensitivity of 97.3% with a specificity of 54.9%, respectively, in the test cohort.

**Conclusion** CT indices using liver and spleen attenuations have higher performance for diagnosing HS than indices using liver attenuation alone. The  $CT_{L-S}$  cut-off values in this study may have utility for diagnosing HS in clinical practice and research.

## Key Points

- CT indices based on both liver attenuation and spleen attenuation ( $CT_{L-S}$  and  $CT_{L/S}$ ) have higher diagnostic performance than  $CT_L$  based on liver attenuation alone in diagnosing HS using various CT techniques.
- The CT index cut-off values determined in this study can be utilized for reliable diagnosis or to rule out subjects with moderate to severe HS in clinical practice and research, including the selection of living liver donors and the development of cohorts with HS or healthy controls.

**Keywords** Fatty liver · Non-alcoholic fatty liver disease · Tomography, X-ray computed

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00330-018-5905-1>) contains supplementary material, which is available to authorized users.

---

✉ Seung Soo Lee  
seungsoolee@amc.seoul.kr

<sup>1</sup> Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, South Korea

<sup>2</sup> Department of Diagnostic Pathology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, South Korea

<sup>3</sup> Division of Hepatobiliary Surgery and Liver Transplantation, Department of Surgery, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, South Korea

## Abbreviations

AUROC	Areas under the receiver operating characteristic curve
CI	Confidence interval
CT	Computed tomography
CT <sub>L</sub>	Hepatic attenuation
CT <sub>L-S</sub>	Hepatic attenuation minus splenic attenuation
CT <sub>L/S</sub>	Hepatic attenuation divided by splenic attenuation
HS	Hepatic steatosis
NAFLD	Non-alcoholic fatty liver disease
ROC	Receiver operating characteristic
ROI	Region of interest

## Introduction

The diagnosis of hepatic steatosis (HS) is an important clinical issue, as HS is a component of non-alcoholic fatty liver disease (NAFLD), a condition that may progress to liver cirrhosis and hepatocellular carcinoma [1–4]. HS is associated with metabolic syndrome [5, 6], and HS in donor livers can affect the prognosis of liver transplantation recipients [7]. Due to its high prevalence and asymptomatic presentation, however, HS is often overlooked in healthy controls selected for clinical trials, an omission that may affect the validity of study findings [8]. Because of the invasiveness of liver biopsy, imaging methods are non-invasive alternatives for assessing HS. In particular, MR is currently considered the most accurate imaging technique for the diagnosis of HS [9, 10].

Computed tomography (CT) can also detect HS. Frequently used indices include hepatic attenuation (CT<sub>L</sub>), hepatic attenuation minus splenic attenuation (CT<sub>L-S</sub>), and hepatic attenuation divided by splenic attenuation (CT<sub>L/S</sub>) [9–20]. Although CT is not accurate in assessing mild HS [10], it may allow for highly specific diagnosis of moderate to severe HS [16, 18]. Since CT is widely available and frequently performed in clinical practice for various clinical indications, pre-existing retrospective CT data may be used to develop a large patient cohort with HS. Therefore, CT has been used to detect moderate to severe HS in living liver donor candidates [10, 20] and to identify subjects with clinically significant HS in cohort studies and in clinical trials evaluating the prevalence, natural history, prognosis, and treatment of NAFLD [11, 21–26].

However, until now, neither the accuracy of CT indices nor the cut-off values have been clearly determined. Recent studies have reported normal reference ranges for CT<sub>L-S</sub> and CT<sub>L</sub> cut-off values, resulting in a specific diagnosis of moderate to severe HS [17, 18]; however, they included relatively small samples and lacked external validation, limiting their generalization. Given a limited overall accuracy of CT [10], CT would not allow for the diagnosis of HS with both high sensitivity and specificity. Thus, for CT to be used for ruling-in or ruling-out HS, cut-off values of CT indices that result in

highly specific or highly sensitive diagnoses of HS would be more clinically relevant than CT indices that diagnose HS through a balance of sensitivity and specificity.

The rationale of this study was to evaluate the diagnostic performance of three CT indices in a large cohort of adults with biopsy-proven HS, as well as determining and validating cut-off values of these CT indices for the highly sensitive or highly specific diagnosis of HS.

## Materials and methods

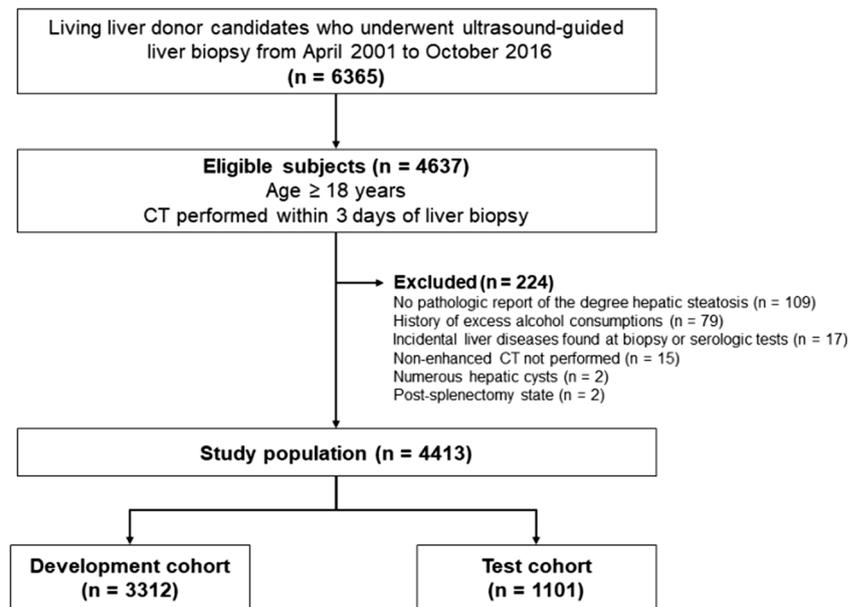
This study was approved by the institutional review board of our institution, which waived the requirement for informed consent due to the retrospective nature of this study.

### Study population

The medical records of the 6365 living liver donor candidates who underwent donor work-up at our institution between April 2001 and October 2016 were reviewed retrospectively to identify subjects aged  $\geq 18$  years who underwent ultrasound-guided percutaneous liver biopsy, along with CT within 3 days of liver biopsy. Of the 4637 eligible subjects, 224 were excluded (109 because their pathology reports did not include the degree of HS, 79 because of a history of excess alcohol consumption, 17 because liver disease was incidentally detected on biopsy or serologic tests, and 19 because of conditions that precluded measurement of CT indices, including 15 subjects with unavailable non-enhanced CT images, two with numerous hepatic cysts, and two who had undergone splenectomy). The study population consisted of 4413 subjects (2939 men and 1474 women; mean age,  $31.4 \pm 9.4$  years). The study population was randomly divided 3:1 into development ( $n = 3312$ ) and test ( $n = 1101$ ) cohorts. The flow diagram for the study population is shown in Fig. 1.

### CT protocol and image analysis

Because the CT data in this study were collected over a long period, various CT techniques were used. CT scans were performed using four-channel (LightSpeed Qx/I; GE Medical Systems;  $n = 410$ ), 16-channel (LightSpeed 16; GE Medical Systems or Sensation 16; Siemens;  $n = 3093$ ), 64-channel (Definition AS; Siemens;  $n = 846$ ), and 128-channel (Definition Flash; Siemens;  $n = 64$ ) scanners. Non-enhanced CT images were obtained at beam collimations of  $4 \times 2.5$  mm (LightSpeed Qx/I),  $8 \times 2.5$  mm (LightSpeed 16),  $16 \times 1.5$  mm (Sensation 16),  $24 \times 1.2$  mm (Definition AS), and  $64 \times 0.6$  mm (Definition Flash); a spiral pitch of 1 to 1.5; tube voltages of 120 kVp ( $n = 3621$ ) and 100 kVp ( $n = 792$ ); and tube currents of 200 mAs (GE scanners) or variable mAs (Siemens scanners) with an automatic exposure control (Care Dose 4D; Siemens; a

**Fig. 1** Flow diagram of the study population

maximum effective dose of 200 mAs). Axial images were reconstructed at section thicknesses of 3 mm ( $n = 81$ ), 5 mm ( $n = 4207$ ), and 10 mm ( $n = 125$ ), with no gaps.

### Image analysis

Liver and spleen attenuation values on non-enhanced CT images were measured by one of two radiology technicians using in-house software built as a plug-in to ImageJ (National Institutes of Health, Bethesda). Liver attenuation was measured by averaging the Hounsfield units (HU) of eight 1.5-cm<sup>2</sup> circular regions of interest (ROIs) placed in the right hepatic lobe. Splenic attenuation was obtained by averaging HU values of three 1.5-cm<sup>2</sup> circular ROIs placed in the upper, middle, and lower thirds of the spleen. The CT images with the ROIs were screen captured and re-evaluated by an abdominal imaging fellow (B.J.) to reconfirm the adequacy of ROI locations. Detailed methods of image analysis are presented in Figure E1 and the Appendix. The three CT indices were calculated from the attenuation values of the liver ( $CT_L$ ) and spleen ( $CT_S$ ), with  $CT_L =$  the attenuation value of the liver,  $CT_{L-S} = CT_L - CT_S$ , and  $CT_{L/S} = CT_L/CT_S$ .

### Reference standard

The pathologic degree of HS was defined as the reference standard. Briefly, all subjects underwent US-guided percutaneous liver biopsy using an 18-gauge needle (Stericut 18G coaxial; TSK Laboratory), with at least two biopsy specimens measuring approximately 1.5 cm in length each obtained from different sites in the central portion of the right hepatic lobe, approximately between hepatic segments V, VI, VII, and VIII. The biopsy specimens were stained with hematoxylin and eosin, and the degree of HS was visually assessed as the percentage of liver

parenchyma replaced by macrovesicular steatotic droplets. Based on the percentage of macrovesicular steatosis in pathologic reports, HS was retrospectively graded as none (< 5%), mild (5–33%), moderate (34–66%), or severe (> 66%) according to the non-alcoholic steatohepatitis Clinical Research Network scoring system [27] by one radiologist (S.S.L.).

### Statistical analysis

Continuous variables in the development and test cohorts were compared using the Mann-Whitney  $U$  test because of their non-normal distribution, and categorical variables were compared using chi-square tests. The ability of the three CT indices ( $CT_L$ ,  $CT_{L-S}$ , and  $CT_{L/S}$ ) to diagnose all degrees of HS ( $HS \geq 5\%$ ) and moderate to severe HS ( $HS > 33\%$ ) was evaluated in the entire study population by receiver operating characteristic (ROC) analysis, with areas under the ROC curves (AUROC) compared using Delong's method [28]. The correlation between CT indices and the pathologic degree of HS was evaluated by linear regression analysis, with correlation coefficients compared using the  $z$ -test. The confounding effects of CT techniques on the CT indices were assessed by multivariable linear regression analysis, which included the pathologic degree of HS, CT vendors, CT tube voltage, and CT slice thickness as independent variables.

In our study, two cut-off values were selected for each CT index: one for reliable, highly specific diagnosis of HS and the other for reliable elimination of most subjects with HS. Therefore, CT index cut-off values for diagnosing  $HS \geq 5\%$  and  $> 33\%$  with 95% specificity or 95% sensitivity were determined, along with 95% confidence intervals (CIs), by ROC analysis with 1000-fold bootstrapping in the development cohort. These cut-off values were used to calculate the

sensitivities, specificities, and accuracies in the development and test cohorts. To estimate the possibility of HS at a given  $CT_{L-S}$  value, the proportions of subjects with  $HS \geq 5\%$  and  $HS > 33\%$  and their 95% CIs were assessed in the development cohort according to 12  $CT_{L-S}$  ranges (ten ranges evenly divided from a  $CT_{L-S}$  value of  $-15$  to a  $CT_{L-S}$  value of  $15$ ,  $CT_{L-S} \geq 15$ , and  $CT_{L-S} < -15$ ) and compared with the proportions of subjects with HS observed in the test cohort. The Bonferroni correction (i.e., multiplying the  $p$  value by the number of comparisons) was utilized for multiple comparisons. All statistical analyses were performed using SPSS software (version 21.0; IBM Corp.) and R studio software (version 1.1.383; R Foundation for Statistical Computing), with  $p$  values less than 0.05 considered statistically significant.

## Results

### Baseline characteristics of subjects in the development and test cohorts

The characteristics of the study population are summarized in Table 1. Of the 4413 subjects, 1916 (43.4%) had clinically relevant HS (i.e.,  $HS \geq 5\%$ ), with 1534 (34.8%), 335 (7.6%),

and 47 (1.1%) subjects having mild, moderate, and severe HS, respectively. These percentages were similar in the development and test cohorts, and there were no significant differences in any clinical characteristic between the development and test cohorts ( $p \geq 0.07$ ).

### Comparison of CT indices in the development cohort

Among the three CT indices,  $CT_{L-S}$  showed the highest performance for the diagnosis of HS (Fig. 2, Table E1). The AUROC values of  $CT_L$ ,  $CT_{L-S}$ , and  $CT_{L/S}$  were 0.707, 0.737, and 0.732, respectively, for diagnosing  $HS \geq 5\%$  and 0.880, 0.926, and 0.925, respectively, for diagnosing  $HS > 33\%$ . The AUROC values for  $CT_{L-S}$  were significantly higher than those for  $CT_L$  in diagnosing  $HS \geq 5\%$  ( $p < 0.001$ ) and  $HS > 33\%$  ( $p < 0.001$ ) and for  $CT_{L/S}$  in diagnosing  $HS \geq 5\%$  ( $p < 0.001$ ).  $CT_{L/S}$  also showed significantly higher AUROC values than  $CT_L$  for diagnosing  $HS \geq 5\%$  ( $p < 0.002$ ) and  $HS > 33\%$  ( $p < 0.001$ ).  $CT_{L-S}$  ( $r = 0.58$ ) showed the strongest correlation with the pathologic degree of HS, followed by  $CT_{L/S}$  ( $r = 0.559$ ) and  $CT_L$  ( $r = 0.526$ ), with the difference between  $CT_{L-S}$  and  $CT_L$  being statistically significant ( $p < 0.001$ ).

$CT_{L-S}$  and  $CT_{L/S}$  appeared to be less influenced by CT techniques than  $CT_L$ . Multivariable linear regression analyses showed

**Table 1** Characteristics of the study population

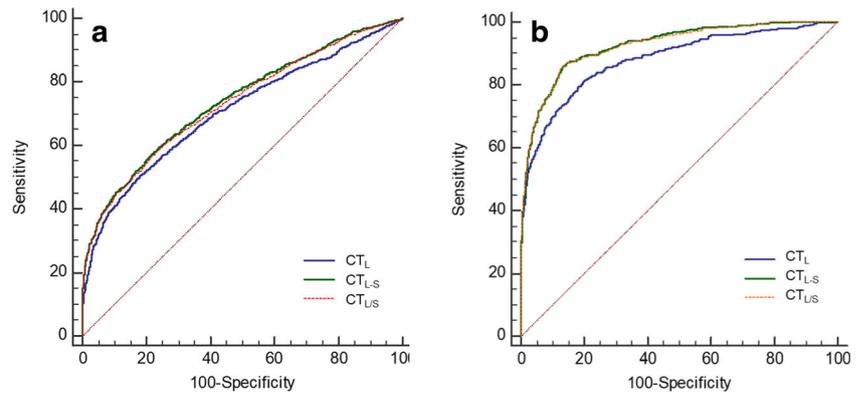
Characteristics	Total	Development cohort	Test cohort	$p$ value
No. of subjects	4413	3312	1101	
Age (years)	31.4 ± 9.4 (18, 62)	31.6 ± 9.3 (18, 62)	31.1 ± 9.5 (18, 60)	0.07 <sup>a</sup>
Sex (male)	2939 (66.6%)	2206 (66.6%)	733 (66.6%)	0.98 <sup>b</sup>
Hepatic steatosis				1.00 <sup>b</sup>
No	2497 (56.6%)	1873 (56.6%)	624 (56.7%)	
Mild	1534 (34.8%)	1151 (34.8%)	383 (34.8%)	
Moderate	335 (7.6%)	252 (7.6%)	83 (7.5%)	
Severe	47 (1.1%)	36 (1.1%)	11 (1.0%)	
CT vendor				0.36 <sup>b</sup>
Siemens	2873 (65.1%)	2153 (65.0%)	720 (65.4%)	
GE	1540 (34.9%)	1159 (35.0%)	381 (34.6%)	
CT tube voltage				0.17 <sup>b</sup>
120 kVp	3621 (82.1%)	2733 (82.5%)	888 (80.7%)	
100 kVp	792 (17.9%)	579 (17.5%)	213 (19.3%)	
CT slice thickness				0.56 <sup>b</sup>
3 mm	81 (1.8%)	58 (1.8%)	23 (2.1%)	
5 mm	4207 (95.3%)	3164 (95.5%)	1043 (94.7%)	
10 mm	125 (2.8%)	90 (2.7%)	35 (3.2%)	
CT index values				
$CT_L$	59.0 ± 7.7 (−1.8, 91.5)	59.0 ± 7.8 (−1.8, 91.5)	59.0 ± 7.7 (17.5, 80.5)	0.79 <sup>a</sup>
$CT_{L-S}$	6.6 ± 7.4 (−53.1, 26.7)	6.6 ± 7.5 (−53.1, 26.7)	6.8 ± 7.2 (−34.8, 25.7)	0.41 <sup>a</sup>
$CT_{L/S}$	1.1 ± 0.1 (−0.04, 1.7)	1.1 ± 0.2 (−0.04, 1.7)	1.1 ± 0.1 (0.3, 1.6)	0.37 <sup>a</sup>

Results are reported as mean ± standard deviation (range) or the number (%) of subjects

<sup>a</sup>  $p$  value was obtained using the Mann-Whitney  $U$  test

<sup>b</sup>  $p$  value was obtained using the chi-square test

**Fig. 2** Receiver operating characteristics curves showing the performance of  $CT_L$ ,  $CT_{L-S}$ , and  $CT_{L/S}$  in diagnosing hepatic steatosis  $\geq 5\%$  (a) and hepatic steatosis  $> 33\%$  (b)



that  $CT_L$  was significantly influenced by CT vendor ( $p < 0.001$ ), tube voltage ( $p < 0.001$ ), slice thickness ( $p < 0.001$ ), and the pathologic degree of HS ( $p < 0.001$ ). By contrast,  $CT_{L-S}$  and  $CT_{L/S}$  were affected only by tube voltage ( $p = 0.001$  and  $p = 0.003$ , respectively) and the pathologic degree of HS ( $p < 0.001$  for both) but not by CT vendor ( $p = 0.159$  and  $p = 0.206$ , respectively) or slice thickness ( $p = 0.11$  and  $p = 0.329$ , respectively).

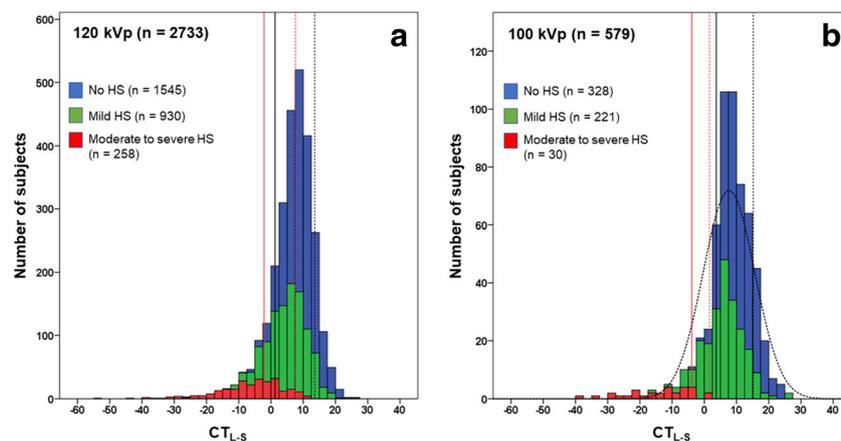
Because of its superior diagnostic accuracy and relative robustness to different CT techniques than the other CT indices,  $CT_{L-S}$  was selected as the main CT index for further analysis in this study.

**$CT_{L-S}$  cut-off values and corresponding diagnostic performances in the development cohort**

Figure 3 shows the histograms of  $CT_{L-S}$  values in the development cohort according to the pathologic grades of HS, presented separately for subjects with CT performed using 120 kVp ( $n = 2733$ ) and 100 kVp ( $n = 579$ ) tube voltages. At both tube voltages,  $CT_{L-S}$  values overlapped considerably for different pathologic grades of HS, especially between no HS and mild HS. The mean  $CT_{L-S}$  values in subjects with no HS

were significantly higher for 100 kVp ( $10.6 \pm 4.9$ ) than those for 120 kVp ( $9.0 \pm 4.6$ ,  $p < 0.001$ ).

Table 2 summarizes the  $CT_{L-S}$  cut-off values for diagnosing HS with 95% specificity (i.e., cut-off values for reliable diagnosis of HS) and with 95% sensitivity (i.e., cut-off values for reliable elimination of HS), and their corresponding diagnostic performances in the development cohort. Because CT tube voltage had a significant effect on  $CT_{L-S}$  values,  $CT_{L-S}$  cut-off values were determined separately for 120 kVp and 100 kVp. The  $CT_{L-S}$  criteria for the diagnosis of HS  $\geq 5\%$  with 95% specificity,  $\leq 1.3$  for 120 kVp and  $\leq 3.7$  for 100 kVp, had sensitivities of 33.9% and 38.7%, respectively. The  $CT_{L-S}$  criteria for the diagnosis of HS  $\geq 5\%$  with 95% sensitivity,  $\leq 13.6$  for 120 kVp and  $\leq 15.2$  for 100 kVp, resulted in specificities of 15.7% for 120 kVp and 19.2% for 100 kVp; this means that the use of  $CT_{L-S}$  criteria of  $> 13.6$  for 120 kVp and  $> 15.2$  for 100 kVp can rule out 95% of subjects with HS  $\geq 5\%$  while excluding 84.3% or 80.8% (i.e., 1-specificity) of non-steatotic subjects. The  $CT_{L-S}$  criteria for the diagnosis of HS  $> 33\%$  with 95% specificity,  $\leq -2.1$  for 120 kVp and  $\leq -3.9$  for 100 kVp, had sensitivities of 64.0% and 90.0%, respectively, whereas the criteria for the diagnosis of HS  $> 33\%$  with 95% sensitivity,  $\leq 7.6$  for 120 kVp and  $\leq 1.6$  for



**Fig. 3** Histogram of  $CT_{L-S}$  values according to the pathologic grades of HS in subjects in the development cohort who underwent CT at tube voltages of 120 kVp (a) and 100 kVp (b). Vertical reference lines indicate  $CT_{L-S}$  cut-off values for diagnosing hepatic steatosis  $\geq 5\%$  with

95% specificity (solid black line) and 95% sensitivity (dashed black line) and for diagnosing hepatic steatosis  $> 33\%$  with 95% specificity (solid red line) and 95% sensitivity (dashed red line)

**Table 2** CT<sub>L-S</sub> cut-off values and their ability to diagnose hepatic steatosis in the development cohort

Diagnostic settings	Subjects who underwent CT at 120 kVp ( <i>n</i> = 2733)				Subjects who underwent CT at 100 kVp ( <i>n</i> = 579)			
	CT <sub>L-S</sub> cut-off values*	Sensitivity	Specificity	Accuracy	CT <sub>L-S</sub> cut-off values*	Sensitivity	Specificity	Accuracy
Diagnosis of HS ≥ 5%								
With 95% specificity	1.3 (0.8, 2.1)	33.9% (403/1188)	95.0% (1469/1545)	69.0% (1872/2733)	3.7 (2.7, 4.2)	38.7% (97/251)	95.1% (312/328)	70.6% (409/579)
With 95% sensitivity	13.6 (12.9, 14.0)	95.0% (1129/1188)	15.7% (242/1545)	50.2% (1371/2733)	15.2 (14.1, 16.6)	95.2% (239/251)	19.2% (63/328)	52.2% (307/579)
Diagnosis of HS > 33%								
With 95% specificity	-2.1 (-2.9, -1.2)	64.0% (165/258)	95.0% (2358/2475)	92.0% (2523/2733)	-3.9 (-4.4, 0.8)	90.0% (27/30)	97.6% (536/549)	97.2% (563/579)
With 95% sensitivity	7.6 (6.0, 9.0)	95.3% (246/258)	54.3% (1344/2475)	58.2% (1590/2733)	1.6 (-4.4, 1.8)	96.7% (29/30)	91.1% (500/549)	91.3% (529/579)

Unless otherwise specified, data are percentages, with the number of subjects used to calculate the percentage in parentheses. Percentages were rounded  
HS hepatic steatosis

\*Data are CT<sub>L-S</sub> cut-off values, and data in parentheses are 95% confidence intervals. Dual CT<sub>L-S</sub> cut-off values were determined for either a highly specific diagnosis of HS (with 95% specificity) or a reliable elimination of HS (with 95% sensitivity)

100 kVp, had specificities of 54.3% and 91.1%, respectively. For comparison, the cut-off values for CT<sub>L</sub> and CT<sub>L/S</sub> are shown in Tables E2 and E3, respectively.

### Diagnostic performances of CT<sub>L-S</sub> cut-off values in the test cohort

Table 3 summarizes the results of validation of the CT<sub>L-S</sub> cut-off values in the test cohort. At a tube voltage of 120 kVp (*n* = 888), the CT<sub>L-S</sub> cut-off values for diagnosing HS ≥ 5% and HS > 33% in the development cohort performed similarly in the test cohort, with the intended 95% sensitivity or 95% specificity achieved in all diagnostic tasks except for diagnosing HS ≥ 5% with 95% sensitivity (i.e., the actual sensitivity was 93.6%). However, the validation results at a tube voltage of 100 kVp were less consistent in the test cohort (*n* = 213) than in the development cohort, with the intended 95% sensitivity or 95% specificity not achieved in diagnosing HS ≥ 5% with 95% specificity (i.e., the actual specificity was 92%) and in diagnosing HS > 33% with 95% sensitivity (i.e., the actual sensitivity was 89.5%).

### Probability of hepatic steatosis according to CT<sub>L-S</sub> value

Figure 4 depicts the plots of the estimated possibilities of HS ≥ 5% and HS > 33% as a function of CT<sub>L-S</sub> in the development cohort and the proportion of the subjects with HS in the test cohort (data presented in Tables E4 and E5). At a tube voltage of 120 kVp, the observed proportions of subjects with HS ≥ 5% and HS > 33% in the test cohort were within the 95% CIs of the estimated possibilities of HS for 11 of the 12 CT<sub>L-S</sub> ranges. However, the results obtained at a tube voltage of 100 kVp were less reliable than those for 120 kVp, with the proportions of subjects with HS ≥ 5% and HS > 33% in the test cohort being outside the 95% CIs of the estimated possibilities of HS in two of the 12 CT<sub>L-S</sub> ranges.

### Discussion

This study demonstrated that CT indices combining liver attenuation and spleen attenuation (CT<sub>L-S</sub> and CT<sub>L/S</sub>) were superior to liver attenuation alone in diagnosing HS and were less affected by technical settings. A previous study of 88 patients, reporting that CT<sub>L</sub> was more strongly correlated with the pathologic degree of HS than either CT<sub>L-S</sub> or CT<sub>L/S</sub> [29], was conducted on a single CT machine, whereas our study used many different devices. The relative robustness of CT<sub>L-S</sub> and CT<sub>L/S</sub> to different CT techniques compared with CT<sub>L</sub>, as shown in our results, may have resulted in better performance of CT<sub>L-S</sub> and CT<sub>L/S</sub> than CT<sub>L</sub> in our study and may explain the discrepant results between our study and the previous study. In our study, CT<sub>L-S</sub> had significantly higher performance for

**Table 3** Diagnostic performance of CT<sub>L-S</sub> to diagnose hepatic steatosis in the test cohort

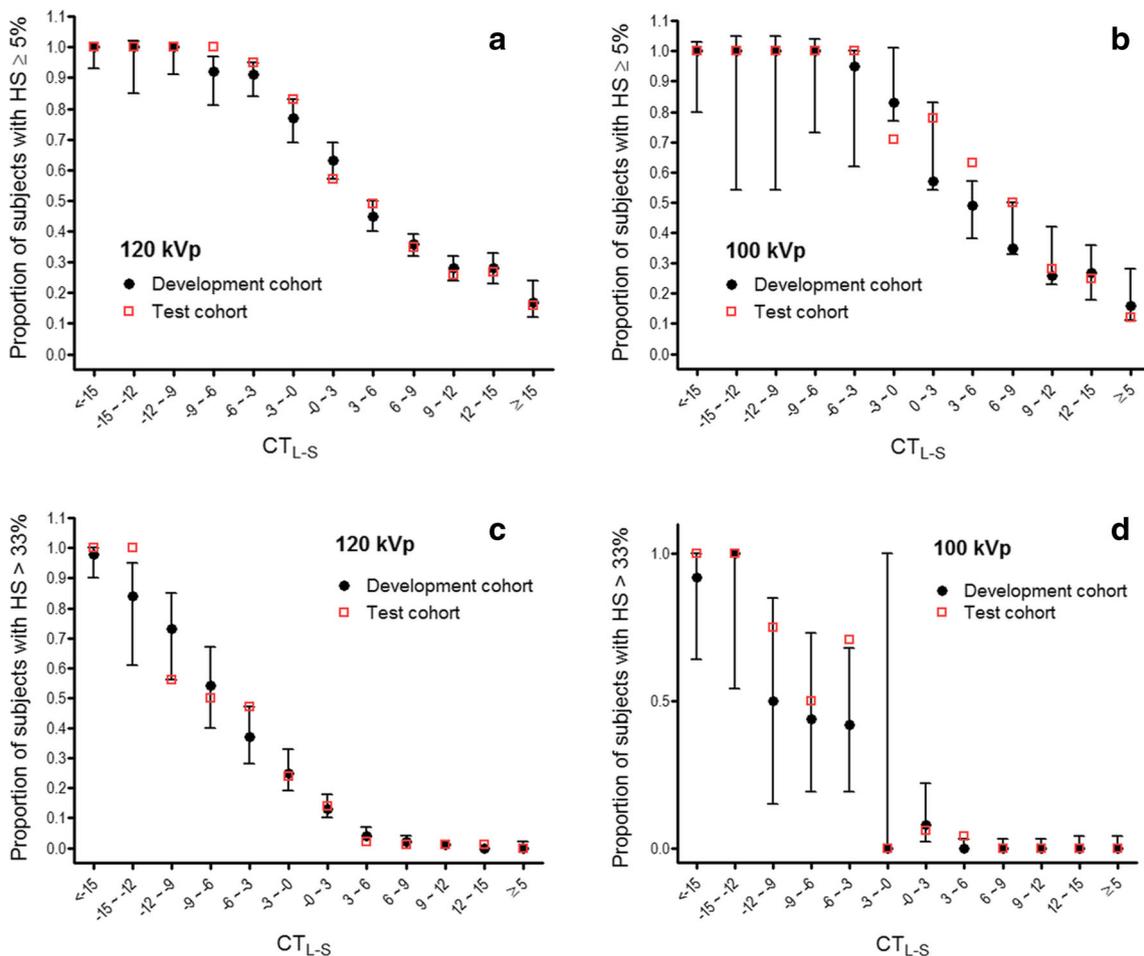
Diagnostic settings	CT performed at 120 kVp (n = 888)			CT performed at 100 kVp (n = 213)		
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
Diagnosis of HS ≥ 5%						
With 95% specificity	32.7% (123/376)	97.0% (495/512)	70.0% (618/888)	46.5% (47/101)	92.0% (103/112)	70.0% (150/213)
With 95% sensitivity	93.6% (352/376)	17.0% (87/512)	49.4% (439/888)	98.0% (99/101)	10.7% (12/112)	52.1% (111/213)
Diagnosis of HS > 33%						
With 95% specificity	64.0% (48/75)	96.4% (784/813)	93.7% (832/888)	73.7% (14/19)	97.9% (190/194)	95.8% (204/213)
With 95% sensitivity	97.3% (73/75)	54.9% (446/813)	58.4% (519/888)	89.5% (17/19)	92.3% (179/194)	92.0% (196/213)

Unless otherwise specified, data are percentages, with the number of subjects used to calculate the percentage in parentheses. Percentages were rounded. Diagnostic performances in the test cohort were evaluated using the CT<sub>L-S</sub> cut-off values determined in the development cohort  
 HS hepatic steatosis

diagnosing HS ≥ 5% than CT<sub>L-S</sub>. Because both CT indices incorporate spleen attenuation as the internal reference, the reason for the higher performance of CT<sub>L-S</sub> than CT<sub>L/S</sub> remains unclear. One possible explanation for this finding may be a wider dynamic range of CT<sub>L-S</sub> (i.e., -53.1~26.7) than

that of CT<sub>L/S</sub> (i.e., -0.04 to 1.7) values, which may have resulted in CT<sub>L-S</sub> being better than CT<sub>L/S</sub> at distinguishing mild HS from no HS.

Our results suggested that CT is not reliable to detect or rule out mild HS (HS ≥ 5%) because of its low overall accuracy



**Fig. 4** Proportions of subjects with hepatic steatosis ≥ 5% (a, b) and ≥ 33% (c, d) according to CT<sub>L-S</sub> ranges for CT scans performed at tube voltages of 120 kVp (a, c) and 100 kVp (b, d) in the development (black

circles) and test (red squares) cohorts; error bars indicate the 95% confidence intervals of the proportions

and combination of either very low sensitivity with high specificity or low specificity with high sensitivity. However, CT may be used to identify subjects with moderate to severe HS. The  $CT_{L-S}$  cut-off values  $-2.1$  for a tube voltage of 120 kVp and  $-3.9$  for a tube voltage of 100 kVp resulted in the diagnosis of moderate to severe HS with 95% specificity and 64–90% sensitivity. These criteria may be used to screen out donor candidates who are not appropriate for liver donation or to identify a cohort of subjects with moderate to severe HS for clinical research. By contrast, the  $CT_{L-S}$  cut-off values for the diagnosis of moderate to severe HS with 95% sensitivity were 7.6 for a tube voltage of 120 kVp and 1.6 for a tube voltage of 100 kVp in our study, which was associated with 54.3–91.7% specificity. Thus, by excluding 95% of subjects with moderate to severe HS, these  $CT_{L-S}$  cut-off values may be utilized to identify a normal control group for clinical studies.

The effects of CT tube voltage on  $CT_{L-S}$  values suggest the need for kVp-specific  $CT_{L-S}$  cut-off values for diagnosing HS. Although our validation results for 100 kVp in the test cohort were not as strong as those for 120 kVp, the  $CT_{L-S}$  cut-off values specific to 100 kVp would be clinically useful because low-voltage CT imaging is increasingly popular for dose reduction purpose [30, 31]. In addition, we estimated the probability of HS  $\geq 5\%$  and HS  $> 33\%$  according to  $CT_{L-S}$  range in the development cohort. At 120 kVp, the actual proportion of subjects with HS  $\geq 5\%$  and HS  $> 33\%$  according to  $CT_{L-S}$  range in the test cohort closely matched the estimated probability of HS in the development (Fig. 4, Table E4), suggesting that these data may be useful in determining the likelihood of HS at a given  $CT_{L-S}$  value.

This study had several limitations. First, the study population was derived from hepatic donor candidates, mostly young, healthy, and lean, and, therefore, may not fully represent the general population. Second, although we validated the diagnostic accuracy of  $CT_{L-S}$  cut-off values in an independent test cohort, the development and test cohorts were obtained from the same population of hepatic donor candidates. External validation in a completely different population may have yielded more conclusive results. Lastly, percutaneous needle biopsy may be subject to some degree of sampling error. Because we obtained CT indices based on liver attenuation values averaged over multiple ROIs in the right hepatic lobe, our results may have been influenced by incomplete co-registration between CT index and biopsy, especially in subjects with uneven HS [32].

In conclusion, CT indices based on both liver and spleen attenuation ( $CT_{L-S}$  and  $CT_{L/S}$ ) perform better in diagnosing HS and are less dependent on CT technique than liver attenuation alone. The  $CT_{L-S}$  cut-off values for the highly specific or sensitive diagnosis of moderate to severe HS determined in this study may have utility in the diagnosis or elimination of HS in clinical practice and research. It should also be remembered that the cut-off values are different according to tube voltage; namely, the cut-off value defined at 120 kV is not applicable as such to 100 kV.

**Funding** This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT, and Future Planning (NRF-2017R1A2B4003114), the Bio and Medical Technology Development Program of the NRF funded by the Ministry of Science and ICT (NRF-2016M3A9A7918706), a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI18C2383), and a grant (2014-444) from the Asan Institute for Life Sciences, Asan Medical Center, Seoul, Korea.

## Compliance with ethical standards

**Guarantor** The scientific guarantor of this publication is Seung Soo Lee.

**Conflict of interest** The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

**Statistics and biometry** One of the authors has significant statistical expertise.

**Informed consent** Written informed consent was waived by the institutional review board.

**Ethical approval** Institutional review board approval was obtained.

## Methodology

- retrospective
- diagnostic or prognostic study
- performed at one institution

## References

1. Teli MR, James OF, Burt AD, Bennett MK, Day CP (1995) The natural history of nonalcoholic fatty liver: a follow-up study. *Hepatology* 22:1714–1719
2. de Alwis NM, Day CP (2008) Non-alcoholic fatty liver disease: the mist gradually clears. *J Hepatol* 48(Suppl 1):S104–S112
3. Adams LA, Sanderson S, Lindor KD, Angulo P (2005) The histological course of nonalcoholic fatty liver disease: a longitudinal study of 103 patients with sequential liver biopsies. *J Hepatol* 42:132–138
4. Ascha MS, Hanounch IA, Lopez R, Tamimi TA, Feldstein AF, Zein NN (2010) The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *Hepatology* 51:1972–1978
5. Korenblat KM, Fabbrini E, Mohammed BS, Klein S (2008) Liver, muscle, and adipose tissue insulin action is directly related to intrahepatic triglyceride content in obese subjects. *Gastroenterology* 134:1369–1375
6. Gastaldelli A, Kozakova M, Højlund K et al (2009) Fatty liver is associated with insulin resistance, risk of coronary heart disease, and early atherosclerosis in a large European population. *Hepatology* 49:1537–1544
7. Marsman WA, Wiesner RH, Rodriguez L et al (1996) Use of fatty donor liver is associated with diminished early patient and graft survival. *Transplantation* 62:1246–1251
8. Takyar V, Nath A, Beri A, Gharib AM, Rotman Y (2017) How healthy are the “healthy volunteers”? Penetration of NAFLD in the biomedical research volunteer pool. *Hepatology* 66:825–833
9. Lee SS, Park SH (2014) Radiologic evaluation of nonalcoholic fatty liver disease. *World J Gastroenterol* 20:7392–7402

10. Lee SS, Park SH, Kim HJ et al (2010) Non-invasive assessment of hepatic steatosis: prospective comparison of the accuracy of imaging examinations. *J Hepatol* 52:579–585
11. Boyce CJ, Pickhardt PJ, Kim DH et al (2010) Hepatic steatosis (fatty liver disease) in asymptomatic adults identified by unenhanced low-dose CT. *AJR Am J Roentgenol* 194:623–628
12. Hahn L, Reeder SB, Muñoz del Rio A, Pickhardt PJ (2015) Longitudinal changes in liver fat content in asymptomatic adults: hepatic attenuation on unenhanced CT as an imaging biomarker for steatosis. *AJR Am J Roentgenol* 205:1167–1172
13. Iwasaki M, Takada Y, Hayashi M et al (2004) Noninvasive evaluation of graft steatosis in living donor liver transplantation. *Transplantation* 78:1501–1505
14. Kan H, Kimura Y, Hyogo H et al (2014) Non-invasive assessment of liver steatosis in non-alcoholic fatty liver disease. *Hepatol Res* 44:E420–E427
15. Limanond P, Raman SS, Lassman C et al (2004) Macrovesicular hepatic steatosis in living related liver donors: correlation between CT and histologic findings. *Radiology* 230:276–280
16. Park SH, Kim PN, Kim KW et al (2006) Macrovesicular hepatic steatosis in living liver donors: use of CT for quantitative and qualitative assessment. *Radiology* 239:105–112
17. Park YS, Park SH, Lee SS et al (2011) Biopsy-proven nonsteatotic liver in adults: estimation of reference range for difference in attenuation between the liver and the spleen at nonenhanced CT. *Radiology* 258:760–766
18. Pickhardt PJ, Park SH, Hahn L, Lee SG, Bae KT, Yu ES (2012) Specificity of unenhanced CT for non-invasive diagnosis of hepatic steatosis: implications for the investigation of the natural history of incidental steatosis. *Eur Radiol* 22:1075–1082
19. Piekarski J, Goldberg HI, Royal SA, Axel L, Moss AA (1980) Difference between liver and spleen CT numbers in the normal adult: its usefulness in predicting the presence of diffuse liver disease. *Radiology* 137:727–729
20. Rogier J, Roullet S, Coméris F et al (2015) Noninvasive assessment of macrovesicular liver steatosis in cadaveric donors based on computed tomography liver-to-spleen attenuation ratio. *Liver Transpl* 21:690–695
21. Mellinger JL, Pencina KM, Massaro JM et al (2015) Hepatic steatosis and cardiovascular disease outcomes: an analysis of the Framingham Heart Study. *J Hepatol* 63:470–476
22. Pickhardt PJ, Hahn L, Muñoz del Rio A, Park SH, Reeder SB, Said A (2014) Natural history of hepatic steatosis: observed outcomes for subsequent liver and cardiovascular complications. *AJR Am J Roentgenol* 202:752–758
23. Hong HC, Hwang SY, Choi HY et al (2014) Relationship between sarcopenia and nonalcoholic fatty liver disease: the Korean Sarcopenic Obesity Study. *Hepatology* 59:1772–1778
24. Puchner SB, Lu MT, Mayrhofer T et al (2015) High-risk coronary plaque at coronary CT angiography is associated with nonalcoholic fatty liver disease, independent of coronary plaque and stenosis burden: results from the ROMICAT II trial. *Radiology* 274:693–701
25. Bae JC, Lee WY, Yoon KH et al (2015) Improvement of nonalcoholic fatty liver disease with Carnitine-Orotate Complex in Type 2 Diabetes (CORONA): a randomized controlled trial. *Diabetes Care* 38:1245–1252
26. Bos D, Leening MJG (2018) Leveraging the coronary calcium scan beyond the coronary calcium score. *Eur Radiol* 28:3082–3087
27. Kleiner DE, Brunt EM, Van Natta M et al (2005) Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 41:1313–1321
28. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44:837–845
29. Kodama Y, Ng CS, Wu TT et al (2007) Comparison of CT methods for determining the fat content of the liver. *AJR Am J Roentgenol* 188:1307–1312
30. Kaul A, Bauer B, Bernhardt J, Nosske D, Veit R (1997) Effective doses to members of the public from the diagnostic application of ionizing radiation in Germany. *Eur Radiol* 7:1127–1132
31. Nakayama Y, Awai K, Funama Y et al (2005) Abdominal CT with low tube voltage: preliminary observations about radiation dose, contrast enhancement, image quality, and noise. *Radiology* 237:945–951
32. Tran TT, Changsri C, Shackleton CR et al (2006) Living donor liver transplantation: histological abnormalities found on liver biopsies of apparently healthy potential donors. *J Gastroenterol Hepatol* 21:381–383