의학석사학위논문

비조영강증강 전산화 단층촬영과 임상 요인에 기반한 비알콜성 지방간 질환 예측 지수의 개발과 검증

Development and validation of a simple index based on non-enhanced CT and clinical factors for prediction of nonalcoholic fatty liver disease

울산대학교 대학원

의 학 과

안 유 라

비조영강증강 전산화 단층촬영과 임상 요인에 기반한 비알콜성 지방간 질환 예측 지수의 개발과 검증

지도교수 이 승 수

이 논문을 의학석사학위 논문으로 제출함

2019 년 12 월

울산대학교 대학원

의 학 과

안 유 라

안유라의 의학석사학위 논문을 인준함

심사위원 김 형 중 인

심사위원 이 승 수 인

심사위원 김 소 연 인

울 산 대 학 교 대 학 원

2019 년 12 월

Abstract

Objective: A widely applicable, non-invasive screening method for the nonalcoholic fatty liver disease (NAFLD) is needed. We aimed to develop and validate an index combining computed tomography (CT) and routine clinical data for the reliable diagnosis and exclusion of pathologically proven NAFLD in a large cohort of adults.

Materials and Methods: This retrospective study included 2218 living liver donors who underwent liver biopsy and CT within 3 days. Donors were randomized 2:1 into development and test cohorts. CT_{L-S} was measured by subtracting splenic attenuation from hepatic attenuation on non-enhanced CT. Multivariable logistic regression analysis of the development cohort was utilized to develop a clinical-CT index predicting pathologically proven NAFLD. The diagnostic performance was evaluated by analysis of areas under the receiver operating characteristic curve (AUC). The cutoffs for clinical-CT index were determined for 90% sensitivity and 90% specificity in the development cohort, and their diagnostic performance was evaluated in the test cohort.

Results: The clinical-CT index included CT_{L-S}, body mass index, and aspartate transaminase and triglyceride concentrations. In the test cohort, the clinical-CT index (AUC, 0.81) outperformed CT_{L-S} (0.74; $P < 0.001$) and the clinical indices (0.73–0.75; $P < 0.001$) for diagnosing NAFLD. A cutoff of ≥46 had a sensitivity of 89% and a specificity of 41%, whereas a cutoff of ≥56.5 had a sensitivity of 57% and a specificity of 89%.

Conclusion: The clinical-CT index is more accurate than CT_{L-S} and clinical indices alone for the diagnosis of NAFLD and may be clinically useful in diagnosing or excluding NAFLD.

목차

표 및 그림 차례

Introduction

Nonalcoholic fatty liver disease (NAFLD) is a common cause of chronic liver disease, affecting up to 40% of the general population in developed countries (1, 2). NAFLD may progress to cirrhosis and hepatocellular carcinoma (3, 4) and is also associated with metabolic syndrome (5). Due to its high prevalence and asymptomatic presentation, however, NAFLD is often overlooked in healthy controls selected for clinical trials, possibly hampering the validity of study findings (6). Thus, there is a need for a widely applicable, non-invasive screening method allowing for the reliable diagnosis or exclusion of NAFLD.

Liver biopsy is regarded as the gold standard for assessing NAFLD, especially for evaluating inflammation and fibrosis associated with NAFLD. However, the invasiveness of biopsy limits its use in clinical practice and research. Among imaging methods, magnetic resonance (MR) spectroscopy and imaging are most accurate for quantifying liver fat contents and detecting NAFLD (7, 8). However, these techniques are often unavailable in general practice. Although grayscale ultrasonography (US) is commonly used to screen for NAFLD, it may be subject to inter-observer variability and has limited accuracy in detecting NAFLD (7, 8).

Hepatic steatosis may be quantitatively assessed by measuring hepatic attenuation on non-enhanced computed tomography (CT). CT is more widely available and less expensive than MR imaging and may provide an objective assessment of NAFLD. Thus, non-enhanced CT has been used to assess NAFLD in candidates for living liver donation (7, 9), as well as in cohort studies and clinical trials (10-12). Although CT allows for a highly specific diagnosis of moderate to severe fatty liver disease, it is not accurate in detecting a mild degree of fatty liver (7, 13). Therefore, it may not be reliable for selecting or excluding subjects with NAFLD. Several clinical prediction models based on demographic, anthropometric, and laboratory characteristics may be used to screen for NAFLD (14-16) and have been successfully applied to large-scale cohort studies (5, 17). Thus, we hypothesized that combinations of clinical parameters and CT results may enable more accurate detection of NAFLD than clinical indices or CT alone. A simple index based on CT results and routinely accessible clinical data may be useful for detecting NAFLD in clinical practice and research. Furthermore, this index can be applied to pre-existing CT and clinical data to conduct large-scale retrospective cohort studies. The present study was designed to develop and validate a simple index combining CT and standard clinical data for the reliable diagnosis and exclusion of NAFLD in a large cohort of adults with pathologically proven NAFLD.

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 study population included living liver donor candidates who underwent ultrasound-guided percutaneous liver biopsy as part of routine donor work-up at our institution between April 2001 and October 2016. Subjects were included if they were aged ≥18 years, underwent CT scanning within 3 days of liver biopsy, and underwent clinical and laboratory examinations within 7 days of liver biopsy. Of the 2787 consecutive living liver donor candidates evaluated during the study period, 569 were excluded, including 438 with missing laboratory results (will be discussed later); 34 with pathology reports that did not include the degree of hepatic steatosis; 65 with a history of excess alcohol consumption (i.e. over the 20g of ethanol/day) (18); 13 with liver disease incidentally detected on biopsy or serologic tests; and 19 with conditions that precluded measurement of CT indices, including 15 lacking non-enhanced CT images, two with numerous hepatic cysts, and two with prior splenectomy. The remaining 2218 subjects (1439 men and 779 women; mean age, 31.0 ± 9.0 years; range, 18–62 years) were randomly divided 2:1 into development (n = 1480) and test (n = 738) cohorts. The flow diagram for the study population is shown in Fig. 1. The CT data in the study population have been reported previously (13); in that study, the data were used to evaluate the performance of CT indices and to determine cutoff values for diagnosing hepatic steatosis.

CT protocol

Because the CT data in this study were collected over a long period, various CT techniques were used. CT scans were obtained using 4-channel (Lightspeed Qx/I; GE Medical Systems, Milwaukee, WI, USA; n = 2), 16-channel (Lightspeed 16; GE Medical Systems or Sensation 16; Siemens, Erlangen, Germany; n = 1611), 64-channel (Definition AS, Siemens; n = 564), and 128-channel (Definition Flash, Siemens; n = 41) scanners. Non-enhanced CT images were obtained at beam collimations of 4 × 2.5 mm (Lightspeed Qx/I), 8 × 2.5 mm (Lightspeed 16), 16 × 1.5 mm (Sensation 16), 24 × 1.2 mm (Definition AS), and 64 × 0.6 mm (Definition Flash); at a spiral pitch of 1 to 1.5; at tube voltages of 120 kVp (n = 1672) and 100 kVp (n = 546); and at tube currents of 200 mAs (GE scanners) or variable

mAs (Siemens scanners) with an automatic exposure control (Care Dose 4D, Siemens; maximum effective dose, 200 mAs). Axial images were reconstructed at section thicknesses of 3 mm (n = 45) and 5 mm (n = 2173), with no gaps. The mean interval between CT and liver biopsy was 0.4 ± 0.7 days (range, 0–3 days) with 1710 (74.8%) subjects undergoing CT scanning and liver biopsy on the same day.

CT image analysis

The quantitative CT index used in this study was CT_{L-S} because it was reported to be the most accurate and robust CT index for assessing NAFLD (13). CT_{L-S} was calculated as mean liver attenuation minus mean spleen attenuation. Liver and spleen attenuation values on non-enhanced CT images were measured by one of two radiology technicians using in-house software plugged into ImageJ (National Institutes of Health, Bethesda, MD, USA). Liver attenuation was calculated as the average number of Hounsfield units (HU) of eight 1.5 cm^2 circular regions of interest (ROIs) of the right hepatic lobe. Splenic attenuation was calculated as the average HU of three 1.5 cm^2 circular ROIs of the upper, middle, and lower thirds of the spleen (Fig. 2). The CT images with the ROIs were screen-captured and re-evaluated by an abdominal imaging fellow with 2 years of experience in abdominal imaging (*BLINDED*) to reconfirm the adequacy of ROI locations.

Clinical parameters

Clinical variables included body mass index (BMI), calculated as body weight (kg)/height $(m)^2$; age; sex; and serum concentrations of aspartate transaminase (AST), alanine aminotransferase (ALT), bilirubin, alkaline phosphatase (ALP), triglyceride (TG), cholesterol, high-density lipoprotein (HDL) cholesterol, glucose, and albumin. These parameters were selected based on variables previously included in clinical models for the diagnosis of NAFLD (14-16, 19). TG, cholesterol, and HDL-cholesterol concentrations were missing from the records of 438 (15.7%) of the 2787 eligible subjects who met the inclusion criteria. Because these variables were frequently included in previous clinical indices and because their rates of absence were too high for reliable data imputation, these subjects were excluded from this study. All laboratory tests were performed after a 12 hour overnight fast. The mean interval between laboratory examination and liver biopsy was 1.4 ± 2.2 days (range, 0–7 days).

Reference standard

A pathologic diagnosis of NAFLD was defined as the reference standard. All subjects underwent US-guided percutaneous liver biopsy using an 18-gauge needle (Stericut 18G coaxial; TSK Laboratory, Tochigi, Japan), with at least two biopsy specimens measuring approximately 1.5 cm in length each, obtained from different sites in the right hepatic lobe. The biopsy specimens were stained with hematoxylin and eosin, and the degree of parenchymal involvement of macrovesicular steatosis was graded as none (<5%), mild (5–33%), moderate (34–66%), or severe (>66%), as defined by the nonalcoholic steatohepatitis Clinical Research Network scoring system (20). NAFLD was defined as the presence of ≥5% macrovesicular steatosis (20).

Clinical indices for diagnosing NAFLD

Two previous described clinical indices for diagnosing NAFLD, the hepatic steatosis index (HS-I) and the fatty liver disease index (FLD-I), were calculated for each subject (14, 16). The HS-I was calculated as 8·(ALT/AST) + BMI (+2 if diabetic and +2 if female), and the FLD-I was calculated as BMI + TG + 3 (ALT/AST) (+2 if hyperglycemic, with hyperglycemia defined as a fasting plasma glucose concentration ≥126 mg/dl).

Statistical analysis

Continuous variables in the development and test cohorts were compared using t-tests or Mann–Whitney U-tests, whereas categorical variables were compared using the χ^2 test or Fisher's exact test. The normality of continuous variables was tested using the Kolmogorov–Smirnov test. Variables not normally distributed were log-transformed. Variables in subjects in the development cohort with and without NAFLD were compared by univariable logistic regression analysis. Candidate predictors were selected among all variables by multivariable logistic regression analysis with 1000-fold bootstrap resampling; variables selected in more than 50% of bootstrap logistic models were chosen as candidate predictors. Variables independently associated with NAFLD were identified by multivariable logistic regression analysis with backward elimination. To construct a simplified predictive model, logistic models that included an increasing number of variables were sequentially developed by one-by-one addition of independent variables to $CT_{L\text{-}S}$. The diagnostic performance of each logistic model was assessed by calculating the area under the receiver operating characteristic curve (AUC). This procedure

was continued until model performance did not improve by an AUC of 0.005 upon the inclusion of an additional variable. A formula for the clinical-CT index was derived using the variables in the final logistic model and the proportions of the corresponding regression coefficients. The diagnostic performances of the clinical-CT index and the CT_{L-S} were evaluated by comparing the AUCs using Delong's method (21). The incremental difference between the clinical-CT index and CT_{L-S} alone was evaluated by calculating the net reclassification improvement (NRI) and the integrated discrimination improvement (IDI) (22, 23). The NRI is used to evaluate the net proportion of subjects reclassified correctly using the new model (i.e., the clinical-CT index) relative to the baseline model (i.e., CT_{L-S}), whereas the IDI measures the improvement in sensitivity of the new model relative to the baseline model without a loss in specificity (22, 23). Positive NRI and IDI values indicate the superiority of the new relative to the baseline model for correct classification. Cutoffs were selected for the clinical-CT index and the CT_{L-S} at points of 90% sensitivity and 90% specificity for diagnosing NAFLD, thus reliably detecting and ruling out NAFLD, and the corresponding sensitivities, specificities, and accuracies were calculated. The diagnostic performance of the clinical-CT index and CT_{L-S} in the test cohort were compared using the AUCs, NRI, and IDI, whereas the diagnostic performances of the clinical-CT index and CT_{L-S} were compared with those of the HS-I and FLD-I using the AUC. The sensitivities, specificities, and accuracies in diagnosing NAFLD were evaluated in the test cohort using the cutoff values for the clinical-CT index and CT_{L-S} determined in the development cohort. A P-value less than .05 was considered statistically significant. All statistical analyses were performed using SAS software version 9.3 (SAS Institute Inc., Cary, NC, USA)

Results

Patient characteristics

Table 1 summarizes the characteristics of the study population. NAFLD was present in 620 (41.9%) of the 1480 subjects in the development cohort and in 310 (42.0%) of the 738 in the test cohort. None of the clinical and laboratory characteristics assessed differed significantly in the development and test cohorts.

Development of a simple clinical-CT index in the development cohort

Univariable logistic regression analysis showed that all variables analyzed, except for bilirubin, were significantly associated with NAFLD (Table 2). Multivariable logistic regression analysis followed by candidate predictor selection showed that BMI, ALT, TG, cholesterol, and HDL-cholesterol, and $CT_{L\text{-}S}$ were independently associated with NAFLD (Supplementary Table 1 and 2). These independent variables were utilized to construct logistic models containing different numbers of variables predictive of NAFLD (Supplementary Table 2). The final logistic models included $CT_{L\text{-}S}$, BMI, TG, and ALT (Table 3). The relationship of the regression coefficients of the final logistic models resulted in a clinical-CT index predictive of NAFLD.

Clinical-CT index for the prediction of NAFLD = $5 \cdot \mathrm{Log_e}\,(\mathrm{ALT} \cdot \mathrm{TG}) + \mathrm{BMI} - CT_{L\text{-}S}$.

Diagnostic performances of the clinical-CT index and $CT_{L\text{-}S}$ in the development cohort

In the development cohort, the AUC for the clinical-CT index was 0.82 (95% confidence interval [CI], 0.80–0.84), which was significantly higher than that of $CT_{L\text{-}S}$ (0.74; 95% CI, 0.72–0.76; $P < 0.001$) (Fig. 3A). The clinical-CT index also showed significant improvement in reclassification (NRI, 0.75; $P < 0.001$) and discrimination (IDI, 0.12; $P < 0.001$) than $CT_{L\text{-}S}$. Table 4 summarizes the dual cutoff values of the clinical-CT index and $CT_{L\text{-}S}$ for diagnosing NAFLD. A clinical-CT index cutoff of ≥46 diagnosed NAFLD with a sensitivity of 90%, a specificity of 44%, and an accuracy of 64%, whereas a $CT_{L\text{-}S}$ cutoff of ≤8.4 showed a sensitivity of 90%, a specificity of 28%, and an accuracy of 54%. Alternatively, a clinical-CT index cutoff of ≥56.5 diagnosed NAFLD with a sensitivity of 37%, a specificity of 90%, and an accuracy of 76%, whereas a $CT_{L\text{-}S}$ cutoff of ≤3.9 had a sensitivity of 44%, a specificity of 90%, and an accuracy of 71%.

Diagnostic performances of the clinical-CT index and $CT_{L\text{-}S}$ in the test cohort

The AUC for diagnosing NAFLD in the test cohort was significantly higher for the clinical-CT index (0.81; 95% CI, 0.78–0.84) than for $CT_{L\text{-}S}$ (0.74; 95% CI, 0.71–0.78; $P < 0.001$) (Fig. 3B). In addition, the clinical-CT index had a significant incremental value in reclassification (NRI, 0.61; $P < 0.001$) and discrimination (IDI, 0.09; $P < 0.001$) than $CT_{L\text{-}S}$. Compared with clinical indices, the clinical-CT index significantly outperformed HS-I (AUC, 0.73; 95% CI, 0.70–0.76; $P < 0.001$) and FLD-I (AUC, 0.75; 95% CI, 0.72–0.78; $P < 0.001$) for diagnosing NAFLD, whereas the AUC for $CT_{L\text{-}S}$ did not differ significantly from the AUCs for HS-I ($P = 0.64$) and FLD-I ($P = 0.19$). The cutoff values for the clinical-CT index and $CT_{L\text{-}S}$ from the development cohort performed similarly in the test cohort. A clinical-CT index cutoff of ≥46 diagnosed NAFLD with a sensitivity of 89% and a specificity of 41%, whereas a cutoff of ≥56.5 had a sensitivity of 57% and a specificity of 89%. By comparison, a $CT_{L\text{-}S}$ cutoff of ≤12.5 had a sensitivity of 89% and a specificity of 27%, whereas a cutoff of ≤3.9 had a sensitivity of 46% and a specificity of 90%.

Discussion

The present study describes the development of a simple index, combining the CT index and routinely tested blood and anthropometric parameters, to predict NAFLD. The clinical-CT index outperformed the $CT_{L\text{-}S}$ for diagnosing NAFLD with significantly higher AUCs and significant improvements in reclassification and discrimination in both the development (AUCs, 0.82 vs. 0.74, $P < 0.001$; NRI, 0.747, $P < .001$; IDI, 0.121, $P < 0.001$) and test (AUCs, 0.81 vs. 0.74, $P < 0.001$; NRI, 0.609, $P < .001$; IDI, 0.089, $P < 0.001$) cohorts. These findings indicate that adding clinical parameters to the $CT_{L\text{-}S}$ improved the diagnosis of NAFLD. The clinical-CT index also performed better than the clinical indices for diagnosing NAFLD in the test cohort, whereas the AUC for $CT_{L\text{-}S}$ did not differ significantly from the AUCs for the clinical indices. These results suggest that using CT alone for the diagnosis or exclusion of NAFLD would not be a reasonable approach, as the clinical indices that can be determined more easily and at lower cost than the $CT_{L\text{-}S}$ showed performances similar to that of the $CT_{L\text{-}S}$. However, the clinical-CT index, incorporating both CT and clinical parameters, may have clinical utility, allowing for better detection of NAFLD than the CT index and the individual clinical indices.

Dual cutoff values of the clinical-CT index for diagnosing NAFLD were selected. One of these cutoffs was based on 90% sensitivity, which could be used to define a normal control group by excluding most subjects with NAFLD. In the test cohort, cutoffs of ≥46 for the clinical-CT index and ≤12.5 for $CT_{L\text{-}S}$ resulted in the diagnosis of NAFLD with sensitivities approximating 90% and specificities of 40.7% and 26.9%, respectively. The other cutoff, based on 90% specificity, could be used to identify a cohort of subjects with NAFLD for clinical research. In the test cohort, cutoffs of ≥56.5 for the clinical-CT index and ≤3.9 for $CT_{L\text{-}S}$ resulted in the diagnosis of NAFLD with specificities of approximately 90% and sensitivities of 57.1% and 46.1%, respectively.

Several predictive models based on clinical and laboratory parameters have been developed to distinguish subjects with and without NAFLD. Some of these models, however, included parameters not always routinely measured in clinical practice. For example, models have included serum concentrations of insulin (24), uric acid (25), hemoglobin A1C (19), and haptoglobin (26), as well as waist circumference (15, 27), which are parameters that may not be easily retrieved from patient databases. This may limit the use of these models in retrospective analyses. Furthermore, many predictive models have been based on the diagnosis of NAFLD by grayscale US (14-16, 25, 27), despite US diagnosis of NAFLD being operator-dependent and having limited accuracy (7). By contrast, the clinical-CT index for diagnosing NAFLD in our study was based on pathologic proof of NAFLD in a large population.

Given a potential radiation hazard of CT and the availability of other imaging methods for the diagnosis of NAFLD such as MR imaging, MR spectroscopy, quantitative US methods, and the controlled attenuation parameter of transient elastography (8, 28, 29) our clinical-CT index may not be an optimal method for identifying patients with NAFLD in clinical practice and in a prospective research. However, because of the widespread use of CT, the clinical-CT index described in our study incorporating routinely measured laboratory and anthropometric parameters may be useful for constructing large cohorts of subjects with NAFLD and normal controls using preexisting retrospective CT and clinical data, and which may be used to conduct large-scale retrospective cohort studies for investigating the natural history and outcome of NAFLD.

This study had several limitations. First, the study population was derived from liver donor candidates, most of whom were young and healthy, and therefore may not fully represent the general population. Second, split-sample validation of the diagnostic accuracy of the clinical-CT indices was performed. External validation in a different test population may have been more conclusive. Third, we excluded subjects with excessive alcohol intake in our study to avoid confounding effects of alcohol on clinical, CT, and pathologic findings. However, in clinical practice and in retrospective research, the information on alcohol consumption may not be always available, and which may potentially affect the performance of the clinical-CT index. Finally, although percutaneous needle biopsy is a well-accepted reference method for the diagnosis NAFLD, it may be subject to some degree of sampling error and inter-observer variability.

In conclusion, a clinical-CT index combining clinical parameters and CT_{L-S} was more accurate in the diagnosis of NAFLD than CT_{L-S} or clinical indices alone. This clinical-CT index may have utility in the diagnosis or elimination of NAFLD in clinical practice and research.

References

1. Wong VW-S, Wong GL-H, Yeung DK-W, Lau TK-T, Chan CK-M, Chim AM-L, et al. Incidence of non-alcoholic fatty liver disease in Hong Kong: a population study with paired proton-magnetic resonance spectroscopy. *Journal of hepatology* 2015;62:182-189
2. Younossi ZM, Blissett D, Blissett R, Henry L, Stepanova M, Younossi Y, et al. The economic and clinical burden of nonalcoholic fatty liver disease in the United States and Europe. *Hepatology* 2016;64:1577-1586
3. Adams LA, Sanderson S, Lindor KD, Angulo P. The histological course of nonalcoholic fatty liver disease: a longitudinal study of 103 patients with sequential liver biopsies. *Journal of hepatology* 2005;42:132-138
4. Teli MR, James OF, Burt AD, Bennett MK, Day CP. The natural history of nonalcoholic fatty liver: a follow up study. *Hepatology* 1995;22:1714-1719
5. Gastaldelli A, Kozakova M, Højlund K, Flyvbjerg A, Favuzzi A, Mitrakou A, et al. Fatty liver is associated with insulin resistance, risk of coronary heart disease, and early atherosclerosis in a large European population. *Hepatology* 2009;49:1537-1544
6. Takyar V, Nath A, Beri A, Gharib AM, Rotman Y. How healthy are the "healthy volunteers"? Penetrance of NAFLD in the biomedical research volunteer pool. *Hepatology* 2017;66:825-833
7. Lee SS, Park SH, Kim HJ, Kim SY, Kim M-Y, Kim DY, et al. Non-invasive assessment of hepatic steatosis: prospective comparison of the accuracy of imaging examinations. *Journal of hepatology* 2010;52:579-585
8. Lee SS, Park SH. Radiologic evaluation of nonalcoholic fatty liver disease. *World journal of gastroenterology: WJG* 2014;20:7392
9. Rogier J, Roullet S, Cornélis F, Biais M, Quinart A, Revel P, et al. Noninvasive assessment of macrovesicular liver steatosis in cadaveric donors based on computed tomography liver to spleen attenuation ratio. *Liver Transplantation* 2015;21:690-695
10. Mellinger JL, Pencina KM, Massaro JM, Hoffmann U, Seshadri S, Fox CS, et al. Hepatic steatosis and cardiovascular disease outcomes: an analysis of the Framingham Heart Study. *Journal of hepatology* 2015;63:470-476
11. Pickhardt PJ, Hahn L, del Rio AM, Park SH, Reeder SB, Said A. Natural history of hepatic steatosis: observed outcomes for subsequent liver and cardiovascular complications. *American Journal of Roentgenology* 2014;202:752-758

12. Bae JC, Lee WY, Yoon KH, Park JY, Son HS, Han KA, et al. Improvement of nonalcoholic fatty liver disease with Carnitine-Orotate Complex in Type 2 Diabetes (CORONA): a randomized controlled trial. *Diabetes care* 2015;38:1245-1252
13. Byun J, Lee SS, Sung YS, Shin Y, Yun J, Kim HS, et al. 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. *European radiology* 2018:1-9
14. Fuyan S, Jing L, Wenjun C, Zhijun T, Weijing M, Suzhen W, et al. Fatty liver disease index: a simple screening tool to facilitate diagnosis of nonalcoholic fatty liver disease in the Chinese population. *Digestive diseases sciences* 2013;58:3326-3334
15. Bedogni G, Bellentani S, Miglioli L, Masutti F, Passalacqua M, Castiglione A, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. *BMC gastroenterology* 2006;6:33
16. Lee J-H, Kim D, Kim HJ, Lee C-H, Yang JI, Kim W, et al. Hepatic steatosis index: a simple screening tool reflecting nonalcoholic fatty liver disease. *Digestive Liver Disease* 2010;42:503-508
17. Choi YJ, Lee DH, Han K-D, Yoon H, Shin CM, Park YS, et al. Is nonalcoholic fatty liver disease associated with the development of prostate cancer? A nationwide study with 10,516,985 Korean men. *PloS one* 2018;13:e0201308
18. Neuschwander-Tetri BA, Caldwell SH. Nonalcoholic steatohepatitis: summary of an AASLD Single Topic Conference. *Hepatology* 2003;37:1202-1219
19. Yip TF, Ma A, Wong VS, Tse YK, Chan HY, Yuen PC, et al. Laboratory parameter based machine learning model for excluding non alcoholic fatty liver disease (NAFLD) in the general population. *Alimentary pharmacology therapeutics* 2017;46:447-456
20. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005;41:1313-1321
21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-845
22. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1-73

23. Moons KG, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio) marker. *Heart* 2012;98:683-690
24. Kotronen A, Peltonen M, Hakkarainen A, Sevastianova K, Bergholm R, Johansson LM, et al. Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors. *Gastroenterology* 2009;137:865-872
25. Zhang Z, Wang G, Kang K, Wu G, Wang P. Diagnostic accuracy and clinical utility of a new noninvasive index for hepatic steatosis in patients with hepatitis B virus infection. *Scientific reports* 2016;6:32875
26. Poynard T, Ratziu V, Naveau S, Thabut D, Charlotte F, Messous D, et al. The diagnostic value of biomarkers (SteatoTest) for the prediction of liver steatosis. *Comparative hepatology* 2005;4:10
27. Bedogni G, Kahn HS, Bellentani S, Tiribelli C. A simple index of lipid overaccumulation is a good marker of liver steatosis. *BMC gastroenterology* 2010;10:98
28. Lin SC, Heba E, Wolfson T, Ang B, Gamst A, Han A, et al. Noninvasive Diagnosis of Nonalcoholic Fatty Liver Disease and Quantification of Liver Fat Using a New Quantitative Ultrasound Technique. *Clin Gastroenterol Hepatol* 2015;13:1337-1345.e1336
29. Lee DH. Imaging evaluation of non-alcoholic fatty liver disease: focused on quantification. *Clin Mol Hepatol* 2017;23:290-301

Table 1. Characteristics of the study population

	Developmental cohort (n = 1480)	Test cohort (n = 738)	*P*-value
Age (years)	30.9 ± 9.1 (18–62)	31.3 ± 8.9 (18–58)	0.40
Sex (female)†	534 (36.1%)	245 (33.2%)	0.18
Pathologic steatosis grade†			0.95
None (≤5%)	860 (58.1%)	428 (58.0%)	
Mild (>5%)	517 (34.9%)	256 (34.7%)	
Moderate/severe (≥33%)	103 (7.0%)	54 (7.3%)	
BMI (kg/m^2)	23.3 ± 3.2 (15.4–41.3)	23.4 ± 3.0 (15.4–34.9)	0.69
Laboratory findings			
AST (IU/mL)	21.3 ± 11.7 (10–365)	21.4 ± 8.2 (10–129)	0.77
ALT (IU/mL)	20.3 ± 13.6 (1–181)	20.9 ± 13.1 (6–121)	0.29
Bilirubin (ng/mL)	0.98 ± 0.37 (0.2–3.5)	1.00 ± 0.39 (0.3–4.1)	0.23
ALP (IU/mL)	62.87 ± 18.22 (20–186)	62.70 ± 18.30 (10–182)	0.84
Triglyceride (mg/dL)	103.81 ± 77.22 (17–935)	108.02 ± 88.85 (21–1304)	0.25
Cholesterol (mg/dL)	173.74 ± 32.83 (68–320)	174.77 ± 32.51 (98–302)	0.48
HDL-cholesterol (mg/dL)	50.70 ± 13.27 (20–106)	50.40 ± 12.95 (19–94)	0.61
Glucose (mg/dL)	94.54 ± 15.02 (58–370)	95.19 ± 17.05 (63–374)	0.36
Albumin (g/dL)	4.37 ± 0.29 (3.2–5.2)	4.37 ± 0.27 (3.5–5.2)	0.97
Hepatic steatosis index	31.45 ± 4.91 (20–54)	31.62 ± 4.50 (21–47)	0.43
Fatty liver disease index	27.28 ± 4.19 (24–30)	27.45 ± 4.03 (25–30)	0.25

Note: Unless otherwise indicated, data are mean ± standard deviation; data in parentheses are range. BMI = body mass index, AST = aspartate transaminase, ALT = alanine aminotransferase, ALP = alkaline phosphatase, HDL = high-density lipoprotein. †Number (percentages) of patients.

Table 2. Univariable and multivariable logistic regression analysis of factors associated with a diagnosis of NAFLD in the development cohort

Variables	Univariable analysis		Multivariable analysis[†]	
	Odds ratio (95% CI)	*P*-value	Adjusted odds ratio (95% CI)	*P*-value
Age	1.03 (1.02, 1.04)	<0.001		
Sex (female)	0.44 (0.35, 0.55)	<0.001		
BMI	1.34 (1.28, 1.40)	<0.001	1.17 (1.12, 1.23)	<0.001
AST[†]	5.40 (3.57, 8.16)	<0.001		
ALT[†]	5.57 (4.32, 7.17)	<0.001	1.75 (1.31, 2.36)	<0.001
Bilirubin	1.18 (0.89, 1.57)	0.25		
ALP	1.02 (1.01, 1.02)	<0.001		
Triglyceride[†]	3.5 (2.85, 4.30)	<0.001	1.58 (1.31, 2.08)	<0.001
Cholesterol	1.02 (1.01, 1.02)	<0.001	1.01 (1.00,1.02)	<0.001
Glucose[†]	8.40 (3.46, 20.37)	<0.001		
Albumin	1.68 (1.17, 2.40)	0.005		
HDL-cholesterol	0.96 (0.95, 0.97)	<0.001	0.98 (0.97, 0.99)	<0.001
CT_{L-S}	0.84 (0.83, 0.86)	<0.001	0.87 (0.85, 0.89)	<0.001

Note: Data in parentheses are range. CI = confidence interval, BMI = body mass index, AST = aspartate transaminase, ALT = alanine aminotransferase, ALP = alkaline phosphatase, HDL = high-density lipoprotein, CT_{L-S} = mean liver attenuation – mean spleen attenuation on non-enhanced CT.

[†]Analyzed after log transformation.

[‡]Multivariable analysis included candidate variables selected in more than 50% of logistic models from 1000-fold bootstrap resampling.

Table 3. Final logistic model for the diagnosis of nonalcoholic fatty liver disease

Parameters	Coefficient (95% CI)	*P*-value
CT_{L-S}	-0.14 (-0.16, -0.11)	<0.001
BMI	0.18 (0.13, 0.23)	<0.001
Log (triglyceride)	0.79 (0.56, 1.02)	<0.001
Log (ALT)	0.64 (0.35, 0.93)	<0.001

Note: CI = confidence interval, BMI = body mass index, ALT = alanine aminotransferase.

Table 4. Cutoff values for clinical-CT index and CT_{L-S} and their corresponding diagnostic performances in the development and test cohorts

		Development cohort				Test cohort			
Indices	Cutoffs	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Clinical-CT index	≥46	90.3 (560/620)	44.2 (380/860)	53.8 (560/1040)	86.4 (380/440)	89.4 (277/310)	40.7 (174/428)	52.2 (277/531)	84.1 (174/207)
	≥56.5	57.4 (356/620)	90.2 (776/860)	80.9 (356/440)	74.6 (776/1040)	57.1 (177/310)	88.6 (379/428)	78.3 (177/226)	74.0 (379/512)
CT_{L-S}	≤12.5	90.2 (559/620)	27.6 (237/860)	47.3 (559/1182)	79.5 (237/298)	89.0 (276/310)	26.9 (115/478)	46.9 (276/589)	77.2 (115/149)
	≤3.9	44.0 (273/620)	90.0 (774/860)	76.0 (273/359)	69.0 (774/1121)	46.1 (143/310)	90.2 (386/478)	77.3 (143/185)	69.8 (386/553)

Note: The upper cutoff values are those for 90% sensitivity, and the lower cutoff values are those for 90% specificity. Results are presented as percentages (number of patients/total number of patients assessed). PPV = positive predictive value, NPV = negative predictive value.

Supplementary Table 1. Results of bootstrap multivariable logistic regression analysis for selecting candidate predictors

Variables	Frequency*
Age	535 (53.5%)‡
Sex (female)	60 (6.0%)
BMI	1000 (100%)‡
AST	241 (24.1%)
ALT	904 (90.4%)‡
Bilirubin	156 (15.6%)
ALP	425 (42.5%)
Triglyceride	912 (91.2%)‡
Cholesterol	932 (93.2%)‡
Glucose	842 (84.2%)‡
Albumin	108 (10.8%)
HDL-cholesterol	801 (80.1%)‡
CT_{L-S}	1000 (100%)‡

Note: Data in parentheses are percentages. BMI = body mass index, AST = aspartate transaminase, ALT = alanine aminotransferase, ALP = alkaline phosphatase, HDL = high-density lipoprotein, CT_{L-S} = mean liver attenuation – mean spleen attenuation on non-enhanced CT.

†Number of variables included in 1000 bootstrap logistic models.

‡Variables selected as candidate predictors

Supplementary Table 2. Diagnostic performance of logistic models with clinical variables sequentially added to CT_{L-S} for predicting nonalcoholic fatty liver disease

Number of variables	Variables in model	AUC
One	CT_{L-S}	0.794
Two	Baseline: CT_{L-S}	0.794
	+ BMI*	0.794*
	+ Log (ALT)	0.783
	+ Log (triglyceride)	0.789
	+ Cholesterol	0.772
	+ HDL-cholesterol	0.769
Three	Baseline: CT_{L-S} and BMI	0.794
	+ Log (ALT)	0.806
	+ Log (triglyceride)*	0.817*
	+ Cholesterol	0.810
	+ HDL-cholesterol	0.805
Four	Baseline: CT_{L-S}, BMI, and log (triglyceride)	0.817
	+ Log (ALT)*	0.823*
	+ Cholesterol	0.822
	+ HDL-cholesterol	0.821
Five	Baseline: **CT_{L-S}, BMI, log (triglyceride), and log (ALT)**†	0.823
	+ Cholesterol	0.827
	+ HDL-cholesterol	0.825

Note: Logistic models including an increasing number of variables were sequentially developed by adding the independent variables to CT_{L-S} in a one-by-one manner. Among the models with the same number of variables, the model with the highest AUC was selected and utilized in the next step. This procedure continued until addition of an additional variable did not improve model performance by an AUC of 0.005. BMI = body mass index, AST =

aspartate transaminase, ALT = alanine aminotransferase, ALP = alkaline phosphatase, HDL = high-density lipoprotein.

*Model selected at each step.

†Variables for the final logistic model.

Figures and Figure legends

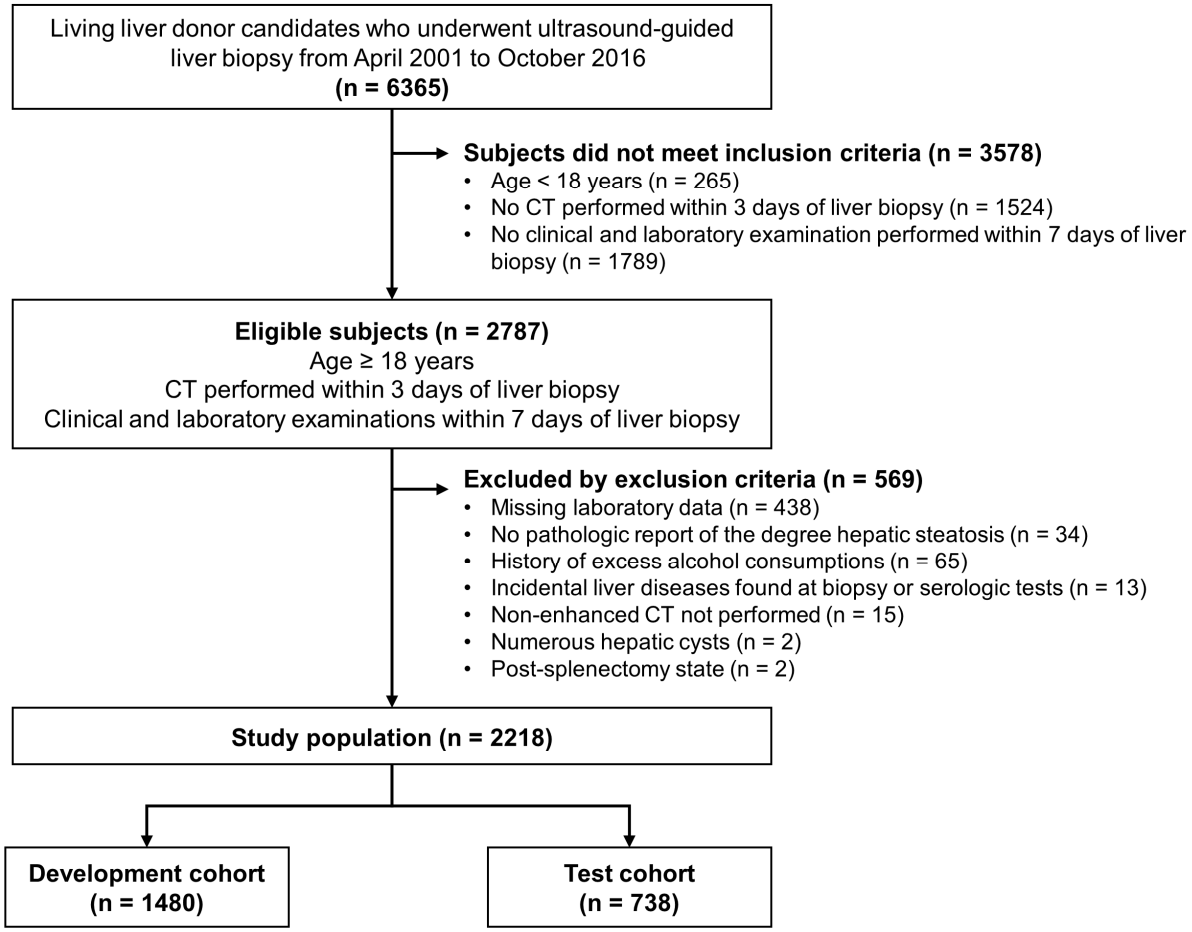

Fig 1. Flow diagram of the study population.

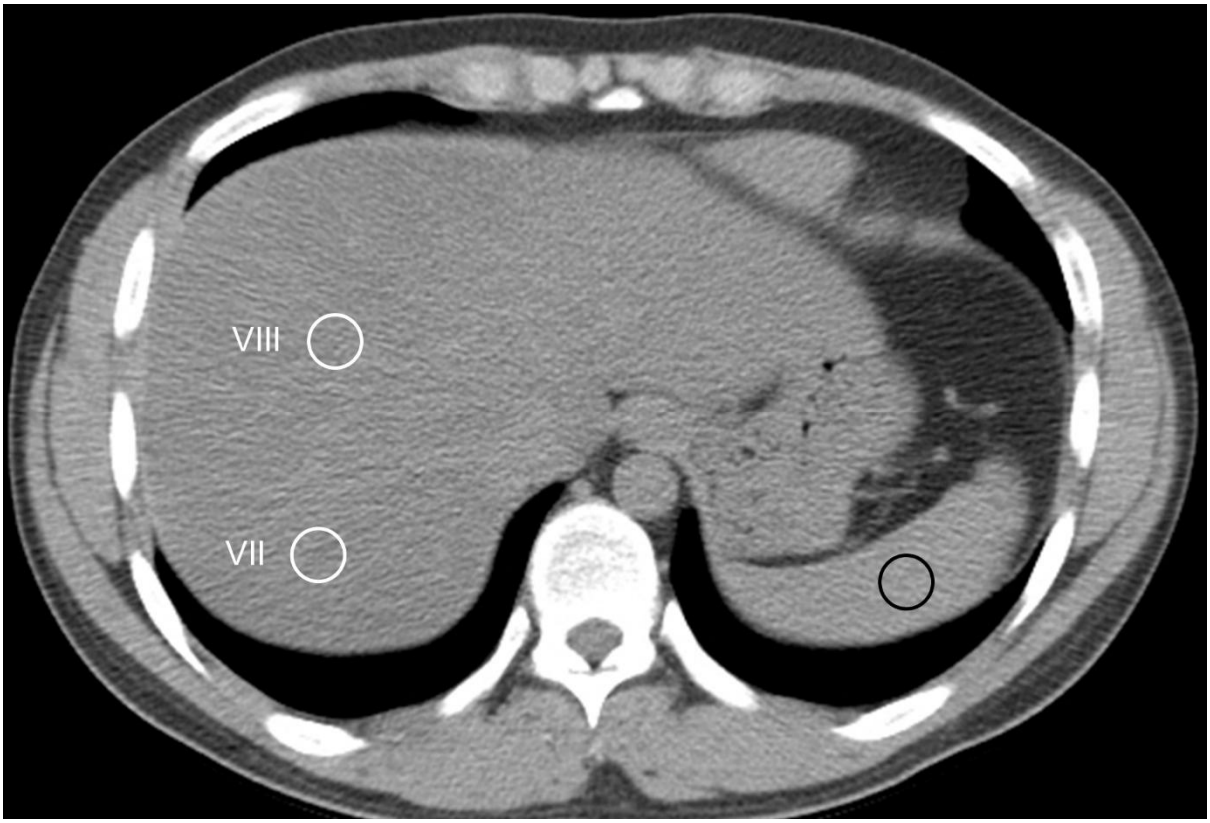

Fig 2. Measurement of liver and spleen attenuation on non-enhanced axial CT images. Two 1.5 cm circular ROIs (white circles) were placed on hepatic segments VIII and VII, which were devoid of macroscopic vessels. A 1.5 cm circular ROI was positioned on the spleen (black circle).

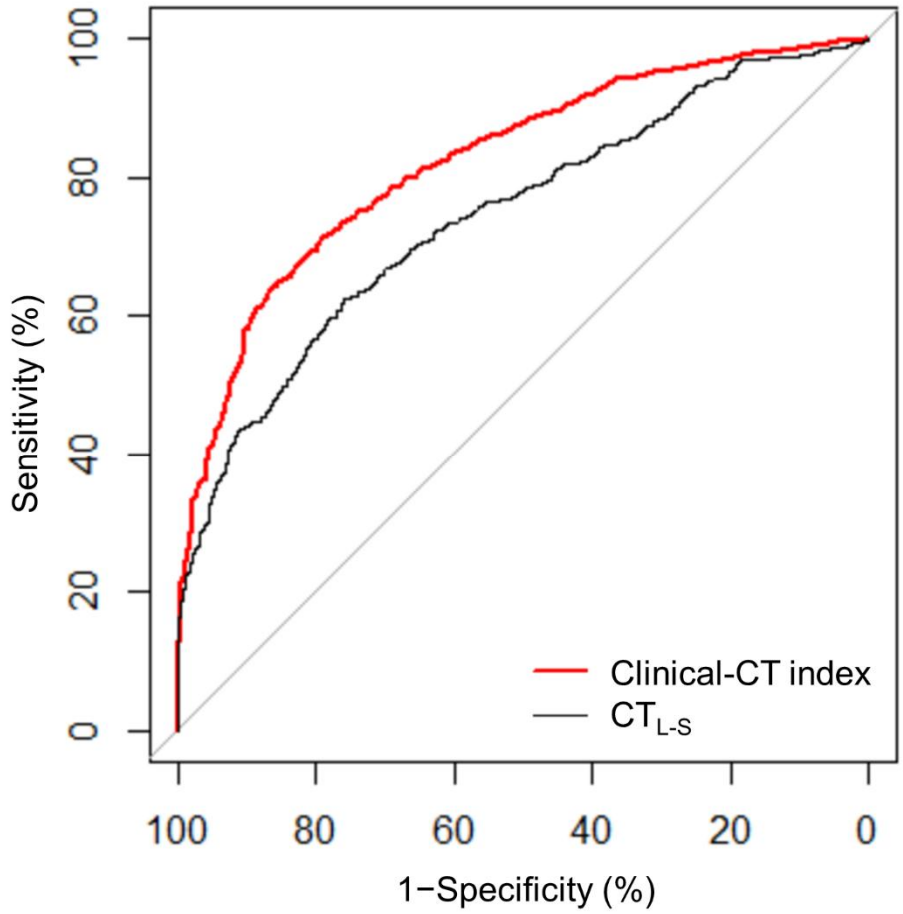

Fig 3A. Receiver operating characteristic (ROC) curves of the performance of the clinical-CT index in diagnosing nonalcoholic fatty liver disease, comparison with $CT_{L\text{-}S}$ in the development cohort (A) and compared with $CT_{L\text{-}S}$, hepatic steatosis index (HS-I), and fatty liver disease index (FLD-1) in the test cohort (B).

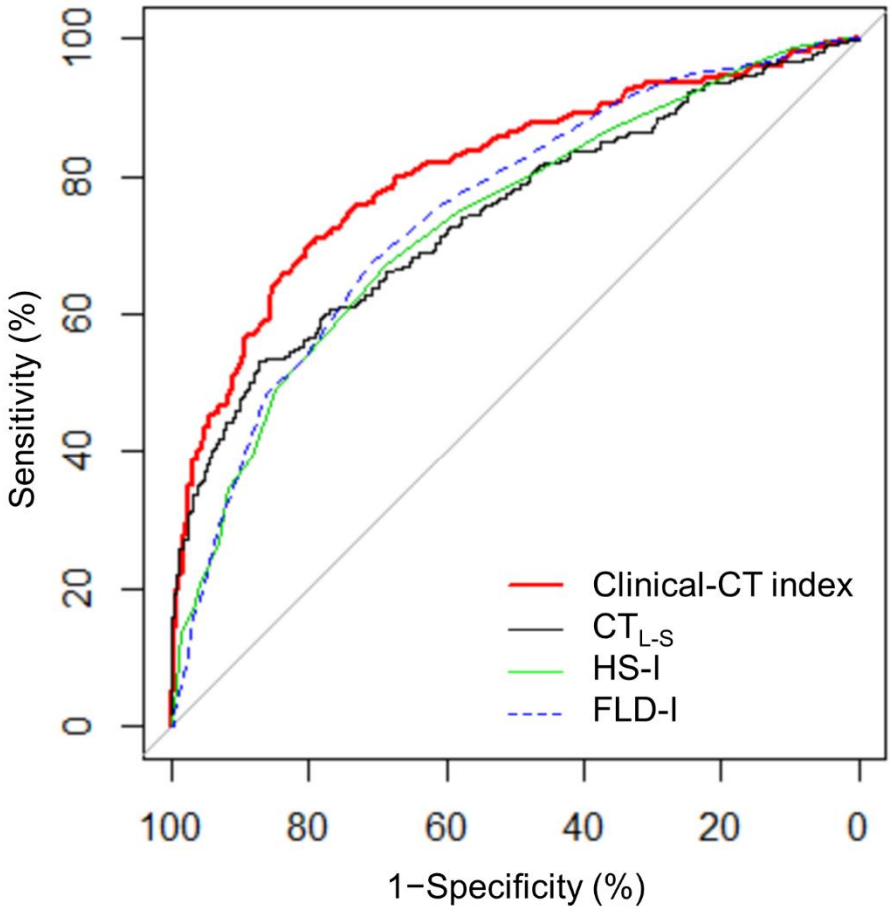

Fig 3B. Receiver operating characteristic (ROC) curves of the performance of the clinical-CT index in diagnosing nonalcoholic fatty liver disease, comparison with $CT_{L\text{-}S}$ in the development cohort (A) and compared with $CT_{L\text{-}S}$, hepatic steatosis index (HS-I), and fatty liver disease index (FLD-1) in the test cohort (B).

국문요약

비조영강증강 전산화 단층촬영과 임상 요인에 기반한 비알콜성 지방간 질환 예측 지수의 개발과 검증

안유라, 이승수
울산대학교 의과대학 서울아산병원 영상의학과

목적: 전산화 단층촬영(CT) 소견과 기본 임상 데이터를 이용하여 대규모 성인 코호트에서 비알코올성 지방간 질환(NAFLD)의 신뢰성 있는 진단과 병리학적으로 진단된 NAFLD 를 배제할 수 있는 예측 지수를 개발하고 검증하려 한다.

연구방법: CT 와 간조직검사를 3 일 이내에 시행한 2218 명의 성인 간 기증자를 대상으로 후향적 연구를 진행하였다. 간 기증자들을 2:1 의 비율로 개발군과 실험군으로 무작위 배정하였다. 비조영증강 CT 에서 간의 Hounsfield unit (HU) 로부터 비장의 HU 과 차이를 구하여 CTL-S 를 구하였다. 개발군에서 임상 요인 중 후보 예측 요인들을 multivariable logistic regression analysis 를 이용하여 선정하여, 병리학적으로 진단된 NAFLD 를 예측하는 Clinical-CT index 를 개발하였다. Areas under the receiver operating characteristic curve (AUC) 분석으로 진단능을 평가하였으며 개발군에서 90%의 민감도와 90%의 특이도에 해당하는 clinical-CT index 의 cutoff 를 설정하였고 실험군에서 진단능을 평가하였다.

연구결과: 임상 요인으로는 body mass index, aspartate transaminase, 그리고 triglyceride concentrations 가 clinical-CT index 에 포함되었다. 실험군에서 clinical-CT index 의 NAFLD 의 진단능은 (AUC, 0.81) CTL-S 단독과 (0.74; $P < 0.001$) 기존 clinical indices (0.73–0.75; $P < 0.001$) 를 능가하였다. Cutoff ≥46 에서 민감도와 특이도는 각각 89%와 41% 였으며 cutoff ≥56.5 에서 민감도와 특이도는 각각 57%와 89% 였다.

결론: Clinical-CT index 는 NAFLD 를 진단하는 데 있어 CTL-S 또는 clinical index 단독보다 정확하며 NAFLD 를 진단하거나 배제하는데 있어 임상적으로 유용할 것으로 기대된다.