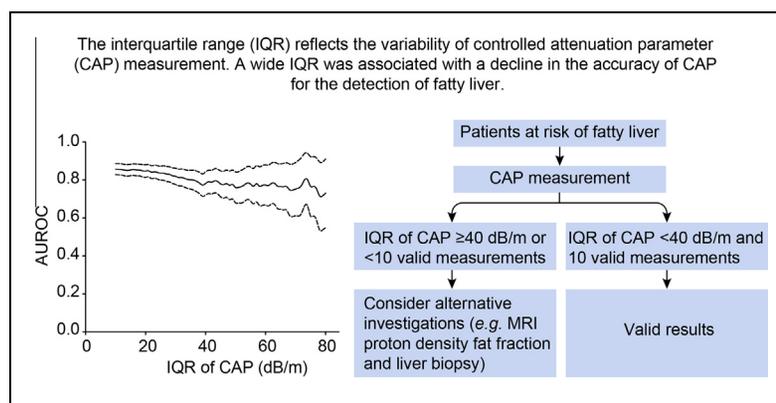


Validity criteria for the diagnosis of fatty liver by M probe-based controlled attenuation parameter

Graphical abstract



Highlights

- Controlled attenuation parameter (CAP) can detect fatty liver with moderate accuracy.
- The interquartile range (IQR) of CAP reflects its measurement variability.
- The accuracy of CAP declines when its IQR exceeds 40 dB/m.
- The accuracy of CAP is not affected by high transaminase or bilirubin levels.

Authors

Vincent Wai-Sun Wong, Salvatore Petta, Jean-Baptiste Hiriart, ..., Umberto Arena, Antonio Craxì, Victor de Lédinghen

Correspondence

wongv@cuhk.edu.hk (V.W.-S. Wong), victor.deledinghen@chu-bordeaux.fr (V. de Lédinghen)

Lay summary

Controlled attenuation parameter (CAP) is measured by transient elastography (TE) for the detection of fatty liver. In this large study, using liver biopsy as a reference, we show that the variability of CAP measurements based on its interquartile range can reflect the accuracy of fatty liver diagnosis. In contrast, other clinical factors such as adiposity and liver enzyme levels do not affect the performance of CAP.

Validity criteria for the diagnosis of fatty liver by M probe-based controlled attenuation parameter

Vincent Wai-Sun Wong^{1,2,*†}, Salvatore Petta^{3,†}, Jean-Baptiste Hiriart⁴, Calogero Cammà³, Grace Lai-Hung Wong^{1,2}, Fabio Marra⁵, Julien Vergniol⁴, Anthony Wing-Hung Chan⁶, Antonino Tuttolomondo⁷, Wassil Merrouche⁴, Henry Lik-Yuen Chan^{1,2}, Brigitte Le Bail^{8,9}, Umberto Arena⁵, Antonio Craxi³, Victor de Lédینگhen^{4,8,*}

¹Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong; ²State Key Laboratory of Digestive Disease, The Chinese University of Hong Kong, Hong Kong; ³Sezione di Gastroenterologia, Di.Bi.M.I.S., University of Palermo, Palermo, Italy; ⁴Centre d'Investigation de la Fibrose Hépatique, Hôpital Haut-Lévêque, Bordeaux University Hospital, Pessac, France; ⁵Dipartimento di Medicina Sperimentale e Clinica, Università degli Studi di Firenze, Florence, Italy; ⁶Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Hong Kong; ⁷Sezione di Medicina Interna e Cardioangiologia, Di.Bi.M.I.S., University of Palermo, Palermo, Italy; ⁸INSERM U1053, Bordeaux University, Bordeaux, France; ⁹Service de Pathologie, Hôpital Pellegrin, Bordeaux University Hospital, Bordeaux, France

Background & Aims: Controlled attenuation parameter (CAP) can be performed together with liver stiffness measurement (LSM) by transient elastography (TE) and is often used to diagnose fatty liver. We aimed to define the validity criteria of CAP.

Methods: CAP was measured by the M probe prior to liver biopsy in 754 consecutive patients with different liver diseases at three centers in Europe and Hong Kong (derivation cohort, n = 340; validation cohort, n = 414; 101 chronic hepatitis B, 154 chronic hepatitis C, 349 non-alcoholic fatty liver disease, 37 autoimmune hepatitis, 49 cholestatic liver disease, 64 others; 277 F3-4; age 52 ± 14 ; body mass index 27.2 ± 5.3 kg/m²). The primary outcome was the diagnosis of fatty liver, defined as steatosis involving $\geq 5\%$ of hepatocytes.

Results: The area under the receiver-operating characteristics curve (AUROC) for CAP diagnosis of fatty liver was 0.85 (95% CI 0.82–0.88). The interquartile range (IQR) of CAP had a negative correlation with CAP ($r = -0.32$, $p < 0.001$), suggesting the IQR-to-median ratio of CAP would be an inappropriate validity parameter. In the derivation cohort, the IQR of CAP was associated with the accuracy of CAP (AUROC 0.86, 0.89 and 0.76 in patients with IQR of CAP < 20 [15% of patients], 20–39 [51%], and ≥ 40 dB/m [33%], respectively). Likewise, the AUROC of CAP in the validation cohort was 0.90 and 0.77 in patients with IQR

of CAP < 40 and ≥ 40 dB/m, respectively ($p = 0.004$). The accuracy of CAP in detecting grade 2 and 3 steatosis was lower among patients with body mass index ≥ 30 kg/m² and F3-4 fibrosis.

Conclusions: The validity of CAP for the diagnosis of fatty liver is lower if the IQR of CAP is ≥ 40 dB/m.

Lay summary: Controlled attenuation parameter (CAP) is measured by transient elastography (TE) for the detection of fatty liver. In this large study, using liver biopsy as a reference, we show that the variability of CAP measurements based on its interquartile range can reflect the accuracy of fatty liver diagnosis. In contrast, other clinical factors such as adiposity and liver enzyme levels do not affect the performance of CAP.

© 2017 European Association for the Study of the Liver. Published by Elsevier B.V. All rights reserved.

Introduction

Non-alcoholic fatty liver disease (NAFLD) is currently the most common chronic liver disease worldwide and has become an important cause of end-stage liver disease and hepatocellular carcinoma.^{1–4} The presence of fatty liver and metabolic syndrome in patients with chronic viral hepatitis is also associated with increased risk of cirrhosis and hepatocellular carcinoma.^{5–8} Abdominal ultrasonography is commonly used to diagnose fatty liver, but it cannot reliably diagnose mild steatosis, and its performance is suboptimal in obese patients. Recently, the controlled attenuation parameter (CAP) was developed as a new test for fatty liver. It is based on the physical phenomenon that the amplitude of ultrasound waves is attenuated more rapidly when they traverse across a steatotic liver. In previous studies, CAP had moderate to good accuracy for fatty liver detection, when compared to histology or magnetic resonance spectroscopy.^{9–13} CAP is measured simultaneously with liver stiffness (LSM) using

Keywords: FibroScan; Liver stiffness measurement; Non-alcoholic fatty liver disease; Hepatic steatosis; Liver biopsy; Diagnostic accuracy.

Received 3 January 2017; received in revised form 3 April 2017; accepted 1 May 2017; available online 12 May 2017

* Corresponding authors. Addresses: Centre d'Investigation de la Fibrose Hépatique, Service d'Hépatogastroentérologie, Hôpital Haut-Lévêque, 33604 Pessac, France. Tel.: +33 557656439; fax: +33 557656445 (V. de Lédینگhen) or Department of Medicine and Therapeutics, 9/F, Clinical Sciences Building, Prince of Wales Hospital, 30-32 Ngan Shing Street, Shatin, Hong Kong. Tel.: +852 26321205; fax: +852 26373852 (V. Wong).

E-mail addresses: wongv@cuhk.edu.hk (V.W.-S. Wong), victor.deledinghen@chu-bordeaux.fr (V. de Lédینگhen).

[†] These authors contributed equally to this work.



ELSEVIER

Research Article

transient elastography (TE). It is thus possible to diagnose fatty liver and assess the disease severity at the same time.

A TE examination involves 10 measurements, with the median CAP and LSM taken as the estimates for liver fat and fibrosis, respectively, and the interquartile range (IQR) as the dispersion or fluctuation of the measurements. For LSM, the IQR-to-median ratio is a well-recognized parameter for determining the validity of measurement.¹⁴ A high IQR-to-median ratio reflects inconsistent results from the 10 measurements and is associated with less accurate results. However, the validity criteria for CAP are undefined. It is therefore difficult for clinicians to interpret CAP results.

In this study, we aim to determine factors associated with less accurate CAP measurements and define the validity criteria for CAP.

Patients and methods

Patients

This was a cross-sectional study of a prospective cohort of adult patients aged 18 years or above who underwent liver biopsy for the evaluation of chronic liver diseases at two European centers and one Hong Kong center. We excluded patients with less than 10 successful CAP/LSM acquisitions and those with liver biopsy specimens shorter than 15 mm. The study protocol was approved by all participating centers. All patients provided informed written consent. Part of this cohort (94 patients from Hong Kong and 124 patients from Bordeaux) was described in previous publications^{15,16} and was included in a recent meta-analysis on the performance of CAP.¹⁷

To determine factors associated with the accuracy of CAP and to test the validity criteria, we divided the patients into the derivation (France) and validation (Italy and Hong Kong) cohorts. The latter allowed validation of the study findings in a multi-ethnic cohort.

Since the main application of CAP is to diagnose NAFLD, we further evaluated the performance of CAP and the validity criteria by comparing NAFLD patients with healthy controls, who participated in a population screening project using proton-magnetic resonance spectroscopy and TE.¹⁸ The controls had no fatty liver, as evidenced by an intrahepatic triglyceride content of <5% on two occasions, three to five years apart.

Clinical assessment

Medical history was obtained and physical examination was performed for all patients. Body mass index (BMI) was calculated as body weight (kg) divided by body height (m) squared. The liver etiologies were determined by history, viral hepatitis serology (hepatitis B surface antigen and anti-hepatitis C virus antibody), autoimmune markers and liver histology. Liver biochemistry, complete blood count and international normalized ratio were checked less than one week before liver biopsy.

Histological assessment

Liver biopsy was performed using 16G trucut needles and the specimens were evaluated by experienced pathologists. Steatosis grade was determined according to the Non-alcoholic Steatohepatitis Clinical Research Network (NASH CRN) scoring system, based on parenchymal involvement by steatosis under low- to medium-power evaluation: S0 = <5%, S1 = 5–33%, S2 = >33–66%, and S3 = >66%.¹⁹ Patients with ≥5% steatosis were considered to have fatty liver. Because it would be inappropriate to use the same fibrosis staging system in patients with different liver diseases, we classified the patients as having no to moderate fibrosis (F0–2 by the METAVIR or NASH CRN systems), bridging fibrosis (F3) and cirrhosis (F4).

Transient elastography (TE)

TE (FibroScan, Echosens, Paris, France) was performed less than one week before liver biopsy, after fasting for at least 6 h. All operators had undergone formal

training and performed at least 100 examinations before this study. Measurements were performed using the M probe on the right lobe of the liver, through intercostal spaces, with the patient lying in a dorsal position with the right arm in maximal abduction.²⁰ Only cases with 10 or more successful acquisitions were evaluated. The success rate was calculated as the number of successful measurements divided by the total number of measurements. The operators were blinded to the clinical data and diagnosis.

Statistical analysis

Statistical tests were performed using the IBM SPSS Statistics 23 (IBM Corporation, Armonk, NY). Continuous variables were expressed as mean ± standard deviation, or median (interquartile range) if the data were not normally distributed. The primary outcome was fatty liver, defined as hepatic steatosis of 5% or more by histology. Receiver-operating characteristics curves were constructed to assess the accuracy of CAP for detecting fatty liver. This study focused on the detection of fatty liver rather than individual steatosis grades, because the severity of steatosis correlates poorly with liver injury and adverse clinical outcomes,^{21,22} and previous studies showed considerable overlap in CAP values across different steatosis grades.^{9–13} The area under the receiver-operating characteristics curves (AUROC) was compared using the method by Hanley and McNeil.²³ The correlation between CAP and the potential quality indicators of TE examination was determined using the Pearson's coefficients. All statistical tests were two-sided. Significance was taken as $p < 0.05$. With a sample size of 750 patients, the AUROC of CAP could be determined at a standard error of 0.01 to 0.04.

For further details regarding the materials used, please refer to the [Supplementary material](#) and the [CTAT table](#).

Results

From May 2009 to September 2016, 1,036 patients underwent paired TE and liver biopsy. After excluding patients with suboptimal liver biopsy specimens and TE examination, 754 patients were included in the final analysis (340 from the French cohort, 203 from the Italian cohort, and 211 from the Hong Kong cohort) ([Table 1](#), [Fig. S1](#)). The excluded patients were older, more likely to have diabetes, hypertension and high BMI, but had lower albumin, alanine aminotransferase (ALT) and platelet count ([Table S1](#)). The mean age of the patients included was 52 ± 14 years, and the BMI was 27.2 ± 5.3 kg/m² ([Table S1](#)). One hundred and thirty-six (18%) and 141 (19%) patients had bridging fibrosis and cirrhosis, respectively. The Hong Kong cohort was enriched with patients with NAFLD and diabetes. In contrast, patients from the Italian cohort had less severe hepatic steatosis and fibrosis. Overall, 101 patients had chronic hepatitis B, 154 had chronic hepatitis C, 349 had NAFLD, 37 had autoimmune hepatitis, 49 had cholestatic liver disease, and 64 had other liver diseases. [Table S2](#) shows the clinical characteristics of patients with different liver diseases.

Performance of CAP

CAP had good overall accuracy in detecting fatty liver, with an AUROC of 0.85 ([Fig. S2](#)). The accuracy of CAP was highest in Hong Kong ($p = 0.008$ vs. France and $p = 0.002$ vs. Italy), likely reflecting the difference in the proportion of patients with moderate to severe steatosis. Based on previously published cut-offs of 222 and 290 dB/m, CAP had a sensitivity of 87.8% and a specificity of 90.5% for detecting fatty liver, respectively ([Table 2](#)).

Among 435 patients with recorded abdominal ultrasonography results, the AUROC of ultrasonography for diagnosing fatty liver was 0.76 (95% confidence interval [CI] 0.72–0.81). The sensitivity, specificity, positive predictive value and negative predictive value were 65%, 88%, 92% and 54%, respectively. Using

Table 1. Patient characteristics.

Characteristics	All	Derivation cohort		Validation Cohort	
		France	Italy	Hong Kong	
N	754	340	203	211	
Age (years)	52 ± 14	53 ± 14	51 ± 14	52 ± 13	
Male sex, n (%)	415 (55)	193 (57)	104 (51)	118 (56)	
Diabetes, n (%) [†]	224 (30)	85 (25)	34 (17)	105 (50)	
Hypertension, n (%) [†]	288 (38)	130 (38)	58 (29)	100 (47)	
Body mass index (kg/m ²) [†]	27.2 ± 5.3	27.0 ± 5.5	30.0 ± 5.9	26.6 ± 4.6	
Liver disease, n (%) [†]					
Chronic hepatitis B	101 (13)	46 (14)	15 (7)	40 (19)	
Chronic hepatitis C	154 (20)	105 (31)	44 (22)	5 (2)	
Non-alcoholic fatty liver disease	349 (46)	132 (39)	79 (39)	138 (65)	
Autoimmune hepatitis	37 (5)	14 (4)	18 (9)	5 (2)	
Cholestatic liver disease	49 (7)	18 (5)	24 (12)	7 (3)	
Others	64 (9)	25 (7)	23 (11)	16 (8)	
Albumin (g/L) [†]	43 ± 5	41 ± 5	43 ± 4	44 ± 4	
Bilirubin (μmol/L)	12 (9–16)	11 (9–15)	12 (8–16)	12 (9–16)	
Alanine aminotransferase (IU/L) [†]	64 (39–108)	73 (44–117)	59 (39–105)	56 (34–86)	
Platelet count (x10 ⁹ /L) [†]	213 ± 76	199 ± 68	222 ± 78	226 ± 84	
INR [†]	1.0 ± 0.2	1.1 ± 0.3	1.0 ± 0.1	1.0 ± 0.1	
Length of liver biopsy specimen (mm) [†]	24 (20–28)	27 (20–33)	20 (19–26)	23 (18–27)	
Steatosis grade, n (%) [†]					
0	221 (29)	86 (25)	80 (39)	55 (26)	
1	245 (33)	116 (34)	63 (31)	66 (31)	
2	145 (19)	57 (17)	31 (15)	57 (27)	
3	143 (19)	81 (24)	29 (14)	33 (16)	
Fibrosis, n (%) [†]					
No to moderate	477 (63)	207 (61)	155 (76)	115 (55)	
Bridging fibrosis	136 (18)	77 (23)	26 (13)	33 (16)	
Cirrhosis	141 (19)	56 (17)	22 (11)	63 (30)	
Success rate of TE measurements (%) [†]	100 (83–100)	100 (91–100)	91 (71–100)	100 (83–100)	
Liver stiffness (kPa)					
Median	8.4 (6.1–12.7)	8.6 (6.1–12.0)	7.9 (5.9–12.0)	8.9 (6.3–14.1)	
IQR	1.3 (0.8–2.3)	1.3 (0.8–2.3)	1.4 (0.8–2.5)	1.1 (0.7–2.1)	
IQR/M [†]	0.16 (0.10–0.22)	0.16 (0.11–0.22)	0.18 (0.12–0.25)	0.13 (0.09–0.19)	
Controlled attenuation parameter (dB/m)					
Median [†]	279 (218–327)	273 (218–325)	252 (200–312)	300 (238–341)	
IQR [†]	31 (22–44)	32 (23–43)	36 (24–50)	26 (20–40)	
IQR/M [†]	0.12 (0.07–0.18)	0.12 (0.08–0.17)	0.13 (0.09–0.22)	0.10 (0.06–0.16)	

Continuous variables were expressed in mean ± standard deviation or median (interquartile range).

IQR, interquartile range; IQR/M, interquartile range-to-median ratio; TE, transient elastography.

* p = 0.001 among the centers.

† p < 0.001 among the centers.

Table 2. Diagnostic accuracy of CAP for detecting fatty liver (hepatic steatosis ≥5%) according to the cut-offs proposed by Sasso *et al.*⁹

CAP	Fatty liver	No fatty liver	Total
Total	533	221	754
<222 dB/m, n	65	137	202
≥222 dB/m, n	468	84	552
Sensitivity (%)	87.8 (85.0–90.6)		
Specificity (%)	62.0 (55.6–68.4)		
Positive predictive value (%)	84.8 (81.8–87.8)		
Negative predictive value (%)	67.8 (61.4–74.3)		
<290 dB/m, n	212	200	412
≥290 dB/m, n	321	21	342
Sensitivity (%)	60.2 (56.1–64.4)		
Specificity (%)	90.5 (86.6–94.4)		
Positive predictive value (%)	93.9 (91.3–96.4)		
Negative predictive value (%)	48.5 (43.7–53.4)		

Numbers in brackets represent 95% confidence intervals.

ultrasonography as the reference standard, the AUROC of CAP for detecting radiological fatty liver was 0.79 (95% CI 0.75–0.83).

Because the main clinical application of CAP is to diagnose NAFLD, we also tested the performance of CAP in patients with biopsy-proven NAFLD, with reference to healthy controls with intrahepatic triglyceride content of <5%, by proton-magnetic resonance spectroscopy (Table S3). As expected, patients with NAFLD had higher BMI values and were more likely to have diabetes and hypertension. They also had higher LSM and CAP values. Overall, the AUROC of CAP for diagnosing NAFLD was 0.88 (95% CI 0.85–0.90). At the cut-off of 222 dB/m, the sensitivity, specificity, positive predictive value and negative predictive value of CAP, for diagnosing NAFLD, were 93.1%, 45.1%, 57.5% and 89.1%, respectively. The corresponding values at the cut-off of 290 dB/m were 73.6%, 89.9%, 85.4% and 81.0%, respectively.

Research Article

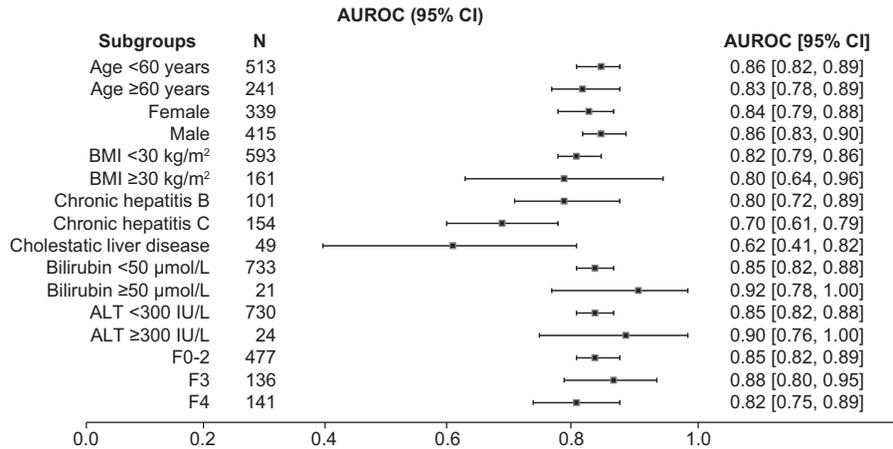


Fig. 1. AUROC of CAP for detecting fatty liver in different subgroups. AUROC, area under receiver-operating characteristic curve; CAP, controlled attenuation parameter.

Factors affecting the performance of CAP

The performance of CAP was consistent across various subgroups. The AUROC of CAP for detecting fatty liver was similar in patients of different ages and gender (Fig. 1). In addition, traditional factors affecting the performance of LSM, like BMI, bilirubin and serum alanine aminotransferase (ALT) level did not affect the performance of CAP. The performance of CAP was also stable across fibrosis stages. On the other hand, the accuracy of CAP appeared lower in patients with chronic hepatitis C and cholestatic liver disease, though the number of subjects was small in the latter group.

In patients with chronic hepatitis B and C, age, gender, BMI, bilirubin, ALT and fibrosis stage did not affect the performance of CAP (Table S4). While the same factors did not affect the diagnosis of NAFLD, using healthy subjects with an intrahepatic triglyceride content of <5% as controls, the accuracy of CAP in detecting grade 2 and 3 steatosis was lower among patients with BMI ≥30 kg/m² and F3-4 fibrosis (Table S5). At each steatosis grade, patients with BMI ≥30 kg/m² had higher CAP values than those <30 kg/m² (S1: 321 [272–352] vs. 278 [231–302], $p = 0.003$; S2: 335 [302–361] vs. 320 [296–343], $p = 0.060$; 350 [326–365] vs. 327 [311–348], $p = 0.002$). On the other hand, fibrosis stage did not affect CAP values after adjusting for steatosis grade (data not shown).

The IQR-to-median ratio of LSM is the most important validity indicator of LSM.¹⁴ We therefore evaluated the meaning of the

IQR of CAP. In the overall population, CAP had a mild negative correlation with the IQR of CAP (Fig. 2A). As a result, adjustment using the IQR-to-median ratio of CAP potentiated the negative correlation with CAP (Fig. 2B). This contrasted with LSM, in which the IQR of LSM increased with increasing LSM (Fig. 2C). Adjustment using the IQR-to-median ratio of LSM normalized the association with LSM, making it a reasonable quality indicator across the entire range of LSM (Fig. 2D). Thus, the IQR of CAP instead of its IQR-to-median ratio would be a better marker of measurement variability.

In the derivation cohort, the accuracy of CAP was lower in patients with IQR of CAP ≥40 dB/m (AUROC 0.76 vs. 0.86 in those with IQR of CAP <20 dB/m and 0.89 in those 20–39 dB/m; Table 3). In contrast, the success rate of measurements did not affect the accuracy of CAP. Because the IQR-to-median ratio of LSM is used to determine the validity of LSM, we also tested if it could reflect the validity of the entire examination, including CAP measurements. However, a higher IQR-to-median ratio of LSM did not affect the validity of CAP.

Using the cut-off of 222 dB/m, the sensitivity, specificity, positive and negative predictive values of CAP for fatty liver detection, in the derivation cohort, were 78.1%, 65.0%, 80.3% and 61.9%, respectively in patients with an IQR of CAP ≥40 dB/m (Table 4). The corresponding figures increased to 89.5%, 65.2%, 91.0% and 61.2%, respectively, in patients with an IQR of CAP <40 dB/m. Likewise, using the cut-off of 290 dB/m, the

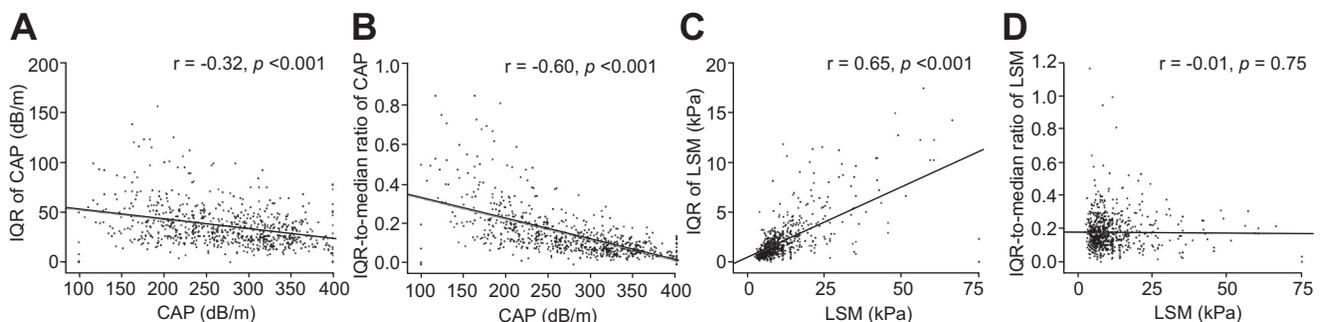


Fig. 2. Scattered plots showing the correlation between TE parameters. (A) CAP and the IQR of CAP, (B) CAP and the IQR-to-median ratio of CAP, (C) LSM and the IQR of LSM, and (D) LSM and the IQR-to-median ratio of LSM. CAP, controlled attenuation parameter; IQR, interquartile range; LSM, liver stiffness measurement.

Table 3. Performance of CAP according to potential quality indicators in the derivation cohort.

Quality indicators	n	AUROC
Success rate of measurements		
<60%	13	0.96 (0.83–1.00)
≥60%	327	0.83 (0.78–0.88)
<i>p</i> value (<60% vs. ≥60%)		0.071
IQR of CAP		
<20 dB/m	52	0.86 (0.76–0.96)
20–39 dB/m	175	0.89 (0.83–0.94)
≥40 dB/m	113	0.76 (0.68–0.86)
<i>p</i> (<20 vs. 20–39)		0.68
<i>p</i> (<20 vs. ≥40)		0.17
<i>p</i> (20–39 vs. ≥40)		0.027
IQR-to-median ratio of LSM		
<0.10	68	0.85 (0.75–0.95)
0.10–0.29	248	0.83 (0.77–0.89)
≥0.30	24	1.00 (1.00–1.00)
<i>p</i> (<0.10 vs. 0.10–0.29)		0.72
<i>p</i> (<0.10 vs. ≥0.30)		0.004
<i>p</i> (0.10–0.29 vs. ≥0.30)		<0.001

AUROC, area under the receiver-operating characteristics curves; CAP, controlled attenuation parameter; IQR, interquartile range; LSM, liver stiffness measurement.

corresponding figures were 46.6%, 87.5%, 87.2% and 47.3 in patients with IQR of CAP ≥40 dB/m; and 60.8%, 93.5%, 97.3% and 37.7% in patients with IQR of CAP <40 dB/m, respectively.

In the validation cohort, the AUROC of CAP for detecting fatty liver was 0.90 (95% CI 0.86–0.94) in 275 patients with an IQR of CAP of <40 dB/m, and 0.77 (95% CI 0.70–0.85) in 139 patients with an IQR of CAP of ≥40 dB/m (*p* = 0.004). The sensitivity of CAP values increased from 75.0% in patients with an IQR of CAP ≥40 dB/m, to 94.2% in patients with an IQR of CAP <40 dB/m, at a CAP cut-off of 222 dB/m (Table 4). The positive predictive value also increased from 72.0% to 85.5% at the same cut-off, and from 81.8% to 95.5% at the high cut-off of 290 dB/m.

As shown in Fig. 3, the AUROC of CAP decreased with increasing IQR of CAP, with the AUROC dropping below 0.80 when the IQR of CAP was over 40 dB/m.

Consistent findings were observed in the individual centers of the validation cohort. In the Italian cohort, the AUROC of CAP for detecting fatty liver was 0.83 (95% CI 0.75–0.90) in patients with an IQR of CAP of <40 dB/m, and 0.75 (95% CI 0.65–0.86) in those with an IQR of CAP of ≥40 dB/m. Likewise, in the Hong Kong cohort, the AUROC of CAP for detecting fatty liver was 0.96 (95% CI 0.92–0.99) in patients with an IQR of CAP of <40 dB/m, and 0.80 (95% CI 0.68–0.91) in those with an IQR of CAP of ≥40 dB/m.

The IQR of CAP did not correlate with fibrosis ($\rho = -0.010$, *p* = 0.79) in the overall population. Among NAFLD patients, the IQR of CAP had a borderline negative correlation with lobular inflammation ($\rho = -0.103$, *p* = 0.052) and ballooning ($\rho = -0.149$, *p* = 0.001).

Validation of the validity criteria in NAFLD

Similar to the main analysis, the AUROC of CAP for diagnosing NAFLD, with reference to healthy controls with intrahepatic triglyceride content of <5%, was higher among patients with an IQR of CAP <40 dB/m (0.89 [95% CI 0.86–0.92] vs. 0.81 [95% CI 0.75–0.87]). The sensitivity of CAP values increased from 84.9%

Table 4. Diagnostic accuracy of CAP to detect fatty liver according to the interquartile range (IQR) of CAP in the derivation and validation cohorts.

CAP	Fatty liver	No fatty liver	Total
Derivation cohort	254	86	340
IQR of CAP <40 dB/m	181	46	227
<222 dB/m, n	19	30	49
≥222 dB/m, n	162	16	178
Sensitivity (%)	89.5 (85.0–94.0)		
Specificity (%)	65.2 (51.5–79.0)		
Positive predictive value (%)	91.0 (86.8–95.2)		
Negative predictive value (%)	61.2 (47.6–74.9)		
<290 dB/m, n	71	43	114
≥290 dB/m, n	110	3	113
Sensitivity (%)	60.8 (53.7–67.9)		
Specificity (%)	93.5 (86.3–100.0)		
Positive predictive value (%)	97.3 (94.4–100.0)		
Negative predictive value (%)	37.7 (28.8–46.6)		
IQR of CAP ≥40 dB/m	73	40	113
<222 dB/m, n	16	26	42
≥222 dB/m, n	57	14	71
Sensitivity (%)	78.1 (68.6–87.6)		
Specificity (%)	65.0 (50.2–79.8)		
Positive predictive value (%)	80.3 (71.0–89.5)		
Negative predictive value (%)	61.9 (47.2–76.6)		
<290 dB/m, n	39	35	74
≥290 dB/m, n	34	5	39
Sensitivity (%)	46.6 (35.1–58.0)		
Specificity (%)	87.5 (77.3–97.7)		
Positive predictive value (%)	87.2 (76.7–97.7)		
Negative predictive value (%)	47.3 (35.9–58.7)		
Validation cohort	279	135	414
IQR of CAP <40 dB/m	207	68	275
<222 dB/m, n	12	35	47
≥222 dB/m, n	195	33	228
Sensitivity (%)	94.2 (91.0–97.4)		
Specificity (%)	51.5 (39.6–63.3)		
Positive predictive value (%)	85.5 (81.0–90.1)		
Negative predictive value (%)	74.5 (62.0–86.9)		
<290 dB/m, n	57	61	118
≥290 dB/m, n	150	7	157
Sensitivity (%)	72.5 (66.4–78.5)		
Specificity (%)	89.7 (82.5–96.9)		
Positive predictive value (%)	95.5 (92.3–98.8)		
Negative predictive value (%)	51.7 (42.7–60.7)		
IQR of CAP ≥40 dB/m	72	67	139
<222 dB/m, n	18	46	64
≥222 dB/m, n	54	21	75
Sensitivity (%)	75.0 (65.0–85.0)		
Specificity (%)	68.7 (57.5–79.8)		
Positive predictive value (%)	72.0 (61.8–82.2)		
Negative predictive value (%)	71.9 (60.9–82.9)		
<290 dB/m, n	45	61	106
≥290 dB/m, n	27	6	33
Sensitivity (%)	37.5 (26.3–48.7)		
Specificity (%)	91.0 (84.2–97.9)		
Positive predictive value (%)	81.8 (68.7–95.0)		
Negative predictive value (%)	57.5 (48.1–67.0)		

Numbers in brackets represent 95% confidence intervals.

to 95.3% cut-off in patients with an IQR of CAP <40 dB/m, at the CAP cut-off of 222 dB/m (Table S6). The positive predictive value increased from 44.3% to 61.9% at the same cut-off, and from 76.5% to 87.2% at the high cut-off of 290 dB/m.

Research Article

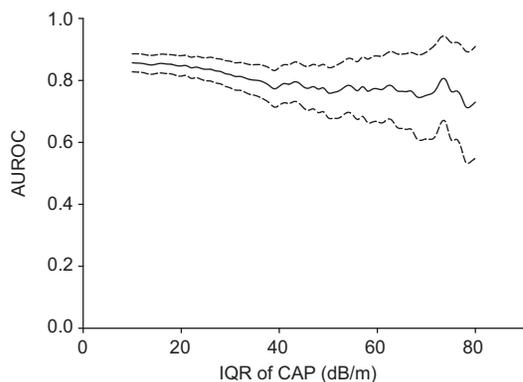


Fig. 3. Association between IQR of CAP and its accuracy in detecting fatty liver in the entire population. The solid line represents the AUROC for detecting fatty liver in patients with IQR of CAP above different levels. The dotted lines represent the 95% confidence interval of the AUROC. The 95% confidence interval widens with increasing IQR of CAP because fewer patients had such high IQR. AUROC, area under receiver-operating characteristic curve; CAP, controlled attenuation parameter; IQR, interquartile range.

Sensitivity analysis

The validity criteria were applied similarly across subgroups, except in those with small patient numbers (Table 5). Using different CAP cut-offs in the literature, the diagnostic accuracy of CAP was consistently better among patients with an IQR of CAP <40 dB/m (Table S7). Besides, if all 845 patients with liver biopsy length of ≥ 10 mm were included, an IQR of CAP <40 dB/m remained predictive of better diagnostic accuracy (AUROC 0.85 [95% CI 0.81–0.88] vs. 0.79 [95% CI 0.73–0.84]).

Discussion

CAP is performed together with LSM and can be used to detect fatty liver and assess the degree of liver injury simultaneously and conveniently. However, compared with LSM, the overall accuracy of CAP is lower, with a significant proportion of patients being misclassified.^{9–13,17} Head-to-head comparisons also showed that CAP is inferior to magnetic resonance imaging-based proton density fat fraction.²⁴ Nevertheless, because of

cost and availability, TE will likely remain a commonly used non-invasive test in clinical practice, and it is all the more important for identifying factors associated with the accuracy of CAP.

The IQR-to-median ratio has been the most important validity indicator for LSM.^{14,20} Normalization for the median LSM is important because the IQR of LSM increases with median LSM values. In stark contrast with LSM, we found that the IQR of CAP declined slightly with increasing median CAP values in this large biopsy cohort (Fig. 3). Thus, the IQR-to-median ratio of CAP cannot be used for normalization. Instead, it potentiates the negative correlation with median CAP and would only spuriously make CAP examination appear more accurate in patients with high CAP values. For this reason, a previous study involving 153 patients with BMI ≥ 28 kg/m² failed to show any impact of the IQR-to-median ratio on the accuracy of CAP.¹³

Among the quality indicators, the IQR of CAP turns out to be the most significant factor associated with the validity of CAP. The IQR of CAP reflects the variability of measurements and how well the examination is performed. An IQR of CAP of ≥ 40 dB/m was consistently associated with inferior accuracy in both the derivation and validation cohorts. Importantly, it also affected the accuracy of diagnosing NAFLD, which is the main clinical use of CAP.¹⁵ Besides, the new validity criteria performed similarly across different subgroups.

Although this study adopted the CAP cut-offs suggested by Sasso and colleagues,⁹ we acknowledge the selection of cut-offs is arbitrary and represents a compromise between sensitivity and specificity. As such, we further tested the validity criteria across a wide range of suggested cut-offs in the literature and confirmed that the validity criteria worked well regardless of the cut-offs (Table S5).

Furthermore, a few factors are well known to confound LSM. False-positive LSM can occur in patients with high ALT,²⁵ biliary obstruction,²⁶ congested liver,²⁷ and extreme BMI.²⁸ Patients with some of these characteristics were included in this study, but ALT, bilirubin and BMI did not affect the accuracy of CAP to detect fatty liver (Fig. 2). On the other hand, the accuracy of CAP to detect moderate to severe steatosis was lower in NAFLD patients with BMI ≥ 30 kg/m² and F3–4 fibrosis (Table S2). Since fibrotic stage did not influence the CAP values after adjusting for steatosis grade, the apparently poorer accuracy of CAP in patients with advanced fibrosis was probably due to the small number of

Table 5. AUROC of CAP for the diagnosis of fatty liver according to the IQR of CAP in selected subgroups.

Subgroups	n	AUROC		p value
		IQR of CAP <40 dB/m	IQR of CAP ≥ 40 dB/m	
Age <60	513	0.90 (0.86–0.93)	0.78 (0.71–0.85)	0.005
Age ≥ 60	241	0.87 (0.81–0.93)	0.76 (0.65–0.86)	0.073
Female	339	0.89 (0.84–0.93)	0.75 (0.65–0.84)	0.009
Male	415	0.89 (0.85–0.94)	0.79 (0.72–0.87)	0.020
BMI <30 kg/m ²	593	0.87 (0.83–0.90)	0.75 (0.68–0.81)	0.002
BMI ≥ 30 kg/m ²	161	0.88 (0.74–1.00)	0.50 (0.33–0.67)	<0.001
Chronic hepatitis B	101	0.85 (0.75–0.96)	0.70 (0.54–0.87)	0.134
Chronic hepatitis C	154	0.75 (0.65–0.86)	0.63 (0.49–0.78)	0.190
Cholestatic liver disease	49	0.62 (0.32–0.92)	0.61 (0.31–0.92)	0.979
Bilirubin <50 μ mol/L	733	0.89 (0.86–0.92)	0.77 (0.71–0.83)	<0.001
Bilirubin ≥ 50 μ mol/L	21	0.88 (0.65–1.00)	1.00 (1.00–1.00)	0.317
ALT <300 IU/L	730	0.89 (0.86–0.92)	0.77 (0.71–0.83)	<0.001
ALT ≥ 300 IU/L	24	0.94 (0.81–1.00)	0.89 (0.66–1.00)	0.680

ALT, alanine aminotransferase; BMI, body mass index.

subjects in this subgroup. In contrast, high BMI was consistently associated with higher CAP values after adjusting for steatosis grade. Previously, the skin capsular distance has also been shown to be associated with increased CAP values.²⁹ Since BMI and skin capsular distance are both surrogate markers of adiposity, it is difficult to determine the mechanism underlying the association. In any case, while CAP remains robust in the diagnosis of fatty liver across different subgroups, one should exercise caution in using CAP to determine steatosis grading in obese patients.

Our study has the strength of a large sample size, the use of histological correlation and the inclusion of patients from different ethnic backgrounds. The consistent findings across subgroups and participating centers add weight to our conclusions. Nonetheless, there are also a few limitations. Firstly, although the histological slides were scored by one experienced pathologist at each center, there was no central pathologist to review all of the cases. However, compared with other histological features, the assessment of hepatic steatosis tends to be robust.³⁰ We also included only satisfactory liver samples to minimize bias. Secondly, the study included patients with different etiologies. However, previous studies did not show a significant impact of specific liver diseases on the CAP value,^{10–13} and our subgroup analysis in the NAFLD cohort yielded consistent results. Thirdly, for historical reasons, CAP was measured using the M probe. Since patients with fatty liver have higher BMI, the applicability of the M probe is lower. To this end, the XL probe has been developed to cater for obese patients.^{31–33} CAP can also be measured by the XL probe now; its performance and validity criteria warrant further studies.³⁴ Finally, NAFLD was overrepresented in this cohort. Although consistent findings were observed across subgroups, one should exercise caution when extrapolating the results to other liver diseases. That said, in clinical practice, the diagnosis of NAFLD is the most important use of CAP, and we have specifically evaluated CAP and the validity criteria in the NAFLD population.

Furthermore, while the diagnosis of fatty liver based on $\geq 5\%$ steatosis by liver biopsy was in accordance with international recommendations, newer magnetic resonance imaging (MRI)-based techniques such as MRI-proton density fat fraction (MRI-PDF) and proton-magnetic resonance spectroscopy have been shown to be highly accurate and even more sensitive than liver biopsy in detecting changes in liver fat over time.³⁵ In head-to-head comparison, MRI-PDF has shown higher applicability and accuracy than CAP in the United States and Japan.^{24,36} It would be important to validate the validity criteria of CAP against MRI in future studies. Similarly, the performance of XL probe-based CAP measurement should also be validated against MRI. Quantitative ultrasound has also shown good correlation with MRI-PDF and deserves further evaluation.³⁷

In conclusion, the validity of CAP for the diagnosis of fatty liver is lower if the IQR of CAP is ≥ 40 dB/m. Traditional factors affecting the performance of LSM have little impact on the validity of CAP. Our findings provide guidance on the interpretation of CAP results in routine clinical practice.

Financial support

This project was partially supported by a grant from the Research Grant Council of the Hong Kong SAR Government (Project no. CUHK 477813).

Conflict of interest

Vincent Wong, Grace Wong, Henry Chan and Victor de Lédighen have served as speakers for Echosens.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Study design: Vincent Wong. Data collection: Vincent Wong, Salvatore Petta, Jean-Baptiste Hiriart, Calogero Cammà, Grace Wong, Fabio Marra, Julien Vergniol, Anthony Chan, Antonino Tuttolomondo, Wassil Merrouche, Brigitte Le Bail, Umberto Arena, Antonio Craxì, Victor de Lédighen. Data analysis: Vincent Wong.

Administrative support: Henry Chan, Antonio Craxì. Manuscript drafting: Vincent Wong, Salvatore Petta. All authors read and approved the final version of the manuscript.

Acknowledgements

We would like to thank Mr Yee-Kit Tse for his statistical support.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jhep.2017.05.005>.

References

- [1] Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease—Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology* 2016;64:73–84.
- [2] Wong VW, Chu WC, Wong GL, Chan RS, Chim AM, Ong A, et al. Prevalence of non-alcoholic fatty liver disease and advanced fibrosis in Hong Kong Chinese: a population study using proton-magnetic resonance spectroscopy and transient elastography. *Gut* 2012;61:409–415.
- [3] Wong RJ, Aguilar M, Cheung R, Perumpail RB, Harrison SA, Younossi ZM, et al. Nonalcoholic steatohepatitis is the second leading etiology of liver disease among adults awaiting liver transplantation in the United States. *Gastroenterology* 2015;148:547–555.
- [4] Younossi ZM, Otgonsuren M, Henry L, Venkatesan C, Mishra A, Erario M, et al. Association of nonalcoholic fatty liver disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004 to 2009. *Hepatology* 2015;62:1723–1730.
- [5] Wong GL, Wong VW, Choi PC, Chan AW, Chim AM, Yiu KK, et al. Metabolic syndrome increases the risk of liver cirrhosis in chronic hepatitis B. *Gut* 2009;58:1111–1117.
- [6] Veldt BJ, Chen W, Heathcote EJ, Wedemeyer H, Reichen J, Hofmann WP, et al. Increased risk of hepatocellular carcinoma among patients with hepatitis C cirrhosis and diabetes mellitus. *Hepatology* 2008;47:1856–1862.
- [7] Chen CL, Yang HI, Yang WS, Liu CJ, Chen PJ, You SL, et al. Metabolic factors and risk of hepatocellular carcinoma by chronic hepatitis B/C infection: a follow-up study in Taiwan. *Gastroenterology* 2008;135:111–121.
- [8] Chan AW, Wong GL, Chan HY, Tong JH, Yu YH, Choi PC, et al. Concurrent fatty liver increases risk of hepatocellular carcinoma among patients with chronic hepatitis B. *J Gastroenterol Hepatol* 2016.
- [9] Sasso M, Tengher-Barna I, Ziol M, Miette V, Fournier C, Sandrin L, et al. Novel controlled attenuation parameter for noninvasive assessment of steatosis using Fibroscan((R)): validation in chronic hepatitis C. *J Viral Hepat* 2012;19:244–253.
- [10] de Lédighen V, Vergniol J, Foucher J, Merrouche W, le Bail B. Non-invasive diagnosis of liver steatosis using controlled attenuation parameter (CAP) and transient elastography. *Liver Int* 2012;32:911–918.
- [11] Chon YE, Jung KS, Kim SU, Park JY, Park YN, Kim DY, et al. Controlled attenuation parameter (CAP) for detection of hepatic steatosis in patients

Research Article

- with chronic liver diseases: a prospective study of a native Korean population. *Liver Int* 2014;34:102–109.
- [12] Shen F, Zheng RD, Mi YQ, Wang XY, Pan Q, Chen GY, et al. Controlled attenuation parameter for non-invasive assessment of hepatic steatosis in Chinese patients. *World J Gastroenterol* 2014;20:4702–4711.
- [13] Myers RP, Pollett A, Kirsch R, Pomier-Layrargues G, Beaton M, Levstik M, et al. Controlled Attenuation Parameter (CAP): a noninvasive method for the detection of hepatic steatosis based on transient elastography. *Liver Int* 2012;32:902–910.
- [14] Boursier J, Zarski JP, de Ledinghen V, Rousselet MC, Sturm N, Lebaill B, et al. Determination of reliability criteria for liver stiffness evaluation by transient elastography. *Hepatology* 2013;57:1182–1191.
- [15] Kwok R, Choi KC, Wong GL, Zhang Y, Chan HL, Luk AO, et al. Screening diabetic patients for non-alcoholic fatty liver disease with controlled attenuation parameter and liver stiffness measurements: a prospective cohort study. *Gut* 2016;65:1359–1368.
- [16] de Ledinghen V, Vergniol J, Capdepon M, Chermak F, Hiriart JB, Cassinotto C, et al. Controlled attenuation parameter (CAP) for the diagnosis of steatosis: a prospective study of 5323 examinations. *J Hepatol* 2014;60:1026–1031.
- [17] Karlas T, Petroff D, Sasso M, Fan JG, Mi YQ, de Ledinghen V, et al. Individual patient data meta-analysis of controlled attenuation parameter (CAP) technology for assessing steatosis. *J Hepatol* 2017;66:1022–1030.
- [18] Wong VW, Wong GL, Yeung DK, Lau TK, Chan CK, Chim AM, et al. Incidence of non-alcoholic fatty liver disease in Hong Kong: a population study with paired proton-magnetic resonance spectroscopy. *J Hepatol* 2015;62:182–189.
- [19] 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.
- [20] Wong VW, Vergniol J, Wong GL, Foucher J, Chan HL, Le Bail B, et al. Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology* 2010;51:454–462.
- [21] Younossi ZM, Stepanova M, Rafiq N, Makhlof H, Younoszai Z, Agrawal R, et al. Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality. *Hepatology* 2011;53:1874–1882.
- [22] Ekstedt M, Hagstrom H, Nasr P, Fredrikson M, Stal P, Kechagias S, et al. Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology* 2015;61:1547–1554.
- [23] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36.
- [24] Imajo K, Kessoku T, Honda Y, Tomeno W, Ogawa Y, Mawatari H, et al. Magnetic resonance imaging more accurately classifies steatosis and fibrosis in patients with nonalcoholic fatty liver disease than transient elastography. *Gastroenterology* 2016;150:626–637 e627.
- [25] Wong GL, Wong VW, Choi PC, Chan AW, Chim AM, Yiu KK, et al. Increased liver stiffness measurement by transient elastography in severe acute exacerbation of chronic hepatitis B. *J Gastroenterol Hepatol* 2009;24:1002–1007.
- [26] Millonig G, Reimann FM, Friedrich S, Fonouni H, Mehrabi A, Buchler MW, et al. Extrahepatic cholestasis increases liver stiffness (FibroScan) irrespective of fibrosis. *Hepatology* 2008;48:1718–1723.
- [27] Millonig G, Friedrich S, Adolf S, Fonouni H, Golriz M, Mehrabi A, et al. Liver stiffness is directly influenced by central venous pressure. *J Hepatol* 2010;52:206–210.
- [28] Wong GL, Chan HL, Choi PC, Chan AW, Lo AO, Chim AM, et al. Association between anthropometric parameters and measurements of liver stiffness by transient elastography. *Clin Gastroenterol Hepatol* 2013;11:295–302 e291–e293.
- [29] Shen F, Zheng RD, Shi JP, Mi YQ, Chen GF, Hu X, et al. Impact of skin capsular distance on the performance of controlled attenuation parameter in patients with chronic liver disease. *Liver Int* 2015;35:2392–2400.
- [30] Merriman RB, Ferrell LD, Patti MG, Weston SR, Pabst MS, Aouizerat BE, et al. Correlation of paired liver biopsies in morbidly obese patients with suspected nonalcoholic fatty liver disease. *Hepatology* 2006;44:874–880.
- [31] Myers RP, Pomier-Layrargues G, Kirsch R, Pollett A, Duarte-Rojo A, Wong D, et al. Feasibility and diagnostic performance of the FibroScan XL probe for liver stiffness measurement in overweight and obese patients. *Hepatology* 2012;55:199–208.
- [32] de Ledinghen V, Wong VW, Vergniol J, Wong GL, Foucher J, Chu SH, et al. Diagnosis of liver fibrosis and cirrhosis using liver stiffness measurement: comparison between M and XL probe of FibroScan(R). *J Hepatol* 2012;56:833–839.
- [33] Wong VW, Vergniol J, Wong GL, Foucher J, Chan AW, Chermak F, et al. Liver stiffness measurement using XL probe in patients with nonalcoholic fatty liver disease. *Am J Gastroenterol* 2012;107:1862–1871.
- [34] Sasso M, Audiere S, Kemgang A, Gaouar F, Corpechot C, Chazouilleres O, et al. Liver steatosis assessed by controlled attenuation parameter (CAP) measured with the XL probe of the fibroscan: a pilot study assessing diagnostic accuracy. *Ultrasound Med Biol* 2016;42:92–103.
- [35] Noureddin M, Lam J, Peterson MR, Middleton M, Hamilton G, Le TA, et al. Utility of magnetic resonance imaging vs. histology for quantifying changes in liver fat in nonalcoholic fatty liver disease trials. *Hepatology* 2013;58:1930–1940.
- [36] Park CC, Nguyen P, Hernandez C, Bettencourt R, Ramirez K, Fortney L, et al. Magnetic resonance elastography vs transient elastography in detection of fibrosis and noninvasive measurement of steatosis in patients with biopsy-proven nonalcoholic fatty liver disease. *Gastroenterology* 2017;152:598–607 e592.
- [37] 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.