Chapter 5

Ultrasonography to quantify hepatic fat content: validation by ^{1}H Magnetic Resonance Spectroscopy

Obesity (Silver Spring) 2009; 17(12):2239-2244

Mireille A. Edens [1]

Peter M. van Ooijen [2]

Wendy J. Post [1]

Mark J. Haagmans [2]

Wisnumurti Kristanto [2]

Paul E. Sijens [2]

Erik J. van der Jagt [2]

Ronald P. Stolk [1]

Department of Epidemiology [1]
Department of Radiology [2]
University Medical Center Groningen
University of Groningen
Groningen, the Netherlands

ABSTRACT

An abundance of fat stored within the liver, or steatosis, is the beginning of a broad hepatological spectrum, usually referred to as fatty liver disease (FLD). For studies on FLD, quantitative hepatic fat ultrasonography would be an appealing study modality. Objective of the present study was to developed a technique for quantifying hepatic fat content by ultrasonography and validate this using proton magnetic resonance spectroscopy (^{1}H MRS) as gold standard. Eighteen White volunteers (BMI range 21.0 to 42.9) were scanned by both ultrasonography and ^{1}H MRS. Altered ultrasound characteristics, present in the case of FLD, were assessed using a specially developed software program. Various attenuation and textural based indices of FLD were extracted from ultrasound images. Using linear regression analysis, the predictive power of several models (consisting of both attenuation and textural based measures) on log 10-transformed hepatic fat content by ^{1}H MRS were investigated. The best quantitative model was compared with a qualitative ultrasonography method, as used in clinical care. A model with four ultrasound characteristics could modestly predict the amount of liver fat (adjusted explained variance 43.2%, $p=0.021$). Expanding the model to seven ultrasound characteristics increased adjusted explained variance to 60% ($p=0.015$), with $r=0.789$ ($p<0.001$). Comparing this quantitative model with qualitative ultrasonography revealed a significant advantage of the quantitative model in predicting hepatic fat content ($p<0.001$). This validation study shows that a combination of computer-assessed ultrasound measures from routine ultrasound images can be used to quantitatively assess hepatic fat content.

INTRODUCTION

A continuous accumulation of lipids in the liver may result in a broad hepatological spectrum, usually referred to as fatty liver disease (FLD) [1, 2]. An abundance of fat within the liver, or steatosis, can progress to steatohepatitis (fat and inflammation, with or without fibrosis) and cirrhosis (maximum fibrosis score) [3], and has also been associated with hepatocarcinoma [4]. Additionally, FLD, particularly non-alcoholic FLD (NAFLD), is an underlying condition for cardiovascular disease [5, 6]. As alcoholic FLD (AFLD) and NAFLD are histologically indistinct [7], distinction between both is neither possible nor relevant in relation to measurement of hepatic fat content. The estimated FLD prevalence is one third of the general adult Western population [8-10], and may have been increasing in parallel with the global increase of obesity [11].

Hepatic fat content can be determined by histological [2] or biochemical [12, 13] analysis of liver tissue by biopsy, magnetic resonance techniques [14], computed tomography [15] and ultrasonography [16, 17]. Ultrasonography is, in contrast to other diagnostic modalities, an appealing method for large population studies on FLD, as is it non-invasive (painless, no harmful radiation), portable and relatively inexpensive. In the case of parenchymal liver disease, reflections of liver tissue by ultrasonography are altered [16, 17]. In clinical care, ultrasonography is the most often used diagnostic modality, but in a qualitative way. Steatosis can be qualitatively assessed by: i] hyperechogenity of liver tissue ('bright liver') as often compared to hypoechogenity of the kidney cortex, ii] fine, tightly packed echoes, iii] fall of echo amplitude with depth (posterior beam attenuation), iiii] loss of echoes from the walls of the portal veins (featureless appearance) [16, 17]. As this is a qualitative scoring method and also subjective [18], quantitative approaches for identification of liver disease have been suggested. However, these methods have never been validated by an appropriate quantitative gold standard. The purpose of the present study was to develop and validate quantitative analysis of ultrasonography images, for assessment of hepatic fat content, using proton magnetic resonance spectroscopy (^{1}H MRS) as gold standard.

METHODS AND PROCEDURES

VOLUNTEERS

Volunteers were recruited by advertisement, and a heterogenic study population was strived after. Exclusion criteria were current presence of hepatic pathology, previous hepatic or renal surgery, and standard MR-contraindications. The volunteers underwent both hepatic ultrasonography and ^{1}H MRS, and a short physical examination. All volunteers gave written informed consent. This study was approved by the Medical Ethics Committee of the University Medical Center Groningen.

ULTRASONOGRAPHY

Ultrasonography was performed using a Philips ATL ultrasound machine (Philips, Best, the Netherlands), with a 5 – 2 MHz curved array transducer.

Quantitative ultrasonography

Imaging

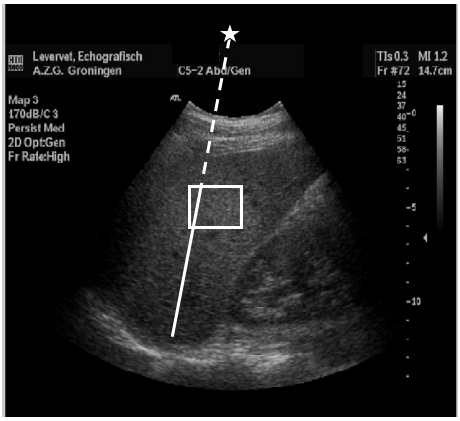
In one ultrasound image both liver and right kidney were visualised [17], as shown by figure I. Imaging was performed by an experienced radiologist. One standard image, with 'persist' (med), '2D opt' (gen), 'frame rate' (high), 'gain' (40) and 'image depth' (14.7 cm), was used for analysis.

Analysis

Images were analysed by an operator (operator 1) twice, with a one month interval, and the average values by operator 1 were used in this study. In order to study inter-operator reliability, another operator (operator 2) analysed the images, while untrained for the method and blinded for all study outcomes.

Data extraction and data

Data were extracted from ultrasound images using a modified version of a specially developed software program (dept. of BME, Technion IIT, Haifa, Israel) in the MATLAB programming environment, previously described by Gaitini et al. [19]. Figure I shows an example of data extraction. According to a standard protocol, regions of interest and attenuation lines were interactively placed in the liver images in order to calculate several attenuation indices and several textural indices. Figure II shows a scheme on quantitative ultrasonography measures, including the presently validated indices in the white boxes.



Both a region of interest (quadrangle) and an attenuation line (closed line) were placed according to a standard protocol.

The region of interest had to be placed in a bright area, while avoiding large artefacts like rib shadows and large blood vessels, at a depth of 4 to 6 cm. The attenuation line had to be placed in a bright pathway, while avoiding large artefacts, at a straight line from the 'origin' of ultrasound (intermittent line).

The region of interest served for the determination of several textural indices. The attenuation line was used for determining attenuation estimates [19].

Figure I. Ultrasonography image analysis

Qualitative ultrasonography

In addition to the quantitative approach, the radiologist made an ultrasound image with optimum settings, as used in clinical care. This image was qualitatively scored by the radiologist, according to standard qualitative criteria [16, 17, 26], while blinded for all study outcomes.

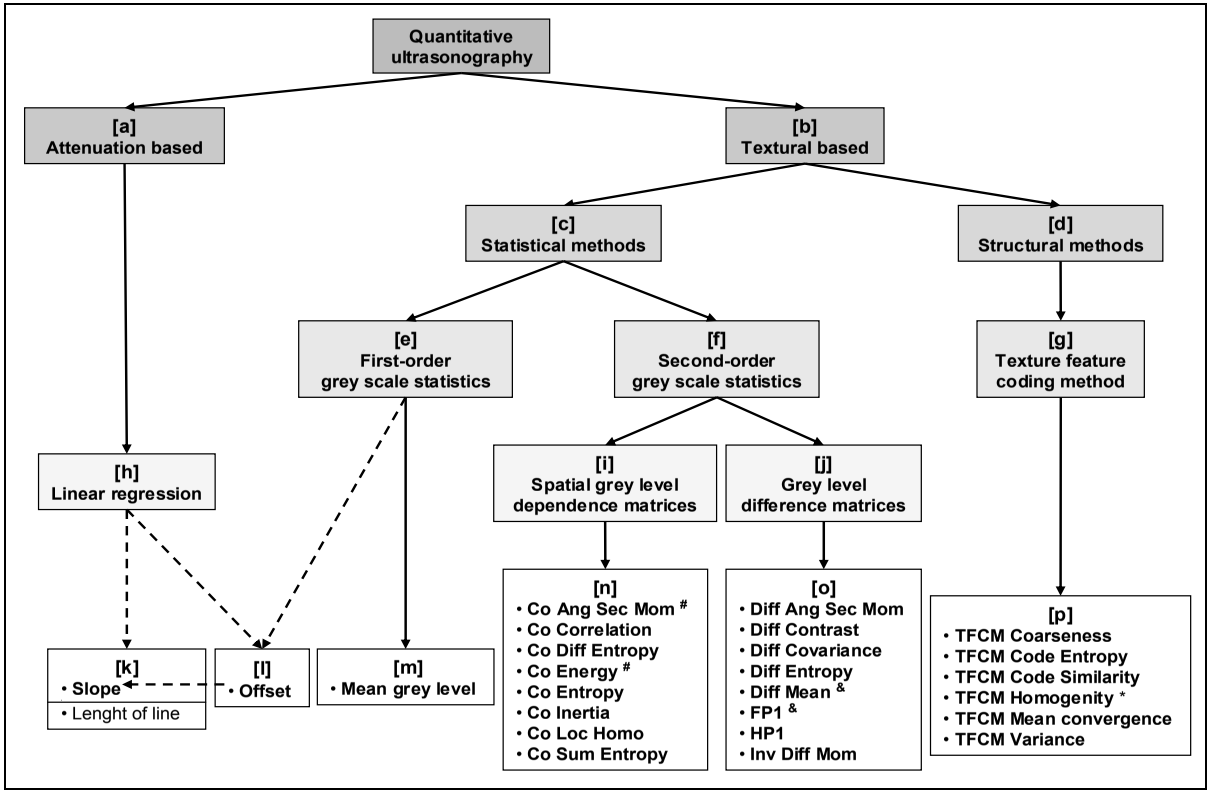


Figure II. Scheme and classification on quantitative ultrasonography, including the presently validated measures in the white boxes [k] to [p]

Measures sharing superscripts had the same outcome, * = outcome was always 0. For a detailed explanation on the presently validated quantitative ultrasonography measures [19], the reader is referred to the following references: [a] [19, 20], [b] [21], [c] [19, 21, 22], [d] [19, 21, 22], [e] [22], [f] [22-24], [g] [23], [h] appendix I, [i] [25], [j] [25], [k]and[l] appendix I, [m] [19], [n] appendix II, [o] appendix II, [p] [23].

MULTI-VOXEL PROTON MAGNETIC RESONANCE SPECTROSCOPY

In general, by means of radiofrequency transmission and reception, a magnetic resonance scanner detects resonance signals of both hepatic lipids (mainly methylene, i.e. CH_2, from fatty acyl chains) and hepatic water [27]. As previously described in detail [14, 28], ^{1}H MRS was performed, using a 1.5 Tesla whole-body scanner (MAGNETOM Avanto; Siemens Medical Solutions, Erlangen, Germany) equipped with gradients of up to 40 mTm^{-1} (maximal slew rate = 200 mT m^{-1}/ms) and a six-channel spine array coil. Subjects were in supine position with a large flex coil placed over the liver, which was simultaneously used with the spine array coil as receiver. T_1-weighted gradient-echo images were recorded to assess the anatomy of liver. Using a field of view of 16×16 cm^2 and a volume of interest of 5×8×4 cm^3 positioned within the liver, hybrid 2D-spectroscopic imaging (chemical shift imaging or CSI), point resolved spectroscopy (PRESS) with a repetition time (TR) of 5000 ms and an echo time (TE) of 30 ms was performed. The CSI measurement lasted 16×16×5 = 1280 s, corresponding to approximately 21 min. Shimming was automated and water suppression was not applied in order to be able to calculate the fat-water ratio distributions in the liver directly. At the used TR of 5 s, T_1 saturation of the water and fat signals is negligible, i.e. TR > $5T_1$. At the used TE of 30 ms the correction applied to our data to compensate for the fact that the fat signal has a longer T_2 (78 ms) than water (60 ms) was 12.2 %. Hepatic fat content was calculated by the peak CH_2 signal (at 1.3 parts/ million) divided by the sum of the peak CH_2 signal and peak H_2O signal (at 4.7 parts/ million), using water as an internal reference [14, 28]. ^{1}H MRS has been validated, by comparison with both histological and biochemical analysis of liver tissue by biopsy [27, 29, 30].

A hepatic fat content of 5.56% by ^{1}H MRS is used as cut-off value for diagnosing FLD, based on the 95th percentile hepatic fat distribution of a low risk group [10].

STATISTICS

Univariate analysis and multiple regression analysis

As distribution of hepatic fat content by ^{1}H MRS was skewed, values were log 10 transformed. Plotting and correlation (Pearson) was used to explore univariate concordance with log 10 ^{1}H MRS. The classification of variables in figure II (white boxes), followed by 'backward selection', was used for variable selection in a linear regression model. Firstly, the variables from separate boxes of figure II were assessed, i.e. separate ultrasonography

aspects. Secondly, variables from combinations of boxes of figure II were assessed, i.e. information from several ultrasonography aspects.

Evaluation and bootstrap

Models were evaluated on adjusted explained variance (adj. R^2) and explained variance (R^2). By means of bootstrap, 95% confidence intervals were estimated for regression coefficients, and adj. R^2 and R^2. Moreover, a 95% prediction interval was calculated.

Quantitative versus qualitative ultrasonography

The Chi-square test was used to test differences between the two methods. Additionally, sensitivity and specificity of both methods were calculated. In addition to the 95% prediction interval of the quantitative method, a 95% prediction interval was calculated for the qualitative method as well.

Reliability

Both intra-observer and inter-observer reproducibility of algorithms were studied, using the Bland & Altman method [31].

Statistical analysis was performed using software programs SPSS version 14 and R version 2.6.2.

RESULTS

STUDY POPULATION

Twenty apparently healthy White volunteers were examined. One volunteer was excluded because of hepatic pathology (haemangioma), and one volunteer was excluded because of rib shadows over the liver-kidney image.

The study population (n = 18) consisted of 10 males and 8 females, with a mean ±sd age of 46.0 ±14.1 year, body mass index of 28.7 ±6.4 kg/m^2 (range 21.0 to 42.9), and waist to hip ratio of 0.93 ±0.11. Hepatic fat content by ^{1}H MRS ranged from 0.32% – 18.55%, with a median of 1.75%.

UNIVARIATE ANALYSIS

Plots and correlation coefficients revealed no associations. Only slope and co-entropy were borderline significantly associated with log 10 ^{1}H MRS, with r=−0.423 (p=0.081) and r=−0.418 (p=0.084), respectively.

MULTIPLE ANALYSIS

Information from separate ultrasonography aspects, i.e. separate boxes from figure II, revealed no associations with log 10 ^{1}H MRS (data not shown). Combining information from different ultrasonography aspects, i.e. by combining boxes from figure II, was associated with hepatic fat content by log 10 ^{1}H MRS (model 1, table I). Including more ultrasonography characteristics, further improved the association (model 2, table I).

Algorithms

The algorithms derived from these models are:

Algorithm 1:

$$\log 10\ {}^1\text{H MRS}_{\text{pred.}} = -37.67 - 0.07*\text{offset} - 0.78*\text{slope} - 3.85*\text{co entropy} + 3.56*\text{co sum entropy}.$$

Algorithm 2:

$$\log 10\ {}^1\text{H MRS}_{\text{pred.}} = -72.68 - 0.07*\text{offset} - 0.81*\text{slope} - 3.63*\text{co entropy} + 3.34*\text{co sum entropy} - 0.20*\text{diff contrast} + 84.84*\text{inv diff mom} + 5.98*\text{FP1}.$$

A scatter plot of algorithm 2 with log 10 ^{1}H MRS, including a prediction interval, is shown in figure IIIa. This means that by applying our algorithm to ultrasound images of new volunteers, 95% of their predicted values will fall within this interval.

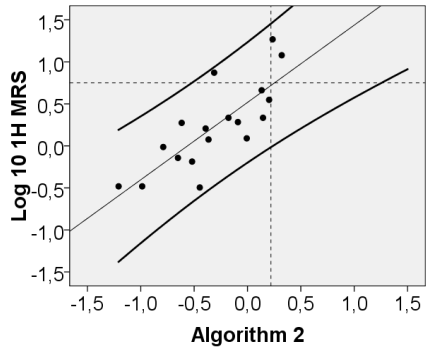
Table I. Prediction by combinations of ultrasound characteristics

Model ID	Independent variables [a]	B [95% CI[b]]	p-value	Model p-value	Model adj. R^2 [95% CI[b]]	Model R^2 [95% CI[b]]
1	(constant)	−37.67 [−57.45 – −18.07]	.007	.021	43.2% [−2.9% – 97.2%]	56.5% [21.3% – 97.9%]
	offset	−.07 [−0.10 – −0.03]	.009			
	slope	−.78 [−1.19 – −0.37]	.007			
	co entropy	−3.85 [−5.95 – −1.79]	.008			
	co sum entropy	3.56 [0.81 – 6.31]	.052			
2	(constant)	−72.68 [−96.70 – −48.76]	.001	.015	60.0% [49.1% – 99.6%]	76.5% [70.0% – 99.8%]
	offset	−.07 [−0.10 – −0.04]	.012			
	slope	−.81 [−1.17 – −0.44]	.009			
	co entropy	−3.63 [−5.33 – −1.93]	.011			
	co sum entropy	3.34 [1.24 – 5.40]	.039			
	diff contrast	−.20 [−0.30 – −0.09]	.019			
	inv diff mom	84.84 [40.42 – 129.65]	.019			
	FP1	5.98 [2.97 – 9.01]	.016			

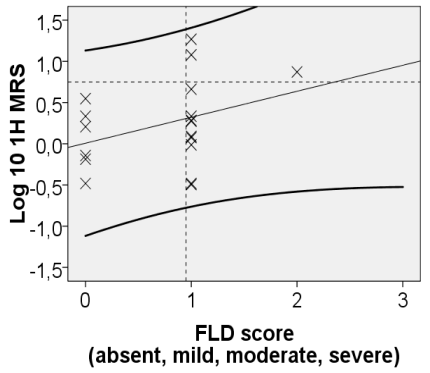
[a] = dependent variable is log 10 ^{1}H MRS, [b] = bootstrapped 95% confidence interval. Adj. R^2 = adjusted explained variance, B = regression coefficients, R^2 = explained variance.

QUANTITATIVE VERSUS QUALITATIVE ULTRASONOGRAPHY

Validity of algorithm 2, in comparison with a qualitative ultrasonography method used in clinical care, is shown in figure III. Using cut-off value 0.75 log 10 ^{1}H MRS, which is 5.56% by ^{1}H MRS [10], 3 people had FLD. Quantitative ultrasonography was significantly better associated with the presence of FLD than qualitative ultrasonography ($\chi^2 = 32.8$, with df = 1, $p<0.001$). Sensitivity and specificity were 66.7% and 100%, respectively, for quantitative ultrasonography. Sensitivity and specificity were 100% and 40%, respectively, for qualitative ultrasonography. Additionally, the prediction interval of the quantitative method was much narrower in comparison to the prediction interval of the qualitative method (figure III).



a. Quant. ultrasonography
r = .789 (p<.001)
R^2 = 62.3%



b. Qual. ultrasonography
r = .353 (p=.151)
R^2 = 12.4%

		^{1}H MRS	
		FLD	no FLD
Quant.	FLD	2	0
	no FLD	1	15
Qual.	FLD	3	9
	no FLD	0	6

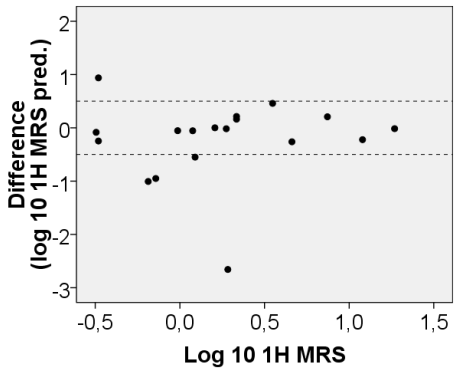
c. Comparison
Quantitative cut-off: ≥.22 Qualitative cut-off: ≥1

Figure III. Quantitative *versus* qualitative ultrasonography, using correlation coefficients, sensitivity and specificity, and prediction intervals

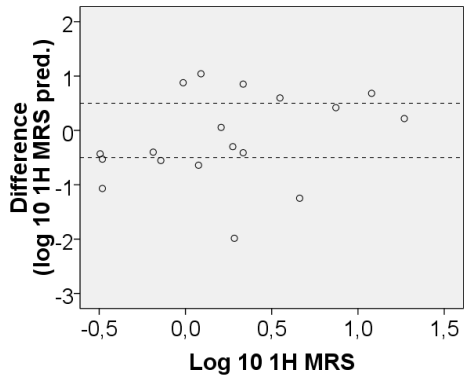
Qual. = qualitative, Quant. = quantitative, r = Pearson's correlation coefficient, R^2 = explained variance.

RELIABILITY

Intra-operator and inter-operator reproducibility, regarding algorithm 2, are shown in figure IV. If we tolerate an operator difference smaller than 0.5, as shown by the interrupted lines, 5 people had a larger intra-operator difference of 0.55 to 2.66 log 10 ^{1}H MRS_{pred}. Eleven people had a larger inter-operant difference of 0.53 to 1.98 log 10 ^{1}H MRS_{pred}. These differences were independent of hepatic fat content (figure IV).



a. 1st attempt by op. 1 *minus* 2nd attempt by op. 1



b. 1st attempt by op. 1 *minus* attempt by op. 2

Figure IV. Reliability

op. = operator

DISCUSSION

This study shows that combinations of quantitative ultrasonography measures are significantly associated with hepatic fat content by (log 10) ^{1}H MRS (table I), and even better than a qualitative method currently used in clinical care (figure III). Reliability was reasonably well (figure IV).

These results suggest that combinations of computer-assessed ultrasonography measures quantify the ultrasonographic characteristics of FLD [16, 17], i.e. i] hyperechogenity, by offset in a certain degree, ii] fine, tightly packed echoes in the case of hepatic fat, and coarse pin head echoes in the case of fibrosis, by the textural based measures, and iii] fall of echo amplitude with depth, by slope, and by offset in lesser degree (table I).

VALIDITY

Previously, 'slope' [19, 20] and 'offset', 'mean grey level', 'co entropy' and 'co sum entropy' [19], revealed discriminative power in the FLD spectrum. In the present study, none of the measures were univariately associated with hepatic fat content by ^{1}H MRS. The attenuation based measure 'slope' did not significantly predict hepatic fat content, whereas the slope previously did reveal power for discriminating pure fatty livers (steatosis) from healthy livers, with an area under the curve of 1 [19, 20]. However, the slope lost discriminative power in the total FLD spectrum [19, 20]. It is known that both hepatic fat content [3, 32, 33] and (therefore) attenuation [17] are decreased in the case of (advanced) fibrosis and cirrhosis, which might have caused the fall in discriminative power of the slope in the total FLD spectrum [19, 20]. Fibrosis itself does not produce attenuation [34]. This may also explain why, in models, the attenuation based indices must be accompanied by textural indices of coarseness/ fibrosis, in particular 'co entropy' [19, 35].

In this apparently healthy study population, we obviously did not perform hepatic biopsy for histological scoring of fibrosis stage, nor magnetic resonance elastography (MRE), which determines liver stiffness as a marker of fibrosis [36]. Therefore, it was not possible to verify the effect of fibrosis on ultrasonography algorithms. Additionally, because of both inclusion of the right kidney [17] and rib shadows in ultrasound images, it was not always possible to draw the attenuation line to the bottom of the liver for realization of a 'far field' slope [19, 20]. However, inclusion of length of the attenuation line in models on slope did not

lead to an improvement in the prediction of hepatic fat content by ^{1}H MRS (data not shown). Adjusting for focus depth and frequency was not possible, because of the small ranges.

A limitation of the present study is the small study population (n = 18), however, bootstrap of model 2 revealed good 95% confidence intervals (table I). Additionally, model 1 with only 4 variables already showed significant results.

Because the performance of a test depends on the prevalence of an underlying disorder, e.g. FLD, the sensitivity of quantitative ultrasonography may be lower in a clinical population.

RELIABILITY

Intra-operator difference was reasonably well, but inter-operator difference was less (figure IV). This may be explained by operator 1 being experienced, while operator 2 was not. Retrospective analysis of the outliers from figure IV (print screens' of analysed images were saved), revealed insight in the differences in analysis.

FUTURE RESEARCH

While histological scoring of liver tissue by biopsy is often considered the gold standard for diagnosis of FLD, ^{1}H MRS is more reliable, and may be more valid for quantification of hepatic fat content. Reproducibility of qualitative histological analysis regarding steatosis grade is good as weighted kappa scores range from 0.64 to 0.90 [2, 37], whereas reproducibility regarding hepatic fat content by ^{1}H MRS is excellent with reported correlation coefficients of up to 0.99 ($p<0.001$) [10]. For future studies, ^{1}H MRS combined with MRE [36] would be an interesting gold standard.

CONCLUSION

This is the first in vivo validation study on quantitative hepatic fat ultrasonography, using an excellent quantitative gold standard, i.e. multi-voxel ^{1}H MRS [14, 28]. Therefore, we feel that the method needs to be improved before used as clinical diagnosis modality.

This validation study shows that a combination of computer-assessed ultrasound measures from routine ultrasound images can be used to quantitatively assess hepatic fat content. Reliability should be improved by more protocolized procedures and training of operators.

Please also see 'additional remarks and recommendations for future research'

ACKNOWLEDGEMENTS

The authors thank dr. H Azhari, who enabled this study by sharing the software program, developed by the Department of Biomedical Engineering, Technion Israel Institute of Technology, Haifa, Israel [19].

The authors thank I Willeboordse, AM van Tienhoven, JH Potze, and P Kappert for magnetic resonance scanning.

Reference List

(1) Matteoni CA, Younossi ZM, Gramlich T, Boparai N, Liu YC, McCullough AJ. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. *Gastroenterology* 1999 June;116(6):1413-9.

(2) Kleiner DE, Brunt EM, Van NM, Behling C, Contos MJ, Cummings OW, Ferrell LD, Liu YC, Torbenson MS, Unalp-Arida A, Yeh M, McCullough AJ, Sanyal AJ. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005 June;41(6):1313-21.

(3) Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, Kechagias S. Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology* 2006 October;44(4):865-73.

(4) Smedile A, Bugianesi E. Steatosis and hepatocellular carcinoma risk. *Eur Rev Med Pharmacol Sci* 2005 September;9(5):291-3.

(5) Kotronen A, Yki-Jarvinen H. Fatty liver: a novel component of the metabolic syndrome. *Arterioscler Thromb Vasc Biol* 2008 January;28(1):27-38.

(6) Edens MA, Kuipers F, Stolk RP. Non-alcoholic fatty liver disease is associated with cardiovascular disease risk markers. *Obes Rev* 2009 July;10(4):412-9.

(7) Mills SJ, Harrison SA. Comparison of the natural history of alcoholic and nonalcoholic fatty liver disease. *Curr Gastroenterol Rep* 2005 February;7(1):32-6.

(8) Bedogni G, Miglioli L, Masutti F, Tiribelli C, Marchesini G, Bellentani S. Prevalence of and risk factors for nonalcoholic fatty liver disease: the Dionysos nutrition and liver study. *Hepatology* 2005 July;42(1):44-52.

(9) Volzke H, Robinson DM, Kleine V, Deutscher R, Hoffmann W, Ludemann J, Schminke U, Kessler C, John U. Hepatic steatosis is associated with an increased risk of carotid atherosclerosis. *World J Gastroenterol* 2005 March 28;11(12):1848-53.

(10) Szczepaniak LS, Nurenberg P, Leonard D, Browning JD, Reingold JS, Grundy S, Hobbs HH, Dobbins RL. Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. *Am J Physiol Endocrinol Metab* 2005 February;288(2):E462-E468.

(11) Friedman JM. Obesity in the new millennium. *Nature* 2000 April 6;404(6778):632-4.

(12) Araya J, Rodrigo R, Videla LA, Thielemann L, Orellana M, Pettinelli P, Poniachik J. Increase in long-chain polyunsaturated fatty acid n - 6/n - 3 ratio in relation to

hepatic steatosis in patients with non-alcoholic fatty liver disease. *Clin Sci (Lond)* 2004 June;106(6):635-43.

(13) Puri P, Baillie RA, Wiest MM, Mirshahi F, Choudhury J, Cheung O, Sargeant C, Contos MJ, Sanyal AJ. A lipidomic analysis of nonalcoholic fatty liver disease. *Hepatology* 2007 October;46(4):1081-90.

(14) Irwan R, Edens MA, Sijens PE. Assessment of the variations in fat content in normal liver using a fast MR imaging method in comparison with results obtained by spectroscopic imaging. *Eur Radiol* 2008 April;18(4):806-13.

(15) Duman DG, Celikel C, Tuney D, Imeryuz N, Avsar E, Tozun N. Computed tomography in nonalcoholic fatty liver disease: a useful tool for hepatosteatosis assessment? *Dig Dis Sci* 2006 February;51(2):346-51.

(16) Saverymuttu SH, Joseph AE, Maxwell JD. Ultrasound scanning in the detection of hepatic fibrosis and steatosis. *Br Med J (Clin Res Ed)* 1986 January 4;292(6512):13-5.

(17) Joseph AE, Saverymuttu SH, al Sam S, Cook MG, Maxwell JD. Comparison of liver histology with ultrasonography in assessing diffuse parenchymal liver disease. *Clin Radiol* 1991 January;43(1):26-31.

(18) Smith-Levitin M, Blickstein I, Albrecht-Shach AA, Goldman RD, Gurewitsch E, Streltzoff J, Chervenak FA. Quantitative assessment of gray-level perception: observers' accuracy is dependent on density differences. *Ultrasound Obstet Gynecol* 1997 November;10(5):346-9.

(19) Gaitini D, Baruch Y, Ghersin E, Veitsman E, Kerner H, Shalem B, Yaniv G, Sarfaty C, Azhari H. Feasibility study of ultrasonic fatty liver biopsy: texture vs. attenuation and backscatter. *Ultrasound Med Biol* 2004 October;30(10):1321-7.

(20) Graif M, Yanuka M, Baraz M, Blank A, Moshkovitz M, Kessler A, Gilat T, Weiss J, Walach E, Amazeen P, Irving CS. Quantitative estimation of attenuation in ultrasound video images: correlation with histology in diffuse liver disease. *Invest Radiol* 2000 May;35(5):319-24.

(21) Zhang J, Tan T. Brief review of invariant texture analysis methods. *Pattern Recognition* 2002;35:735-47.

(22) Haralick RM. Statistical and structural approaches to texture. *Proc IEEE* 1979 May;67(5):786-804.

(23) Horng MH, Sun YN, Lin XZ. Texture feature coding method for classification of liver sonography. *Comput Med Imaging Graph* 2002 January;26(1):33-42.

(24) Gotlieb C, Kreyszig H. Texture descriptors based on co-occurence matrices. *Computer vision, graphics, and image processing* 1990;51:70-86.

(25) Prakash KN, Ramakrishnan AG, Suresh S, Chow TW. An investigation into the feasibility of fetal lung maturity prediction using statistical textural features. *Ultrason Imaging* 2001 January;23(1):39-54.

(26) Mendler MH, Bouillet P, Le Sidaner A, Lavoine E, Labrousse F, Sautereau D, Pillegand B. Dual-energy CT in the diagnosis and quantification of fatty liver: limited clinical value in comparison to ultrasound scan and single-energy CT, with special reference to iron overload. *J Hepatol* 1998 May;28(5):785-94.

(27) Szczepaniak LS, Babcock EE, Schick F, Dobbins RL, Garg A, Burns DK, McGarry JD, Stein DT. Measurement of intracellular triglyceride stores by H spectroscopy: validation in vivo. *Am J Physiol* 1999 May;276(5 Pt 1):E977-E989.

(28) Sijens PE, Smit GP, Borgdorff MA, Kappert P, Oudkerk M. Multiple voxel (1)H MR spectroscopy of phosphorylase-b kinase deficient patients (GSD IXa) showing an accumulation of fat in the liver that resolves with aging. *J Hepatol* 2006 December;45(6):851-5.

(29) Thomsen C, Becker U, Winkler K, Christoffersen P, Jensen M, Henriksen O. Quantification of liver fat using magnetic resonance spectroscopy. *Magn Reson Imaging* 1994;12(3):487-95.

(30) Vuppalanchi R, Cummings OW, Saxena R, Ulbright TM, Martis N, Jones DR, Bansal N, Chalasani N. Relationship among histologic, radiologic, and biochemical assessments of hepatic steatosis: a study of human liver samples. *J Clin Gastroenterol* 2007 February;41(2):206-10.

(31) Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986 February 8;1(8476):307-10.

(32) Clark JM, Diehl AM. Nonalcoholic fatty liver disease: an underrecognized cause of cryptogenic cirrhosis. *JAMA* 2003 June 11;289(22):3000-4.

(33) 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. *J Hepatol* 2005 January;42(1):132-8.

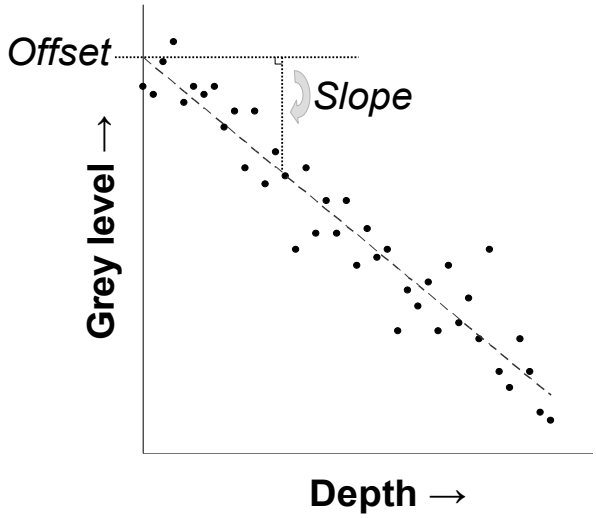
(34) Joseph AE, Saverymuttu SH. Ultrasound in the assessment of diffuse parenchymal liver disease. *Clin Radiol* 1991 October;44(4):219-21.

(35) Mazzone AM, Urbani MP, Picano E, Paterni M, Borgatti E, De Fabritiis A, Landini L. In vivo ultrasonic parametric imaging of carotid atherosclerotic plaque by videodensitometric technique. *Angiology* 1995 August;46(8):663-72.

(36) Yin M, Woollard J, Wang X, Torres VE, Harris PC, Ward CJ, Glaser KJ, Manduca A, Ehman RL. Quantitative assessment of hepatic fibrosis in an animal model with magnetic resonance elastography. *Magn Reson Med* 2007 August;58(2):346-53.

(37) Merriman RB, Ferrell LD, Patti MG, Weston SR, Pabst MS, Aouizerat BE, Bass NM. Correlation of paired liver biopsies in morbidly obese patients with suspected nonalcoholic fatty liver disease. *Hepatology* 2006 October;44(4):874-80.

Online Appendix I. Realization of attenuation based indices



Schematic display of offset and slope realization

Along the attenuation line (figure I), each pixel was automatically selected.

For each pixel under the line, the grey level value was averaged with the grey level values of 3 pixels to the left and 3 pixels to the right horizontally, and stored together with depth information. A linear regression line was applied, using the least-squares approximation, and its corresponding slope (grey level units/ mm) and offset (grey level) were generated [19].

Additionally, the length of the attenuation line (number of pixels) was generated.

Online Appendix II. Realization of textural based indices

Indices of the spatial grey level dependence matrix [19, 21, 22, 24, 25, 35]_ The aim of co-occurrence features is to capture texture characteristics, i.e. heterogeneity. Elements of the co-occurrence matrix (algorithm [1]) designate the probability that two pixels located within a region, separated by distance *d* along direction θ, have grey level values of *i* and *j*:

Co-occurrence: $P(i,j \mid d,\theta) = \frac{\mathrm{N}_{d\theta}((k,l),(m,n))}{\mathrm{N}}((k,l),(m,n)) \in [L_x, L_y],$ [1]

where θ = 0º and *d* = 4 pixels [19], $\mathrm{N}_{d\theta}$ is the number of pixel pairs, and N is the number of grey level transitions, in a region (L_x, L_y).

- Co-occurrence entropy: $\sum_{i=0}^{N_G-1}\sum_{j=0}^{N_G-1} P(i,j \mid d,\theta)\cdot\log(P(i,j \mid d,\theta)),$ [2]
- Co-occurrence sum entropy: $\sum_{k} P_{sum}(k)\cdot\log(P_{sum}(k)),$ [3]

where $P_{sum}(k) = \sum_{i=0}^{N_G-1}\sum_{j=0}^{N_G-1} P(i,j \mid d,\theta)$ *for i + j = k.*

Indices of the grey level difference matrix [22, 24, 25]_ The aim of difference features is to capture texture characteristics, i.e. homogeneity. Elements of the difference matrix (algorithm [4]) designate the probability that after a displacement along vector δ within a region, pixels will have grey level value *i*:

- Difference: $f'(i \mid \delta) = P(I_\delta(x,y) = i) \in [L_x, L_y],$ [4]

where $\delta = (\Delta x, \Delta y)$ and $I_\delta(x,y) = | I(\mathrm{x},\mathrm{y}) - I(x+\Delta x, y+\Delta y) |$, in a region ($L_x$, L_y).

- Difference contrast: $\sum_{i=0}^{N_G-1} i^2 f'(i \mid \delta)$ [5]
- Inverse difference moment: $\sum_{i=0}^{N_G-1} \frac{f'(i \mid \delta)}{i^2+1}$ [6]
- FP1: $\frac{\left(\sum_{x=1}^{M}\sum_{y=1}^{N}|I_\delta(x,y)|\right)}{MN}$, where *M* and *N* are column and row, respectively. [7]