Ultrasound in Med. & Biol., Vol. ■, No. ■, pp. 1–14, 2017

0301-5629/$ - see front matter

http://dx.doi.org/10.1016/j.ultrasmedbio.2017.05.002

● *Original Contribution*

A MACHINE-LEARNING ALGORITHM TOWARD COLOR ANALYSIS FOR CHRONIC LIVER DISEASE CLASSIFICATION, EMPLOYING ULTRASOUND SHEAR WAVE ELASTOGRAPHY

Ilias Gatos,* Stavros Tsantis,* Stavros Spiliopoulos,† Dimitris Karnabatidis,‡ Ioannis Theotokas,§ Pavlos Zoumpoulis,§ Thanasis Loupas,¶ John D. Hazle,‖ and George C. Kagadis*‖

*Department of Medical Physics, School of Medicine, University of Patras, Rion, Greece; †2nd Department of Radiology, School of Medicine, University of Athens, Athens, Greece; ‡Department of Radiology, School of Medicine, University of Patras, Patras, Greece; §Diagnostic Echotomography SA, Kifissia, Greece; ¶SuperSonic Imagine SA, Aix-en-Provence, France; and ‖Department of Imaging Physics, University of Texas MD Anderson Cancer Center, Houston, Texas, USA

(*Received* 5 *December* 2016; *revised* 26 *April* 2017; *in final form* 1 *May* 2017)

Abstract—The purpose of the present study was to employ a computer-aided diagnosis system that classifies chronic liver disease (CLD) using ultrasound shear wave elastography (SWE) imaging, with a stiffness value-clustering and machine-learning algorithm. A clinical data set of 126 patients (56 healthy controls, 70 with CLD) was analyzed. First, an RGB-to-stiffness inverse mapping technique was employed. A five-cluster segmentation was then performed associating corresponding different-color regions with certain stiffness value ranges acquired from the SWE manufacturer-provided color bar. Subsequently, 35 features (7 for each cluster), indicative of physical characteristics existing within the SWE image, were extracted. A stepwise regression analysis toward feature reduction was used to derive a reduced feature subset that was fed into the support vector machine classification algorithm to classify CLD from healthy cases. The highest accuracy in classification of healthy to CLD subject discrimination from the support vector machine model was 87.3% with sensitivity and specificity values of 93.5% and 81.2%, respectively. Receiver operating characteristic curve analysis gave an area under the curve value of 0.87 (confidence interval: 0.77–0.92). A machine-learning algorithm that quantifies color information in terms of stiffness values from SWE images and discriminates CLD from healthy cases is introduced. New objective parameters and criteria for CLD diagnosis employing SWE images provided by the present study can be considered an important step toward color-based interpretation, and could assist radiologists' diagnostic performance on a daily basis after being installed in a PC and employed retrospectively, immediately after the examination. (E-mail: gkagad@gmail.com) **© 2017 World Federation for Ultrasound in Medicine & Biology.**

Key Words: **Fibrosis, Shear wave elastography, Computer-aided diagnosis, Classifier design, Ultrasonics.**

INTRODUCTION

Chronic liver disease (CLD) is considered the 12th leading cause of death in the United States, with a continuously increasing rate in recent years. Especially at the final stage, cirrhosis, mortality rates are significant for those 45–85 y of age (Murphy et al. 2013). In 2013, cirrhosis was responsible for more than 1.2 million deaths worldwide, which represented a 50% increase from the early 1990s (GBD Mortality and Causes of Death Collaborators 2015). In addition, hepatocellular carcinoma (HCC) caused by cirrhosis is the third leading cause of cancer mortality worldwide (Altekruse et al. 2009).

Chronic liver disease is caused by hepatitis viruses (HAV, HBV, HCV, HDV and HEV), extensive alcohol consumption, unhealthy dietary patterns, which may lead to non-alcoholic fatty liver disease (NAFLD) or non-alcoholic steatohepatitis (NASH), autoimmune hepatitis (AIH) and genetic metabolic disorders such as Wilson's disease and haemochromatosis. It can be also caused by primary biliary cirrhosis (PBC), secondary biliary cirrhosis (SBC) or primary sclerosing cholangitis (PSC). CLD causes liver tissue inflammation, which in turn leads to repetitive injury and fibrosis. The disease, if not treated, leads in several years to cirrhosis and partial

Address correspondence to: George C. Kagadis, Associate Professor of Medical Physics–Medical Informatics, Department of Medical Physics, School of Medicine, University of Patras, Rion, GR 265 04, Greece. E-mail: gkagad@gmail.com

or total liver malfunction. A major challenge for clinicians today is to monitor the disease progression and estimate the patient's fibrosis stage to select the best treatment. In the case of cirrhosis and HCC, liver transplantation and surgical removal of the malignant lesion are to be decided.

Fibrosis staging is based on the Metavir classification system, according to liver biopsy (LB). The Metavir scale consists of five stages ranging from 0 to 4 (F0 = no fibrosis, F1 = mild fibrosis, F2 = significant fibrosis, F3 = severe fibrosis, F4 = cirrhosis) (Goodman 2007). Although LB is still considered the gold standard for estimation of fibrosis stage, it carries serious limitations. It is invasive and costly, and nearly 30% of patients have post-surgical effects such as substantial pain, pneumothorax, bleeding, infection, septicemia, biloma, haemobilia, accidental injury to adjacent structures, biliary peritonitis and even death (Gilmore et al. 1995).

Another drawback of LB is the lack of fibrosis uniformity across the liver tissue. Biopsy needles provide a small volume of tissue (1/50,000th of the total mass of the liver), resulting in staging variability between samples derived from different liver areas (Carey and Carey 2010; Goodman 2007).

New and non-invasive approaches to fibrosis stage assessment have been developed in the last few years. These include biochemical serum markers (BSMs) and imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound (US). BSMs have exhibited adequate performance in estimating significant fibrosis (F ≥ F2), and cirrhosis (F4), but failed in early-stage estimation (Carey and Carey 2010; Frulio and Trillaud 2013; Martinez et al. 2011). With respect to CT imaging, Romero-Gómez et al. (2008) reported area under the curve (AUC) values of 0.83 for F ≥ F2 and 0.86 for F ≥ F3 in patients with chronic hepatitis C. Magnetic resonance elastography (MRE) is a contemporary imaging technique exhibiting great accuracy at all stages. Although very accurate in the few studies in which it has been tested (Carey and Carey 2010; Huwart et al. 2008), MRE is expensive and requires extensive validation. B-Mode ultrasound (US) only predicts diffuse abnormalities (Sanford et al. 1985) and confirms the absence of cirrhosis (Giorgio et al. 1986; Harbin et al. 1980).

Over the last decade, US elastography (USE) has been introduced to correlate liver stiffness to fibrosis stages. Several elastographic approaches have been introduced for CLD evaluation, such as transient elastography (TE), acoustic radiation force impulse (ARFI), real-time elastography (RTE) and shear wave elastography (SWE) imaging. All clinical studies dealing with USE are based on stiffness cutoff values that correspond to fibrosis stages. TE carried out with the Fibroscan US system (Echosens, Paris, France) is the oldest and most validated elastographic method employed for liver fibrosis evaluation. The mean diagnostic accuracy of TE has an AUC value of 0.85. with average sensitivity and specificity values of 0.78 and 0.79 for F ≥ F2, F ≥ F3 and F ≥ F4 (Bota et al. 2013; Chung et al. 2013; Frulio and Trillaud 2013; Tsochatzis et al. 2011). TE suffers from optimal stiffness cutoff value validation because of the overlap between different fibrosis stages in several studies. RTE, developed by Hitachi, also has a mean AUC value of 0.85 and mean sensitivity and specificity values of 0.83 and 0.77 for all stages above F2 (Chung et al. 2013; Hong et al. 2014; Kobayashi et al. 2015). Meta-analyses have indicated that the qualitative nature of RTE provides only useful visualization of stiffness contrasts and not quantification of fibrosis stages (Hong et al. 2014; Kobayashi et al. 2015). ARFI elastography, recently introduced by Siemens, is considered a useful tool for fibrosis evaluation, with results similar to those for TE (0.85 mean AUC value) (Bota et al. 2013; Chung et al. 2013; Frulio and Trillaud 2013; Nierhoff et al. 2013). Nevertheless, ARFI imaging carries several limitations; elasticity measurements are not given in real time, and only one acquisition can be made each time in a small pre-determined and constant area. Until today, only few ARFI-based studies have been published, and the validity and utility of the method have not been established. SWE is yet another recently introduced technique that offers real-time elasticity imaging as well as stiffness quantification (in units of kPa) over a 2-D region of interest; it is commercially available on the Aixplorer US scanner (SuperSonic Imagine, Aix-en-Provence, France). SWE offers good classification results (mean AUC values of 0.81 for F ≥ F1 and 0.86 for F ≥ F2), but, like the previously mentioned techniques, also has limitations such as stiffness cutoff threshold variability between different studies, uncertainties in the selection of the optimum quantification region and a lack of clear guidelines on how to avoid artifacts and areas of poor image quality (Bota et al. 2015; Deffieux et al. 2015; Ferraioli et al. 2012a, 2015; Gerber et al. 2015; Sporea et al. 2014). Only few elastographic studies (TE, RTE, ARFI, and SWE) provide mild fibrosis (F ≥ F1) classification results, compared with significant fibrosis (F ≥ F2) and cirrhosis (F = F4). For SWE clinical studies, only Sporea et al. (2014) provided liver stiffness cutoff values for predicting mild fibrosis using TE as the gold standard. The cutoff value for F ≥ F1 was set at 7.1 kPa, achieving an AUC value of 0.825 and total accuracy of 76% with sensitivity and specificity values of 74.5%, and 78%, respectively (Sporea et al. 2014).

Several attempts that quantify sonographic findings by means of pattern recognition algorithms to classify CLD have been reported so far. Regarding B-mode US

imaging, Huet et al. (2015) recently evaluated the diagnostic accuracy of an automatic software applied to a cohort of 114 patients with CLD divided into two classes (F0–F1), and (F2–F4) with an AUC value of 0.80. Various image analysis studies have been conducted on RTE images for liver fibrosis assessment. The approach was to extract first- and second-order textural features from the RTE elastograms to calculate the liver fibrosis index (LFI) *via* multiple regression analysis (MRA). LFI AUC values for each of the fibrosis stages ranged from 0.66 to 0.93 (Colombo et al. 2012; Ferraioli et al. 2012b; Fujimoto et al. 2013; Tatsumi et al. 2010; Wang et al. 2012; Yada et al. 2013). No image analysis studies have been reported so far for TE and ARFI elastographic methods.

For SWE imaging, there are two automatic quantification studies so far, by Pellot-Barakat et al. (2015) and Gatos et al. (2016a). In the first study, factors such as temporal stability of an area and low heterogeneity across frames toward optimized ROI selection were evaluated to assess whether they can improve the reliability and reproducibility of the SWE method for CLD evaluation. No classification accuracies were reported in that study. Gatos et al. (2016a) proposed a computer-aided diagnosis (CAD) system that automatically quantified factors that are important in differentiating healthy persons from patients with CLD (F0–F1 and F2–F4) by extracting and evaluating gray-tone co-occurrence and run-length matrix features from the RGB color-box elastogram. The proposed CAD scheme involved automatic segmentation of the color box, inverse mapping of the RGB values to stiffness values based on the SWE color map used by the Aixplorer and then feature extraction and evaluation by means of support vector machine (SVM) classification. Maximum classification accuracy between healthy and CLD patients was 87.0%, with sensitivity and specificity values of 83.3% and 89.1%, respectively, employing the mean gradient magnitude, gray-level non-uniformity and run-length non-uniformity feature combination. Receiver operating characteristic (ROC) curve analysis gave an area under the ROC curve (AUC) of 0.85 with a confidence interval (CI) of 0.77–0.89. All of these features have established heterogeneity as an important factor in CLD assessment.

Lo et al. (2015) extracted a series of quantitative features from the SWE color box to evaluate breast tumor elasticity automatically. These features were derived from histogram color distributions and were fed to a logistic regression model for breast tumor classification. No actual tissue stiffness information was employed within this study. Tissue stiffness was indirectly estimated by setting a basic color value (blue, green, yellow and red) to each pixel of the color box that is closer to one of the four color clusters. This method is limited by the lack of opacity parameter information, which could produce blended B-mode and stiffness information from each SWE image. Only a value of 100% provides exact correspondence between RGB color values of the SWE color box and the color map used by the machine for stiffness-to-color mapping.

In the present study, an automatic machine-learning algorithm was introduced to further investigate additional factors in CLD evaluation. Weights of importance were assigned to all stiffness value clusters present in the SWE image employing features derived from the color box set by the built-in software by Aixplorer's manufacturer. It is considered the second stage of a three-stage ongoing research effort of our group toward CLD evaluation in SWE images by means of machine-learning algorithms. The primary research stage described by Gatos et al. (2016a) evaluated the elastogram in terms of one cluster, considering only the elastogram values and not the color map provided by the ultrasound machine.

In the study described here, an attempt was made, in a larger data set, to engage the physician drastically in the clinical decision procedure, exploiting stiffness value clusters and the features derived from them. This contrasts with the "black box" procedure of complex and difficult-to-evaluate second-order statistical features. The main aim of this study was to define feasible rules and guidelines based on the importance of each stiffness value cluster on the elastogram. These rules are defined quantitatively with boxplots of the important features and descriptively with the depiction of the behavior of these features in either healthy persons or CLD patients. We strongly believe that the examiner's engagement in SWE imaging evaluation with the provided set of rules could enhance her or his diagnostic performance by altering the evaluation procedure and thus improving the study's clinical impact.

With respect to the third stage, a new protocol of SWE examination has been defined: acquisition of a set of SWE images of the same frame through time to find regions that maintain relative stability over time (Gatos et al. 2016b). The purpose of this study will be to exclude areas in the elastogram that present high stiffness variability and thus may be unreliable to be included in the analysis. A new acquisition protocol will be needed for this study (a set of SWE images instead of a single SWE image). To date, a limited number of patients have been examined using this approach, and no reliable results are currently available. The proposed machine-learning scheme starts with automatic segmentation of the color-bar and color-box objects on the image, followed by an inverse mapping technique (RGB-to-stiffness value space). Then a five-stiffness-value cluster segmentation procedure is employed, followed by feature extraction from all cluster areas in the color box.

All features calculated are fed to a classification scheme based on SVMs to classify healthy *versus* CLD patients. The classification accuracy of the SVM classifier was evaluated by means of the leave-one-out (LOO) method and ROC curves analysis. In addition, a thorough boxplot analysis of the most discriminant features is introduced to provide more color-based details that are indicative of CLD occurrence. Figure 1 is a schematic of the steps in the proposed algorithm.

METHODS

Clinical data set

The clinical data set comprised 126 SWE liver images from 126 individuals (62 men and 64 women) acquired from March 2014 to April 2016; mean age was 56.7 ± 16.0 y (range: 18–79 y). These liver cases comprised 56 healthy persons (F0) and 70 with CLD (23 with F1, 11 with F2, 13 with F3 and 23 with F4). Inclusion criteria for the healthy group were normal biochemical markers, the absence of clinical history and normal liver findings in their US examination. For patients with obesity or ascites, which often cause variation in SWE imaging, only patients with an acoustic window that produced acceptable SWE image quality (images with large areas of stable-through-time stiffness information and large color-filling percentage in the color box) were included in this study.

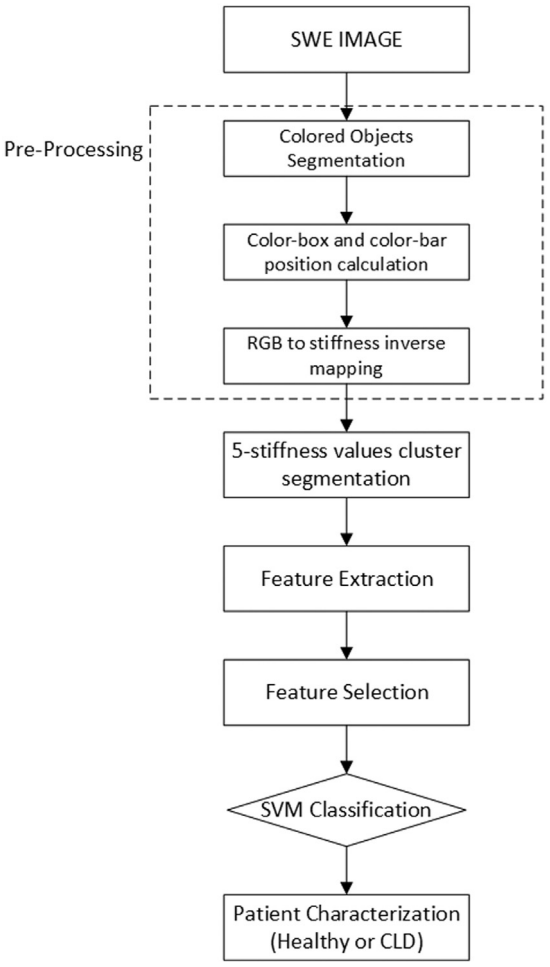

Fig. 1. Schematic of the steps of the proposed algorithm. CLD = chronic liver disease; SVM = support vector machine; SWE = shear wave elastography.

All patients from the CLD group underwent liver biopsy and histologic examination for fibrosis stage estimation by a senior pathologist. Histologic examination results were employed as the gold standard to evaluate the performance of the proposed classification system. Regardless of the lack of registration between biopsy region and SWE measurement, LB is considered the gold standard for all image analysis studies presented in the current literature mainly because of the uniform progression of diffuse disease within the liver. The maximum interval between SWE examinations and the corresponding histologic examinations was approximately 2 mo.

The present study was conducted in accordance with the ethical guidelines of the Helsinki Declaration and was approved by our institutional review board; written informed consent was obtained from each person participating in the study.

SWE imaging protocol

In both groups (healthy and CLD), SWE imaging was performed on the Aixplorer ultrasonic system (SuperSonic Imagine, Aix-en-Provence, France) with an SC6-1 curvilinear transducer by an expert radiologist with more than 10 y of experience in SWE and more than 20 y of experience in US examinations. All SWE acquisitions were performed after obtaining a homogenous harmonic B-mode US image with no visible vessels, by employing the Aixplorer's Abdominal/Liver system preset intended for liver fibrosis assessment exams. Elastography mode is set in Aixplorer to show up the color box, superimposed on B-mode. The imaging parameters used were those specified by the Abdominal/Liver preset by default, except for the SWE opacity parameter, which was set to 100% to allow the inverse mapping procedure described later under Pre-processing. The depth of the SWE quantification region of interest (ROI) (subsequently referred to as the "Q-Box") ranged between 3 and 6 cm. The SWE image size used for the subsequent analysis was 1440 × 1080 pixels.

Each patient was asked to inspire normally to achieve diaphragm stabilization and liver immobility. Two SWE images per patient were acquired for the aims of the present study: one with no markings on the elastogram used in the subsequent image analysis

procedure, and another that included a circular Q-Box (default diameter of 1 cm) within the elastogram. This was placed by the radiologist to manually measure the mean tissue stiffness within the Q-Box. A comparative study has been made to evaluate the classification accuracy of the proposed CAD system against the results derived from the radiologists.

In addition, a second experienced radiologist evaluated the same clinical data set after the acquisition protocol based on European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) guidelines (Cosgrove et al. 2013). For clinical evaluation (F0 vs. CLD), both radiologists employed Sporea et al.'s cutoff value of 7.1 kPa for mild fibrosis (F ≥ F1) (Sporea et al. 2014). A measurement >7.1 kPa is classified as F0, whereas a measurement ≥7.1 kPa is classified as CLD. Two confusion matrices were constructed: (i) one from the comparison of the radiologists' measurement versus LB and (ii) one from the comparison of the binary outcome of the SVM classifier versus LB. Total accuracies were then calculated to evaluate both the cutoff value set by Sporea et al. and the performance of the proposed machine-learning algorithm.

Pre-processing

The pre-processing step involved inverse mapping from the RGB color map to the stiffness (kPa) value space to obtain quantitative data suitable for the subsequent image analysis procedure. The inverse mapping procedure comprises the automatic detection of the color bar positioned at the top right of the SWE image (which provides the color map used to convert stiffness estimates [in units of kPa] into RGB color triplets) and the color box containing the SWE estimates, followed by the actual RGB-to-kPa inverse mapping as previously described by Gatos et al. (2016a). More specifically, the colored components present within the SWE image (Fig. 2a) are first isolated (Fig. 2b). RGB values (header, footer, color bar and color box; distinct components) satisfying the condition R, G, B exist in each SWE image. Color-bar isolation involves the centroid computation of every non-connected component where the one with the maximum abscissa (closest component located on the right border of the image) indicates the color bar (Fig. 2c). The color map used by the Aixplorer scanner to map stiffness estimates (in units of kPa) to the RGB color triplets (Fig. 2d) is extracted by an image profile

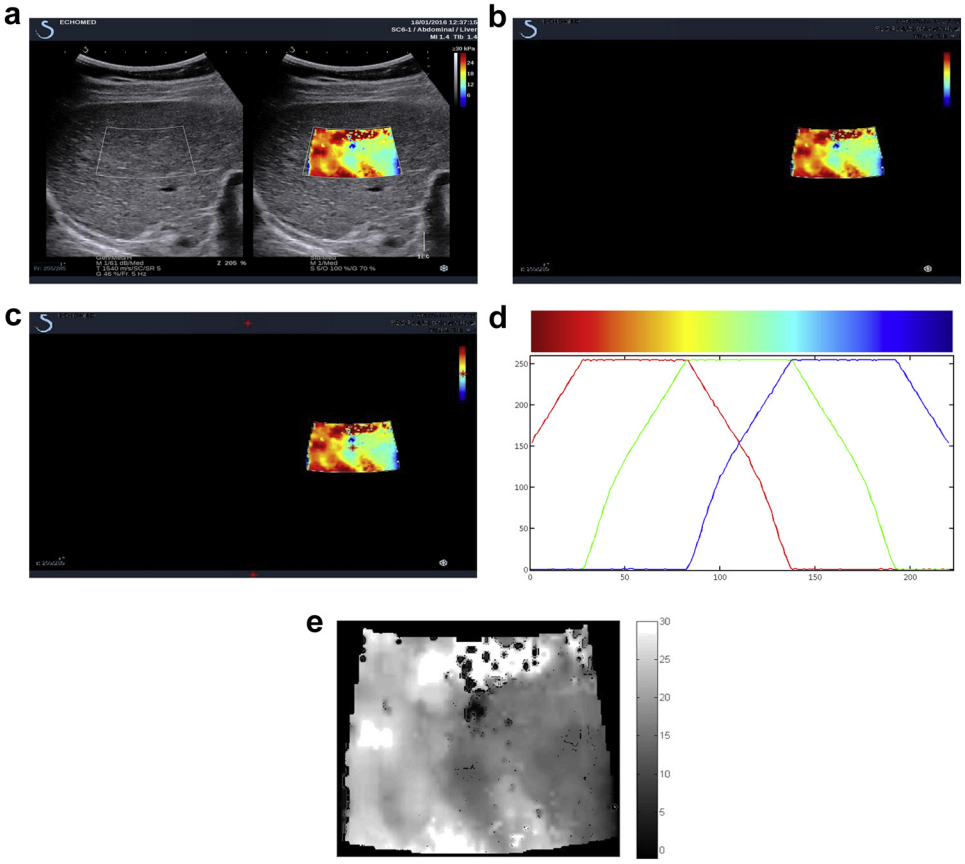

Fig. 2. (a) Initial SWE image. (b) Isolation of colored objects. (c) Calculation of centroids of colored objects (*red stars*). (d) Image profile vector through the first column of the RGB color bar. (e) Stiffness map calculated from RGB- to-stiffness inverse mapping.

vector through the first column of the RGB color bar. The resulting stiffness matrix (Fig. 2e) can be presented as a gray-scale image; black and white values correspond to 0 and maximum stiffness values (MaxVal) in kilopascals, respectively.

Stiffness value clustering procedure

To evaluate the impact of specific stiffness value regions in CLD discrimination and fibrosis staging, a stiffness values clustering procedure was implemented. The number of clusters chosen for the present study was associated with the corresponding different-color regions of equal stiffness value range from the Aixplorer's color bar. The stiffness value scale 0–MaxVal was divided uniformly into five intervals, each corresponding to each cluster employed in this study. The color-bar division is based on the built-in software built into the Aixplorer by the manufacturer and is employed in daily clinical practice worldwide.

Each stiffness value range represents a color characterization provided by the color bar (blue, aqua, green, yellow, and red) (Table 1).

As an example, the results of the stiffness value clustering procedure on a patient with cirrhosis (F4) are illustrated in Figure 3, in which each stiffness value cluster is defined by the equal stiffness value ranges of the color map (the MaxVal for the employed image data set was 30 kPa).

Feature extraction

From each of the five stiffness value clusters detected, a set of seven features is computed, resulting in 35 features for each SWE image. These features are the mean and median values, standard deviation, sample kurtosis, 10th and 90th percentile values and normalized pixel area of each cluster within the elastogram (Table 2). The selected feature set for each cluster represents physical characteristics existing within the SWE image that are noticeable by the radiologist and can be evaluated for CLD classification in the daily clinical routine.

Sample excess kurtosis of the cluster's histogram is calculated with the equation (Joanes and Gill 1998)

$$g_2 = \frac{m_4}{m_2^2} - 3 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2} - 3 \quad (1)$$

where m_4 is the fourth sample moment about the mean, m_2 is the second sample moment about the mean (sample variance), x_i is the ith value and $\bar{x}$ is the sample mean.

The 10th and 90th percentile values are the stiffness values below 10% and 90%, respectively, of all observations detected in each cluster's histogram. Normalized pixel area (NPA) represents each cluster's pixel area within the elastograms divided by the total pixel area:

$$\text{NPA} = \frac{\text{number of pixels with valid stiffness values within cluster}}{\text{total number of pixels with valid stiffness values in elastogram}} \quad (2)$$

Pixels with valid stiffness values refer to those pixels, within each cluster, that have been assigned quantitative stiffness values by the Aixplorer's shear wave speed estimation algorithm, as opposed to pixels without quantitative stiffness values that are depicted by means of the B-mode gray-scale color map.

These features were selected to augment the information already provided by the Aixplorer's Q-Box (mean, standard deviation, minimum and maximum values) in the segmented regions. Median value was selected as an additional and more robust feature compared with mean value, whereas sample kurtosis as an alternative index to standard deviation for heterogeneity and outlier detection. The 10th and 90th percentile values are less sensitive to noise and outliers than the minimum and maximum values. Finally, NPA was selected to define each cluster's extent compared with the total valid stiffness information. With NPA values close to 1, the corresponding cluster dominates the elastogram. On the contrary, NPA values close to 0 are indicative of clusters with considerably small pixel numbers. In rare cases, where certain clusters are absent within the SWE image, the NPA feature was set to 0, and the remaining features for that cluster were removed from further analysis. These cases are solely F0 and F1 cases with absent elastography values representing yellow and red clusters indicative of high fibrosis stages.

Table 1. Color clusters and the corresponding stiffness value ranges

a/a	Cluster color	Stiffness value range (kPa)
1	I. Blue	0.0–6.0
2	II. Aqua	6.1–12.0
3	III. Green	12.1–18.0
4	IV. Yellow	18.1–24.0
5	V. Red	24.1–30.0

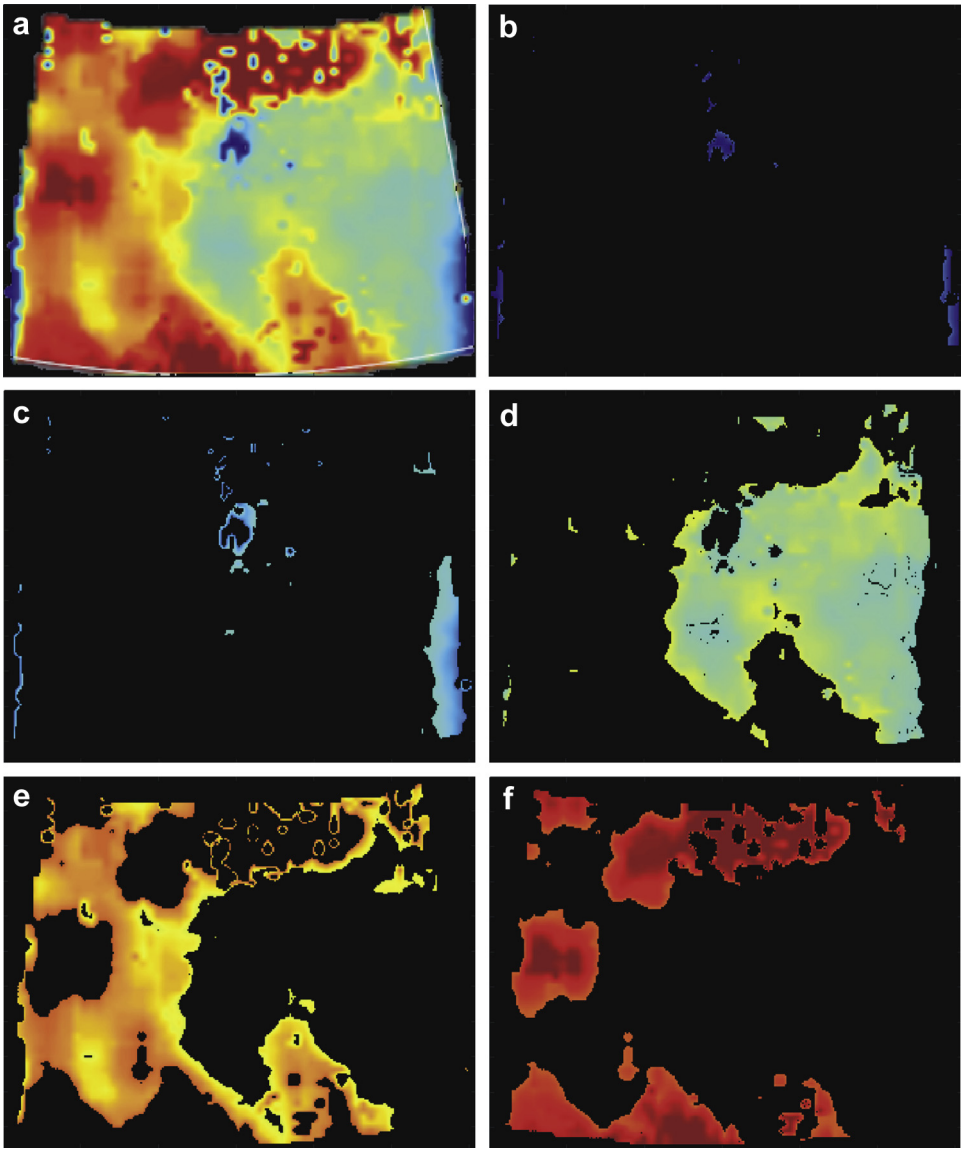

Fig. 3. (a) SWE color box extracted. (b) Cluster I: blue. (c) Cluster II: aqua. (d) Cluster III: green. (e) Cluster IV: yellow. (f) Cluster V: red.

Feature selection and classification

Before classification, the number of features needed to be reduced to avoid excessively long processing times. The feature reduction goal was to provide the most relevant parameter set for CLD discrimination and classification and boost performance of the subsequent machine-learning algorithm. Feature reduction within this study was performed by means of stepwise regression (SR) employing biopsy as the gold standard. SR is considered a sequential feature selection approach, which comprises an objective function as a criterion that aims to be minimized by means of least-squares fitting over all feasible feature subsets, and a sequential search algorithm that adds or removes features from a candidate subset while evaluating the criterion (Draper and Smith 1998). It is based on a scheme that adds or removes features from a multilinear model based on their statistical significance in a regression. An initial model is introduced followed by a sequential comparison between incrementally larger and smaller models in terms of significance power.

Table 2. Feature set computed for each color cluster

a/a	Feature name
1	Mean
2	Median
3	Standard deviation
4	Sample kurtosis
5	10th percentile
6	90th percentile
7	Normalized pixel area

At each step, the p value of an F-statistic is computed to test models with and without a potential term. In the case of statistical significance, the term is added to the model, whereas when the null hypothesis is rejected, the term is removed. The method is terminated when no additional step improves the model. Unlike generalized sequential feature selection, stepwise regression may remove features that have been added or add features that have been removed in previous steps. The feature reduction procedure employed managed to accomplish two goals. First, it defined the ability of each feature to separate the two classes while keeping every class as close-fitted as possible, which in turn determines the discrepancy between them. Second, it reduced computation load in real-time applications such as image classification because of the smaller size of the reduced feature subset. The first goal in employing SR provides a feature subset that has strong discriminant properties. The second goal provides the subsequent classification algorithm with the flexibility to search for the best feature combination to optimize classification performance. The multidimensional best feature combination derived from the SVM classifier indicates the importance of feature diversity in CLD recognition. Filter-based approaches to feature selection such as SR provide identification of the discriminant features, unlike the so-called wrapper approaches such us principal component analysis (PCA), which can only estimate a good feature subset (Huang et al. 2003).

For the aims of the present study, the SVM classifier was chosen. The SVM classifier finds the optimum hyperplane that separates data points between two classes. An optimum hyperplane for an SVM classifier considers as a prerequisite the largest margin between the two classes. The support vectors for both classes are the data points that are closest to the separating hyperplane. The discriminant function of the classifier is given by the equation

$$g(x) = \text{sign}\left(\sum_{i=1}^{N_s} a_i y_i k(x_i, x) + b\right) \quad (3)$$

where a_i are weight parameters, $k(\mathbf{x}_i,\mathbf{x})$ is the kernel function employed for the data transformation into the linearly separable feature space, $\mathbf{x}_i$ are the support vectors (*i.e.,* the training pattern vectors that have their corresponding weights $a_i \neq 0$), N_S is the number of support vectors, $\mathbf{x}$ is the input pattern vector, b is the bias or threshold and $y_i \in \{-1, +1\}$, depending on the class.

In the present work, the SVM classifier was designed by employing various polynomial kernels up to the fourth degree and the radial basis function (RBF) kernel. These kernel functions are described by the relations

$$k_{\text{POLYNOMIAL}}(x_i, x) = \left(\left(x_i^T x\right) + 1\right)^d,\ d = \text{degree} \quad (4)$$

$$k_{\text{RBF}}(x_i, x) = \exp\left(\frac{-(x_i - x)^T (x_i - x)}{2\sigma^2}\right) \quad (5)$$

where σ is the standard deviation.

The best feature combination for maximum performance of the classifier was exploited by means of an exhaustive search that involved designing the classifier by means of every possible feature combination (*i.e.,* 2, 3, 4, …, 9 feature combinations), each time testing the classifier's performance in correctly classifying SWE data and finally selecting that feature combination that had the highest classification accuracy with the smallest number of features.

The classification accuracy of the SVM classifier during this exhaustive search was evaluated by means of the LOO method and ROC curve analysis. In the LOO method, training is performed in all but one SWE feature vector ($N - 1$), and the test is consequently applied to the excluded SWE vector. The training/testing procedure is iterated N times, each time excluding a different SWE vector. The LOO method's iterative training procedure employs essentially all cases while maintaining the independence between training and testing feature vectors.

An ROC curve is a plot of the true-positive rate (sensitivity) versus the false-positive rate (1 − specificity) for different thresholds over the entire range of each classifier discriminant function output value (Lasko et al. 2005). In contrast to the classification accuracies obtained from truth tables, an ROC analysis is independent of class distribution or error costs. In this study, the AUC value was obtained with the empirical non-parametric method to approximate the area. Inter- and intra-observer variability between the two experienced radiologists was made by means of the intra-class correlation index (ICC). The proposed algorithm's running times are 7.9 ± 0.1 s for color-bar and color-box extraction, 4.1 ± 0.6 s for RGB-to-stiffness conversion and less than a second for best feature combination calculation. Generally, the proposed algorithm, after training and testing phases, has running times less than 13 s.

RESULTS

For the SVM model designed for this study, accuracy for each of the nine features individually was inferior (<80.0%) to that of the best feature combination presented, reaching an overall accuracy of 87.3%. Highest

classification accuracy between healthy persons (F0) and patients with chronic liver disease (F1–F4), combined with the minimum number of features by the LOO method, was achieved by the feature combination "blue StD, blue 10th percentile, aqua 10th percentile and aqua NPA" employing the RBF kernel.

The parameter C that controls the error-margin trade-off was set to 150 (approximately $\max(|\bar{y}+3\sigma|, |\bar{y}-3\sigma|)$, where $\bar{y}$ and σ are the training data mean and standard deviation) (Draper and Smith 1998), and the standard deviation in the RBF kernel was set at 0.5, which lies within the [0, 1] range that yields good SVM performance for various training sets (Cherkassky and Ma 2004).

The selected feature subset (nine features) acquired by SR analysis and fed as input into the subsequent SVM model for classification is outlined in Table 3.

Table 4 summarizes the results obtained by the SVM classifier with the leave-one-out method for different kernel functions.

Table 5 gives a detailed account of SVM-RBF kernel classification accuracies obtained with the LOO method, employing the best feature combination.

Fifty-two healthy participants were correctly classified, whereas 4 cases were incorrectly assigned to the CLD group, yielding a negative predictive value (NPV) of 92.9% by the LOO method. In the case of CLD patients, 58 were assigned to the correct group, whereas 12 were erroneously classified in the healthy group, for a positive predictive value (PPV) of 82.9%. Overall, the SVM achieved 87.3% precision in correctly distinguishing healthy participants from CLD patients. The number of support vectors for the best feature combination was 25, corresponding to 20% of the number of training points. Sensitivity and specificity values were 93.5% and 81.2% respectively.

Receiver operating characteristic curve analysis of the SVM classifier results yielded an AUC value of 0.87 with a 0.77–0.92 confidence interval. Figure 4 is the empirical ROC curve of the SVM (RBF kernel).

Table 3. Subset of features selected by stepwise regression analysis

a/a	Feature name
1	Blue mean
2	Blue standard deviation
3	Blue 10th percentile
4	Blue normalized pixel area
5	Aqua mean
6	Aqua 10th percentile
7	Aqua 90th percentile
8	Aqua normalized pixel area
9	Red normalized pixel area

Table 4. Classification accuracies for various SVM kernels using the leave-one-out for the "blue StD, blue 10th percentile, aqua 10th percentile and aqua NPA" best feature combination

	Classification accuracy	
SVM kernel	LOO$^+$ (%)	N_{SV}
Polynomial of 1st degree	79.2	42
Polynomial of 2nd degree	81.7	34
Polynomial of 3rd degree	86.1	29
Polynomial of 4th degree	84.2	32
Radial basis function	87.3	25

SVM = support vector machine; StD = standard deviation; LOO = leave one out; N_{sv} = number of support vectors employed using the LOO method.

The feature subset provided by SR analysis possesses discriminatory attributes, which, in terms of quantification, can be evaluated by the boxplots provided in Figure 5. The low–high parameter values of each cluster for both classes could be employed by the physician to extract certain guidelines and assign weights of importance to each stiffness value cluster. In addition, color boxes from representative cases of the clinical data set are depicted in Figure 6.

With respect to qualitative analysis of the features extracted, descriptive rules of group discrimination are depicted in Figure 7. Figure 7 illustrates the rules derived by the boxplots of Figure 5 that are indicative of CLD existence on a specific color area of the color box. The *arrows* indicate that high or low values of specific features on certain color areas and specific combinations between them are informative of a patient's condition with respect to CLD.

Clinical evaluation by the two radiologists yielded accuracy values of 75.3% and 76.6% with sensitivity/specificity values 72.2/80.1 and 73.8/81.3, respectively. The ICC index from the inter-observer variability study was 0.91, whereas the mean difference between the

Table 5. Confusion matrix of the SVM classifier employing the "blue StD, blue 10th percentile, aqua 10th percentile and aqua NPA" best feature combination employing the RBF kernel

	SVM classification (radial basis function kernel)		
SWE cases	Healthy	CLD	LOO accuracy
Healthy	52	4	(NPV) 92.9%
CLD	12	58	(PPV) 82.9%
Overall accuracy			87.3%

SVM = support vector machine; NPA = normalized pixel area; StD = standard deviation; SWE = shear wave elastrography; CLD = chronic liver disease; LOO = leave one out; NPV = negative predictive value; PPV = positive predictive value.

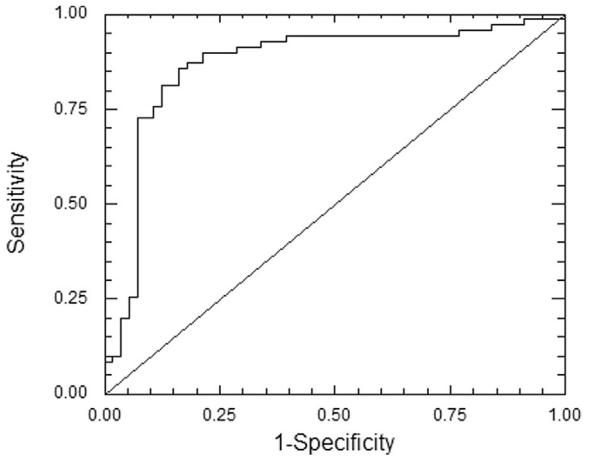

Fig. 4. Receiver operating characteristics curve derived by the empirical method of the support vector machine classifier with the radial basis function kernel employing the best feature combination toward the classification of chronic liver disease.

radiologists (the corresponding ROI mean value) was -0.26 ± 1.69. ICC indices of intra-observer variability study were 0.94 and 0.95 for the two examiners, respectively.

DISCUSSION

A sophisticated machine-learning algorithm for differentiation of healthy participants (F = F0) from CLD patients (F ≥ F1), exploiting features derived from different SWE stiffness values areas, was proposed. The performance of the proposed system was evaluated against histologic results on liver biopsy. The stiffness value segmentation procedure involved automatic detection of the SWE color box, inverse RGB-to-stiffness mapping and stiffness value clustering. This was accomplished by employing four constant thresholds on the stiffness value range (0–MaxVal) and acquiring five different stiffness values regions of equal stiffness value ranges from the Aixplorer's color bar. From each cluster, 7 features (35 in total) were extracted and consequently fed to the SVM classifier to discriminate healthy participants from CLD cases.

The proposed machine-learning algorithm reached an 87.3% classification accuracy (with sensitivity and specificity values of 93.5% and 81.2%, respectively), combined with an AUC value of 0.87 (0.77–0.92 CI). Gatos et al. (2016a) reached 87.0% total accuracy with a sensitivity and specificity of 83.3% and 89.1%, respectively, and an AUC value of 0.85 in a smaller data set (85 patients). The classifier designed for this study has also been evaluated in the new data set and achieved 84.9% accuracy, with sensitivity and specificity values of 77.6 % and 93.2%, respectively. With respect to misclassified patients, a significant number of cirrhotic patients (F = F4) were added to the false-negative cases compared with the previous data set. A significant number of cirrhotic patients have homogeneous high stiffness values, which in turn lowered the sensitivity of the SVM classifier. As far as specificity is concerned, it remained high in the new data set as well. The proposed method outperforms the standard SWE examination protocol employing Sporea's cutoff value of 7.1 kPa for mild fibrosis (F ≥ F1) (mean elasticity value of the manually selected ROI) performance by the two expert radiologists (75.3% and 76.6% classification accuracies, respectively) and several clinical studies published recently with average AUC, sensitivity and specificity values of 0.81, 0.78 and 0.80, respectively (Gerber et al. 2015; Kobayashi et al. 2015; Sporea et al. 2014; Tsochatzis et al. 2011). All other clinical studies (Bota et al. 2013, 2015; Chung et al. 2013; Deffieux et al. 2015; Ferraioli et al. 2012a, 2015; Frulio and Trillaud 2013; Hong et al. 2014; Nierhoff et al. 2013) employing SWE for CLD evaluation, mentioned in the Introduction, have not evaluated mild fibrosis stage (F ≥ F1) to compare results.

Compared with quantification studies employed on other US imaging modalities, the proposed machine-learning algorithm has a higher AUC value than in Colombo et al. (2012) (0.83) and a lower value than in Wang et al. (2012) (0.93). Wang et al. (2012) have reached an inferior specificity value (0.83). However, it must be outlined that both these methods employ the same elasticity index on different clinical data sets. The diverse results derived from these two studies question their performance reproducibility in unknown cases. All other quantification studies (Ferraioli et al. 2012b; Fujimoto et al. 2013; Huet et al. 2015; Tatsumi et al. 2010; Yada et al. 2013) did not evaluate mild fibrosis (F ≥ F1). From the feature subset derived in the feature selection procedure, high values of blue mean, blue 10th percentile, blue NPA and aqua NPA, combined with low values of blue standard deviation, aqua mean, aqua 10th percentile, aqua 90th percentile and red NPA, are suggestive of normal liver tissue (Fig. 6a). On the contrary, the opposite values (low values of blue mean, blue 10th percentile, blue NPA and aqua NPA, combined with high values of blue standard deviation, aqua mean, aqua 10th percentile, aqua 90th percentile and red NPA) are indicative of CLD (Fig. 6b). Employment of the aforementioned general decision rules (also depicted in Fig. 7) could be clinically useful in the differentiation of healthy persons from CLD patients by taking into account all the color information provided to the physician and not only the mean ROI value utilized in the clinical studies. The specific method provides real-time evaluation and is easy to use and perform, with a smooth learning curve.

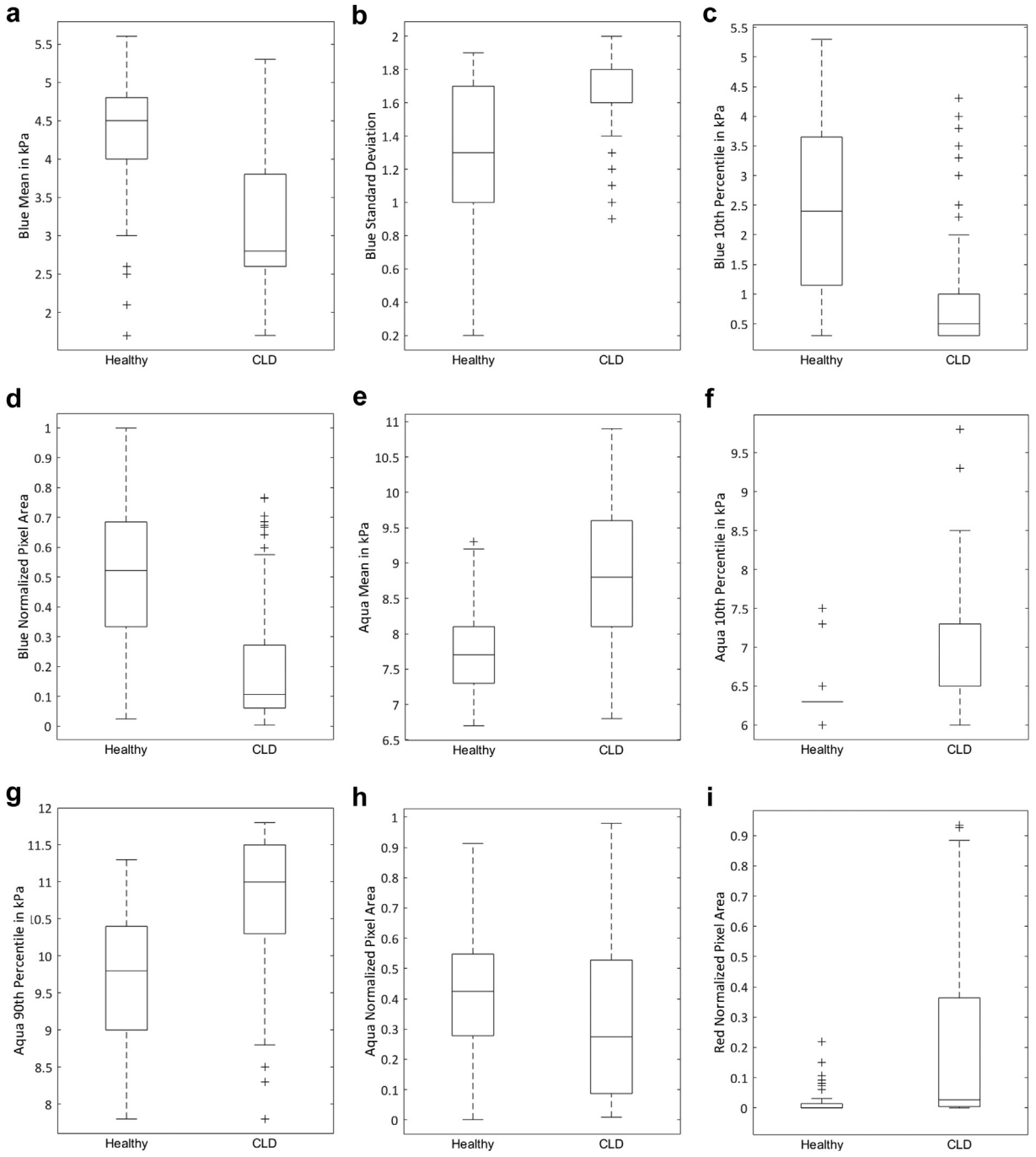

Fig. 5. Boxplots from the selected subset of both groups (chronic liver disease and healthy). (a) Blue mean. (b) Blue standard deviation. (c) Blue 10th percentile. (d) Blue NPA. (e) Aqua mean. (f) Aqua 10th percentile. (g) Aqua 90th percentile. (h) Aqua NPA. (i) Red NPA. NPA = normalized pixel area.

Regarding the proposed method's sensitivity, 4 healthy patients were misclassified as having CLD. A retrospective analysis of their BSMs revealed relatively high values within normal range, though. In one case (Fig. 6c), the BSMs exceeded normal value range in the follow-up examination, and a retrospective LB revealed early-stage NASH, with the fibrosis stage being characterized as mild (Metavir stage between F0 and F1). A future study on an augmented sample of healthy patients with elevated liver-related BSMs could subdivide these patients as low and high risk, respectively, bases on their SWE images and application of the machine-learning algorithm.

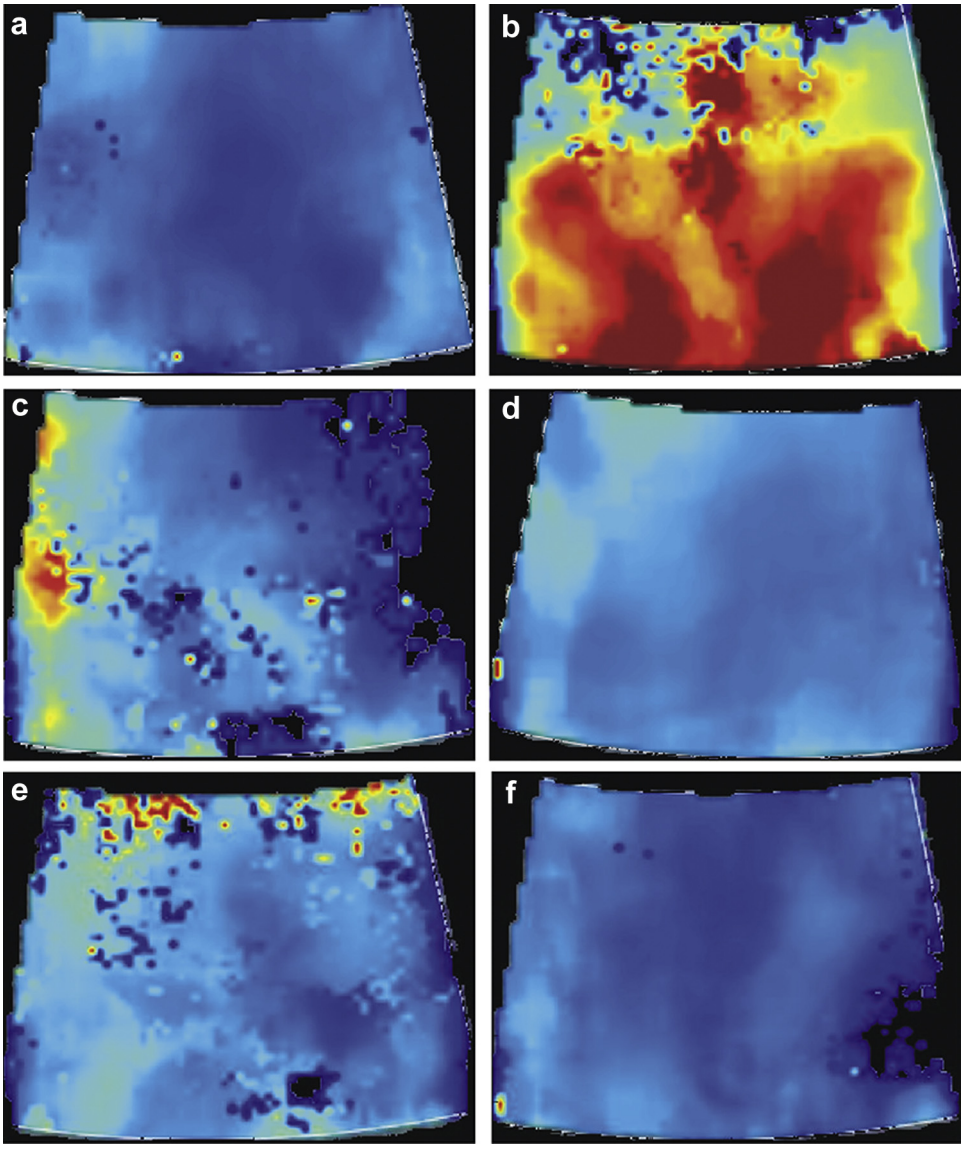

Fig. 6. Various representative cases from the clinical data set and their classification with the proposed system. (a) F0 true negative (healthy classified as healthy). (b) F4 true positive (CLD classified as CLD). (c) F0 false positive (healthy classified as CLD). (d) F3 false negative (CLD classified as healthy). (e) F3 false negative (CLD classified as healthy). (f) F2 false negative (CLD classified as healthy). CLD = chronic liver disease.

With respect to the proposed method's specificity, 12 CLD patients were misclassified as healthy; 9 were staged by LB as F1, 1 as F2 and 2 as F3. F1 is relatively mild fibrosis, where small fibrous tissue areas start to develop on liver tissue. Small tissue differences in terms of elasticity occur between F0 and F1 corresponding to small percentages of diseased liver tissue that could be difficult to detect by SWE. Regarding the misclassified CLD patients with significant or greater fibrosis ($F \geq F2$), a patient with F2 (Fig. 6e) and a patient with F3 (Fig. 6d) had an aqua cluster NPA covering 70%–90% of the elastogram and small areas of green or yellow clusters. The F3 misclassified case in Figure 6f has significantly larger blue and aqua cluster NPAs and no green, yellow or red clusters. These have been proven very rare cases, as SWE images with $F \geq F2$ usually exhibit significant green and yellow areas with much fewer aqua and blue cluster areas. The inferior results of the clinical evaluations by both radiologists, compared with the proposed method, clearly indicate that the standard clinical protocol (ROI mean value) can be considered inadequate in accurate CLD classification.

Diagnosis of CLD, especially at an early stage, remains a major challenge for any quantification approach regardless of the high classification accuracy results obtained with the proposed algorithm. The proposed

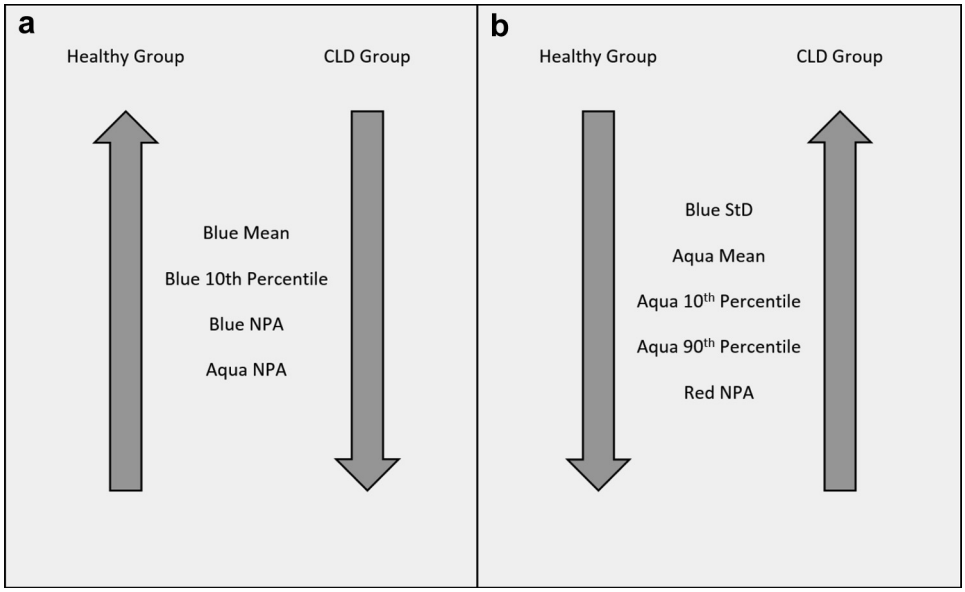

Fig. 7. Discriminative behavior of the selected feature subset. (a) Features (blue mean, blue 10th percentile, blue and aqua NPA) with high values indicative of the healthy group, and low values suggestive of the CLD group. (b) Features (blue standard deviation, aqua mean, 10th and 90th percentiles and red NPA) with low values indicative of the healthy group and high values suggestive of the CLD group. CLD = chronic liver disease; NPA = normalized pixel area.

method does not consider color-bar settings as optimum for disease stage discrimination. Color remapping based on regression analysis for optimized threshold selection to assign fibrosis stages at certain color ranges of elasticity values could be a task for further research. Also, differentiation of all fibrosis stages from F0 to F4 would be a very important future task of this study. Because of the limited number of CLD cases, only binary classification between F0 healthy cases and F1–F4 CLD patients could be performed. Apparently, an augmented data set with more cases (especially for cases with $F1 \geq F \leq F3$) is required for such an analysis.

Limitations of the specific protocol include the relatively small number of individuals investigated and the fact that obese patients and those with ascites, which represent a substantial number of patients examined in everyday clinical practice, were excluded from the study.

CONCLUSIONS

A new machine-learning algorithm that quantifies stiffness value information from SWE images and employs a sophisticated classification algorithm for CLD identification was introduced. The proposed system outperformed all clinical and automated studies, as well as expert radiologists, in the differentiation of healthy persons from CLD patients. The definition of objective parameters and criteria for CLD diagnosis employing SWE images is a significant step toward interpretation of important color-based parameters for assisting clinicians in their daily clinical practice.

REFERENCES

Altekruse SF, McGlynn KA, Reichman ME. Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005. J Clin Oncol 2009;27:1485–1491.

Bota S, Herkner H, Sporea I, Salzl P, Sirli R, Neghina AM, Peck-Radosavljevic M. Meta-analysis: ARFI elastography versus transient elastography for the evaluation of liver fibrosis. Liver Int 2013;33:1138–1147.

Bota S, Paternostro R, Etschmaier A, Schwarzer R, Salzl P, Mandorfer M, Kienbacher C, Ferlitsch M, Reiberger T, Trauner M, Peck-Radosavljevic M, Ferlitsch A. Performance of 2-D shear wave elastography in liver fibrosis assessment compared with serologic tests and transient elastography in clinical routine. Ultrasound Med Biol 2015;41:2340–2349.

Carey E, Carey WD. Noninvasive tests for liver disease, fibrosis, and cirrhosis: Is liver biopsy obsolete? Cleve Clin J Med 2010;77:519–527.

Cherkassky V, Ma Y. Practical selection of SVM parameters and noise estimation for SVM regression. Neural Netw 2004;17:113–126.

Chung JH, Ahn HS, Kim SG, Lee YN, Kim YS, Jeong SW, Jang JY, Lee SH, Kim HS, Kim BS. The usefulness of transient elastography, acoustic-radiation-force impulse elastography, and real-time elastography for the evaluation of liver fibrosis. Clin Mol Hepatol 2013;19:156–164.

Colombo S, Buonocore M, Del Poggio A, Jamoletti C, Elia S, Mattiello M, Zabbialini D, Del Poggio P. Head-to-head comparison of transient elastography (TE), real-time tissue elastography (RTE), and acoustic radiation force impulse (ARFI) imaging in the diagnosis of liver fibrosis. J Gastroenterol 2012;47:461–469.

Cosgrove D, Piscaglia F, Bamber J, Bojunga J, Correas JM, Gilja OH, Klauser AS, Sporea I, Calliada F, Cantisani V, D'Onofrio M, Drakonaki EE, Fink M, Friedrich-Rust M, Fromageau J, Havre RF, Jenssen C, Ohlinger R, Saftoiu A, Schaefer F, Dietrich CF, EFSUMB. EFSUMB guidelines and recommendations on the clinical use of ultrasound elastography: Part 2. Clinical applications. Ultraschall Med 2013;34:238–253.

Deffieux T, Gennisson JL, Bousquet L, Corouge M, Cosconea S, Amroun D, Tripon S, Terris B, Mallet V, Sogni P, Tanter M, Pol S. Investigating liver stiffness and viscosity for fibrosis, steatosis and

activity staging using shear wave elastography. J Hepatol 2015;62: 317–324.
Draper NR, Smith H. Applied regression analysis. Hoboken, NJ: Wiley–Interscience; 1998.
Ferraioli G, Tinelli C, Dal Bello B, Zicchetti M, Filice G, Filice C, Liver Fibrosis Study Group. Accuracy of real-time shear wave elastography for assessing liver fibrosis in chronic hepatitis C: A pilot study. Hepatology 2012a;56:2125–2133.
Ferraioli G, Tinelli C, Malfitano A, Dal Bello B, Filice G, Filice C; Liver Fibrosis Study Group, Above E, Barbarini G, Brunetti E, Calderon W, Di Gregorio M, Lissandrin R, Ludovisi S, Maiocchi L, Michelone G, Mondelli M, Patruno SF, Perretti A, Poma G, Sacchi P, Zaramella M, Zicchetti M. Performance of real-time strain elastography, transient elastography, and aspartate-to-platelet ratio index in the assessment of fibrosis in chronic hepatitis C. AJR Am J Roentgenol 2012b;199:19–25.
Ferraioli G, Filice C, Castera L, Choi BI, Sporea I, Wilson SR, Cosgrove D, Dietrich CF, Amy D, Bamber JC, Barr R, Chou YH, Ding H, Farrokh A, Friedrich-Rust M, Hall TJ, Nakashima K, Nightingale KR, Palmeri ML, Schafer F, Shiina T, Suzuki S, Kudo M. WFUMB guidelines and recommendations for clinical use of ultrasound elastography: Part 3. Liver. Ultrasound Med Biol 2015;41:1161–1179.
Frulio N, Trillaud H. Ultrasound elastography in liver. Diagn Interv Imaging 2013;94:515–534.
Fujimoto K, Kato M, Kudo M, Yada N, Shiina T, Ueshima K, Yamada Y, Ishida T, Azuma M, Yamasaki M, Yamamoto K, Hayashi N, Takehara T. Novel image analysis method using ultrasound elastography for noninvasive evaluation of hepatic fibrosis in patients with chronic hepatitis C. Oncology 2013;84(Suppl 1):3–12.
Gatos I, Tsantis S, Spiliopoulos S, Karnabatidis D, Theotokas I, Zoumpoulis P, Loupas T, Hazle JD, Kagadis GC. A new computer aided diagnosis system for evaluation of chronic liver disease with ultrasound shear wave elastography imaging. Med Phys 2016a; 43:1428–1436.
Gatos I, Tsantis S, Zoumpoulis PS, Theotokas I, Spiliopoulos S, Karnabatidis D, Kagadis GC, Chronic liver disease evaluation with an automatic shear wave elastography optimization technique. In: Proceedings, Radiological Society of North America annual meeting 2016. Chicago, IL: RSNA; 2016b.
GBD 2013 Mortality and Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015;385:117–171.
Gerber L, Kasper D, Fitting D, Knop V, Vermehren A, Sprinzl K, Hansmann ML, Herrmann E, Bojunga J, Albert J, Sarrazin C, Zeuzem S, Friedrich-Rust M. Assessment of liver fibrosis with 2-D shear wave elastography in comparison to transient elastography and acoustic radiation force impulse imaging in patients with chronic liver disease. Ultrasound Med Biol 2015;41:2350–2359.
Gilmore IT, Burroughs A, Murray-Lyon IM, Williams R, Jenkins D, Hopkins A. Indications, methods, and outcomes of percutaneous liver biopsy in England and Wales: An audit by the British Society of Gastroenterology and the Royal College of Physicians of London. Gut 1995;36:437–441.
Giorgio A, Amoroso P, Lettieri G, Fico P, de Stefano G, Finelli L, Scala V, Tarantino L, Pierri P, Pesce G. Cirrhosis: Value of caudate to right lobe ratio in diagnosis with US. Radiology 1986;161:443–445.
Goodman ZD. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. J Hepatol 2007;47:598–607.
Harbin WP, Robert NJ, Ferrucci JT Jr. Diagnosis of cirrhosis based on regional changes in hepatic morphology: A radiological and pathological analysis. Radiology 1980;135:273–283.
Hong H, Li J, Jin Y, Li Q, Li W, Wu J, Huang Z. Performance of real-time elastography for the staging of hepatic fibrosis: A meta-analysis. PLoS One 2014;9:e115702.
Huang K, Velliste M, Murphy RF. Feature reduction for improved recognition of subcellular location patterns in fluorescence microscope images. Proc SPIE 2003;307–318.
Huet N, Denis I, Martino A, Gallix B, Sturm N, Leroy V, Bricault I. Ultrasonographic assessment of liver fibrosis with computer-assisted analysis of liver surface irregularities. Diagn Interv Imaging 2015; 96:941–946.
Huwart L, Sempoux C, Vicaut E, Salameh N, Annet L, Danse E, Peeters F, ter Beek LC, Rahier J, Sinkus R, Horsmans Y, Van Beers BE. Magnetic resonance elastography for the noninvasive staging of liver fibrosis. Gastroenterology 2008;135:32–40.
Joanes DN, Gill CA. Comparing measures of sample skewness and kurtosis. J R Stat Soc Ser D 1998;47:183–189.
Kobayashi K, Nakao H, Nishiyama T, Lin Y, Kikuchi S, Kobayashi Y, Yamamoto T, Ishii N, Ohashi T, Satoh K, Nakade Y, Ito K, Yoneda M. Diagnostic accuracy of real-time tissue elastography for the staging of liver fibrosis: A meta-analysis. Eur Radiol 2015; 25:230–238.
Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. J Biomed Inform 2005;38:404–415.
Lo CM, Lai YC, Chou YH, Chang RF. Quantitative breast lesion classification based on multichannel distributions in shear-wave imaging. Comput Methods Programs Biomed 2015;122:354–361.
Martinez SM, Crespo G, Navasa M, Forns X. Noninvasive assessment of liver fibrosis. Hepatology 2011;53:325–335.
Murphy SL, Xu J, Kochanek D. 2013 deaths: Final data for 2010. Natl Vital Stat Rep 2013;61:1–117.
Nierhoff J, Chavez Ortiz AA, Herrmann E, Zeuzem S, Friedrich-Rust M. The efficiency of acoustic radiation force impulse imaging for the staging of liver fibrosis: A meta-analysis. Eur Radiol 2013;23: 3040–3053.
Pellot-Barakat C, Lefort M, Chami L, Labit M, Frouin F, Lucidarme O. Automatic assessment of shear wave elastography quality and measurement reliability in the liver. Ultrasound Med Biol 2015;41: 936–943.
Romero-Gomez M, Gomez-Gonzalez E, Madrazo A, Vera-Valencia M, Rodrigo L, Perez-Alvarez R, Perez-Lopez R, Castellano-Megias VM, Nevado-Santos M, Alcon JC, Sola R, Perez-Moreno JM, Navarro JM, Andrade RJ, Salmeron J, Fernandez-Lopez M, Aznar R, Diago M. Optical analysis of computed tomography images of the liver predicts fibrosis stage and distribution in chronic hepatitis C. Hepatology 2008;47:810–816.
Sanford NL, Walsh P, Matis C, Baddeley H, Powell LW. Is ultrasonography useful in the assessment of diffuse parenchymal liver disease? Gastroenterology 1985;89:186–191.
Sporea I, Bota S, Gradinaru-Tascau O, Sirli R, Popescu A, Jurchis A. Which are the cut-off values of 2D-shear wave elastography (2D-SWE) liver stiffness measurements predicting different stages of liver fibrosis, considering transient elastography (TE) as the reference method? Eur J Radiol 2014;83:e118–e122.
Tatsumi C, Kudo M, Ueshima K, Kitai S, Ishikawa E, Yada N, Hagiwara S, Inoue T, Minami Y, Chung H, Maekawa K, Fujimoto K, Kato M, Tonomura A, Mitake T, Shiina T. Non-invasive evaluation of hepatic fibrosis for type C chronic hepatitis. Intervirology 2010;53:76–81.
Tsochatzis EA, Gurusamy KS, Ntaoula S, Cholongitas E, Davidson BR, Burroughs AK. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: A meta-analysis of diagnostic accuracy. J Hepatol 2011;54:650–659.
Wang J, Guo L, Shi X, Pan W, Bai Y, Ai H. Real-time elastography with a novel quantitative technology for assessment of liver fibrosis in chronic hepatitis B. Eur J Radiol 2012;81:e31–e36.
Yada N, Kudo M, Morikawa H, Fujimoto K, Kato M, Kawada N. Assessment of liver fibrosis with real-time tissue elastography in chronic viral hepatitis. Oncology 2013;84(Suppl 1):13–20.