Magnetic Resonance Elastography for the Noninvasive Staging of Liver Fibrosis

LAURENT HUWART,* CHRISTINE SEMPOUX,‡ ERIC VICAUT,§ NAJAT SALAMEH,* LAURENCE ANNET,* ETIENNE DANSE,* FRANK PEETERS,* LEON C. TER BEEK,‖ JACQUES RAHIER,‡ RALPH SINKUS,¶ YVES HORSMANS,# and BERNARD E. VAN BEERS

**Diagnostic Radiology Unit, ‡Department of Pathology, and #Department of Gastroenterology, Université Catholique de Louvain, St-Luc University Hospital, Brussels, Belgium; §Department of Nuclear Medicine, Hôpital Lariboisière, and ¶Ecole Supérieure de Physique et de Chimie Industrielles, Université Denis Diderot, Paris, France; and ‖Philips Medical Systems, Best, The Netherlands*

See Jacqueminet S et al on page 828 in *CGH*; See editorial on page 299.

Background & Aims: **The purpose of our study was to prospectively compare the success rate and diagnostic accuracy of magnetic resonance elastography, ultrasound elastography, and aspartate aminotransferase to platelets ratio index (APRI) measurements for the noninvasive staging of fibrosis in patients with chronic liver disease.** ***Methods:*** **We performed a prospective blind comparison of magnetic resonance elastography, ultrasound elastography, and APRI in a consecutive series of patients who underwent liver biopsy for chronic liver disease in a university-based hospital. Histopathologic staging of liver fibrosis according to the METAVIR scoring system served as the reference.** ***Results:*** **A total of 141 patients were assessed. The technical success rate of magnetic resonance elastography was higher than that of ultrasound elastography (133/141 [94%] vs 118/141 [84%]; $P = .016$). Magnetic and ultrasound elastography, APRI measurements, and histopathologic analysis of liver biopsy specimens were technically successful in 96 patients. The areas under the receiver operating characteristic curves of magnetic resonance elasticity (0.994 for $F \geq 2$; 0.985 for $F \geq 3$; 0.998 for $F = 4$) were larger ($P < .05$) than those of ultrasound elasticity, APRI, and the combination of ultrasound elasticity and APRI (0.837, 0.709, and 0.849 for $F \geq 2$; 0.906, 0.816, and 0.936 for $F \geq 3$; 0.930, 0.820, and 0.944 for $F = 4$, respectively).** ***Conclusions:*** **Magnetic resonance elastography has a higher technical success rate than ultrasound elastography and a better diagnostic accuracy than ultrasound elastography and APRI for staging liver fibrosis.**

Liver biopsy is the current reference examination for the assessment of liver fibrosis. However, it is a costly procedure that carries a small risk of severe complications and is difficult to accept for patients. In addition, its accuracy remains debated because of sampling variability caused by the small size of the hepatic samples and the heterogeneity of liver fibrosis.[1] The noninvasive assessment of liver fibrosis has become a real challenge given that chronic liver diseases affect hundreds of millions of patients worldwide. Multiple data now emphasize that fibrosis is dynamic and, with effective intervention, reversible.[2] Successful treatment of viral hepatitis, autoimmune liver disease, alcohol-related disease, and other chronic liver diseases results not only in clinical improvement but also in decreased histologic fibrosis. Although experimental studies have revealed targets to prevent fibrosis progression in rodents, the efficacy of most treatments has not been proven in humans. The development of reliable noninvasive markers of liver fibrosis is essential to assess the prognosis of the disease and the response to treatment.[2,3]

Several noninvasive methods have been proposed to stage liver fibrosis, including biochemical tests and imaging methods. The biochemical tests are composite scores (aspartate aminotransferase to platelets ratio index [APRI], FibroTest [BioPredictive, Paris, France], and so on) or serum markers of fibrosis such as hyaluronic acid. However, the value of these diagnostic methods remains debated.[2,3]

Abbreviations used in this paper: APRI, aspartate aminotransferase to platelets ratio index; CI, confidence interval; MR, magnetic resonance; ROC, receiver operating characteristics.

0016-5085/08/$34.00
doi:10.1053/j.gastro.2008.03.076

Among the imaging methods, elastography has been shown to be a reliable method to stage liver fibrosis. It is based on the observation that fibrosis leads to increased tissue stiffness. Most clinical studies have been performed with ultrasound elastography.[4–6] Recently, magnetic resonance (MR) elastography has emerged as an alternative method to assess liver elasticity.[7–9] To the best of our knowledge, no study has been reported about the comparison of MR and ultrasound elastography for the assessment of liver fibrosis.

Therefore, the purpose of our study was to prospectively compare the technical success rate and diagnostic accuracy of MR elastography, ultrasound elastography, and APRI measurements for staging hepatic fibrosis in patients who underwent liver biopsy for chronic liver disease.

Patients and Methods

Patients

This study was a single-center, prospective, blind comparison of MR elastography, ultrasound elastography, and APRI in a consecutive series of patients who underwent liver biopsy in the Department of Gastroenterology of St-Luc University Hospital, Université Catholique de Louvain, for suspicion of chronic diffuse liver disease between November 2005 and February 2007. The study protocol was in accordance with the Declaration of Helsinki and was approved by the ethics committee of our institution. Patients were enrolled after giving written informed consent. MR elastography, ultrasound elastography, and APRI measurements were performed within 2 days of liver biopsy. The order of the examinations was not randomized because it was considered that this order should not have influenced the results in patients with chronic liver diseases. For organizational reasons (access to the MR scanner), MR elastography was performed immediately before ultrasound elastography. A blood test, including measurements of aspartate aminotransferase levels and platelet counts, was systematically performed the day before the liver biopsy.

To assess the reproducibility of MR and ultrasound elastography, these examinations were repeated within 1 month in all patients in whom MR elastography, ultrasound elastography, APRI measurements, and analysis of liver biopsy specimens were technically successful and who agreed to have repeated examinations.

MR Elastography

The 3-dimensional MR elastography method has been described in detail previously.[7,10,11] Briefly, low-frequency longitudinal mechanical waves of 65 Hz were transmitted into the right liver by a transducer placed against the last ribs at the back of the patient in supine position. The reconstruction of the local shear elasticity requires the presence of shear waves inside the organ. Shear waves were obtained by utilizing the fact that longitudinal waves generate shear waves due to mode conversion at interfaces everywhere inside the liver. The reason to excite in the first place with longitudinal waves is that these get less attenuated by tissue, leading to an efficient way for shear wave generation deep inside the body. Shear waves afterward become separated from the longitudinal contribution by applying the curl operator on the total displacement vector field.

Images were obtained on a 1.5-T whole-body MR scanner (Gyroscan Intera; Philips Medical Systems, Best, The Netherlands) using a 4-elements torso coil. Five sagittal slices through the right liver were acquired with a slice thickness of 4 mm, field of view of 250 mm, matrix size of 64^2, echo time of 61 milliseconds, repetition time of 431 milliseconds, and 2 signal averages. The patients breathed freely, and respiratory gating was performed with a navigator on the right hemidiaphragm. Four dynamics were obtained by changing the phase offset between the mechanical excitation and the MR sequence to assess the amplitude and phase of the displacement after Fourier transformation. The motion-encoding gradients were applied successively in the 3 orthogonal directions to capture all the components of the 3-dimensional displacement vector. The total acquisition time was about 20 minutes, depending on the efficiency of the respiratory gating navigator.

The phase images were analyzed with the Voigt model to obtain shear elasticity maps, that is, the real part of the complex shear modulus is attributed to the solid component of the material while the imaginary part accounts for losses. The shear elasticity (kPa) of the liver was measured as the mean value within the largest rectangular region of interest that fitted into the liver on the elasticity map of the central slice. The first series of measurements (ie, within 2 days of liver biopsy) was performed by a junior radiologist with 4 years of experience in MR imaging. The second series of measurements (within the month) was performed by a junior physicist with 2 years of experience in MR imaging. These 2 observers were blinded to the clinical, biochemical, and ultrasound data of the patient and to the results of histopathologic analysis.

Ultrasound Elastography

One-dimensional transient ultrasound elastography measurements were performed with a FibroScan (EchoSens, Paris, France). The technique and examination procedure have been described previously.[5] This method measures the velocity of the shear wave, which is directly related to Young's elastic modulus (kPa). It should be noted that, within tissues, the Young's modulus equals 3 times the shear elasticity modulus measured with the 3-dimensional MR elastographic method.

Measurements were performed in the right lobe of the liver with the tip of the transducer probe placed in the intercostal space. The measurement depth was between 25 and 65 mm. Ten validated measurements were performed in each patient. The median value was considered representative of the Young's elastic modulus of the liver. The interquartile range of the 10 measurements was also recorded for each patient, and the ratio interquartile range/median value of liver stiffness was calculated. Ratios <0.2 indicate reliable measurements.[12] The whole examination lasted less than 5 minutes. The first and second series of measurements were performed by 2 different senior radiologists who had performed more than 100 examinations with the FibroScan. These 2 observers were blinded to the clinical, biochemical, and MR data of the patient and to the results of histopathologic analysis.

APRI

Aspartate aminotransferase levels and platelet counts were measured with a Synchron Clinical System LX20 autoanalyzer (Beckman Coulter, Fullerton, CA) and an Advia 120 Hematology System autoanalyzer (Bayer, Leverkusen, Germany), respectively. The APRI index was calculated as follows: Aspartate Aminotransferase (/Upper Limit of Normal) $\times$ 100/Platelet Count (10^9/L).

Histologic Analysis

Percutaneous liver biopsy was performed by senior operators using the Menghini technique with a 1.4-mm-diameter needle (Hepafix; Braun, Melsungen, Germany). In 20 patients with ascites and/or trouble of blood crasis, liver biopsy was performed through a transjugular approach using a Ross-modified Colapinto catheter needle with a diameter of 1.5 mm (Cook, Bjaeverskov, Denmark).

After biopsy, the liver samples were fixed in formalin, paraffin embedded, and stained with H&E and Masson's trichrome. All biopsy specimens were analyzed by 2 senior hepatopathologists blinded to the biological and clinical data and to the results of MR and ultrasound elastography. Liver biopsy specimens were considered suitable for fibrosis staging when they contained at least 10 portal tracts or obvious regenerating nodules.

The stage of fibrosis was evaluated semiquantitatively on Masson's trichrome–stained slides according to the METAVIR scoring system.[13] The METAVIR scoring system was initially described for chronic hepatitis C and was later applied to other chronic liver diseases.[14] With this score, F0 represents no fibrosis; F1, portal fibrosis without septa; F2, portal fibrosis and few septa; F3, numerous septa without cirrhosis; and F4, cirrhosis. The fibrosis stage was assessed independently by each pathologist. In case of discrepancies, a consensus was obtained. However, in patients with nonalcoholic steatohepatitis and alcoholic liver disease, the use of the Brunt classification has been recommended because of the presence of a specific perisinusoidal fibrosis in these diseases.[15] This scoring system differs from the METAVIR scoring system mainly for the staging of mild fibrosis. It is defined as follows: stage 0, no fibrosis; stage 1, perisinusoidal or periportal fibrosis; stage 2, perisinusoidal and periportal fibrosis; stage 3, bridging fibrosis; and stage 4, cirrhosis. The consequence is that in the case of nonalcoholic steatohepatitis and alcoholic liver disease, a METAVIR F0 score can already be a Brunt stage 1 and a F1 score can be a Brunt stage 2. To evaluate the possible influence of using a common staging system (METAVIR) for all patients in our study, a secondary and unblinded analysis for the presence of perisinusoidal fibrosis was performed in patients with nonalcoholic steatohepatitis or alcoholic liver disease who were staged F0 or F1 with the METAVIR scoring system.

The grade of necroinflammatory activity was evaluated semiquantitatively on H&E-stained slides according to the Ludwig scoring system with its 2 main criteria: portal and lobular inflammation.[16] The portal inflammation was graded from 0 to 4, with P0 representing none or minimal inflammation; P1, portal inflammation; P2, mild limiting plate necrosis; P3, moderate limiting plate necrosis; and P4, severe limiting plate necrosis. The lobular inflammation was graded from 0 to 4, with L0 representing no inflammation; L1, inflammation but no necrosis; L2, focal necrosis or acidophilic bodies; L3, severe focal cell damage; and L4, damage including bridging necrosis.

Steatosis was graded according to the Brunt scoring system as follows: 0, none; 1, steatosis in 1%–33% of hepatocytes; 2, steatosis in 33%–66% of hepatocytes; 3, steatosis in 66%–100% of hepatocytes.[15] The length of each liver biopsy was established in millimeters, and the number of portal tracts was counted.

Statistical Analysis

The sample size was fixed to allow at least 80% power to detect by the McNemar's test at a 2-sided 5% significance level, a difference in sensitivity between MR elastography and ultrasound elastography of 0.10 considering a sensitivity of ultrasound elastography for $F \geq 2$ close to 0.67 (similar to that found in the study by Castera et al[4]), and considering a prevalence of patients with $F \geq 2$ around 60%. Under these conditions, a sample size of 95 patients evaluated by the 2 tests was required. The difference in technical success between MR elastography and ultrasound elastography was assessed with the χ^2 test.

Further analysis was performed in the patients in whom MR elastography, ultrasound elastography, APRI measurements, and histologic analysis of liver biopsy specimens were technically successful. The agreement between the pathologists for the scoring of liver fibrosis was assessed with weighted κ coefficients. The reproducibility of MR and ultrasound elastography was evaluated with intraclass correlation coefficients and coefficients of repeatability. The coefficients of repeatability were calculated as 1.96

times the standard deviations of the differences between the 2 measurements made by the 2 observers.

Diagnostic performance of MR and ultrasound elastography was assessed on the first series of measurements obtained within 2 days of liver biopsy. The relationships between fibrosis stage and MR elasticity, ultrasound elasticity, and APRI measurements were tested with the nonparametric Kendal's coefficients of correlation. Magnetic resonance elastography, ultrasound elastography, and APRI were evaluated for their predictive performance using areas under the receiver operating characteristic (ROC) curves. Areas under ROC curves of MR elastography, ultrasound elastography, and APRI were compared using the method proposed by DeLong et al.[17] To assess our results in a population of patients with moderate fibrosis, we performed a complementary sensitivity analysis comparing the areas under the ROC curves in the subpopulation of patients with $F \leq 2$.

In addition, MR elastography, ultrasound elastography, and APRI were introduced in a stepwise multivariate model to identify the best predictor(s) of fibrosis and the possible gain in prediction when several methods were used. A well-known problem of predictive models is the frequent overestimation of their performance because the models are evaluated on the samples used for their construction. Such a phenomenon called "optimism" should be assessed for appropriate choice and validation of the multivariate model.

To assess the quality of the model (and thus the robustness of our conclusions), we used 10-fold cross-validation and bootstrapping procedure ($N = 1000$ bootstrap samples).[18] Briefly, for K-fold cross-validation, the original sample was partitioned into K subsamples. Of the K subsamples, a single subsample was retained as the validation data for testing the model, and the remaining $K - 1$ subsamples were used as training data. The cross-validation process was then repeated several times, with each of the K subsamples used as the validation data.

Bootstrapping (ie, based on repeatedly analyzing subsamples of the data) allows calculation of several parameters estimating the optimism of the model. As proposed by Steyerberg et al,[18] bootstrapping was used to evaluate the quality of the model by calculation of calibration slopes (well-calibrated models have a slope equal or close to 1) and Brier scores (models scores range from 0 [perfect] to 0.25 [worthless]).[18]

Sensitivity, specificity, and positive and negative predictive values for the classification of F0 versus F1 − F4 ($F \geq 1$), F0 − F1 versus F2 − F4 ($F \geq 2$), F0 − F2 versus F3 − F4 ($F \geq 3$), and F0 − F3 versus F4 ($F = 4$) were computed for MR elastography at the maximum total of sensitivity and specificity. In addition, the misclassification rates were given for each individual fibrosis stage.

Moreover, the possible effects on the MR elasticity measurements of the type of liver biopsy (percutaneous vs transjugular), the length of the hepatic samples (with 3 groups: <15 mm, 15–25 mm, >25 mm), and the cause of liver disease (chronic viral hepatitis B and C vs alcoholic and nonalcoholic steatohepatitis) were studied using analysis of variance. Lastly, possible effects of steatosis or portal and/or lobular inflammation on the MR elasticity measurements were studied using 3-way analysis of variance on ranks.[19] Results are given as mean ± SD. Statistical analysis was performed with SAS 9.13 software (SAS Institute Inc, Cary, NC).

Results

A total of 146 patients underwent liver biopsy for chronic liver disease between November 2005 and February 2007. Five patients refused to participate, and 141 patients entered the study. Ultrasound elastography could be performed in 118 of 141 patients (84%). Ultrasound elastography was unsuccessful in 13 patients with ascites. Ascites was subclinical and detected with imaging in 5 of these patients. Ten other failures were caused by obesity. The body mass index of these patients was 32.8 ± 1.8 kg/m² (29.8–35.1 kg/m²). Magnetic resonance elastography could be performed in 133 of 141 patients (94%) ($P = .016$ vs ultrasound elastography). The 8 failures were caused by claustrophobia in 3 patients, low hepatic signal related to hemochromatosis in 3 patients, and obesity in 2 patients. These 2 patients could not fit into the magnet bore, which was narrowed by the transducer. Measurements of APRI were obtained in 141 of 141 patients (100%). Liver biopsy specimens were suitable for fibrosis staging in 127 of 141 patients (90%).

Further analysis was performed in the 96 patients in whom MR elastography, ultrasound elastography, APRI measurements, and histologic analysis were successful (Figure 1). There were 45 men and 51 women. Their age was 54 ± 13 years (range, 22–83 years), and their mean body mass index was 25.9 ± 4.0 kg/m² (range, 18.7–33.1 kg/m²). Body mass index was >28 kg/m² in 12 of 96 patients (12%). The cause of chronic liver disease was chronic viral hepatitis in 65 patients (chronic hepatitis C in 60 and chronic hepatitis B in 5), alcohol abuse in 14, nonalcoholic steatohepatitis in 8, α_1-antitrypsin deficiency in 1, drug toxicity in 2, and unknown in 6.

The length of the liver biopsy specimens was 30 ± 11 mm (range, 12–66 mm) and was ≥25 mm in 70 of 96 patients (73%). The number of portal tracts was 15 ± 6 (range, 6–33) and was ≥11 in 87 of 96 patients (91%). The distribution of fibrosis stage at the consensus reading was F0 in 22 of 96 patients (23%), F1 in 22 (23%), F2 in 19 (20%), F3 in 15 (15%), and F4 in 18 (19%), showing a homogeneous distribution of fibrosis stages. The 2 pathologists were initially in agreement for 81 of the 96

liver biopsy specimens analyzed (weighted κ coefficient, 0.90; 95% confidence interval [CI], 0.84–0.95) with no significant rating bias. Among the patients with alcoholic or nonalcoholic steatohepatitis, one who was staged F1 according to METAVIR had portal and perisinusoidal fibrosis and would have been classified as stage 2 according to Brunt. The elasticity was 2.49 ± 0.17 kPa in this patient.

Portal inflammation was graded P0 in 17 of 96 patients (17%), P1 in 43 (45%), P2 in 35 (37%), P3 in 1 (1%), and P4 in 0 (0%). Lobular inflammation was graded L0 in 32 of 96 patients (33%), L1 in 34 (35%), L2 in 30 (32%), L3 in 0 (0%), and L4 in 0 (0%). Steatosis was graded 0 in 40 of 96 patients (42%), 1 in 32 (34%), 2 in 18 (19%), and 3 in 6 (5%).

At ultrasound elastography, the ratio interquartile range/median value was 0.16 for F0, 0.17 for F1, 0.16 for F2, 0.19 for F3, and 0.19 for F4. The reproducibility of MR and ultrasound elastography could be assessed in 56 patients. The intraclass correlation coefficients of MR elasticity and ultrasound elasticity were 0.97 (95% CI, 0.92–0.99) and 0.94 (95% CI, 0.51–0.97), respectively. The 95% CIs of the interobserver differences of measurements were −0.016 to 0.088 for MR elasticity and 0.028 to 2.192 for ultrasound elasticity. The coefficients of repeatability were 0.385 for MR elastography and 8.149 for ultrasound elastography.

In the 96 patients in whom elastography, APRI measurements, and histologic analysis were successful, MR elasticity, ultrasound elasticity, and APRI increased according to the stage of liver fibrosis (Figure 2). The MR elasticity was 2.15 ± 0.19 kPa (range, 1.89–2.43 kPa) for F0, 2.41 ± 0.11 kPa (range, 2.19–2.63 kPa) for F1, 2.85 ± 0.23 kPa (range, 2.49–3.18 kPa) for F2, 3.49 ± 0.53 kPa (range, 2.84–4.46 kPa) for F3, and 5.25 ± 0.65 kPa (range, 4.13–6.73 kPa) for F4. The measurements were correlated to the fibrosis stage: $r = 0.84$, $P < .0001$ for MR elasticity; $r = 0.56$, $P < .0001$ for ultrasound elasticity; and $r = 0.36$, $P < .0001$ for APRI.

Table 1 shows the areas under the ROC curves of MR elasticity, ultrasound elasticity, and APRI and the combinations of the noninvasive methods for the different fibrosis thresholds. The areas under the ROC of MR elastography were significantly larger than those of ultrasound elastography, APRI, and the combination of ultrasound elastography and APRI ($P = .003$, $P < .0001$, and $P = .005$ for F ≥ 1; $P = .0001$, $P < .0001$, and $P = .0002$ for F ≥ 2; $P = .01$, $P = .0005$, and $P = .02$ for F ≥ 3; and $P = .01$, $P = .008$, and $P = .04$ for F = 4). Moreover, the areas under the ROC curves of MR elastography were not significantly different from those of the combinations including MR elastography, that is, MR elastography and ultrasound elastography, MR elastography and APRI, and MR elastography, ultrasound elastography, and APRI ($P > .05$). The sensitivity analysis showed that the areas under the ROC curves for MR elastography were 0.93 and 0.98 for F0 versus F1–F2 and F0–F1 versus F2, respectively. The corresponding figures were 0.70 and 0.68 for ultrasound elastography ($P = .002$ and $P < .0001$ vs MR elastography) and 0.53 and 0.52 for APRI ($P < .0001$ and $P < .0001$ vs MR elastography).

All calculated parameters indicated good performances of the MR elasticity measurements for the prediction of fibrosis. The areas under the ROC curves were 0.962–0.998, the calibration slopes were 0.924–0.996, and the Brier scores were 0.015–0.067. The corresponding values for ultrasound elasticity were 0.803–0.930, 0.351–0.643, and 0.097–0.162. For APRI, the values were 0.676–0.821, 0.351–0.643, and 0.099–0.210. When the 3 variables were introduced in a stepwise multivariable logistic analysis, the only variable remaining in the model was MR elastography. We checked the validity of the models relating MR elastography and the different fibrosis thresholds and found that the degrees of optimism were very moderate in all models (ie, <0.1% for the areas under the ROC curves, <0.002 for the calibration slopes, and <0.08 for the Brier scores).

The most discriminating cutoff values of MR elasticity were 2.4 kPa for F ≥ 1, 2.5 kPa for F ≥ 2, 3.1 kPa for F ≥ 3, and 4.1 kPa for F = 4. The corresponding sensitivities, specificities, and positive and negative predictive values are detailed in Table 2. When individual stages of fibrosis were considered, the number of patients who were misclassified with MR elastography was 24 of 96 (25%), including 2 patients with histologic stage F0, 14 patients with F1, 2 patients with F2, and 6 patients with F3 (Table 3). In these misclassified patients, there was only one stage difference

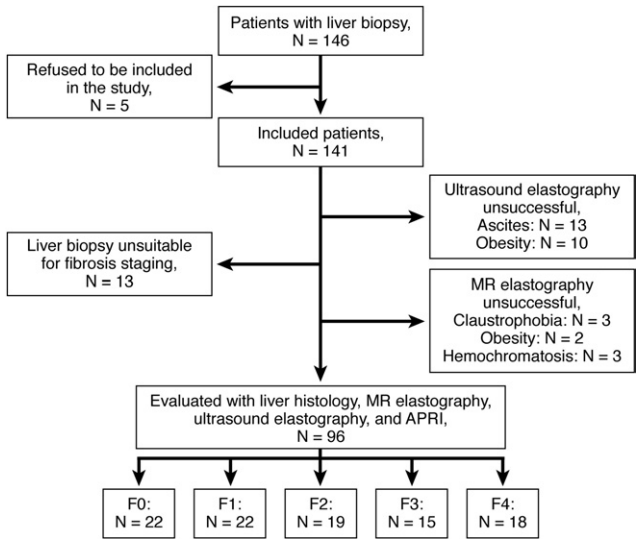

Figure 1. Flow diagram of patients who underwent liver biopsy for chronic liver disease in the Department of Gastroenterology during the 16-month study period.

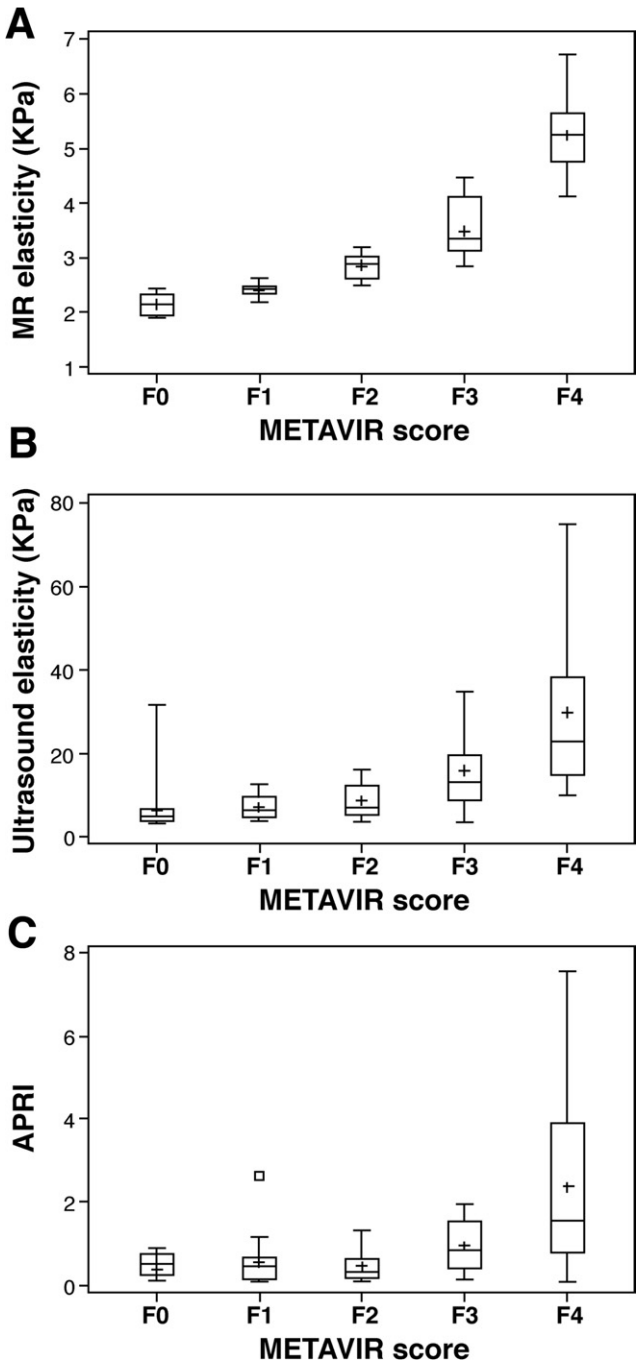

Figure 2. Box plots of (*A*) MR elasticity, (*B*) ultrasound elasticity, and (*C*) APRI for each METAVIR fibrosis stage. The boundary of boxes closest to 0 indicates the 25th percentile, the line within boxes shows the median, and the boundary of boxes furthest from 0 indicates the 75th percentile. The crosses within boxes indicate the mean. Error bars indicate the smallest and the largest values that are within 1.5 box lengths of the 25th and 75th percentiles. Outliers are represented as individual points. In *C*, one outlier has not been represented in the F4 group to maintain the clarity of the graph.

between MR elastography and histology. Two thirds of these misclassifications (13/24; 54%) occurred in patients with no or minimal fibrosis (F0–F1).

No significant effect of type of biopsy, length of hepatic sample, and cause of liver disease was found on the results of MR elastography. When exploring the possible effect of fibrosis, inflammation, and steatosis, fibrosis was the only factor significantly affecting MR elasticity ($P < .001$). In addition, no significant interaction was found.

Discussion

The results of our study show that among 3 noninvasive methods (MR elastography, ultrasound elastography, and APRI), MR elastography offered the best diagnostic performance. The technical success rate of MR elastography was significantly higher than that of ultrasound elastography. The areas under the ROC curves of MR elastography were significantly larger than those of ultrasound elastography, APRI, and the combination of ultrasound elastography and APRI. Lastly, the coefficient of repeatability for MR elastography was better than that of ultrasound elastography. These results suggest that MR elastography is reliable for the evaluation and follow-up of liver fibrosis. This strategy might avoid liver biopsy in most patients with chronic liver disease.

Looking at the ROC curves, the results of MR elastography appear to be very high. This is explained by the fact that the ROC analysis performs a binary categorization of the patients (ie, F ≥ given value) and does not consider the misclassifications between the individual stages of the METAVIR score. Indeed, when individual fibrosis stages were taken into account, MR elastography misclassified 25% of the patients relative to the histopathologic staging. As expected, most of these misclassifications occurred in patients with no or minimal fibrosis. Moreover, the areas under the ROC curves for ultrasound elastography and APRI were similar in our study and in previously reported meta-analyses. For ultrasound elastography, these values were 0.84 for F ≥ 2 and 0.93 for F = 4 in our study versus 0.87 and 0.96 in the meta-analysis of Talwalkar et al.[20] For APRI, the corresponding figures were 0.71 and 0.82 in our study versus 0.76 and 0.82 in the meta-analysis of Shaheen and Myers.[21]

In our study, the better results obtained with MR elastography relative to ultrasound elastography might be explained by the following reasons. First, with MR elastography, the whole 3-dimensional displacement vector is assessed.[7,11] This improves the reconstruction of the shear elastic parameter of the liver relative to ultrasound elastography, which is a one-dimensional method.[5]

Second, MR elastography analyzes a volume that includes several liver sections in contrast to ultrasound elastography, which probes a hepatic sample approximated by a cylinder of 20-mm or 40-mm length.[4–6,14] The volume analyzed with MR elastography is thus far more representative of the hepatic parenchyma, which is essential to avoid the sampling variability caused by the het-

Table 1. Areas Under ROC Curves With 95% CIs for MR Elastography, Ultrasound Elastography, APRI, and Combinations of Methods According to METAVIR Stages

	F ≥ 1	F ≥ 2	F ≥ 3	F = 4
MR elastography	0.962 (0.929–0.995)	0.994 (0.985–1.0)	0.985 (0.968–1.0)	0.998 (0.993–1.0)
Ultrasound elastography	0.803 (0.701–0.904)	0.837 (0.756–0.918)	0.906 (0.838–0.975)	0.930 (0.877–0.982)
APRI	0.676 (0.565–0.787)	0.709 (0.603–0.814)	0.816 (0.717–0.915)	0.820 (0.688–0.952)
MR elastography and ultrasound elastography	0.961 (0.928–0.995)	0.993 (0.984–1.0)	0.985 (0.966–1.0)	0.997 (0.991–1.0)
MR elastography and APRI	0.959 (0.923–0.995)	0.992 (0.982–1.0)	0.989 (0.972–1.0)	0.998 (0.995–1.0)
Ultrasound elastography and APRI	0.814 (0.714–0.915)	0.849 (0.772–0.926)	0.936 (0.884–0.988)	0.944 (0.891–0.997)
MR elastography, ultrasound elastography, and APRI	0.963 (0.930–0.996)	0.992 (0.981–1.0)	0.988 (0.970–1.0)	0.999 (0.995–1.0)

NOTE. Values are expressed as area under the ROC curve (95% CI).

erogeneity of advanced fibrosis. This may be a reason for the better reproducibility observed with MR elastography than with ultrasound elastography. This better reproducibility of MR elastography was observed despite the fact that the 10 measurements performed in each patient with ultrasound elastography showed satisfactory reproducibility as shown by interquartile range/median value of stiffness measurements <0.2.

Third, liver elasticity can be assessed with MR elastography in patients who are obese or have ascites because the compressional waves have good penetration throughout the liver. The only practical problem for obese patients is that the patients should fit into the magnet bore, which diameter is reduced by the presence of the transducer. In contrast, the penetration of the shear waves used with ultrasound elastography is poor in obese patients.[22] Foucher et al showed that a body mass index >28 kg/m^2 was a cause of failure of ultrasound elastography.[22] Moreover, as explained by Sandrin et al,[5] elasticity measurements with ultrasound elastography using the FibroScan approach cannot be performed in patients with ascites, because shear waves generated at the surface of the patient do not propagate in liquids. This limitation applies even if ascites is mild and clinically undetectable. This problem limits the capability of studying the severity of cirrhosis with FibroScan.[14]

Moreover, MR elastography can be integrated into a comprehensive hepatic MR imaging examination including morphologic and perfusion imaging for the detection of hepatocellular carcinomas and steatosis and the assessment of liver function. In contrast, transient ultrasound elastography of the liver is performed with a dedicated machine that only permits elasticity measurements. The feasibility of elasticity measurements in combination with conventional real-time sonography has been reported recently but should be further validated.[23]

MR elastography has some disadvantages. It cannot be performed in livers with high iron overload because of signal-to-noise limitations. The examination time of MR elastography is longer than that of ultrasound elastography. The future replacement of spin-echo acquisition sequences by echo-planar sequences should allow obtaining of much faster acquisition times at MR elastography. Lastly, MR imaging is a more costly procedure than sonography. However, the accuracy of MR elastography being higher than that of ultrasound elastography, it would be interesting to compare the patient outcome efficacy and the cost-effectiveness of the 2 methods.

In our study, the correlation between liver elasticity and fibrosis stage was not affected by steatosis or inflammation, which is in agreement with the previous studies assessing liver elasticity with ultrasound.[5,6,9] It should, however, be acknowledged that most patients in our study did not have steatosis or had only mild steatosis. Moreover, it has been recently reported that the liver elasticity measured with ultrasound elastography was increased during exacerbations of hepatitis in patients with chronic hepatitis B and C infections.[24] Further studies are needed to investigate the effect of severe steatosis and inflammation on elasticity measurements.

When assessing the accuracy of diagnostic methods to stage liver fibrosis, the prevalence of each fibrosis stage is known to influence the areas under the ROC curves with artificially large areas observed when only extreme stages

Table 2. Most Discriminating MR Elasticity Cutoff Values and Corresponding Sensitivities, Specificities, and Positive and Negative Predictive Values With 95% CIs for METAVIR Scores F ≥ 1, F ≥ 2, F ≥ 3 and F = 4

	F ≥ 1	F ≥ 2	F ≥ 3	F = 4
Cutoff (*kPa*)	2.42	2.49	3.13	4.13
Sensitivity (95% CI)	0.85 (0.75–0.92)	1.0 (0.93–1.0)	0.91 (0.76–0.98)	1.0 (0.81–1.0)
Specificity (95% CI)	0.91 (0.71–0.99)	0.91 (0.78–0.97)	0.97 (0.89–1.0)	0.96 (0.89–0.99)
Positive predictive value (95% CI)	0.97 (0.89–1.0)	0.93 (0.83–0.98)	0.94 (0.79–0.99)	0.86 (0.64–0.97)
Negative predictive value (95% CI)	0.64 (0.45–0.81)	1.0 (0.91–1.0)	0.95 (0.87–0.99)	1.0 (0.95–1.0)

Table 3. Distribution of Fibrosis Stages at MR Elastography Versus Fibrosis Stages at Histology

Fibrosis stage at histology	Fibrosis stage at MR elastography					
	F0	F1	F2	F3	F4	Total
F0	20	2				22
F1	11	8	3			22
F2			17	2		19
F3			3	9	3	15
F4					18	18

of fibrosis (ie, F0 vs F4) are included in the study population.[25] In our study, this artificial effect was avoided because the fibrosis stages were uniformly distributed. Moreover, the sensitivity analysis showed similar findings in the subpopulation with $F \leq 2$ than in the whole study population, with statistically larger areas under the ROC curves for MR elastography than for ultrasound elastography and APRI.

Our study is limited by the lack of precise correlation between the hepatic samples analyzed by the pathologists and the volumes of interest at MR elastography or the regions measured at ultrasound elastography. The heterogeneity of liver fibrosis may explain some discrepancies in the stages of fibrosis between the elastographic methods and the histologic analysis. It should be remembered that liver biopsy is not an optimal reference examination (only 1/50,000 of the organ is analyzed), especially when the length of the samples is shorter than 25 mm.[1] We do not know if discordant results between MR elastography and histopathology in our study were specifically caused by problems of inadequate biopsy sampling. These problems were minimized by excluding 13 patients with unsuitable samples containing <10 portal tracks or obvious regenerating nodules.

APRI scoring is only one of the biochemical tests that have been proposed to stage liver fibrosis. We used this test because of its simplicity. The use of more elaborate tests such as the FibroTest may yield better results. However, the areas of MR elastography in our study were larger than those of the combination of ultrasound elastography and FibroTest (0.88 for $F \geq 2$, 0.95 for $F \geq 3$, and 0.95 for $F = 4$) previously reported by Castera et al.[4]

Because we performed a prospective study including consecutive patients who underwent liver biopsy for suspicion of chronic liver disease, our series included various causes of liver fibrosis. Because staging at histology may be influenced by the cause of fibrosis and the type and length of the biopsy specimen, these variables were included in the statistical analysis but were not found to significantly influence the results of MR elastography. This suggests that MR elastography is a robust method. However, a more detailed analysis in homogeneous groups of patients should be performed in future studies.

The use of a common staging system of fibrosis for all types of chronic liver diseases, as in our study, remains controversial. More particularly, use of the Brunt classification has been recommended for patients with nonalcoholic steatohepatitis and alcoholic liver disease. The Brunt scoring system differs from METAVIR by taking perisinusoidal fibrosis into account in mild fibrosis. We do not believe that using the Brunt classification in patients with steatohepatitis would have changed our results significantly because we did not find an influence of the cause of chronic liver disease on the MR elasticity in the analysis of variance and because only one patient classified as F1 with METAVIR would have been reclassified as stage 2 with the Brunt classification. Further studies should be performed to assess the value of MR elastography in patients with nonalcoholic steatohepatitis and alcoholic liver disease.

In conclusion, the results of our study showed that MR elastography is a reproducible imaging method that is superior to ultrasound elastography and APRI measurements for staging liver fibrosis. This suggests that MR elastography should be the preferred noninvasive method for accurate assessment and follow-up of liver fibrosis. Further studies are needed to confirm the reproducibility of MR elastography between centers and to assess the performance of MR elastography in staging patients with specific chronic liver diseases or predicting the response to treatment.

References

1. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology 2003;38:1449–1457.
2. Rockey DC, Bissell DM. Noninvasive measures of liver fibrosis. Hepatology 2006;43:S113–S120.
3. Friedman SL. Liver fibrosis—from bench to bedside. J Hepatol 2003;38(Suppl 1):S38–S53.
4. Castera L, Vergniol J, Foucher J, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. Gastroenterology 2005;128:343–350.
5. Sandrin L, Fourquet B, Hasquenoph JM, et al. Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol 2003;29:1705–1713.
6. Ziol M, Handra-Luca A, Kettaneh A, et al. Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. Hepatology 2005;41:48–54.
7. Huwart L, Peeters F, Sinkus R, et al. Liver fibrosis: non-invasive assessment with MR elastography. NMR Biomed 2006;19:173–179.
8. Huwart L, Sempoux C, Salameh N, et al. Liver fibrosis: noninvasive assessment with MR elastography versus aspartate aminotransferase-to-platelet ratio index. Radiology 2007;245:458–466.
9. Yin M, Talwalkar JA, Glaser KJ, et al. Assessment of hepatic fibrosis with magnetic resonance elastography. Clin Gastroenterol Hepatol 2007;5:1207–1213.
10. Muthupillai R, Lomas DJ, Rossman PJ, et al. Magnetic resonance elastography by direct visualization of propagating acoustic strain waves. Science 1995;269:1854–1857.

11. Sinkus R, Tanter M, Catheline S, et al. Imaging anisotropic and viscous properties of breast tissue by magnetic resonance-elastography. Magn Reson Med 2005;53:372–387.
12. Lucidarme D, Foucher J, Le Bail B, et al. The ratio interquartile range/median value of liver stiffness measurement is a key factor of accuracy of transient elastography (FibroScan) for the diagnosis of liver fibrosis. Hepatology 2007;46:S181.
13. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. The French METAVIR Cooperative Study Group. Hepatology 1994;20:15–20.
14. Foucher J, Chanteloup E, Vergniol J, et al. Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study. Gut 2006;55:403–408.
15. Brunt EM, Janney CG, Di Bisceglie AM, et al. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol 1999;94:2467–2474.
16. Ludwig J. The nomenclature of chronic active hepatitis: an obituary. Gastroenterology 1993;105:274–278.
17. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–845.
18. Steyerberg EW, Harrell FE Jr, Borsboom GJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54:774–781.
19. Akritas M, Arnold SF, Brunner E. Nonparametric hypothesis and rank statistics for unbalanced factorial designs. J Am Stat Assoc 1997;92:258–265.
20. Talwalkar JA, Kurtz DM, Schoenleber SJ, et al. Ultrasound-based transient elastography for the detection of hepatic fibrosis: systematic review and meta-analysis. Clin Gastroenterol Hepatol 2007;5:1214–1220.
21. Shaheen AA, Myers RP. Diagnostic accuracy of the aspartate aminotransferase-to-platelet ratio index for the prediction of hepatitis C-related fibrosis: a systematic review. Hepatology 2007;46:912–921.
22. Foucher J, Castera L, Bernard PH, et al. Prevalence and factors associated with failure of liver stiffness measurement using FibroScan in a prospective study of 2114 examinations. Eur J Gastroenterol Hepatol 2006;18:411–412.
23. Bercoff J, Tanter M, Fink M. Supersonic shear imaging: a new technique for soft tissue elasticity mapping. IEEE Trans Ultrason Ferroelectr Freq Control 2004;51:396–409.
24. Coco B, Oliveri F, Maina AM, et al. Transient elastography: a new surrogate marker of liver fibrosis influenced by major changes of transaminases. J Viral Hepat 2007;14:360–369.
25. Poynard T, Halfon P, Castera L, et al. Standardization of ROC curve areas for diagnostic evaluation of liver fibrosis markers based on prevalences of fibrosis stages. Clin Chem 2007;53:1615–1622.

Received December 21, 2007. Accepted March 27, 2008.

Address requests for reprints to: Laurent Huwart, MD, PhD, Diagnostic Radiology Unit, Université Catholique de Louvain, St-Luc University Hospital, Avenue Hippocrate 10, B-1200 Brussels, Belgium. e-mail: huwart.laurent@wanadoo.fr; fax: (32) 2-7705574.

Supported by grants FRSM 3.4578.00 and 3.4580.06 from the Fonds National de la Recherche Scientifique and by a grant from the Fondation St-Luc, Belgium.

L.C.T. is an employee of Philips Medical Systems.