

## Original Article

# Automated morphometry provides accurate and reproducible virtual staging of liver fibrosis in chronic hepatitis C

Paul Calès<sup>1,2</sup>, Julien Chaigneau<sup>1</sup>, Gilles Hunault<sup>1</sup>, Sophie Michalak<sup>1,3</sup>, Christine Cavaro-Menard<sup>4</sup>, Jean-Baptiste Fasquel<sup>4</sup>, Sandrine Bertrais<sup>1</sup>, Marie-Christine Rousselet<sup>1,3</sup>

<sup>1</sup>HIFIH Laboratory, Unité Propre de Recherche de l'Enseignement Supérieur 3859, Structure Fédérative de Recherche 4208, LUNAM University, Departments of <sup>2</sup>Liver-Gastroenterology and <sup>3</sup>Cellular and Tissue Pathology, CHU Angers, <sup>4</sup>LARIS Laboratory, Unité Propre de Recherche de l'Enseignement Supérieur 7315, LUNAM University, Angers, France

E-mail: \*Prof. Marie-Christine Rousselet - [mcrousset@chu-angers.fr](mailto:mcrousset@chu-angers.fr)

\*Corresponding author

Received: 01 September 14

Accepted: 16 March 15

Published: 28 May 2015

### This article may be cited as:

Calès P, Chaigneau J, Hunault G, Michalak S, Cavaro-Menard C, Fasquel JB, et al. Automated morphometry provides accurate and reproducible virtual staging of liver fibrosis in chronic hepatitis C. *J Pathol Inform* 2015;6:20.

Available FREE in open access from: <http://www.jpathinformatics.org/text.asp?2015/6/1/20/157782>

Copyright: © 2015 Calès P. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## Abstract

**Background:** Liver fibrosis staging provides prognostic value, although hampered by observer variability. We used digital analysis to develop diagnostic morphometric scores for significant fibrosis, cirrhosis and fibrosis staging in chronic hepatitis C. **Materials and Methods:** We automated the measurement of 44 classical and new morphometric descriptors. The reference was histological METAVIR fibrosis (F) staging (F0 to F4) on liver biopsies. The derivation population included 416 patients and liver biopsies  $\geq 20$  mm-length. Two validation population included 438 patients. **Results:** In the derivation population, the area under the receiver operating characteristic (AUROC) for clinically significant fibrosis (F stage  $\geq 2$ ) of a logistic score combining 5 new descriptors (stellar fibrosis area, edge linearity, bridge thickness, bridge number, nodularity) was 0.957. The AUROC for cirrhosis of 6 new descriptors (edge linearity, nodularity, portal stellar fibrosis area, portal distance, granularity, fragmentation) was 0.994. Predicted METAVIR F staging combining 8 morphometric descriptors agreed well with METAVIR F staging by pathologists:  $\kappa = 0.868$ . Morphometric score of clinically significant fibrosis had a higher correlation with porto-septal fibrosis area ( $r_s = 0.835$ ) than METAVIR F staging ( $r_s = 0.756$ ,  $P < 0.001$ ) and the same correlations with fibrosis biomarkers, e.g., serum hyaluronate:  $r_s = 0.484$  versus  $r_s = 0.476$  for METAVIR F ( $P = 0.862$ ). In the validation population, the AUROCs of clinically significant fibrosis and cirrhosis scores were, respectively: 0.893 and 0.993 in 153 patients (biopsy  $< 20$  mm); 0.955 and 0.994 in 285 patients (biopsy  $\geq 20$  mm). The three morphometric diagnoses agreed with consensus expert reference as well as or better than diagnoses by first-line pathologists in 285 patients, respectively: significant fibrosis: 0.733 versus 0.733 ( $\kappa$ ), cirrhosis: 0.900 versus 0.827, METAVIR F: 0.881 versus 0.865. **Conclusion:** The new automated morphometric scores provide reproducible and accurate diagnoses of fibrosis stages via "virtual expert pathologist."

**Key words:** Cirrhosis, digital histopathology, image analysis, liver biopsy, METAVIR staging

### Access this article online

Website:  
[www.jpathinformatics.org](http://www.jpathinformatics.org)

DOI: 10.4103/2153-3539.157782

### Quick Response Code:



## INTRODUCTION

All chronic liver diseases are characterized by the development of liver fibrosis, which includes the growth of fibrous tissue around hepatocytes (perisinusoidal fibrosis), the enlargement of portal tracts and the development of abnormal connections between portal or venular structures, named bridging or septal fibrosis. These events lead to liver architecture modifications and in turn to major complications responsible for increased mortality.<sup>[1]</sup> Determination of the severity of hepatic fibrosis, especially the diagnosis of bridging fibrosis and cirrhosis, is important for management. Currently, the reference method for staging liver fibrosis is the microscopic examination of a liver specimen, usually obtained by liver biopsy, by a pathologist who yields his conclusion with a semi-quantitative score, such as the METAVIR classification.<sup>[2]</sup> However, fibrosis staging by such a conventional liver biopsy examination is limited by poor inter/intra-observer reproducibility.<sup>[3]</sup> This limit can be partially circumvented by an expert reading, although a recent study showed that single experts may not reach excellent reproducibility and that an expert panel seemed to be the best reading reference.<sup>[4]</sup> Thus, working with centers of expertise would be interesting in liver pathology but carries with it the problem of pathologist availability. In addition, expert pathologists are lacking in some countries.

Another approach to liver evaluation is to quantify fibrosis characteristics with morphometric methods and measurements.<sup>[5]</sup> Area of fibrosis (also referred as collagen proportionate area) is the main descriptor, but studies

have also looked at fibrosis fractal dimension<sup>[6]</sup> and the perimeter and size of collagenous elements.<sup>[7]</sup> These published techniques provide quantitative information but correlate moderately with reference pathological staging.

Our objective in the present study was to develop a computer-assisted fibrosis assessment method that correlates well with expert staging and could, therefore, circumvent observer variability or stand in for expert pathologist when they are unavailable. We first developed several new fibrosis-related morphometric descriptors that quantitatively describe the information evaluated by pathologists when they stage fibrosis. Second, we automated these measurements to provide automated scorings for the diagnosis of significant fibrosis, cirrhosis and METAVIR fibrosis stages. This proof of concept study was performed in patients with chronic viral hepatitis C.

## MATERIALS AND METHODS

### Study Design

The study design is summarized in Figure 1. Briefly, there were four steps (the first two being chronologically independent): Construction of morphometric descriptors of fibrosis and related liver lesions by an engineer guided by a pathologist on digital images; classical fibrosis METAVIR staging independently performed by pathologists; statistical combination of morphometric descriptors allowing the statistical construction of morphometric scores targeted on fibrosis staging; and finally statistical development of morphometric diagnoses reflecting histological diagnoses.

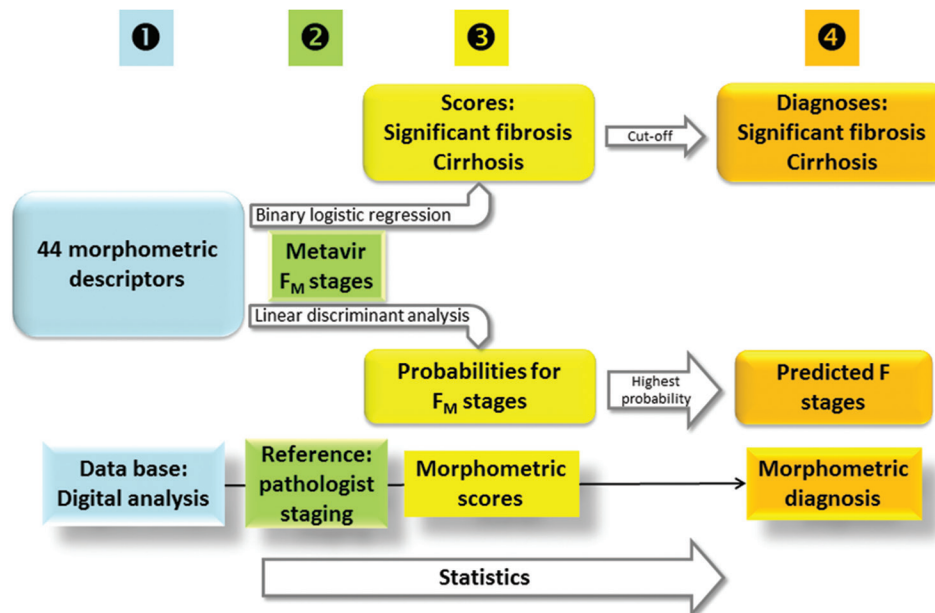


Figure 1: Overall study design. Construction of morphometric diagnoses in 4 steps

## Objectives and Definitions

The observed fibrosis stage by pathologist is called F hereafter. The main objective was to provide automated scorings for the diagnosis of significant fibrosis and cirrhosis. Significant fibrosis was considered when METAVIR stages were  $F \geq 2$ . The dependent (predicted) variable was thus METAVIR stages  $F2 + F3 + F4$  versus  $F0 + F1$ . Significant means clinically significant since long-term survival is decreased by these stages ( $F \geq 2$ ), and consequently anti-viral drugs are indicated. Cirrhosis was considered when METAVIR stage was  $= F4$ . The dependent variable was thus METAVIR stages  $F4$  versus  $F0 + F1 + F2 + F3$ . The independent variables (predictors) were fine descriptors of fibrosis obtained by morphometry. The 44 descriptors measured were called morphometric descriptors. Statistical function provided logistic scores ranging from 0 to 1. This roughly means from 0% to 100% probability of the diagnostic target. The two logistic scores obtained were called morphometric scores and more precisely (morphometric) significant fibrosis score and cirrhosis score.

The automated diagnosis of METAVIR fibrosis stages was a secondary objective since the statistical functions for multinomial targets like METAVIR staging are less performant than those for binary target like significant fibrosis or cirrhosis. The discriminant function directly provided the most probable F stage that was called predicted F stage. Finally, diagnosis of significant fibrosis by significant fibrosis score, diagnosis of cirrhosis by cirrhosis score and diagnosis of Metavir F stage by predicted F stage were all grouped as morphometric diagnoses.

## Population

The two included populations had a liver biopsy scored with METAVIR fibrosis staging. The largest population with adequate ( $\geq 20$  mm or cirrhosis [F4]) liver biopsy was used as the derivation population for morphometric diagnostics validated in other populations. Ultimately, 834 patients were included in the study. All patients had single-etiology, compensated, chronic viral hepatitis C without causal or anti-fibrotic treatment during the 6 months preceding the inclusion. Initial patient groups were originally recruited for studies focused on the evaluation of noninvasive liver fibrosis tests. The only additional selection criterion was the availability of a liver specimen for centralized morphometric analysis. Informed consent was obtained from patients in all studies after Institutional Review Board approval.

### Derivation population

This population included 416 patients with a digitized liver biopsy  $\geq 20$  mm in METAVIR stages F0 to F3 or cirrhosis (liver biopsy length  $\geq 12.1$  mm) provided by a previously published population number 1<sup>[8]</sup> (but excluding the Fibrostar population). The F stage distribution was: 18 F0, 169 F1, 116 F2, 59 F3, and 54 F4.

### Validation population

First, 153 patients with liver biopsy  $< 20$  mm were available in the population number 1. This population included patients with cirrhosis explaining that the sum of the two groups in this population ( $n = 569$ ) was superior to the total number of patients included ( $n = 549$ ). Second, 285 other patients with liver biopsy  $\geq 20$  mm were recruited in the population number 2 from the Fibrostar study.<sup>[9]</sup> Thus, the two validation population included 438 patients.

### Liver Specimen

All liver biopsy were performed by needle (usually suction technique, 1.6 mm diameter) intercostally. Specimens were formalin-fixed and paraffin-embedded.

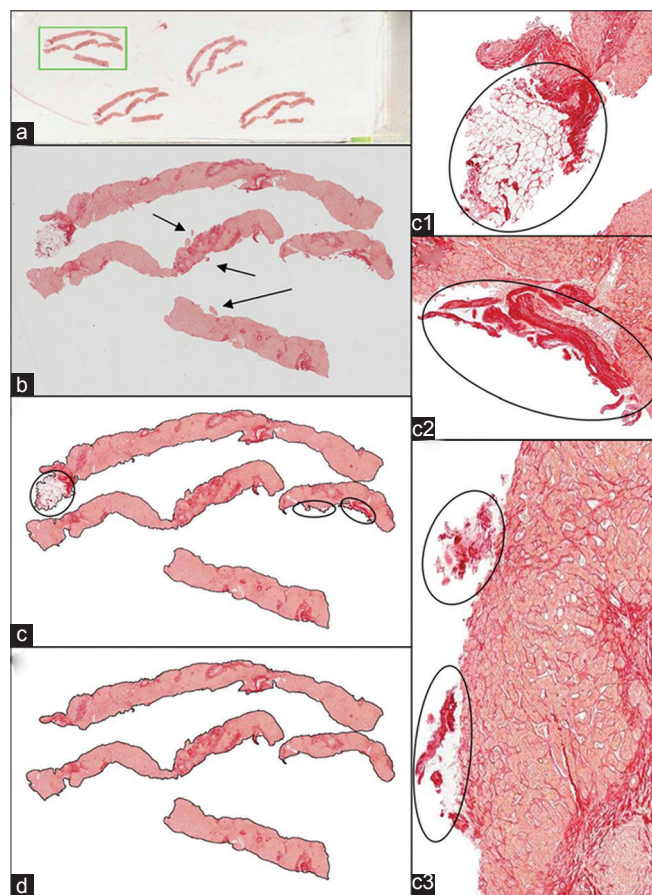
### METAVIR Staging

The METAVIR fibrosis stages were as originally defined:<sup>[2]</sup> F0: No fibrosis, F1: Stellar enlargement of portal tracts without septa, F2: Enlargement of portal tracts with rare septa, F3: Numerous septa without cirrhosis, F4: Cirrhosis. In all population, METAVIR fibrosis staging was performed centrally by the same two senior liver pathologists from the Angers tertiary center with a consensus rereading in cases of initial disagreement. These pathologists (MCR, SM) were experienced in the reproducibility conditions of Metavir staging.<sup>[3]</sup> In the Fibrostar study, first-line diagnoses (i.e. F stages as initially assessed by the pathologists at the tertiary centers) were also available.

### Digitized image analysis

Conditions for liver biopsy specimens were the same in all populations, especially as concerns length measurement and staining. 5  $\mu\text{m}$ -liver sections were stained with picosirius red without hematoxylin counterstain as the target of interest was fibrosis only. Morphometry was centrally performed by a single engineer (JC) in the Angers center. The liver specimen slides were fully scanned with an Aperio digital slide scanner (Scanscope CS System, Aperio Technologies, Vista CA, USA) image processor that provided high quality images with a maximum scanning area capacity of  $120,000 \times 50,000$  pixels (resolution of 0.5  $\mu\text{m}/\text{pixel}$  at magnification  $\times 20$ ). After selecting the scanning area [Figure 2a], a digital selected image was created [Figure 2b]. Images were compressed using JPEG2000 software (quality  $Q = 70$ ) in order to reduce the file size (for example, an image can have a size reduction from 1 GB to 30 MB). An automated preprocessing algorithm was used to clean the background and the tissue to analyze [Figure 2c]. When necessary, the engineer manually removed tissue artefacts or inappropriate stained structures such as folds, dust, large vascular or biliary lumens and liver capsule, which yielded a fully cleaned digital image [Figure 2d]. Further image analysis of the whole liver section included three manipulations. First, the digital color image composed

of three channels red, green and blue, was decomposed to work on the green channel only as it offers the best contrast to visualize fibrosis. Only fibrosis pixels were taken into account, thanks to the preprocessing algorithm and the optional manual cleaning step that deleted artefacts or inappropriate structures. Second, this graded-intensity green image was thresholded according to an automated process:<sup>[10,11]</sup> a fuzzy generalized classification process allowed for the merging of pixel intensities into three classes (fibrosis, steatosis or healthy tissue) using the minimization of an original energy function. This produced a binary black-and-white image where fibrosis appeared in black, and all other structures appeared in white. The third step is the quantitative measure of the black space occupied by fibrosis. An example of the different steps is shown in Figure 3. The present innovation consisted in the distinction of different elementary lesions related to fibrosis (new morphometric descriptors) in addition to classical morphometry.



**Figure 2: Steps of the cleaning process of digital liver specimen image. (a)** Glimpse of the whole slide in order to select the scanning area of a picosirius red stained section of the whole liver specimen (green rectangle). **(b)** First step on the scanned liver specimen: The light-grey background and small artefacts (arrows) have to be removed by the automated process. **(c)** Second step: The liver specimen still displays some inappropriate areas for fibrosis measure (circles). These are, on higher magnification, **c1**: Liver capsule with adipose tissue; **c2** and **c3**: Torn fragments at the edge of the specimen. **(d)** Final image after manual cleaning

### Classical morphometry

Table 1 summarizes the classical morphometric descriptors.

### Area and fractal dimension of fibrosis and steatosis

Area of whole fibrosis [Figure 4a] and area of steatosis measurement has been described elsewhere.<sup>[12]</sup> The “box-counting” method<sup>[13]</sup> has been extensively used for measuring the fractal dimension of many histological objects as a complexity index. This box-counting method provided the fractal dimension of Kolmogorov: Fractal dimension of whole fibrosis or steatosis.

### Area and fractal dimension of porto-septal and perisinusoidal fibrosis

Definition assumptions: Porto-septal fibrosis was defined as fibrosis including at least one of the two following patterns: The portal fibrosis with stellar fine lobular fibrous expansions closely related to portal spaces defining F1; dense fibrous connexions (bridges or septa) between portal spaces or between portal spaces and centrilobular veins defining F2 or F3. Perisinusoidal fibrosis was defined as fibrosis around hepatocytes and around small centrilobular veins; details are provided in Supplement and in a previous study.<sup>[11]</sup> Area of perisinusoidal fibrosis was easily measured by the difference between the areas of whole fibrosis and of porto-septal fibrosis.

The automated mask detection of porto-septal fibrosis [Figure 4b] allowed us to distinguish perisinusoidal fibrosis [Figure 4c] from the whole fibrosis. We thus measured the area and fractal dimension of porto-septal fibrosis, the area and fractal dimension of perisinusoidal fibrosis, the area of lobular perisinusoidal fibrosis and the ratio of perisinusoidal fibrosis area among the whole fibrosis; details on fibrosis metrics are available elsewhere.<sup>[11]</sup> The automated mask detection of porto-septal fibrosis and the new descriptors described were measured at magnification  $\times 5$  since they do not need high-resolution images (details in the Supplement).

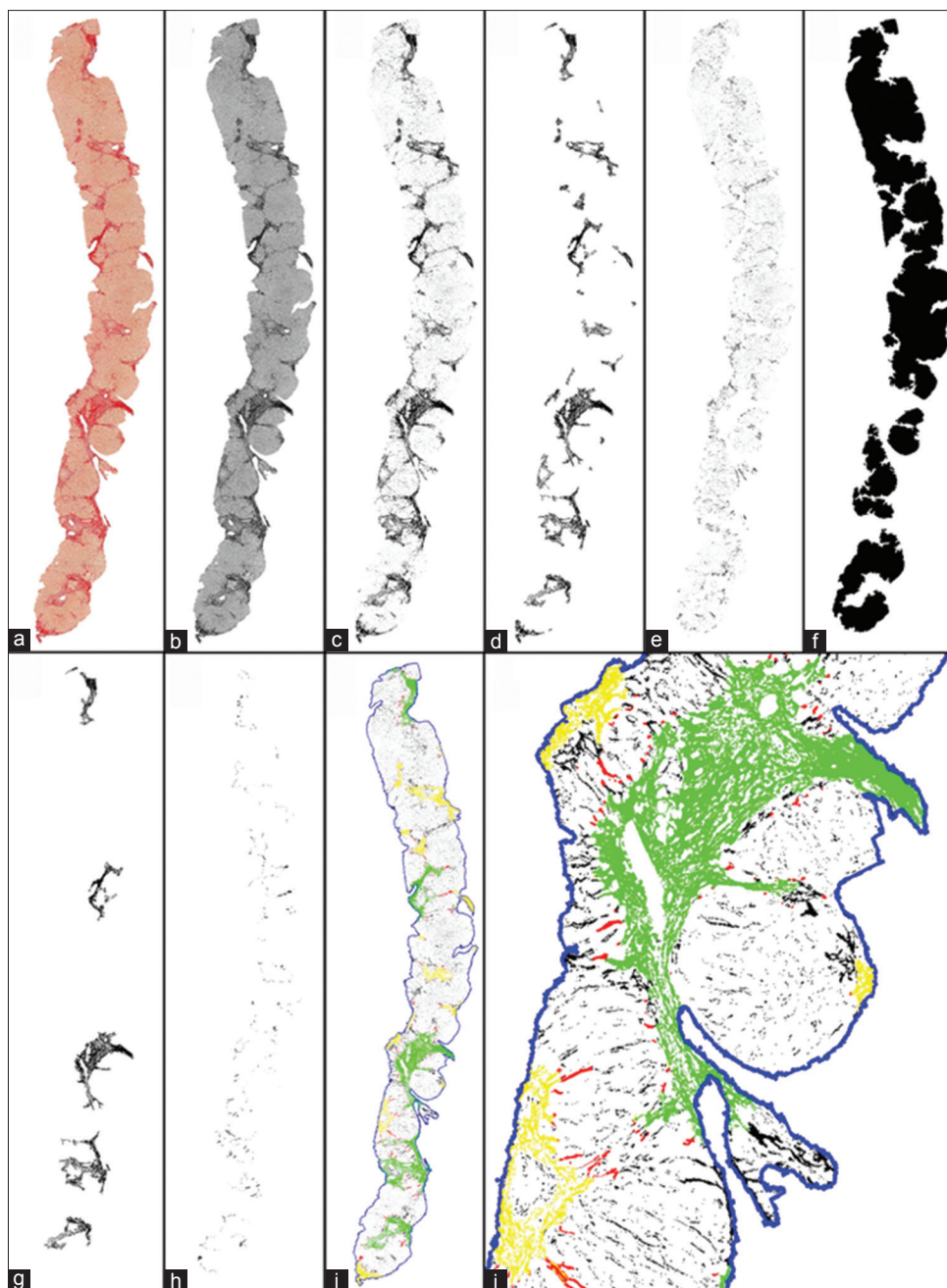
### New morphometric descriptors

Table 1 summarizes the new morphometric descriptors; technical details on the measurement of the following lesions and characteristics are provided in the Supplement.

### Stellar fibrosis

The main feature differentiating F0 from higher stages is the presence of stellar fibrosis [Figure 4d], that is, small, fine fibrils branching from the dense porto-septal regions. Therefore, we automated the quantitation of this stellar fibrosis on the whole liver biopsy specimen surface (whole area of stellar fibrosis) on the surface of porto-septal regions (portal area of stellar fibrosis) and on the surface of lobular regions (lobular area of stellar fibrosis). We also measured the mean area of stellar fibrosis per porto-septal region (mean portal area of stellar fibrosis).





**Figure 3: Steps of the automated analysis of a digital liver specimen for the determination of the morphometric descriptors in a Metavir F3 stage. (a) Picrosirius red stained digital section after cleaning steps. (b) Green channel of the digital liver specimen used to threshold whole fibrosis. (c) Binary image after automated threshold of whole fibrosis (black pixels). (d) Automated detection of porto-septal fibrosis. (e) Automated detection of perisinusoidal fibrosis. (f) Application of the porto-septal mask to measure the granularity (granules in black pixels). (g) Automated detection of fibrous bridges. (h) Automated detection of stellar fibrosis. (i) and (j) (higher magnification). Final image with morphometric descriptors: Edge linearity percentage is highlighted in dark blue pixels, stellar fibrosis in red pixels, porto-septal fibrosis in yellow pixels, perisinusoidal fibrosis in black pixels, and bridges between portal tracts in green pixels**

#### *Porto-septal regions*

We measured the number of porto-septal regions; the mean area of porto-septal regions (mean area of porto-septal fibrosis) and the average distance between the portal tracts (mean portal distance). This last parameter reflects either a normal regular distribution of portal tracts or an architectural distortion with some portal tracts closer to others. In advanced fibrosis, portal

tracts embedded in septa are automatically detected as a unique porto-septal region and consequently the mean portal distance is calculated as zero.

#### *Bridging fibrosis*

Bridging (or septal) fibrosis [Figure 4e] occurs in stages F2 to F4. A bridge was defined with morphometry as a structure lying between two thick elements. There is no

**Table 1: Patterns related to liver fibrosis measured by automated liver morphometry**

Type	Pattern measured	F $\geq$ 2	F4	All F <sup>a</sup>
Fibrosis	Classical patterns (n=10)			
	Area of whole fibrosis (%)			
	Fractal dimension of whole fibrosis			
	Area of porto-septal fibrosis (%)			
	Fractal dimension of porto-septal fibrosis			6
	Area of perisinusoidal fibrosis (%)			
	Fractal dimension of perisinusoidal fibrosis			4
Steatosis	Area of lobular perisinusoidal fibrosis (%)			
	Ratio of perisinusoidal fibrosis area (%)			5
	Area of steatosis (%)			
Stellar fibrosis	Fractal dimension of steatosis			
	New fibrosis patterns (n=23)			
Porto-septal regions	Directly related to fibrosis			
	Whole area of stellar fibrosis (%)	1		
	Portal area of stellar fibrosis (%)		3	
Bridges	Lobular area of stellar fibrosis (%)			
	Mean portal area of stellar fibrosis ( $\mu\text{m}^2$ )			
Bridges	Mean area of porto-septal fibrosis ( $\mu\text{m}^2$ )			
	Number of porto-septal regions			
	Mean portal distance ( $\mu\text{m}$ )		4	
	Number of bridges	4		
	Portal ratio of bridges (%)			3
Nodules	Area of bridging fibrosis (%)			
	Mean bridge thickness ( $\mu\text{m}$ )	3		
	Mean bridge perimeter ( $\mu\text{m}$ )			
	Mean bridge area ( $\mu\text{m}^2$ )			
	Mean granularity percentage		5	1
Fragmentation	Indirectly related to fibrosis			
	Number of nodules			
Edges	Mean nodularity percentage	5	2	8
	Number of fragments			
Specimen dimensions	Fragmentation index (%)		6	7
	Edge linearity percentage	2	1	2
	Fractal dimension of edges			
Single intensity	Liver specimen length (mm)			
	Liver specimen perimeter (mm)			
	Liver specimen surface ( $\text{mm}^2$ )			
	Color intensity (n=11)			
	Mean specimen red intensity			
	Mean specimen green intensity			
	Mean specimen blue intensity			
	Mean fibrosis red intensity			
	Mean fibrosis green intensity			
	Mean fibrosis blue intensity			
Contrast	Mean parenchyma red intensity			
	Mean parenchyma green intensity			
Contrast	Mean parenchyma blue intensity			
	Fibrosis/specimen contrast			
	Fibrosis/parenchyma contrast			

<sup>a</sup>The figures in the right columns indicate the rank of the 13 independent descriptors for the three diagnostic targets evaluated: significant fibrosis (F $\geq$ 2), cirrhosis (F4) and METAVIR fibrosis (F) stage

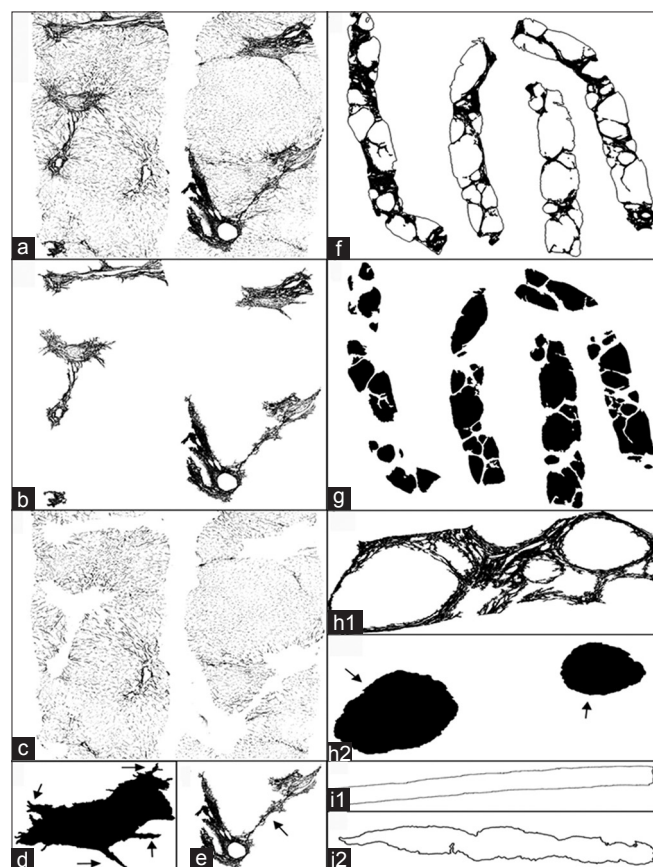
literature reference to distinguish F2 from F3 stages with strict quantitative criteria; pathologists subjectively choose

between a few septa (F2) and numerous septa (F3). The process of automated scoring required a quantitative

definition. Guided by the personal experience of the participating pathologists, we considered that the F3 or F4 stage has been reached when the ratio of bridges was  $>50\%$ . Therefore, we measured the number of bridges, the ratio of bridges in the porto-septal regions (portal ratio of bridges), the area of fibrosis in the bridges (area of bridging fibrosis) and the mean bridge thickness obtained from their perimeter (mean bridge perimeter) and surface (mean bridge area). Liver architecture is modified in high F stages by many fibrous bridges [Figure 4f]. These bridges isolate parenchymal areas among lobules that we called granules [Figure 4g]. Thus, we measured the mean granularity percentage quantifying the liver structure disruption by bridging fibrosis.

### Nodules

Cirrhosis is mainly characterized by the formation of nodules, which we defined in image analysis as circular and non-fibrotic (lacking fibrous septa) areas surrounded by fibrosis [Figure 4h1]. We measured the number of nodules [Figure 4h2] and the mean percentage of fibrosis around nodules (mean nodularity percentage).



**Figure 4: Binary images after automated thresholding of picosirius red-stained liver specimens. (a) Whole fibrosis in a F2 specimen. (b) Porto-septal fibrosis. (c) Perisinusoidal fibrosis. (d) Stellar fibrosis (arrows) expands a portal tract. (e) Fibrous bridge (arrow). (f) Porto-septal fibrosis in cirrhosis. (g) Application of a porto-septal mask to measure granularity. (h) Nodules detection: Circular fibrosis around nodules (1) and count of nodules (2). (i) Edge linearity: Straight (1) and irregular (2) edges**

### Fragmentation

Liver biopsy specimens may be fragmented. This is fibrosis stage dependent and occurs particularly in cirrhosis where there may be several small ( $\leq 5 \text{ mm}^2$ ) fragments. We measured the number of fragments and the fragmentation index; this latter defined as the ratio between the surface of detected small fragments and the total liver biopsy specimen surface.

### Edges

Liver biopsy specimens change in shape according to F stages: Edges are rather straight in low stages (F0, F1, F2) [Figure 4i1] but tend to become more curved and irregular in high stages (F3, F4) [Figure 4i2]. Thus, we automated the measurement of the fractal dimension of edges and the edge linearity percentage.

### Liver specimen dimensions

Pathologists usually consider that a representative liver biopsy specimen must be  $\geq 15$  or  $\geq 20$  mm long. We automated the measurement of liver specimen length, which required specific rules especially for twisted specimens. We also measured liver specimen perimeter and liver specimen surface.

### Staining quality

The performance of morphometric descriptors depends on specimen staining quality (picosirius red in the present study). We measured stain intensity on the three color components of the image (red, green, blue) and obtained a mean color intensity index by averaging all pixel intensities for, respectively, the whole specimen, fibrosis and parenchyma. The term "color intensity" was used to refer to this evaluation of staining quality. The contrast between the mean color intensity of fibrosis and either the liver biopsy specimen or the parenchyma was also recorded.

### Noninvasive Fibrosis Tests

The two following blood tests were calculated: Fibrotest,<sup>[14]</sup> FibroMeter<sup>V2C</sup>,<sup>[15]</sup> Liver stiffness<sup>[14]</sup> evaluation by elastometry (Fibroscan, Echosens, Paris, France) and the combined test Elasto-FibroMeter<sup>2C</sup><sup>[16]</sup> were also performed.

### Statistics

Quantitative variables were expressed as mean  $\pm$  standard deviation. The diagnostic accuracy of each morphometric score was expressed as the area under the receiver operating characteristic (AUROC) and the overall accuracy (rate of well-classified patients according to F stages). Data were reported according to STARD statements<sup>[17]</sup> and analyzed on an intention to diagnose the basis.

### Multivariate analysis

Independent predictors of binary diagnostic targets (significant fibrosis and cirrhosis) were determined by binary logistic regression. Independent predictors of F

stages were determined by a linear discriminant analysis, providing a probability score from 0 to 1 for each F stage. Independent predictors of liver stiffness were determined by multiple linear regression.

**Software**

The main statistical analyses were performed by professional statisticians (GH, SB) using SPSS version 18.0 (IBM, Armonk, NY, USA), SAS 9.1 (SAS Institute Inc., Cary, NC, USA) and R software.<sup>[18]</sup>

**RESULTS**

**Population**

The main characteristics of the population are depicted in Table 2.

**Development of Morphometric Scores**

*Significant fibrosis diagnosis*

In the derivation population, the best logistic model provided a morphometric significant fibrosis score combining 5 independent morphometric descriptors: Whole area of stellar fibrosis ( $P < 0.0001$  in multivariate analysis), edge linearity percentage ( $P < 0.0001$ ), mean bridge thickness ( $P = 0.0018$ ), number of bridges ( $P = 0.0032$ ) and mean nodularity percentage

( $P = 0.0293$ ); descriptor rank is that of stepwise analysis as indicated in Table 1. Examples of score determination are shown in Figure 5. The value of morphometric score, varying from 0 to 1, is proportional to METAVIR F stage [Figure 6], e.g. predicted F1 values are (median, 1<sup>st</sup> and 3<sup>rd</sup> quartiles): 0.011 (0.003–0.042), predicted F2: 0.102 (0.038–0.269). With an *a priori* threshold score at 0.5, the AUROC was 0.957 (95% confidence interval: 0.940–0.973) and 87.3% (84.1–90.5) of the patients were correctly classified for significant fibrosis diagnosis. Table 3 provides the overall performance results.

*Cirrhosis diagnosis*

The best logistic model provided a morphometric cirrhosis score including 6 independent morphometric descriptors: Edge linearity percentage ( $P = 0.0058$ ), mean nodularity percentage ( $P = 0.0029$ ), portal area of stellar fibrosis ( $P < 0.0001$ ), mean portal distance ( $P = 0.0145$ ), mean granularity percentage ( $P = 0.0003$ ) and fragmentation index ( $P = 0.0012$ ). The morphometric score threshold was selected here *a posteriori* at 0.1567 to maximize sensitivity with minimal deterioration of specificity. Thus, sensitivity was 100% and ensuing specificity was 96.4%. The AUROC was 0.994 (0.989–0.999), and 96.9% (95.2–98.5) of the patients were correctly classified for cirrhosis diagnosis. Table 3 provides the overall performance results.

*METAVIR F stage diagnosis*

The final discriminant model included 8 descriptors of predicted F stage: Mean granularity percentage ( $P < 0.0001$ ), edge linearity percentage ( $P < 0.0001$ ), portal ratio of bridges ( $P < 0.0001$ ), fractal dimension of perisinusoidal fibrosis ( $P < 0.0001$ ), ratio of perisinusoidal fibrosis area ( $P < 0.0001$ ), fractal dimension of porto-septal fibrosis ( $P < 0.0001$ ), fragmentation index ( $P = 0.0001$ ) and mean nodularity percentage ( $P = 0.0016$ ). Agreement between observed METAVIR F (by pathologists) and predicted F stage (by automated morphometry) was very good according to the weighted kappa index: 0.868 (0.844–0.891). Thus, there was no significant discordance (difference  $\geq 2$  F stages) between observed METAVIR F and predicted F stage. The rate of correctly classified patients for F

**Table 2: Main characteristics of population**

	Population number 1	Population number 2
Patients (n)	549	285
Age (year)	51.4±11.3	50.7±11.0
Sex (% male)	61.7	60.3
Liver specimen digitized length (mm)	25.2±7.8	28.0±7.1
Adequate liver specimen (n) <sup>b</sup>	416	284
METAVIR		
F0 (%)	4.6	6.0
F1 (%)	40.6	45.3
F2 (%)	27.5	21.1
F3 (%)	16.9	12.3
F4 (%)	10.4	15.4
F-score	1.88±1.08	1.86±1.19

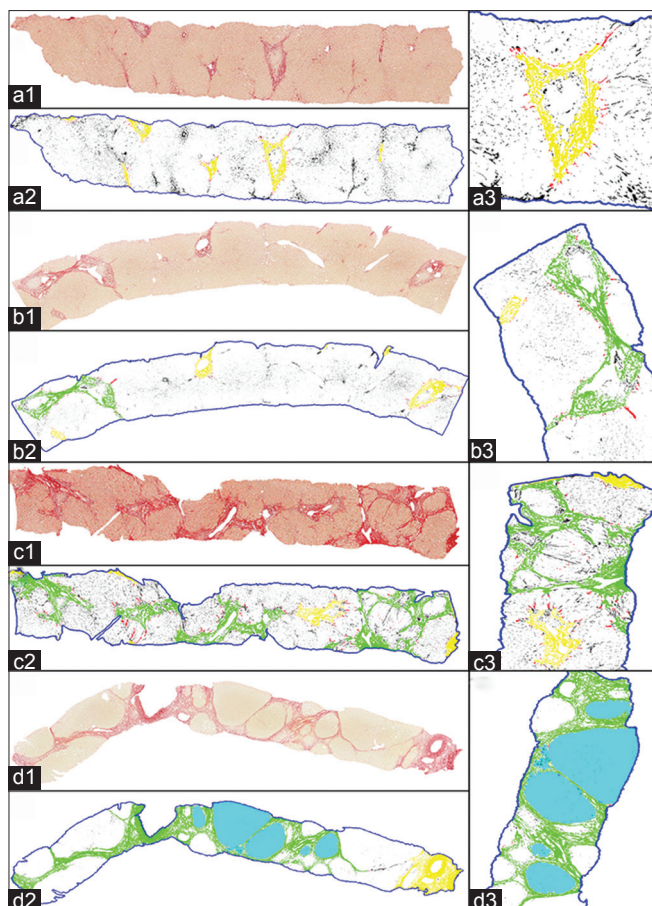
<sup>b</sup>Liver biopsy length  $\geq 20$  mm or cirrhosis

**Table 3: Diagnostic accuracy for significant fibrosis and cirrhosis by respective morphometric scores**

Population	Liver specimen		Significant fibrosis		Cirrhosis	
	Group	n	AUROC	Correctly classified (%)	AUROC	Correctly classified (%)
Number 1	All <sup>a</sup>	549	0.935	84.9	0.991	96.0
	Adequate <sup>b</sup>	416	0.957	87.3	0.994	96.9
	<20 mm	153	0.893	81.0	0.993	94.8
Number 2	All <sup>c</sup>	285	0.955	86.7	0.994	97.2

<sup>a</sup>Derivation population; the sum of adequate and <20 mm liver specimen is >100% since all cirrhosis were included in the former; <sup>b</sup>Adequate indicates a digitized length  $\geq 20$  mm or cirrhosis (corresponds to derivation population number 1); <sup>c</sup>Calculations were not performed in subgroups due to sample size (21 patients with inadequate liver specimen). AUROC: Area under the receiver operating characteristic





**Figure 5: Examples of digital sections of Metavir stages: F1 (a), F2 (b), F3 (c), and F4 (d). First, liver specimens stained with picosirius red (panels 1) then complete automated morphometric analysis (panels 2 and 3). They show the final images with the five morphometric descriptors of the morphometric significant fibrosis score: (a2) edge linearity percentage in dark blue pixels, stellar fibrosis in red pixels, porto-septal fibrosis in yellow pixels, perisinusoidal fibrosis in black pixels; (b2) one fibrous bridge between portal tracts is highlighted in green pixels; (c2) porto-septal fibrosis with many bridges in green pixels; (d2) nodules surrounded by >80% of fibrosis are highlighted in cyan pixels. Higher magnification of morphometric details of the four Metavir stages (a3, b3, c3 and d3)**

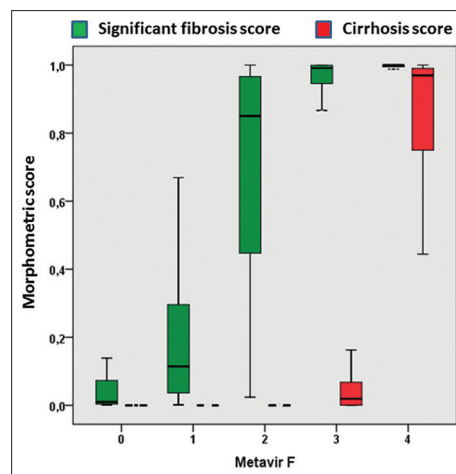
stage diagnosis was, overall: 68.5% (64.0–72.3); and per METAVIR stage: F0: 40.0%, F1: 74.7%, F2: 61.7%, F3: 55.9%, F4: 82.0%.

## Validation of Morphometric Descriptors and Scores

### Morphometric descriptors

#### Correlations

The metrical definition of the new morphometric descriptors was validated by a good correlation with METAVIR F, quantitative classical measures of fibrosis (whole and porto-septal fibrosis area) and noninvasive liver fibrosis tests, especially those quite independent of liver pathology (liver stiffness, hyaluronate, alpha2-macroglobulin) [details in supplement, Table S1]. In addition, the metrical definition of porto-septal fibrosis area was validated by a higher correlation ( $P < 0.0001$ )



**Figure 6: Relationship between morphometric scores and METAVIR F staging. Population number 1 (549 patients). Box plots (median, interquartile range and extremes) of morphometric scores against METAVIR F staging**

with METAVIR F stages ( $r_s$ : 0.756) than did whole area of fibrosis ( $r_s$ : 0.598) [Table 4].

#### Progression

The 13 independent morphometric descriptors for the 3 previous diagnostic targets showed a significant and regular progression (increase or decrease) as a function of METAVIR F stages [Figure 7] except for portal area of stellar fibrosis [Table 5].

#### Diagnostic accuracy

The metrical definition of other morphometric descriptors was indirectly validated by the very significant increase in diagnostic accuracy of new morphometric scores versus classical morphometric descriptors [Table 6].

### Morphometric scores

#### Correlations

Fibrosis morphometric scores were validated by high correlations with other liver fibrosis descriptors in the derivation population, especially the histological references: Area of porto-septal fibrosis (reference independent from their construction) and METAVIR F stage [Table 4]. Morphometric significant fibrosis score was the score the most correlated with area of porto-septal fibrosis ( $r_s$ : 0.835) and METAVIR F ( $r_s$ : 0.851).

#### Comparison with METAVIR F

Significant fibrosis score had a higher correlation with area of porto-septal fibrosis ( $r_s$ : 0.835) than METAVIR F ( $r_s$ : 0.756,  $P < 0.001$ ). Significant fibrosis score allowed morphometry to reach the same high correlations with noninvasive fibrosis tests than did METAVIR F, e.g.  $r_s$  with serum hyaluronate: 0.484 versus 0.476 ( $P = 0.862$ ), respectively, [Table S1] whereas the correlations of other morphometric descriptors were significantly different from those of METAVIR F.

**Table 4: Correlations (Spearman's rho) between fibrosis descriptors. Derivation population number I (416 patients)**

	Whole fibrosis area	Porto-septal fibrosis area	Observed F stage	Predicted F stage	Significant fibrosis score	Cirrhosis score
Whole fibrosis area	-	0.797	0.598	0.641	0.662	0.424
Porto-septal fibrosis area	0.797	-	0.756	0.793	0.835	0.723
Observed F stage	0.598	0.756	-	0.841	0.851	0.697
Predicted F stage	0.641	0.793	0.841	-	0.889	0.748
Significant fibrosis score	0.662	0.835	0.851	0.889	-	0.692
Cirrhosis score	0.424	0.723	0.697	0.748	0.692	-

**Table 5: Values (mean±SD) and correlation of the 13 independent morphometric descriptors as a function of METAVIR F stages. Derivation population number I (416 patients)**

	METAVIR F stages					P <sup>a</sup>	F <sub>A</sub> <sup>b</sup>	r <sub>s</sub> <sup>c</sup>
	0	1	2	3	4			
Fractal dimension of porto-septal fibrosis	1.599±0.101	1.604±0.086	1.634±0.085	1.651±0.087	1.622±0.096	0.003	4	0.178
Fractal dimension of perisinusoidal fibrosis	0.631±0.140	0.773±0.158	0.887±0.164	0.899±0.128	0.868±0.150	<0.001	21	0.359
Ratio of perisinusoidal fibrosis area (%)	69±14	68±14	60±15	46±15	31±15	<0.001	78	-0.582
Whole area of stellar fibrosis (%)	0.03±0.02	0.08±0.04	0.17±0.08	0.26±0.12	0.29±0.14	<0.001	97	0.759
Portal area of stellar fibrosis (%)	2.19±1.17	2.52±0.99	2.82±0.94	2.50±0.97	1.48±0.55	<0.001	20	-0.185
Mean portal distance (µm)	1.3±0.6	0.9±0.4	0.6±0.2	0.5±0.2	0.4±0.1	<0.001	60	-0.667
Number of bridges	0.6±1.0	1.3±1.4	3.8±2.5	6.2±3.4	6.7±3.9	<0.001	78	0.692
Portal ratio of bridges (%)	0.06±0.03	0.07±0.02	0.10±0.03	0.13±0.04	0.27±0.28	<0.001	42	0.621
Mean bridge thickness (µm)	23±30	50±38	84±27	95±28	130±42	<0.001	75	0.655
Mean granularity percentage (%)	3±6	9±10	17±16	38±19	52±25	<0.001	106	0.651
Mean nodularity percentage (%)	0±0	9±20	24±28	45±25	63±13	<0.001	81	0.650
Fragmentation index (%)	6.9±8.9	7.5±13.2	8.3±13.9	9.7±12.4	23.1±30.0	<0.001	10	0.167
Edge linearity percentage (%)	70±7	67±9	64±9	58±8	51±9	<0.001	43	-0.497

<sup>a</sup>By ANOVA, <sup>b</sup>The F value of ANOVA (F<sub>A</sub>) reflects the relative weight of each descriptor for discrimination of METAVIR F stages, <sup>c</sup>Spearman's correlation between morphometric descriptor and METAVIR F stages. SD: Standard deviation

**Table 6: Comparison of diagnostic accuracy (AUROC) of new morphometric scores and classical morphometric descriptors for the two main diagnostic targets. Derivation population number I (416 patients)**

	Significant fibrosis			Cirrhosis		
	AUROC	P versus <sup>a</sup>		AUROC	P versus <sup>a</sup>	
		Whole fibrosis	Porto-septal fibrosis		Whole fibrosis	Porto-septal fibrosis
Morphometric scores	0.957	<0.0001	<0.0001	0.994	<0.0001	<0.0001
Area of whole fibrosis	0.817	-	<0.0001	0.806	-	<0.0001
Area of porto-septal fibrosis	0.893	<0.0001	-	0.903	<0.0001	-

<sup>a</sup>P by Delong test. AUROC: Area under the receiver operating characteristic

**Progression**

The plots of morphometric significant fibrosis and cirrhosis scores versus reference F stages showed a marked progression thus an excellent F stage discrimination by morphometric scores [Figure 6].

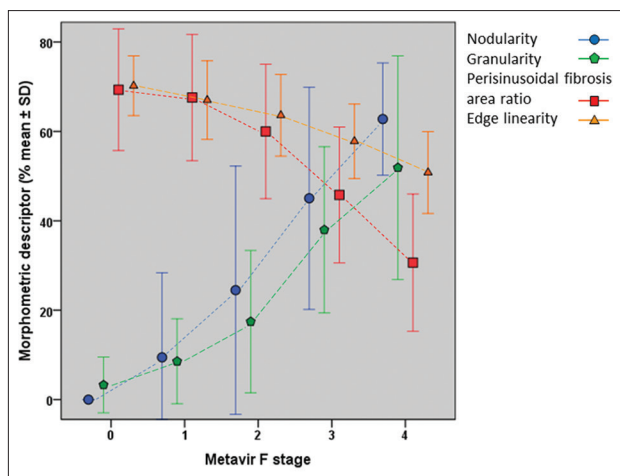
**Complementarity**

The scatter plot of significant fibrosis score versus cirrhosis score showed a perfect complementarity with

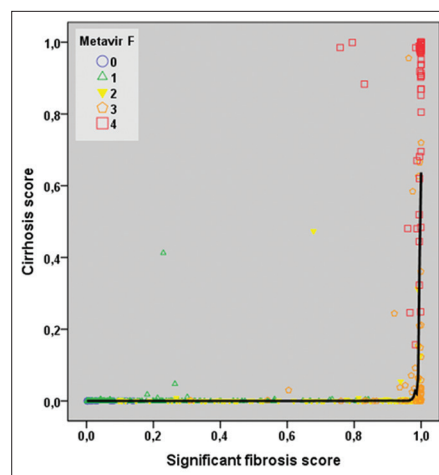
a progression of cirrhosis score from 0 to 1 only when significant fibrosis score reached 1 [Figure 8].

**Diagnostic accuracy**

Accuracies for the 3 morphometric diagnoses were validated in one different population (number 2). AUROCs of morphometric significant fibrosis and cirrhosis scores were, respectively: 0.955 and 0.994. Table 3 presents the overall results for significant fibrosis



**Figure 7: Relationship between main composite descriptors (with comparable % units) of morphometric scores and METAVIR F staging. Population number 1 (549 patients). Box plots (median, interquartile range and extremes) of composite descriptors against METAVIR F staging**



**Figure 8: Scatter plot of significant fibrosis score versus cirrhosis score as a function of METAVIR F stages. Bold line indicates the nonlinear regression by LOWESS. Derivation population number 1 (416 patients)**

and cirrhosis diagnosis. Considering the predicted F stage, there was globally an expected (optimism bias deletion), but small, decrease in accuracy. Thus, the overall rate of correctly classified patients was 66.3% for F stage diagnosis.

**Sensitivity analysis**

We analyzed the effect of liver biopsy length on diagnostic performance. AUROCs of morphometric significant fibrosis and cirrhosis scores were, respectively: 0.893 and 0.993 in 153 patients with biopsy <20 mm; the decrease in accuracy when liver biopsy length was <20 mm was thus minor [Table 3].

**Reproducibility and cleaning effect**

Intra-investigator reproducibility of morphometric scores and section cleaning were separately evaluated in 20 patients (details in Supplement). Intra-investigator reproducibility was excellent with intra-class correlation coefficient  $\geq 0.993$  for the three scores [Table S2]. Cleaning had no statistically significant effect on morphometric scores. However, at the individual level there was a clinically significant impact in about 1 out of 20 patients (details in Supplement).

**Clinical Applications**

**Morphometric descriptors**

The clinical interest (as well as metrical definition) of morphometric descriptors was evaluated by taking liver stiffness as an independent reference in the derivation population. By multivariate analysis including METAVIR F and activity, classical morphometry and the 13 main new independent descriptors [Table 1], liver stiffness was independently predicted (overall  $R^2$ : 0.502) by porto-septal fibrosis area ( $R^2$ : 0.349), METAVIR F ( $R^2$ : 0.069), fractal dimension of porto-septal fibrosis

( $R^2$ : 0.029), area of steatosis ( $R^2$ : 0.012), mean granularity percentage ( $R^2$ : 0.006), fragmentation index ( $R^2$ : 0.007), mean portal distance ( $R^2$ : 0.006), fractal dimension of perisinusoidal fibrosis ( $R^2$ : 0.007), portal area of stellar fibrosis ( $R^2$ : 0.012) and whole area of stellar fibrosis ( $R^2$ : 0.005). Thus, 86.3% of the information were brought by morphometry (among which 14.4% by new descriptors) and 13.7% by METAVIR F.

**Morphometric diagnoses**

In the validation population, the agreement between histological F reference (expert consensus) and the three morphometric diagnoses was equal or superior to the agreement between histological reference (expert consensus) and the histological diagnosis by first-line pathologists from tertiary centers, respectively, morphometry versus first-line pathologist: Significant fibrosis: 0.733 versus 0.733 ( $\kappa$ ), cirrhosis: 0.900 versus 0.827 ( $\kappa$ ) and METAVIR F staging: 0.881 versus 0.865 (weighted  $\kappa$ ).

**DISCUSSION**

**Originalities**

To date, morphometry has been used mainly to measure the surface (area) and subsequently the fractal dimension of liver lesions. Initially, morphometry was based on the manual determination of regions of interest<sup>[19]</sup> before some procedures became automated.<sup>[12]</sup> In the present study, we used a highly automated technique. Dioguardi *et al.* standardized metrical evaluation of the geometric properties of the parenchyma, inflammation, fibrosis, and alterations in liver tissue tectonics, but its application was not evaluated.<sup>[7]</sup>

The present work describes a method for automating the pathological staging of liver fibrosis in chronic

viral hepatitis C based on a histopathological classification frequently used by pathologists. Among the 44 morphometric descriptors, 13 had independent value for the three diagnostic targets (significant fibrosis, cirrhosis, F stage). Most were expected as they reflect patterns linked to METAVIR F staging; however, the most accurate and constant was a new morphometric descriptor, the edge linearity percentage, which underlines the importance of considering indirect signs of liver fibrosis, such as the external deformation of the liver biopsy specimen. Other descriptors were unexpected like those reflecting perisinusoidal fibrosis.

We used consensus expert diagnosis as a reference to construct morphometric scores for the binary diagnosis. The morphometric scores of significant fibrosis and cirrhosis only included new morphometric descriptors, which again emphasizes the originality of the present study. Moreover, the morphometric diagnosis of F stages was based on a different statistical method. This was the sole diagnosis to include several relatively classical descriptors, such as area or fractal dimension of fibrosis.

It should be noted that the binary morphometric diagnosis of significant fibrosis and cirrhosis was based on a numerical score from 0 to 1. These morphometric scores can be used as quantitative descriptors of the fibrosis degree intrinsically. Thus, morphometric significant fibrosis score had the highest correlations with other fibrosis descriptors, especially excellent correlations with area of porto-septal fibrosis or METAVIR F stages. With their precision and reproducibility, morphometric scores might be more sensitive in detecting changes than semi-quantitative scorings, from a statistical point of view. Thus, it has recently been suggested that morphometry of liver fibrosis, with a simple measurement such as area of fibrosis,<sup>[20]</sup> could be the best prognosis predictor. Moreover, liver morphometry might be a sensitive descriptor for predicting or monitoring liver fibrosis progression.<sup>[21,22]</sup> The present new morphometric scores and classical morphometry will need to be compared in these domains. Morphometric significant fibrosis and cirrhosis scores were very complementary [Figure 8]. This complementarity has recently been observed with corresponding blood tests for significant fibrosis and cirrhosis.<sup>[23]</sup> Thus, cirrhosis score can well describe the fibrosis progression within cirrhosis stage whereas significant fibrosis score is the basic score. Indeed, significant fibrosis score was the most accurate score with a significantly higher correlation with porto-septal fibrosis area than METAVIR F staging and the only morphometric diagnosis reaching high correlations with noninvasive fibrosis tests, not significantly different from METAVIR F staging correlations.

To summarize, based on expert consensus in a large, well-defined database, we created algorithms providing automated morphometric diagnoses reflecting that made

by expert pathologists. In this sense, this new technique may be considered as a “virtual expert pathologist”, providing reproducible and accurate automated diagnoses.

### Limits

Although we used a consensus expert reading, the main limit to our study is the reference based on classical histological fibrosis staging, which comprises two sources of variability: The observer and the specimen size/location. Specimen size has been evaluated in the present work, but location bias will be evaluated in further study. Liver biopsy should be considered as a best standard but not a gold standard<sup>[24]</sup> or an imperfect gold standard.<sup>[25]</sup> It has already been shown that blood tests, when constructed on liver biopsy, might offer better prognostication than their reference for construction.<sup>[26]</sup> Comparatively, it remains to be determined if morphometric scores provide better prognostication than histological staging.

Our material included liver biopsy with a pathological diagnosis of cirrhosis irrespective of the liver specimen length. Recently, sample size requirements for the digital analysis of cirrhosis was evaluated and appeared dependent of liver disease etiology.<sup>[27]</sup> Furthermore, no decompensated cirrhosis was included in our study. Morphometry of decompensated cirrhosis might improve cirrhosis prognostication<sup>[20]</sup> and should be compared to a prognostic histological score, e.g. the Laennec score,<sup>[28]</sup> several items of which were included in the present morphometric scores.

The main factors limiting the reliability of image analysis are thresholding, artefacts, staining quality and the representativeness of whole liver fibrosis. A classical limiting factor is thresholding for black-and-white binarization but, in our technique, thresholding was automated.<sup>[10,11]</sup> We observed that cleaning of artefacts had nonstatistically significant effect on morphometric scores, but they induced a large individual variation in about 1 out 20 patients. Since this technical prerequisite is useful, a further study aimed at the automated detection of artefacts is ongoing. The lack of standardization in staining techniques may be another limit. Color intensity parameters may reduce variability, but this needs further testing. In the present study, none of the color intensity parameters influenced diagnosis. This putative limit might be circumvented by nonstaining techniques.<sup>[29]</sup> Staining performed in other centers will be evaluated, and optionally our algorithms will be adapted. Finally, our results show that the present new morphometric scores are less sensitive to specimen length than area of fibrosis.<sup>[30]</sup>

Another limit is the interest of developing morphology compared to that of noninvasive tests, especially in chronic hepatitis C. First, it should be underlined that a weakness of classical histological fibrosis assessment is observer variability, furthermore present in all areas of diagnosis in



medicine like hepatology,<sup>[31]</sup> especially in clinical practice.<sup>[3]</sup> However, this inconvenience was only marginally present in the present automated morphometric diagnosis. Thus, the application of reliable and consistent imaging techniques might renew interest in liver biopsy utilization, particularly when issues of fibrosis and staging are paramount. Moreover, virtual liver biopsy<sup>[32,33]</sup> may become available in the near future, for which automated morphometry would be a natural partner.

Images of the area of portal or perisinusoidal fibrosis were quantitatively validated by experts.<sup>[11]</sup> Images of the new morphometric descriptors were validated qualitatively by experts and quantitatively by correlations with other fibrosis descriptors and accuracies for pathological diagnoses or liver stiffness. Other validation steps are needed: Other reference pathologists, relationship with other staging systems like Ishak's staging, other population and clinical outcomes.

## CONCLUSION

The morphometric descriptors presented here describe the full range of main lesions present in chronic viral hepatitis. The morphometric scores provide accurate qualitative diagnosis of significant fibrosis, cirrhosis and METAVIR fibrosis classification. They can be also used as quantitative descriptors useful for precise staging and follow-up by experienced physicians (e.g., pathologists, hepatologists) for diagnosis, prognosis and drug trials. The results of the present study open a door toward "virtual expert pathology." However, as fibrosis staging is only a part of liver biopsy evaluation, classical histopathology remains a cornerstone in liver diseases.

## ACKNOWLEDGMENTS

We thank all participants from the SNIFF 12 study:<sup>[8]</sup> Jérôme Boursier, Frédéric Oberti, Isabelle Hubert Fouchard; and from the Fibrostar study supported by ANRS:<sup>[9]</sup> Jean Pierre Zarski was the principal investigator with Nathalie Sturm as the principal pathologist. Kevin Erwin provided writing assistance for English proofreading.

## REFERENCES

1. Poynard T, Bedossa P, Opolon P. Natural history of liver fibrosis progression in patients with chronic hepatitis C. The OBSVIRC, METAVIR, CLINIVIR, and DOSVIRC groups. *Lancet* 1997;349:825-32.
2. 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.
3. Rousselet MC, Michalak S, Dupré F, Croué A, Bedossa P, Saint-André JP, et al. Sources of variability in histological scoring of chronic viral hepatitis. *Hepatology* 2005;41:257-64.
4. Boursier J, Bertrais S, Oberti F, Gallois Y, Fouchard-Hubert I, Rousselet MC, et al. Comparison of accuracy of fibrosis degree classifications by liver biopsy and non-invasive tests in chronic hepatitis C. *BMC Gastroenterol* 2011;11:132.
5. Germani G, Burroughs AK, Dhillon AP. The relationship between liver disease stage and liver fibrosis: A tangled web. *Histopathology* 2010;57:773-84.
6. Dioguardi N, Grizzi F, Bossi P, Roncalli M. Fractal and spectral dimension analysis of liver fibrosis in needle biopsy specimens. *Anal Quant Cytol Histol* 1999;21:262-6.
7. Dioguardi N, Grizzi F, Fiamengo B, Russo C. Metrically measuring liver biopsy: A chronic hepatitis B and C computer-aided morphologic description. *World J Gastroenterol* 2008;14:7335-44.
8. Boursier J, de Ledinghen V, Zarski JP, Fouchard-Hubert I, Gallois Y, Oberti F, et al. Comparison of eight diagnostic algorithms for liver fibrosis in hepatitis C: New algorithms are more precise and entirely noninvasive. *Hepatology* 2012;55:58-67.
9. Zarski JP, Sturm N, Guechot J, Paris A, Zafrani ES, Asselah T, et al. Comparison of nine blood tests and transient elastography for liver fibrosis in chronic hepatitis C: The ANRS HCEP-23 study. *J Hepatol* 2012;56:55-62.
10. Roullier V, Cavaro-Ménard C, Guillaume C, Aubé C. Fuzzy algorithms to extract vacuoles of steatosis on liver histological color images. *Conf Proc IEEE Eng Med Biol Soc* 2007;2007:5575-8.
11. Sandrini J, Boursier J, Chaigneau J, Sturm N, Zarski JP, Le Bail B, et al. Quantification of portal-bridging fibrosis area more accurately reflects fibrosis stage and liver stiffness than whole fibrosis or perisinusoidal fibrosis areas in chronic hepatitis C. *Mod Pathol* 2014;27:1035-45.
12. Boursier J, Chaigneau J, Roullier V, Lainé F, Sandrini J, Michalak S, et al. Steatosis degree, measured by morphometry, is linked to other liver lesions and metabolic syndrome components in patients with NAFLD. *Eur J Gastroenterol Hepatol* 2011;23:974-81.
13. Moal F, Chappard D, Wang J, Vuillemin E, Michalak-Provost S, Rousselet MC, et al. Fractal dimension can distinguish models and pharmacologic changes in liver fibrosis in rats. *Hepatology* 2002;36:840-9.
14. Castéra L, Vergniol J, Foucher J, Le Bail B, Chanteloup E, Haaser M, et al. Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. *Gastroenterology* 2005;128:343-50.
15. Leroy V, Halfon P, Bacq Y, Boursier J, Rousselet MC, Bourlière M, et al. Diagnostic accuracy, reproducibility and robustness of fibrosis blood tests in chronic hepatitis C: A meta-analysis with individual data. *Clin Biochem* 2008;41:1368-76.
16. Calès P, Boursier J, Ducancelle A, Oberti F, Hubert I, Hunault G, et al. Improved fibrosis staging by elastometry and blood test in chronic hepatitis C. *Liver Int* 2014;34:907-17.
17. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Clin Chem* 2003;49:7-18.
18. Team RDC. R: A Language and Environment for Statistical Computing; 2011. Available from: <http://www.R-project.org/>.
19. Pilette C, Rousselet MC, Bedossa P, Chappard D, Oberti F, Rifflet H, et al. Histopathological evaluation of liver fibrosis: Quantitative image analysis vs semi-quantitative scores. Comparison with serum markers. *J Hepatol* 1998;28:439-46.
20. Calvaruso V, Dhillon AP, Tsochatzis E, Manousou P, Grillo F, Germani G, et al. Liver collagen proportionate area predicts decompensation in patients with recurrent hepatitis C virus cirrhosis after liver transplantation. *J Gastroenterol Hepatol* 2012;27:1227-32.
21. Calès P, Zarski JP, Chaplain JM, Bertrais S, Sturm N, Michelet C, et al. Fibrosis progression under maintenance interferon in hepatitis C is better detected by blood test than liver morphometry. *J Viral Hepat* 2012;19:e143-53.
22. Manousou P, Burroughs AK, Tsochatzis E, Isgro G, Hall A, Green A, et al. Digital image analysis of collagen assessment of progression of fibrosis in recurrent HCV after liver transplantation. *J Hepatol* 2013;58:962-8.
23. Boursier J, Brochard C, Bertrais S, Michalak S, Gallois Y, Fouchard-Hubert I, et al. Combination of blood tests for significant fibrosis and cirrhosis improves the assessment of liver-prognosis in chronic hepatitis C. *Aliment Pharmacol Ther* 2014;40:178-88.
24. Bedossa P, Carrat F. Liver biopsy: The best, not the gold standard. *J Hepatol* 2009;50:1-3.
25. Poynard T, Ingiliz P, Elkrief L, Munteanu M, Lebray P, Morra R, et al. Concordance in a world without a gold standard: A new non-invasive

- methodology for improving accuracy of fibrosis markers. PLoS One 2008;3:e3857.
26. Ngo Y, Munteanu M, Messous D, Charlotte F, Imbert-Bismut F, Thabut D, et al. A prospective analysis of the prognostic value of biomarkers (FibroTest) in patients with chronic hepatitis C. Clin Chem 2006;52:1887-96.
  27. Hall AR, Tsochatzis E, Morris R, Burroughs AK, Dhillon AP. Sample size requirement for digital image analysis of collagen proportionate area in cirrhotic livers. Histopathology 2013;62:421-30.
  28. Kim SU, Oh HJ, Wanless IR, Lee S, Han KH, Park YN. The Laennec staging system for histological sub-classification of cirrhosis is useful for stratification of prognosis in patients with liver cirrhosis. J Hepatol 2012;57:556-63.
  29. Gailhouse L, Le Grand Y, Odin C, Guyader D, Turlin B, Ezan F, et al. Fibrillar collagen scoring by second harmonic microscopy: A new tool in the assessment of liver fibrosis. J Hepatol 2010;52:398-406.
  30. Bedossa P, Dargère D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. Hepatology 2003;38:1449-57.
  31. Winkfield B, Aubé C, Burtin P, Calès P. Inter-observer and intra-observer variability in hepatology. Eur J Gastroenterol Hepatol 2003;15:959-66.
  32. Campo-Ruiz V, Lauwers GY, Anderson RR, Delgado-Baeza E, González S. *In vivo* and *ex vivo* virtual biopsy of the liver with near-infrared, reflectance confocal microscopy. Mod Pathol 2005;18:290-300.
  33. Fabila D, de la Rosa JM, Stolik S, Moreno E, Suárez-Álvarez K, López-Navarrete G, et al. *In vivo* assessment of liver fibrosis using diffuse reflectance and fluorescence spectroscopy: A proof of concept. Photodiagnosis Photodyn Ther 2012;9:376-82.

# Supplement virtual liver fibrosis staging

## CLASSICAL MORPHOMETRY

### Area and fractal dimension of fibrosis and steatosis

The main principles were described in the main text and in previous papers.<sup>[1]</sup> All the algorithms were specifically developed as plug-ins in ImageJ.<sup>[2]</sup> First, a fuzzy generalized classification<sup>[3]</sup> process was used to merge pixel intensities into three classes (fibrosis, healthy tissue, white areas) using the minimization of an original energy function. Threshold determinations of the fibrosis class and the white class were automated. Then, a specifically-developed expert system<sup>[4]</sup> (+ patent WO2010 058295) was applied on the previously obtained white-labelled areas to extract steatosis vacuoles and eliminate small blood vessels and biliary tracts.

The expert system rules were based on the size of the regions (very small areas, considered as noise, were eliminated), the neighbouring regions (vessels surrounded by fibrosis were eliminated), the circularity of the region ( $4\pi \cdot \text{area}/\text{perimeter}^2$ ) (non-round regions, e.g., biliary tracts, were eliminated), the Hough transform of the region (allowing the detection and retention of vesicle aggregates), and a statistical texture parameter (heterogeneous regions, e.g., blood vessels, were eliminated).

The measure of the area of fibrosis (AOF) or steatosis (AOS) was equal to the ratio of pixels of fibrosis or steatosis divided by the number of pixels in the studied area:

$$\text{AOF} = \text{Pix}_{\text{FIB}}/\text{Pix}_{\text{MASK\_LB}} * 100$$

$$\text{AOS} = \text{Pix}_{\text{STEA}}/\text{Pix}_{\text{MASK\_LB}} * 100$$

where  $\text{Pix}_{\text{FIB}}$  is the number of fibrosis pixels,  $\text{Pix}_{\text{STEA}}$  is the number of steatosis pixels and  $\text{Pix}_{\text{MASK\_PBH}}$  is the number of pixels of the study area.

### Fractal dimension of fibrosis and steatosis

Briefly, a grid of square boxes (with  $\epsilon$  pixels as the side length) resembling a chessboard was superimposed over the histological image of threshold fibrosis. Boxes intersecting with collagen fibres were counted. Another chessboard grid was then used to cover the entire surface of the microscopic field. Thus, the total number ( $N$ ) of boxes of sides ( $\epsilon$ ) required to completely cover the collagen fibres reflected the perimeter examined with the scale ratio  $\epsilon$ . This step was repeated with  $\epsilon$  varying until a size of 14 pixels, and data were plotted on a log-log graph (*i.e.*  $\log [N]$  against  $\log[\epsilon]$ ). Relationships between points were measured by linear regression analysis using the least square method; the slope  $D$  of the regression line corresponded to the fractal dimension  $D$ . We performed the same measurement for the fractal

dimension of steatosis using the images showing steatosis. We called the fractal dimensions of fibrosis and steatosis respectively  $D_F$  and  $D_S$ .

## AREA AND FRACTAL DIMENSION OF PORTO-SEPTAL AND PERISINUSOIDAL FIBROSIS

The creation of this mask requires a large number of morphometric treatments (regarding erosions, dilations, the size of fibrosis areas, etc.) necessitating long computing times and large amounts of memory to process the data; this is why we reduced the image dimensions in order to work at a  $\times 5$  magnification with a  $2 \mu\text{m}/\text{pixel}$  resolution. We do not need high quality resolution for these steps.  $IM_{\text{COL}}$  identified the colour image of the liver biopsy specimen resized with a scale factor  $R_{\text{SCALE}} = 4$ . For example, an image with a size of  $20,000 \times 16,000$  pixels ( $\times 20$  magnification) used for measuring area and fractal dimension of fibrosis was reduced to a size of  $5,000 \times 4,000$  pixels ( $\times 5$  magnification,  $2 \mu\text{m}/\text{pixel}$ ,  $R_{\text{SCALE}} = 4$ ).  $IM_{\text{GREEN}}$  (the green component of  $IM_{\text{COL}}$ ) was thresholded by  $S_{\text{FIB}}$  in order to get a binary image of fibrosis ( $IM_{\text{FIB}}$ ). Perisinusoidal fibrosis lies between the rows of hepatocytes; therefore, it was sufficient to eliminate it by detecting the hepatocytes that had the intensity of healthy tissue. Concerning centrilobular veins, we determined that if their diameter was less than  $200 \mu\text{m}$  and if the surrounding fibrosis was small compared to their size, they would not be considered in the  $\text{MASK}_{\text{PORT}}$ . Dilatations and erosions were applied on  $IM_{\text{FIB}}$  in order to obtain compacted regions to study. Porto-septal regions present large amounts of fibrosis and therefore the elements of fibrosis with large areas were considered in the  $\text{MASK}_{\text{PORT}}$ . The fractal dimension of fibrosis into these regions was also a criterion to determine if it was included in  $\text{MASK}_{\text{PORT}}$ . Thin regions (small ratio between the perimeter of the area and its surface) with nodes (interconnection points in fibrosis filaments) are a sign of branched forms as observed in perisinusoidal fibrosis and thus were not considered as porto-septal fibrosis. Finally, we obtained a  $\text{MASK}_{\text{PORT}}$  that separated perisinusoidal fibrosis from porto-septal fibrosis. The lobular region was the region of the liver biopsy without porto-septal fibrosis.

We measured the area and the fractal dimension of porto-septal fibrosis ( $\text{AO}_{\text{FPORT}}$  and  $\text{DF}_{\text{FPORT}}$ ), the area and fractal dimension of perisinusoidal fibrosis ( $\text{AO}_{\text{FPS}}$  and  $\text{DF}_{\text{FPS}}$ ), the area of lobular perisinusoidal fibrosis ( $\text{AO}_{\text{FPS\_LOB}}$ ) and the ratio of perisinusoidal fibrosis ( $\text{RATIO}_{\text{FPS}}$ ) among the whole fibrosis as follows:

$$\begin{aligned} \text{AO}_{\text{FPORT}} &= \text{Pix}_{\text{PORT}}/\text{Pix}_{\text{MASK\_LB}} * 100 \\ \text{AO}_{\text{FPS}} &= \text{Pix}_{\text{FPS}}/\text{Pix}_{\text{MASK\_LB}} * 100 \end{aligned}$$

$$\begin{aligned} AO\_FPS\_LOB &= Pix_{FPS}/(Pix_{MASK\_LB} - Pix_{Mask\_Port}) * 100 \\ RATIO\_FPS &= Pix_{FPS}/Pix_{FIB\_TOT} * 100 \end{aligned}$$

where  $Pix_{PORT}$  is the number of pixels of porto-septal fibrosis,  $Pix_{FPS}$  is the number of pixels of perisinusoidal fibrosis,  $Pix_{MASK\_LB}$  is the number of pixels of the total area of the biopsy specimen,  $Pix_{Mask\_Port}$  is the number of pixels of the porto-septal area and  $Pix_{FIB\_TOT}$  is the number of pixels representing all the fibrosis (porto-septal and perisinusoidal fibrosis).

## NEW MORPHOMETRY

The image analysis software was the same as with classical morphometry but we developed new algorithms. These measurements are available upon request (PC).

### Edge linearity

First, we detected the edges of the fragments on the liver biopsy specimen and, using the mask ( $MASK_{EDGE}$ ), we combined two methods to optimize the detection of straight edges. Method 1 consisted in applying the Hough transform to detect straight lines on  $MASK_{EDGE}$ . The measurement of the Hough transform is well known in the domain of image processing to detect shapes<sup>[5]</sup>. This gave a mask called  $MASK_{HOUGH}$  containing only the edges of the mask detected as straight by the Hough transform. Method 2 consisted in creating a straight mask  $MASK_{RECT}$  from the edge mask ( $MASK_{EDGE}$ ). For this, we first detected the corners with a Harris detector<sup>[6]</sup> and then we kept the edge points separated by a sampling  $step = 2.4 mm$ . The lines between all these points were drawn and we finally obtained a theoretical straight mask.  $MASK_{RECT}$  represented the edges of  $MASK_{EDGE}$  that were in common with this theoretical straight mask.

By combining the two masks, we attained a  $MASK_{RECTCOMB}$  ( $MASK_{RECTCOMB} = MASK_{HOUGH} + MASK_{RECT}$ ).  $MASK_{RECTCOMB}$  thus contained all the straightest edges of the liver biopsy specimen. They determined the following formulas:

$$PCT\_RECT = Pix_{MaskRectComb}/Pix_{MaskEdge} * 100$$

where  $Pix_{MaskRectComb}$  is the number of pixels of  $MASK_{RECTCOMB}$  and  $Pix_{MaskEdge}$  is the number of pixels of  $MASK_{EDGE}$ .  
 $DF\_EDGE =$  Fractal dimension of the edge of the LB on  $MASK_{EDGE}$ .

### Liver specimen length

We used the mask obtained after the elimination of the artefacts ( $MASK_{LB}$ ) and we applied a morphometric operation (skeletonisation) that provides a skeleton of the biopsy. This was composed of a main skeleton that followed the shape of the biopsy and several small ramifications. The main skeleton had the advantage of being very representative of the length, especially for twisted biopsies. Then the small ramifications were removed in order to keep the main

skeleton ( $MASK_{SKELETON}$ ) representing LB length:

$$LB\_LENGTH = Pix_{LB\_SKELETON} * IM_{Resolution} * R_{SCALE}$$

where  $Pix_{LB\_SKELETON}$  is the number of pixels representing the liver biopsy specimen length on  $MASK_{LB}$ ,  $IM_{Resolution}$  is the resolution of the scanned image (0.5  $\mu m$ ) and  $R_{SCALE} = 4$  is the scale factor used to resize the image.

### Stellar fibrosis

To make this measurement, we detected fibrosis in a distance of 100  $\mu m$  around the porto-septal regions of  $MASK_{PORT}$ . In order to differentiate stellar fibrosis from peri-sinusoidal fibrosis and concentrated fibrosis in the portal tracts, we combined several morphometric operations (successive erosions and dilations) which enabled us to keep only the thin fibrils of fibrosis connected to the porto-septal regions. Stellar fibrosis was measured with the following formulas:

$$AOF\_STELLAR\_TOT = Pix_{Fib\_Stellar}/Pix_{MASK\_LB} * 100$$

$$AOF\_STELLAR\_EP = Pix_{Fib\_Stellar}/Pix_{Mask\_Port} * 100$$

$$AOF\_STELLAR\_LOB = Pix_{Fib\_Stellar}/(Pix_{MASK\_LB} - Pix_{Mask\_Port}) * 100$$

$$MEAN\_STELLAR\_PORT = Pix_{Fib\_Stellar}/NB\_PORT$$

$$MEAN\_AO\_PORT = Pix_{Mask\_Port}/NB\_PORT$$

where  $Pix_{Fib\_Stellar}$  is the number of pixels detected as stellar fibrosis,  $Pix_{MASK\_LB}$  is the number of pixels of the total area on the liver biopsy specimen,  $Pix_{Mask\_Port}$  is the number of pixels of the porto-septal area in  $MASK_{PORT}$  and  $NB\_PORT$  is the number of porto-septal regions in  $MASK_{PORT}$ .

### Bridging fibrosis

For each porto-septal region in  $MASK_{PORT}$ , we applied morphometric operations such as high erosion followed by a small dilation. The aim was to observe a separation of the structures, which determines the presence of bridges. At the end of the morphometric operations, a studied area with at least two elements was considered as a bridge and added in the  $MASK_{BRIDGE}$ . We obtained the following parameters:

$$\begin{aligned} NB\_BRIDGE & \text{ is the number of bridges in } MASK_{BRIDGE} \\ RATIO\_BRIDGE & = Pix_{Mask\_Bridge}/Pix_{Mask\_Port} * 100 \end{aligned}$$

$$AOF\_BRIDGE = Pix_{Fib\_Bridge}/Pix_{MASK\_LB} * 100$$

where  $Pix_{Mask\_Port}$  is the number of pixels in  $MASK_{PORT}$ ,  $Pix_{Mask\_Bridge}$  is the number of pixels in  $MASK_{BRIDGE}$ ,  $Pix_{Fib\_Bridge}$  is the number pixels of fibrosis in the bridges and  $Pix_{MASK\_LB}$  is the number of pixels of the total area on the liver biopsy specimen.

$$MEAN\_THICK\_BRIDGE = MEAN\_SURF\_BRIDGE / MEAN\_PERIM\_BRIDGE * 100$$



where  $MEAN\_SURF\_BRIDGE$  is the mean of pixels representing the surface of bridges and  $MEAN\_PERIM\_BRIDGE$  is the mean of pixels representing the perimeter of bridges.

### Granularity

First, we counted the number of fragments ( $NB\_FRAG$ ) in the mask obtained after the elimination of artefacts ( $MASK_{LB}$ ). Then, we used the porto-septal mask ( $MASK_{PORT}$ ) in which we applied several dilatations to extend the porto-septal areas. We subtracted  $MASK_{PORT}$  from  $MASK_{LB}$  to only observe the granules formed by breaking the fragments.  $Nb\_Granules$  was the number of granules.  $PCT\_GRANULARITY$  was the ratio between the number of fragments without destructure and the number of granules obtained in these fragments after destructure by porto-septal areas:

$$PCT\_GRANULARITY = 100 - (NB\_FRAG/Nb\_Granules * 100)$$

### Fragmentation

As for measuring  $PCT\_GRANULARITY$ , we used  $MASK_{LB}$ . On this mask, we detected the small fragments in order to obtain  $MASK_{FRAG\_SMALL}$ .

Small fragments were those with a surface under two mm<sup>2</sup> or those with a surface under three mm<sup>2</sup> but with circularity up to 0.7. The fragmentation index was the ratio between the surface of small fragments detected and the total surface of the liver biopsy specimen:

$$INDEX\_FRAGMENTATION = \frac{Pix\_Mask\_Frag\_Small}{Pix\_Mask\_Frag} * 100$$

Where  $Pix\_Mask\_Frag$  is the number of pixels in  $MASK_{LB}$  and  $Pix\_Mask\_Frag\_Small$  the number of pixel in  $MASK_{FRAG\_SMALL}$ .

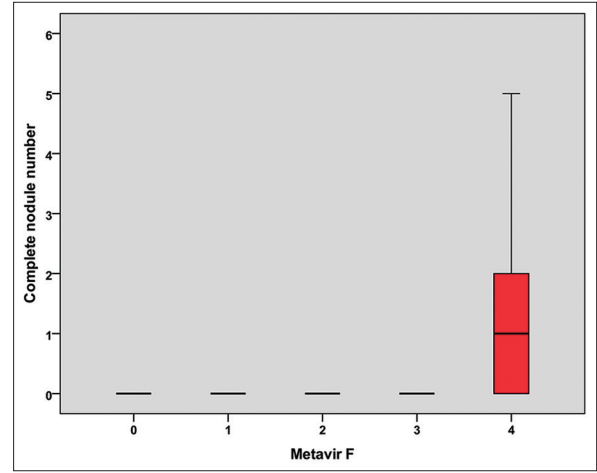
### Nodules

The process was the same as the measure of the  $PCT\_GRANULARITY$  applying  $MASK_{PORT}$  on  $MASK_{LB}$  to study the granules obtained. We only kept the granules if they were circular. The nodules were not always perfectly round, which is why we chose a circularity threshold of 0.45. Among these round granules, we kept those that had at least 30% of fibrosis around them (30% of the external border). We finally obtained a mask with regions tending towards nodularity ( $MASK_{NOD}$ ). Nodules were considered as definitive (or complete nodules) when  $PCT\_NODULARITY$  was  $\geq 80\%$ . We used the following formulas:

$$PCT\_NOD = \text{mean of percentage of fibrosis around areas in } MASK_{NOD}$$

$$NB\_NOD = \text{number of nodules in } MASK_{NOD} \text{ with more than 80\% of fibrosis around.}$$

One can note that the mean nodularity percentage was  $63 \pm 13\%$  in cirrhosis reflecting a substantial proportion of incomplete nodules [Table 5].



**Figure S1: Distribution of nodule number as a function of METAVIR F stages; box plots: Median, interquartile range, extremes**

The following Figure S1 shows that complete nodules, as we defined, were only observed in cirrhosis.

### Portal distance

$NB\_FRAG$  was the number of fragments in  $MASK_{LB}$ .  $Nb\_EP_n$  was the number of porto-septal regions present on the fragment  $n$  ( $n$  lying between 1 and  $NB\_FRAG$ ). For each  $n$  fragment, we measured the minimum distance  $Dmin_n$  between all porto-septal regions present on  $MASK_{PORT}$ .  $Dmoy_n$  was the average distance between regions on the porto-septal fragment  $n$ .

$$Dmoy_n = Dmin_n / (Nb\_EP_n - 1).$$

The average distance between porto-septal regions for all fragments was called  $DIST\_EP\_MEAN$ :

$$DIST\_EP\_MEAN = (Dmoy_1 + Dmoy_2 + \dots + Dmoy_{NB\_FRAG}) / NB\_FRAG.$$

## STAINING QUALITY

The performance of our measures depended on the quality of the staining (colouration) of the biopsy. Indeed, with a pale colour of fibrosis, the detection could miss some porto-septal regions, and result in an underestimation of the lesion diagnosis. Usually, pathologists exclude cases with poor colouration. We thus decided to automatically detect liver biopsy with poor colouration to exclude them. All the measures of color intensity were applied on the three components of the image (red, green and blue: R, G, and B respectively). The color intensity of the biopsy was calculated by averaging all pixel intensities and we obtained a mean intensity for each component: ILbR (the mean intensity of the biopsy on the red component), ILbG and ILbB. We did the same for the intensity of fibrosis (IfibR, IfibG, and IfibB) and for the parenchyma (IparenchymaR, IparenchymaG, and IparenchymaB).

Quality may also be bad when the intensity of the fibrosis is close to that of the parenchyma or the liver biopsy. We thus developed a measure of the contrast between fibrosis and background as follow:

$$\text{Contrast\_Fib\_Parenchyma} = \frac{\sqrt{(\text{IparenchymaR} - \text{IfibR})^2 + (\text{IparenchymaG} - \text{IfibG})^2}}{\sqrt{(\text{IparenchymaB} - \text{IfibB})^2}}$$

$$\text{Contrast\_Fib\_Lb} = \frac{\sqrt{(\text{ILbR} - \text{IfibR})^2 + (\text{ILbG} - \text{IfibG})^2 + (\text{ILbB} - \text{IfibB})^2}}$$

## VALIDATION OF MORPHOMETRIC MEASURES

Correlations between main pathological descriptors and non-invasive liver fibrosis tests are shown in Table S1. New morphometric descriptors were validated by good correlation with METAVIR F, quantitative pathological measures of fibrosis (whole fibrosis area and porto-septal fibrosis area) and non-invasive liver fibrosis tests, especially those quite independent of liver pathology

(liver stiffness, hyaluronate, alpha2-macroglobulin). Only two among the 13 main descriptors had globally poor correlations (fragmentation index and fractal dimension of porto-septal fibrosis).

It should be noted that METAVIR had higher correlations with non-invasive liver fibrosis tests, especially those quite independent of liver pathology, than did all single morphometric descriptors, confirming the reference status of this staging. Nevertheless, the *whole area of stellar fibrosis* had correlations with non-invasive liver fibrosis tests very close to those of METAVIR F. Finally, correlations of *significant fibrosis score* with whole area of fibrosis or porto-septal area of fibrosis were significantly higher than those of METAVIR F ( $P = 0.011$  and  $P < 0.001$ , respectively, by Steigert test) whereas correlations with non-invasive fibrosis test were not significantly different.

## REPRODUCIBILITY

Unchanged liver specimen slides from 20 patients were evaluated for morphometry 12 months apart by the same investigator without knowledge of the first evaluation. Patients had chronic hepatitis C and each

**Table S1: Correlations (Spearman's rho) between main pathological descriptors and non-invasive liver fibrosis tests. Bold figures:  $P < 0.001$ , italic bold figures:  $0.001 \leq P < 0.05$ ; figure underlined: Highest correlation among single pathological descriptors. Derivation population #1 (549 patients)**

	WFA	PSFA	F	FT	FM	FS	EFM	HA	A2M
Whole fibrosis area (WFA)	-	0.786	0.572	0.289	0.368	0.444	0.452	0.317	0.271
Porto-septal fibrosis area (PSFA)	0.786	-	0.719	0.436	0.519	0.553	0.586	0.369	0.398
METAVIR F (F)	0.572	0.719	-	0.541	0.621	0.587	0.680	0.476	0.496
Fibrotest (FT)	0.289	0.436	0.541	-	0.812	0.408	0.770	0.524	0.756
FibroMeter <sup>v2G</sup> (FM)	0.368	0.519	0.621	0.812	-	0.474	0.880	0.650	0.763
Fibroscan (FS)	0.444	0.553	0.587	0.408	0.474	-	0.750	0.436	0.348
Elasto-FibroMeter <sup>2G</sup> (EFM)	0.452	0.586	0.680	0.770	0.880	0.750	-	0.616	0.782
Hyaluronic acid (HA)	0.317	0.369	0.476	0.524	0.650	0.436	0.616	-	0.433
Alpha2-macroglobulin (A2M)	0.271	0.398	0.496	0.756	0.763	0.348	0.782	0.433	-
Edge linearity percentage	-0.192	-0.282	-0.452	-0.317	-0.319	-0.256	-0.338	-0.288	-0.248
Whole area of stellar fibrosis	0.710	0.807	0.739	0.491	0.581	0.562	0.625	0.439	0.450
Portal area of stellar fibrosis	0.025	-0.277	-0.115	-0.036	-0.052	-0.164	-0.109	-0.018	-0.032
Mean portal distance	-0.517	-0.619	-0.637	-0.400	-0.448	-0.407	-0.486	-0.342	-0.339
Number of bridges	0.494	0.642	0.645	0.401	0.471	0.450	0.524	0.378	0.381
Portal ratio of bridges	0.340	0.496	0.560	0.362	0.399	0.384	0.438	0.316	0.314
Mean bridge thickness	0.417	0.634	0.592	0.372	0.430	0.432	0.467	0.345	0.324
Mean granularity percentage	0.463	0.666	0.619	0.383	0.421	0.452	0.473	0.318	0.347
Mean nodularity percentage	0.418	0.599	0.598	0.393	0.428	0.454	0.486	0.338	0.327
Fragmentation index	0.012	0.061	0.157	0.088	0.078	0.039	0.068	0.127	0.060
FD of porto-septal fibrosis	0.648	0.507	0.146	0.048	0.063	0.099	0.110	0.014	0.044
FD of perisinusoidal fibrosis	0.883	0.488	0.348	0.145	0.222	0.284	0.286	0.237	0.150
Ratio of perisinusoidal fibrosis area	-0.136	-0.676	-0.529	-0.384	-0.427	-0.415	-0.437	-0.260	-0.323
Significant fibrosis score	0.645	0.813	0.819	0.533	0.611	0.579	0.662	0.484	0.481
Cirrhosis score	0.401	0.692	0.649	0.397	0.439	0.477	0.508	0.344	0.345

FD: Fractal dimension

**Table S2: Intra-investigator reproducibility of morphometric scores in 20 patients with CHC**

Score	Measure		Difference (p) <sup>a</sup>	Correlation (r <sub>p</sub> ) <sup>b</sup>	Agreement (r <sub>ic</sub> ) <sup>c</sup>
	1 <sup>st</sup>	2 <sup>nd</sup>			
Significant fibrosis	0.539±0.455	0.562±0.441	0.082	0.992	0.996
Cirrhosis	0.188±0.380	0.191±0.368	0.749	0.995	0.997
Predicted F	2.05±1.32	2.00±1.34	0.330	0.986	0.993

<sup>a</sup>By Student t test, <sup>b</sup>By Pearson coefficient, <sup>c</sup>Intra-class correlation coefficient

METAVIR F stage included four patients. Agreement was excellent [Table S2].

## CLEANING IMPACT

The impact of manual cleaning was tested in the same 20 patients with equally distributed METAVIR F stages. The *significant fibrosis score* was  $0.615 \pm 0.399$  before cleaning and  $0.539 \pm 0.455$  after cleaning with a difference of  $0.076 \pm 0.106$  ( $P = 0.005$ ) with  $r_p: 0.978$  ( $r_s: 0.953$ ) and intra-class correlation coefficient: 0.985. AUROCs for significant fibrosis were 1 and 0.990, respectively. The *cirrhosis score* was  $0.219 \pm 0.392$  before cleaning and  $0.188 \pm 0.380$  after cleaning with a difference of  $0.030 \pm 0.188$  ( $P = 0.476$ ) with  $r_p: 0.882$  ( $r_s: 0.943$ ) and intra-class correlation coefficient: 0.937. AUROCs for cirrhosis were 0.969 and 1, respectively.

Thus, cleaning had a significant impact on morphometric scores which was due to large changes in a small patient proportion (one out 20 patients).

## REFERENCES

1. Moal F, Chappard D, Wang J, Vuillemin E, Michalak-Provost S, Rousselet MC, et al. Fractal dimension can distinguish models and pharmacologic changes in liver fibrosis in rats. *Hepatology* 2002;36:840-849.
2. ImageJ. <http://rsbweb.nih.gov/ij/>.
3. Ménard M, Courboulay V, Dardignac P. Possibilistic fuzzy clustering: unification within the framework of the non-extensive thermostatistics. *Pattern Recognition* 2003;36:1325-1342.
4. Roullier V, Cavaro-Menard C, Guillaume C, Aube C. Fuzzy algorithms to extract vacuoles of steatosis on liver histological color images. *Conf Proc IEEE Eng Med Biol Soc* 2007;2007:5575-5578.
5. Duda RO, Hart PE. Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Comm ACM* 1972;15:11-15.
6. Harris C, Stephens M. A combined corner and edge detector. *Proceedings of the 4<sup>th</sup> Alvey Vision Conference* 1988:147-151.