An Algorithm for the Grading of Activity in Chronic Hepatitis C

PIERRE BEDOSSA AND THIERRY POYNARD FOR THE METAVIR COOPERATIVE STUDY GROUP

Histological activity reflects the global assessment of basic necroinflammatory lesions and is a criterion of major importance in chronic hepatitis C. The aim of this study was to propose and test the accuracy of a simple algorithm that generates a single activity score based on basic pathological features. A panel of 10 pathologists reviewed 363 chronic hepatitis C liver biopsies and graded the activity of hepatitis according to their own experience (reference activity). Then, a consensual algorithm on the grading of activity was established by the 10 experts in a panel discussion. Finally, stepwise discriminant analysis was performed to define which basic features had been intuitively used in the reference activity (statistical activity). To test the accuracy of the algorithm, concordance between the activity defined by the algorithm and the reference activity was assessed. It was compared with concordance between the activity defined by the statistical model and the reference activity. The algorithm proposed by the panel for the grading of activity included piecemeal necrosis and lobular necrosis. Concordance between reference activity and activity defined by the algorithm was substantial (305 cases, 84%, κ = .75). Discriminant analysis showed that piecemeal necrosis, lobular necrosis, and portal inflammation were independently used to grade the activity. Concordance between reference activity and activity defined by the statistical model was substantial (300 cases, 83%, κ = .73), virtually identical to the concordance between reference activity and activity defined by algorithm. This study proposes a simple algorithm for the grading of activity in chronic hepatitis. Its accuracy is as high as that obtained using a statistical approach. (HEPATOLOGY 1996;24:289-293.)

Liver biopsy is a gold standard in chronic hepatitis and is the only investigation that can estimate the severity of tissue damage, because clinicopathological correlations are poor in chronic hepatitis C. However, the value of a liver biopsy depends on the possibility of assessing histological features using a standardized reproducible classification.

With the discovery of the hepatitis C virus and its worldwide diffusion, pathologists have been faced with an increasing number of liver biopsies showing chronic hepatitis.[1-3] They have been tempted to discuss the terminology and the histological classification of chronic hepatitis, and several propositions have recently been made by different groups of pathologists.[4-7] A consensual proposition was to separately assess, using scoring systems, the degree of liver fibrosis and that of activity.[4,6] Liver fibrosis is simple to precisely define and thus to reproducibly score.[8] In contrast, the grading of activity, which integrates the different basic necroinflammatory lesions, is more difficult to assess; morever, it is of importance in the decision of whether or not to treat patients in many therapeutic trials. Although not proven, inflammation and liver cell necrosis are the hallmarks of active disease that may predict evolution toward fibrosis and cirrhosis.[9]

Several approaches have been proposed to assess histological activity. One of them is the semiquantitative scoring system of Knodell et al., in which fibrosis and portal, periportal, and lobular necrotic and inflammatory components are assessed separately and their coding values added.[10] The global score appears accurate because it varies over large ranges, but its value may be limited by poor reproducibility, because each feature has its own observer variation.[8] Another possible approach is to consider that periportal and intralobular necroinflammatory lesions are related to the same pathogenic mechanism and that they must be globally assessed. As previously reported, assessment of activity must be separated from staging of fibrosis.[4]

The aim of this study was to propose a simple algorithm that generates a single score of necroinflammatory activity based on basic histological features. The creation of this algorithm drew on the experience of a panel of pathologists, and its validation was performed by comparison with another method for establishing activity based on a statistical approach.

PATIENTS AND METHODS

Liver Biopsies. A total of 363 percutaneous liver biopsies from patients with chronic hepatitis C were included in this study. All patients were positive for antibody to hepatitis C by second-generation tests and had at least a 6-month increase in aminotransferase levels. All biopsies were from patients included in five French therapeutic trials on chronic hepatitis C, the results of which have already been published.[11-15] Biopsy was performed at the time of inclusion, when patients had not received any treatment. Liver biopsies, measuring more than 10 mm in length, were fixed, paraffin-embedded, and stained with hematoxylin-eosin safran and Masson's trichrome, or Picrosirius red for collagen. Biopsies were anonymously coded. Pathologists were not aware of clinical and biochemical data, but only of the presence of chronic hepatitis C.

Panel of Pathologists. A panel of 10 French senior pathologists experienced in the field of liver pathology was established. Pathologists had already worked together in the French METAVIR group,[8,16] and a consensus was thus reached on the definition and coding of the lesions. Intra- and interobserver variations in the assessment of histological lesions in chronic hepatitis C had already been published.[8]

Questionnaire. Histological analysis of the 363 liver biopsies was performed using a questionnaire. It included 27 semiquantitative histological items.[16] All were recorded, but, for the aim of this study, we focused on the following features:

1. Focal lobular necrosis (0 = less than one necroinflammatory foci per lobule, 1 = at least one necroinflammatory foci per lobule, 2

The METAVIR group includes Pierre Bedossa (Le Kremlin-Bicêtre), Paulette Bioulac-Sage (Bordeaux), Patrice Callard (Paris), Michèle Chevallier (Lyon), Claude Degott (Clichy), Yves Deugnier (Rennes), Monique Fabre (Le Kremlin-Bicêtre), Michel Reynès (Villejuif), Jean-Jacques Voigt (Toulouse), Elie Serge Zafrani (Créteil, France).

From the Service d'Anatomie Pathologique, Hôpital de Bicêtre, Le Kremlin-Bicêtre, France.

Received November 1, 1995; accepted March 26, 1996.

Supported in part by grants from Fonds de Recherche de la Société Nationale Française de Gastroentérologie and sponsored by Produit Roche and Schering-Plough.

Address reprint requests to: Pierre Bedossa, M.D., Laboratoire d'Anatomie pathologique, Hôpital de Bicêtre, 78 rue général Leclerc, 94275 Le Kremlin-Bicêtre Cedex, France.

0270-9139/96/2402-0001$3.00/0

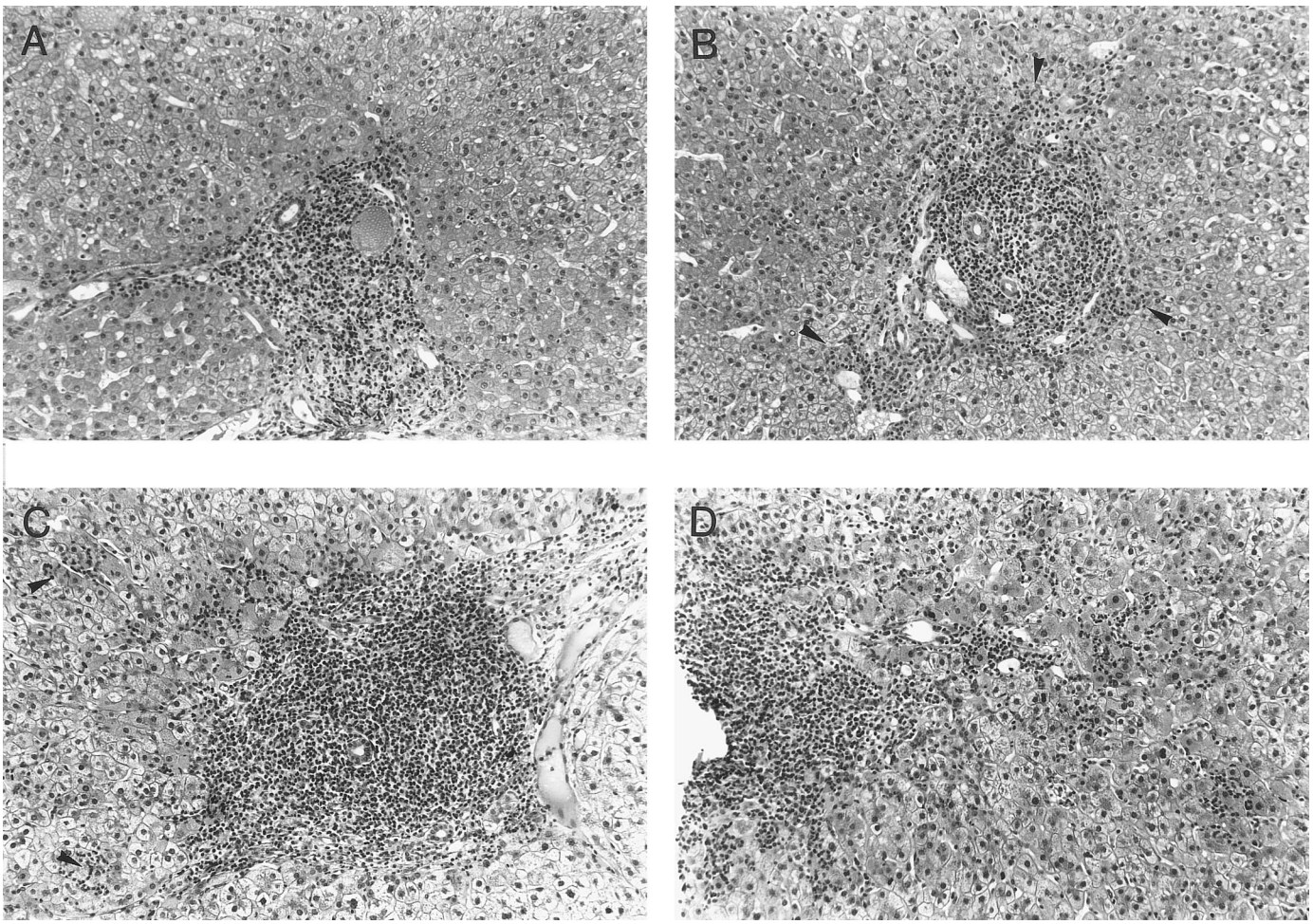

FIG. 1. (A) Chronic hepatitis C with no activity (A0). Portal tract with inflammation but no piecemeal nor lobular necrosis. (B) Chronic hepatitis with mild activity (A1). Piecemeal necrosis (arrowhead) is graded 1 (focal alteration of the periportal plate), and lobular necrosis is graded 0. According to the algorithm, activity is mild. (C) Chronic hepatitis with moderate activity (A2). Piecemeal necrosis is more diffuse (grade 2), and there are foci of lobular necrosis (arrow) (grade 1). (D) Chronic hepatitis with severe activity (A3). Piecemeal necrosis is graded 2, and there are several necroinflammatory foci per lobule (grade 2). According to the algorithm, activity is severe.

= several necroinflammatory foci per lobule or confluent or bridging necrosis).

2. Portal inflammation (0 = absent, 1 = presence of mononuclear aggregates in some portal tracts, 2 = mononuclear aggregates in all portal tracts, 3 = large and dense mononuclear aggregates in all portal tracts).

3. Piecemeal necrosis (0 = absent, 1 = focal alteration of the periportal plate in some portal tracts, 2 = diffuse alteration of the periportal plate in some portal tracts or focal lesions around all portal tracts, 3 = diffuse alteration of the periportal plate in all portal tracts).

4. Bridging necrosis (0 = absent, 1 = present) was also recorded.

The document ended with an item assessing the intensity of necroinflammatory lesions (histological activity). This was indicated as follows: A0 = no histological activity, A1 = mild activity, A2 = moderate activity, and A3 = severe activity. This was related to the classical notions of chronic hepatitis and to the experience of the persons performing the assessment. No precise indications were given to the pathologists concerning the evaluation of this item in correlation with the basic lesions reported before. Representative microphotographs illustrating the different grades of activity are shown in Fig. 1. For each biopsy, the activity evaluated by pathologists in this way was referred to as the "reference activity" and was the gold standard.

According to this procedure, the 363 liver biopsies were simultaneously read by pairs of pathologists who filled out the coding form for each biopsy. Reading sessions were performed over three 1-day periods to minimize observer variations.

Creation and Validation of an Algorithm Based on an Expert Panel Decision. After histological study of the 363 biopsies, a meeting session was organized between the 10 pathologists to discuss and build *a posteriori* an algorithm for the scoring of disease activity. This algorithm was designed on the basis of the pathologists' experience and of their analysis during the reading sessions of this first set of liver biopsies. A consensus was reached on basic features that could be included in the definition of the item "activity" and how they could be associated so as to score its degree. Evaluation of the activity of hepatitis based on a liver biopsy and the algorithm is referred to as "algorithm activity."

Using "reference activity" as the gold standard, the percentage of cases that were graded correctly according to the algorithm was evaluated. The strength of concordance between reference activity and algorithm activity was studied in the 363 liver biopsies using κ statistics.

Statistical Model. Another analysis of methods used by pathologists to generate the score of activity was based on a stepwise discriminant analysis. The aim of this analysis was to select, step by step, the basic lesions that best permitted classifying patients into the four groups of disease activity (A0 to A3) among the 363 liver biopsies. A jacknife procedure was used to reduce bias in the four classification groups. A discriminant function was obtained. Evaluation of the activity of a liver biopsy with this discriminant function was termed its "statistical activity." Using reference activity as the gold standard, the percentage of cases graded correctly according to the discriminant function was evaluated. The strength of concordance between reference activity and statistical activity was studied.

To judge the validity of the proposed algorithm, the concordance

TABLE 1. Distribution of Necroinflammatory Features in the 363 Liver Biopsies of Chronic Hepatitis C

Histological Feature	Score	n (%)
Piecemeal necrosis	0 = No	15 (4%)
	1 = Mild	160 (44%)
	2 = Moderate	159 (44%)
	3 = Severe	29 (8%)
Lobular necrosis	0 = No or mild	140 (39%)
	1 = Moderate	182 (50%)
	2 = Severe	41 (11%)
Portal inflammation	0 = No	4 (1%)
	1 = Mild	112 (31%)
	2 = Moderate	229 (63%)
	3 = Severe	18 (5%)
Bridging necrosis	Yes	1 (0.5%)
	No	362 (99.5%)

between reference activity and algorithm activity was compared with the concordance between reference activity and statistical activity.

Statistical Analysis. The measure of activity was compared by t test to detect a systematic difference. Analysis by κ statistics was performed to assess the degree of concordance. Strength of agreement was considered slight for values between 0 and 0.19, fair for values between 0.20 and 0.39, moderate for values between 0.40 and 0.59, substantial for values between 0.60 and 0.79, and almost perfect if κ values were greater than 0.80.

RESULTS

Prevalence and Correlation Between the Different Necroinflammatory Lesions. The prevalence of necroinflammatory lesions and grading of activity in the 363 liver biopsies are reported in Table 1.

There was a very significant positive correlation between the severity of piecemeal necrosis and portal inflammation ($\kappa^2 = 107.9$, $P < .00001$), as well as between the severity of piecemeal necrosis and the intensity of lobular necrosis ($\kappa^2 = 33$, $P < .0001$).

Panel Algorithm. After study of the first 363 biopsies, pathologists met to propose an algorithm for grading the activity of hepatitis. A consensus was reached by the pathologists on the following point: the panel agreed that the major criterion of decision was the intensity of piecemeal necrosis. Activity could also be modulated by the intensity of lobular necrosis, which was not considered the major criterion. Portal inflammation was not thought to influence the assessment of the degree of activity.

A simple algorithm, including eight possible means of arriving at a decision, was proposed according to the degree of piecemeal necrosis and of lobular necrosis (Fig. 2).

Concordance Between Reference Activity and Algorithm Activity. A review of the histological coding forms of the 363 liver biopsies showed that pathologists intuitively followed the panel algorithm in 305 cases (84%). Concordance between reference activity and algorithm activity is shown in Table 2. Concordance was judged to be substantial as assessed by κ statistics ($\kappa = .75$).

All cases that did not fulfill the algorithm differed by only one grade in the intensity of activity.

To increase the accuracy of the algorithm, we verified whether complementary pathological features would have helped to avoid discordant cases. The 58 cases that did not fulfill the algorithm were reviewed in detail. Neither portal inflammation nor any other basic features led to a better subclassification of these cases.

Statistical Model and Concordance Between Reference Activity and Statistical Activity. Stepwise discriminant analysis showed that the following items, by order of entry, were used in the grading of activity: piecemeal necrosis, lobular necrosis, and portal inflammation. Most of the discriminative information was obtained by piecemeal necrosis. Results are shown in Table 3.

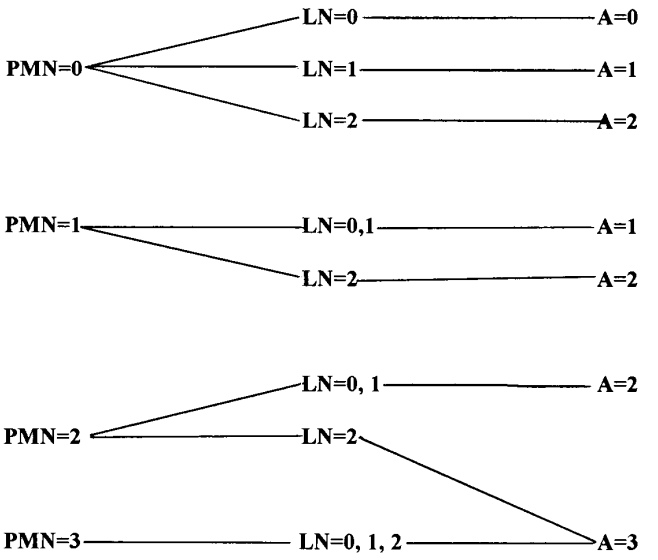

FIG. 2. Algorithm for the evaluation of histological activity. PMN, piecemeal necrosis; 0, none; 1, mild; 2, moderate; 3, severe; LN, lobular necrosis; 0, no or mild; 1, moderate; 2, severe; A, histological activity; 0, none; 1, mild; 2, moderate; 3, severe.

The concordance between the degree of activity defined by the pathologist (reference activity) and the degree of activity defined by the statistical model (statistical activity) is shown in Table 4. Activity defined by the pathologist was concordant with activity defined by the statistical model in 300 cases (83%). Strength of concordance was judged substantial, as assessed by κ statistics ($\kappa = .73$).

Cases that did not fulfill the statistical model differed by only one grade in the intensity of activity.

DISCUSSION

The major aim of studies concerning the prognosis of chronic hepatitis is to try to define groups of patients who are at risk for cirrhosis and later development of hepatocellular carcinoma. Several studies have pointed out biochemical and virological data of prognostic importance,[17] but, to date, liver biopsy findings are still considered valid criteria.

Based on data obtained in chronic hepatitis B, it is presumed that the severity of histological activity is predictive of further development of liver fibrosis. Histological activity of chronic hepatitis refers to a combination of necroinflammatory lesions, but no clear definition of activity is available, and no quantification system has been universally accepted. However, it is important that the same definition of activity

TABLE 2. Concordance Between Reference Activity and Algorithm Activity in the 363 Liver Biopsies

	Algorithm Activity			
Reference Activity	0	1	2	3
0	11	4	–	–
1	4	130	19	–
2	–	9	120	5
3	–	–	17	44

NOTE. 0 = absent, 1 = mild, 2 = moderate, 3 = severe. Concordance is substantial ($\kappa = .75$).

be used by all pathologists. General agreement exists between experienced pathologists concerning the definition of histological activity, as proven by low interobserver variation in the assessment of this feature.[8] However, guidelines must be provided to general pathologists who are increasingly confronted with chronic hepatitis C liver biopsies.

In the present work, we chose as the gold standard the grading of activity intuitively performed by experienced pathologists (reference activity). We asked these pathologists to grade activity without any guidelines as to definition. Then we investigated (1) the basic lesions that pathologists intuitively included in their grading of activity, and (2) the way they established their grading system. Two approaches were used to delineate basic lesions that were intuitively included in this grading system. In the first, the 10 pathologists met to propose a consensual simple algorithm for grading the activity of hepatitis C virus. In another approach, stepwise discriminant analysis was performed to determine the independently associated basic features used to grade the activity.

A panel decided to define activity according to its potential predictive value for the occurrence of liver fibrosis. We chose to include in the algorithm only two features—piecemeal necrosis and lobular necrosis. Portal inflammation was excluded from the algorithm, because this feature is a prerequisite for the definition of chronic hepatitis even without activity. Furthermore, we observed a strong correlation with piecemeal necrosis, making these two features redundant criteria. Piecemeal necrosis was chosen as the first decision criterion because of its proven potential value in other types of chronic hepatitis.[9] However, it is of note that, in all published series on chronic hepatitis C, piecemeal necrosis is a lesion of rather mild severity, whereas fibrosis and cirrhosis are common.[1-3] It was then suspected that another feature, lobular necrosis, was of major importance in the prediction of liver fibrosis. Indeed, it is believed that aggravation of chronic hepatitis C occurs through a burst of lobular necrosis, a lesion that is frequently present in chronic hepatitis C. These two lesions were therefore combined to propose a simple algorithm that defined activity.

For the statistical approach, stepwise discriminant analysis was performed to determine the independently associated basic features intuitively used to grade activity. The major features again were piecemeal necrosis and intralobular necrosis. To a lesser degree, portal inflammation was also used. Using these criteria, the activity of 84% of the liver biopsies evaluated by the pathologists matched the activity defined by this statistical model. Thus, basic features necessary for grading the activity and their respective importance, either proposed by the panel of pathologists or determined by the statistical approach, showed good concordance.

In several existing classifications, the degree of piecemeal and lobular necroses was independently assessed and their scores then added, thus conferring upon each of these two lesions the same weight in the definition of activity. In the present study, the proposed algorithm and the suggested statistical approach included both piecemeal necrosis and lobular necrosis in the definition of activity, but with different values. The rationale for overweighting the piecemeal necrosis item by comparison with lobular necrosis is that piecemeal necrosis is the major discriminating factor used to grade activity, as shown by the stepwise discriminant analysis.

TABLE 3. Stepwise Discriminant Analysis of the Grading of Activity in the 363 Liver Biopsies

	Variables Selected	F	*P*
Step 1	Piecemeal necrosis	261.9	<.0001
Step 2	Lobular necrosis	17.7	<.001
Step 3	Portal inflammation	10.2	<.01

TABLE 4. Concordance Between Reference Activity and Statistical Activity in the 363 Liver Biopsies

Reference Activity	0	1	2	3
0	12	6	–	–
1	3	118	12	–
2	–	19	139	18
3	–	–	5	31

NOTE. 0 = absent, 1 = mild, 2 = moderate, 3 = severe. Concordance is substantial (κ = .73).

We did not include the intensity of portal inflammation in the algorithm decision. Although discriminant analysis suggested that this feature was independently associated with grading of activity, it added little information compared with piecemeal and lobular necrosis. No clear amelioration of algorithm accuracy was observed when including this variable when reanalyzing the questionnaires that did not fulfill the algorithm.

The proposed algorithm has the advantage of simplicity, reproducibility, and application to a large number of biopsies. Because piecemeal necrosis and lobular necrosis are two reproducible features,[8,18] general pathologists can use this algorithm to semiquantitatively assess the histological activity of chronic hepatitis C. In France, the METAVIR scoring system is used for evaluating chronic hepatitis C. This score is composed of a two-letter and two-number coding system: A = histological activity (A0 = no activity, A1 = mild activity, A2 = moderate activity, and A3 = severe activity), which can be assessed using the proposed algorithm that takes into account piecemeal and lobular necrosis, and F = fibrosis (F0 = no fibrosis, F1 = portal fibrosis without septa, F2 = portal fibrosis with rare septa, F3 = numerous septa without cirrhosis, and F4 = cirrhosis). The proposed algorithm provides an easy means of scoring the activity.

REFERENCES

1. Scheuer PJ, Ashrafzadeh P, Sherlock S, Brown D, Dusheiko GM. The pathology of hepatitis C. HEPATOLOGY 1992;15:567-571.
2. Bach N, Thung SH, Schaffner F. The histological features of chronic hepatitis C and autoimmune chronic hepatitis: a comparative analysis. HEPATOLOGY 1992;15:572-577.
3. Lefkowitch JH, Schiff ER, Davis GL, Perrilo RP, Lindsay K, Bodenheimer HC, Balart LA, et al. Pathological diagnosis of chronic hepatitis C. A multicenter comparative study with chronic hepatitis B. Gastroenterology 1993;104:595-603.
4. Scheuer PJ. Classification of chronic viral hepatitis: a need for reassessment. J Hepatol 1991;13:372-374.
5. Ludwig J. The nomenclature of chronic active hepatitis: an obituary. Gastroenterology 1993;105:274-278.
6. Desmet VJ, Gerber M, Hoofnagle JH, Manns M, Scheuer PJ. Classification of chronic hepatitis: diagnosis, grading and staging. HEPATOLOGY 1994; 19:1513-1519.
7. Scheuer PJ. The nomenclature of chronic hepatitis: time for a change. J Hepatol 1995;22:112-114.
8. The METAVIR cooperative group. Inter- and intra-observer variation in the assessment of liver biopsy of chronic hepatitis C. HEPATOLOGY 1994; 20;1:15-20.
9. Liaw Y-F, Tai D-I, Chu C-M, Chen T-J. The development of cirrhosis in patients with chronic type B hepatitis: a prospective study. HEPATOLOGY 1988;8:493-496.
10. Knodell KG, Ishak KG, Black WC, Chen TS, Craig R, Kaplowitz N, Kiernan TW, et al. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. HEPATOLOGY 1981;1:431-435.
11. Poynard T, Bedossa P, Delfraissy JF, Naveau S, Lemaigre G, Chaput JC. Efficacy of low doses, long term alpha recombinant interferon in patients with chronic C hepatitis: clinical, biological, histological and immunohistochemical study. Gastroenterol Clin Biol 1991;15:615-619.
12. Poynard T, Bedossa P, Chevallier M, Mathurin P, Lemonnier C, and a multicenter study group. Improvement of histological lesions by 18 months

therapy with α2B interferon in patients with chronic hepatitis C. A randomized clinical trial N Engl J Med 1995;22:1457-1462.
13. Valla D, Babany G, Ouzan D, Trepo C, Bourliere M, Cales P, Payen JL, et al. Prevention of relapse in patients with chronic non A, non B/C hepatitis who respond to alpha-interferon. A controlled multicenter trial of low-dose maintenance therapy. J Hepatol 1994;21:774-778.
14. Ouzan D, Skaf R, Andreani T, Opolon P, Trepo C, Michel P, Couzigou P. French multicenter control trial of interferon alpha 2a (IFN) in chronic hepatitis C. Does an attack dose (6MU) increase the response rate at 6 and 12 months [Abstract]. J Hepatol 1993;18S:53.
15. Jouet P, Roudot-Thoraval F, Dhumeaux D, Metreau JM. Comparative efficacy of interferon alpha in cirrhotic and noncirrhotic patients with non-A, non-B, C hepatitis. Gastroenterology 1994;106:686-690.
16. METAVIR. Proposition d'une grille de recueil des lésions histopathologiques dans l'hépatite chronique virale C. Ann Pathol 1993;13:260-265.
17. Nousbaum JB, Pol S, Nalpas B, Landais P, Berthelot P, Brechot C. Hepatitis C virus type 1b infection in France and Italy. Collaborative study group. Ann Intern Med 1995;122:161-168.
18. Theodossi A, Skene AM, Portmann B, Knill-Jones RP, Patrick RS, Tate RA, Kealey W, et al. Observer variation in assessment of liver biopsies including analysis by Kappa statistics. Gastroenterology 1980;79:232-241.