LIVER BIOPSY IN CHRONIC HEPATITIS C: AN ASSESSMENT OF INTER AND INTRA-OBSERVER'S VARIABILITY

Alia Zubair, Shahid Jamal, Azhar Mubarik, Tariq Masood Malik, Adeel Arif

Army Medical College, Rawalpindi, Pakistan

ABSTRACT

Objective: To evaluate the inter and intraobserver variability in the histological grading and staging according to modified Knodell scoring system.

Design: A cross-sectional comparative study.

Place and Duration of study: Histopathology department Army Medical College, Rawalpindi, Pakistan from June 2006 to December 2006, at the

Materials and Methods: Slides and original reports of already reported chronic hepatitis C cases were retrieved from the case files. A total of 52 liver biopsies of patients were reevaluated by two pathologists. The inter and intraobserver reproducibility for grade of necroinflammation and stage of fibrosis were calculated by using kappa statistics.

Results: For grades of necroinflammation a substantial level of interobserver (kappa=0.802) and intra-observer (kappa= 0.749) reproducibility was found. Disagreement in the interobserver results was detected in 11.5 % cases, with difference of only one grade in all the cases. Disagreement in the intraobserver diagnosis was noted in 15.4% cases, again with the difference of only one grade of necroinflammation. Similarly for the stage of fibrosis, a substantial level of interobserver (kappa= 0.66) and intra-observer (kappa=0.77) reproducibility was present. Main disagreement for interobserver results was of stage 2 and 3 fibrosis. For intraobserver stage of fibrosis, disagreement was found in 9 cases (17.3%). There was disagreement in 6 of the 9 cases with fibrosis stage 3, where original histological stage was reported 4. There were 3 (5.8 %) cases where presence of steatosis was missed (all in non-tabulated form of reports).

Conclusion: Substantial level of inter and intra-observer agreement can be achieved, both for the necroinflammatory grade and stage of fibrosis, if the scoring system of chronic hepatitis is strictly followed.

Keywords: Necroinflammatory scoring, observer's variability, chronic hepatitis C

INTRODUCTION

In the treatment for hepatitis C interpretation of liver histology is of major importance for assessment of disease severity. It is vital that the pathologists involved in the evaluation render valuable and accurate information to the clinicians for the benefit of the patient. Liver biopsy is the most effective means to assess the condition's process rate (necroinflammatory activity grade) and progression phase (stage). Hematologist and gastroenterologist rely on the histological findings to determine the nature and extent of the hepatic damage in selection of treatment course in patients with liver diseases [1]. Accuracy and reproducibility are essential in the assessment of disease severity in these patients [2].

Various studies have shown great variability in interpretation of liver biopsies among pathologists (interobserver variability) and even in one single pathologist when

Correspondence: Col Shahid Jamal, Associate Professor of Pathology, Army Medical College, Rawalpindi
Email: sjarjawj@yahoo.com

Received Aug 22, 2007: Accepted Nov 19, 2007

assessing one sample at two different times (intraobserver variability). In a recent study, Petz et al [3] showed that interobserver variability was 58%, and intraobserver variability was 56%. A review of the accuracy of liver biopsy in the assessment of liver fibrosis is necessary, as this has been used as the gold standard for almost all studies of noninvasive markers of liver fibrosis [4].

Studies on the inter and intraobserver variability of hepatitis have focused on the different grading systems available for chronic hepatitis [5,6]. The most widely used system is Knodell index [7] and its modification, proposed by Ishak et al [8]. Although these are widely used, there is only limited data regarding the reproducibility of this classification system in our setup. This study was designed to evaluate the inter and intraobserver variation in the histological grading and staging according to modified Knodell scoring system.

MATERIALS AND METHODS

A cross-sectional comparative study was carried out from June 2006 to December 2006, at the histopathology department Army Medical College Rawalpindi, Pakistan.

The material included 60 liver biopsies of consecutive patients with chronic HCV infection. All the patients had abnormal ALT for at least 6 months before the biopsies. A diagnosis of HCV infection was based on positive anti HCV antibodies by third generation enzyme linked immune absorbent assay (ELISA) and a positive HCV RNA by polymerase chain reaction (PCR). Patients with history of alcohol intake and those with serological markers positive for hepatitis B virus infection were excluded from the study.

Formalin fixed paraffin embedded sections were stained with Hematoxylin and Eosin (H&E) and with Reticulin stain. The histological findings were assessed according to the standard grading and staging method based on modified Knodell scoring system proposed by Ishak et al. [8]. The final grades of necroinflammation were classified according to Desmets classification [9].

To determine the interobserver reproducibility the original histological grade of inflammation (OHG) and original histological stage of fibrosis (OHS) of the cases were reviewed by two observers in an independent manner and without knowledge of previous results. The cases were crossed and two pathologists reviewed the cases which they had not reported previously. Intraobserver reproducibility was determined by comparing the reviewed histopathological results of same pathologist with the OHG and OHS of the cases. Data was analyzed using SPSS 11.0. Descriptive statistics was used to describe the data. The kappa score was used to measure inter and intraobserver agreement. The agreement for the kappa estimates was expressed as threshold values: <0.00 (poor); 0.00 to 0.20 (slight); 0.21 to 0.40 (fair); 0.41 to 0.60 (moderate); 0.61 to 0.80 (substantial); 0.81 to 1.00 (almost perfect) [10].

RESULTS

Of the 60 liver biopsies available, we excluded 8 due to inadequacy (containing less than 3 portal tracts). The material subsequently consisted of 52 biopsies from 16 (30.77%) female and 36 (69.23%) male patients. Median age was 36 years.

Interobserver reproducibility:

For evaluation of interobserver reproducibility (pathologist-I x pathologist-II), tables were constructed between histopathological results of the two pathologists, which shows the frequency of distribution for each grade of necroinflammatory activity (table-1) and stage of fibrosis (fig. 1).

Total number of discordant cases for necroinflammatory grades between two pathologists was 6 (11.5%) of 52 cases. Two out of 6 cases were graded as moderate chronic hepatitis C (CHC) by pathologist-I and the grade given by pathologist-II was mild CHC. One case where grade was mild CHC by pathologist-I and the grade given by pathologist-II was moderate CHC and one case of moderate CHC by pathologist-I was graded as severe CHC by pathologist-II. The

kappa value was calculated to be 0.802 meaning substantial level of interobserver reproducibility.

Discrepancies in stage of fibrosis between two pathologists were observed in 15 (28.8%) of the 52 cases (table-2 & fig. 2). Main disagreement (in 7 cases) was observed in stage 2 and 3 of fibrosis. In 4 cases stage of fibrosis was 2 by pathologist-I and observed as stage 3 by pathologist-II, whereas 3 cases of stage 3 by pathologist-I were labeled as stage 2 by pathologist-II. Two cases of stage 2 by pathologist-I were staged as 4 by pathologist-II. Further discrepancies were observed in 2 cases where stage 1 by pathologist-I, was declared as stage 2 by pathologist-II and one case each of stage, 0, 3, 4 and 5 by pathologist-I were one stage upgraded by second pathologist. A substantial level of interobserver reproducibility for stage of fibrosis was observed with kappa value of 0.65.

Intraobserver Reproducibility:

For evaluation of intraobserver reproducibility (pathologist review report x OHG), a table was constructed between OHG and the results were obtained for the same cases reviewed by pathologist again (table-3).

In the intra-observer variability the discrepancies in necroinflammatory grade between the cases reviewed by a pathologist again were observed in 8 (15.4%) cases. In 4 cases where OHG was severe CHC, the reviewed grade was moderate CHC. Two cases of mild CHC were graded as moderate CHC on review, one case of moderate CHC was reviewed as mild CHC and 1 case of mild CHC was reviewed as minimal CHC. A substantial level of intraobserver reproducibility for necroinflammatory grade was observed (kappa value =0.749).

Disagreement in stage of fibrosis between OHS and reviewed cases were observed in 9 (17.3%) cases (table-2 & fig. 2). In 6 cases OHS-3 was reviewed as stage 2. Two cases of OHS-4 were given stage-2 on review and 1 case with OHS of one was reported stage-0 on review. A substantial level of intraobserver reproducibility for stage of fibrosis was observed with kappa value = 0.77.

There were three cases (5.8 %) out of 52 in which steatosis was overlooked in original results but reported by the pathologists in reevaluation of biopsies. It was found that all these cases were from those biopsies where tabulated report was not issued (out of a few cases when department was not using tabulated format of reporting).

DISCUSSION

The approach to assessing and reporting the severity of histopathological lesion in patients with chronic HCV infection has undergone considerable reevaluation in recent years [11-13]. There are studies showing discrepancies of opinion among pathologists when a second review is implemented [14,15]. Inter and intra-observer variability has been reported in liver biopsy interpretation [16,17].

Hepatitis C is very common in our set up but there is no such study available, so in this study the inter and intra observer variation for interpreting liver biopsies from patients with chronic hepatitis C was evaluated. The grading of necroinflammation showed higher inter and intraobserver agreement compared with staging of fibrosis when assessed by kappa statistics. The interobserver agreement for the total HAI was substantial (kappa value = 0.80). Similarly fibrosis stage had substantial level of reproducibility with a kappa value of 0.66. Similar results were found for intraobserver reproducibility with kappa value of 0.74 for necroinflammatory grades and kappa value of 0.77 for stage of fibrosis. The results of some other studies have shown that inter and intraobserver concordance for fibrosis and cirrhosis can be good [5], substantial [18], or essentially perfect [6,16,19]. Grading of the inflammatory activity in chronic hepatitis has also shown variability. Some reports showed 96% agreement [6], whereas in others concordance was weaker [19]. In a recent study, Petz et al.[3] observed 58% inter-observer and 56% intra-observer variability.

Table-1: Results of necroinflammatory grades by two pathologists (interobserver) and their kappa values (n=52).

Pathologist-1	Pathologist-II				
	MinimalCHC	Mild CHC	Moderate CHC	Severe CHC	Total
Minimal CHC	1	1	_	_	2
Mild CHC	_	20	1	_	21
Moderate CHC	_	2	22	1	25
Severe CHC	-	_	1	3	4
Total	1	23	24	4	52

CHC = chronic hepatitis C, (Kappa value = 0.802)

Table-2: Number of cases of inter and intra-observer variability of fibrosis stage.

Inter-observer variability (n=15) kappa value = 0.66			Intra-observer variability (n=9) kappa value = 0.77		
No of case	Stage by Path-I	Stage by Path-II	No of cases	OHS	Stage reviewed by Path
2	Stage-1	Stage-2	1	Stage-1	Stage-0
4	Stage-2	Stage-3	6	Stage-3	Stage-2
3	Stage-3	Stage-2	2	Stage-4	Stage-2
2	Stage-2	Stage-4			
1 each	Stage-0,3,4,5	Stage-1,4,5,6			

Table-3: Original results of necro-inflammatory grades (OHG) and reviewed results by same pathologist (intra-observer OHG.

Pathologist's reviewed results	Minimal CHC	Mild CHC	Moderate CHC	Severe CHC	Total
Minimal CHC	1	1	_	_	2
Mild CHC	_	21	1	_	22
Moderate CHC	_	2	19	4	25
Severe CHC	_	_	_	3	3
Total	1	24	20	7	52

Kappa value = 0.749 OHG, original histological grades of necro-inflammation; CHC, chronic hepatitis C

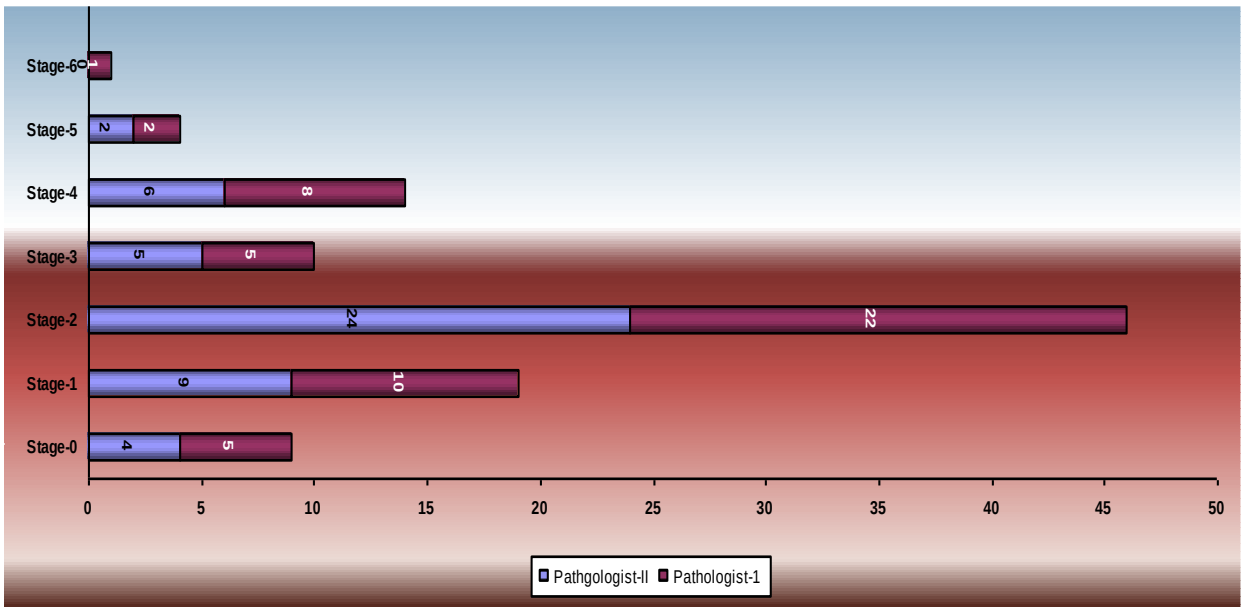

Fig-1: Different stages of fibrosis given by two pathologists.

In the present study inter-observer reproducibility showed one stage of fibrosis in all 13 out of 15 cases. Main disagreement was of stage 2 and 3 fibrosis. As far as intra-observer results are concerned, there were 2 out of 9 discrepant cases whose difference in stage of fibrosis was of 2 points where OHS of 4 were staged as 2. Rest of the 7 cases showed difference of one stage only same was reported by Gronbaek et al [20].

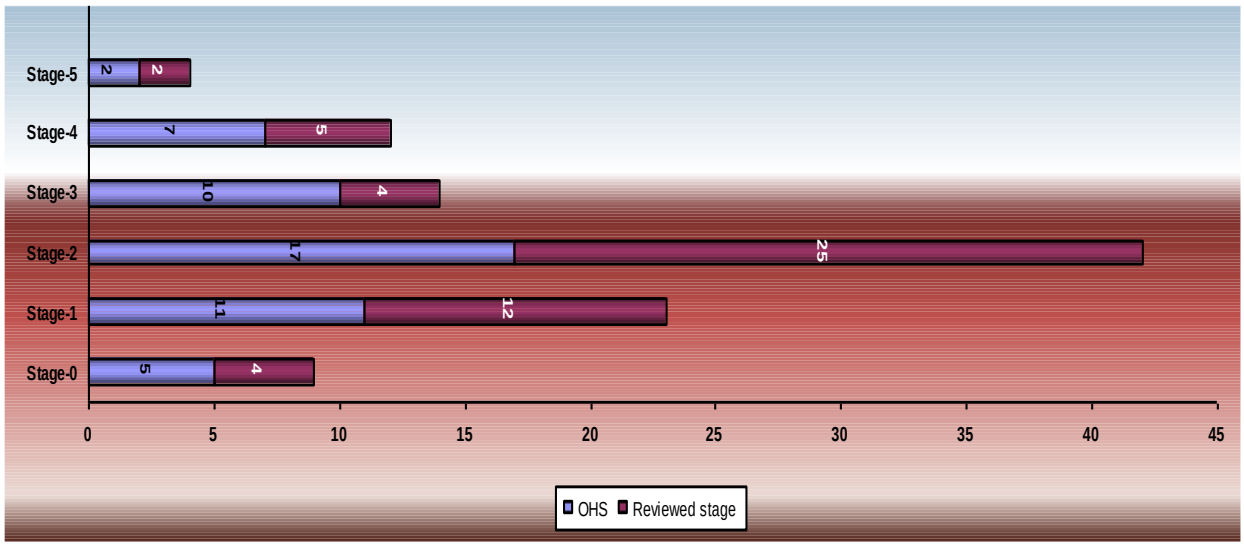

Fig-2: Different stages of fibrosis in original histology stage (OHS) and that which is given on review by same pathologist.

Results of this study highlight that the main limitation of staging of fibrosis in liver biopsies resides in the evaluation of small samples, as presence of less portal tracts does not permit an accurate assessment of stage of fibrosis. Same was found in previous studies, that percutaneous liver biopsy may miss the diagnosis of cirrhosis, secondary to sampling error, in a frequency ranging from 1% to 67% [21,22]. Reducing the amount of tissue, significantly reduces the scores for both necroinflammation and fibrosis. It is said that evaluating small or slender biopsies is likely to lead to underestimation of disease severity and recommendations are that grading and staging should be carried out using specimens at least 20 mm long and 1.4 mm wide [23]. Some authors have argued that sample less than 20 mm in length is inadequate in the assessment of chronic viral hepatitis and will underscore fibrosis [24]. Discrepancy in the stage of fibrosis in the present study can be explained due to the inclusion of small sized biopsies in the calculation of the results.

Some other features of chronic hepatitis C like steatosis and lymphoid follicle formation may have impact on management of the patient. In present study there were minor discrepancies regarding 3 cases, where steatosis was overlooked in original histopathological report and was reported in reevaluated results by the pathologists. This can be explained as the Knodell score does not account for features specific to different types of viral hepatitis, like lymphoid aggregates, bile duct injury, macrovesicular steatosis that are often present in chronic hepatitis C [25]. In our observation there are fewer chances to omit these observations, in tabulated form of reporting.

CONCLUSION

It is concluded from the study that substantial level of inter and intra-observer agreement existed, both for the necroinflammatory grade and stage of fibrosis by the application of scoring system of chronic hepatitis. Some difference in the results obtained by pathologists sample do not mean that one or both of them is wrong ; the differences are to be expected in a subjective analysis. The main discrepancy was observed in the evaluation of stage 3 fibrosis, which is recommended to be reevaluated by the pathologists thus minimizing interobserver variability. Other omissions like presence of steatosis can be avoided if reporting is done in tabulated form.

REFERENCES

1. Bejarano PA, Koehler A, Sherman KE. Second opinion pathology in liver biopsy interpretation. Am J Gastroenterol. 2001; 96: 3158-64.

2. Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, Pyrsopoulos NT, et al. Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. Am J Gastroenterol. 2002; 97: 2614-8.

3. Petz D, Klauck S, Röhl FW, Malfertheiner P, Roessner A, Röcken C, et al. Feasibility of histological grading and staging of chronic viral hepatitis using specimens obtained by thin-needle biopsy. Virchow Arch. 2003; 442: 238-44.

4. Afdhal NH, Nunes D. Evaluation of Liver fibrosis: A concise review. Am J Gastroenterol. 2004; 99: 1160-74.

5. Goldin RD, Goldin JG, Burt AD, Dhillon PA, Hubscher S, Wyatt J, et al. Intra-observer and interobserver variation in the histopathological assessment of chronic viral hepatitis. J Hepatol. 1996; 25: 649-54.

6. Westin J, Lagging LM, Wejstal R, Norkrans G, Dhillon AP. Interobserver study of liver histopathology using the Ishak score in patients with chronic hepatitis C virus infection. Liver. 1999; 19: 183-7.

7. Knodell RG, Ishak KG, Black WC. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. Hepatology 1981; 1: 431-5.

8. Ishak K, Baptista A, Bianchi L, Callea F, De Groote J, Gudat F, et al. Histological grading and staging of chronic hepatitis. J Hepatol. 1995; 22: 696-9.

9. Desmet VJ, Gerber M, Hoofnagle JH, Manns M, Scheur PJ. Classification of chronic hepatitis: Diagnosis, grading and staging. Hepatology. 1994; 19: 1513-20.

10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33: 159-74.

11. Batts KP, Ludwig J. Chronic hepatitis. An update on terminology and reporting. Am J Surg Pathol. 1995; 19: 1409-17.

12. Ludwig J. The nomenclature of chronic active hepatitis: An obituary. Gastroenterology. 1993; 105: 274-8.

13. Brunt EM. Grading and staging the histopathological lesions of chronic hepatitis: The Knodell histology activity index and beyond. Hepatology. 1999; 31: 241-6.

14. Abt AB, Abt LG, Olt GJ. The effect of interinstitution anatomic pathology consultation on patient care. Arch Pathol Lab Med. 1995; 119: 514-7.

15. Epstine JI, Walsh PC, Sanfilippo F. Clinical and cost impact of second-opinion pathology. Review of prostate biopsies prior to radical prostatectomy. Am J Surg Pathol. 1996; 20: 851-7.

16. Bedossa P, Poynard T, Naveau S, Martin ED, Agostini H, Chapat JC. Observer variation in assessment of liver biopsies of alcoholic patients. Alcohol Clin Exp Res. 1988; 12: 173-8.

17. Theodossi A, Skene AM, Portmann B. Observer variation in assessment of liver biopsies including analysis by Kappa statistics. Gastroenterology. 1980;79:232-41

18. Younossi ZM, Gramlich T, Liu YC, Matteoni C, Patrelli M, Goldblum J, et al. Nonalcohlic fatty liver diseases: Assessment of variability in pathologic interpretations. Mod Pathol. 1998; 11:560-5

19. Anonymous. Intra-observer and inter-observer variation in liver biopsy interpretation in patients with chronic hepatitis C. The French metavir cooperative study group. Hepatology. 994; 20: 15-20

20. Gronbaek K, Christensen PB, Hmilton-Dutoit S, Federspiel BH, Hage E, Jensen OJ, et al. Interobserver variation in interpretation of serial liver biopsies from patients with chronic hepatitis C. J Viral Hepat. 2003; 9: 443-9

21. Nord HJ. Biopsy diagnosis of cirrhosis. Blind percutaneous versus guided direct vision techniques- review. Gastrointest Endosc 1982; 28: 102–4.

22. Idorvo V, Dailey PJ, Jeffers LJ, Coelho-Little E, Bernstein D, Bartholomew M, et al. Hepatitis C virus RNA quantification in right and left lobes of the liver in patients with chronic hepatitis C. J Viral Hepat. 1996; 3: 239–46.

23. Colloredo G, Guido M, Sonzogni A, Leandro G. Impact of liver biopsy size on histological evaluation of chronic viral hepatitis: the smaller the sample, the milder the disease. J Hepatol 2003,39:239-44.

24. Guido M, Rugge M. Liver biopsy sampling in chronic viral hepatitis. Semin Liver Dis. 2004; 24: 89-97.

25. Banner BF, Barton AL, Cable EE, Smith L, Bonkovsky HL. A detailed analysis of the Knodell Score and other histological parameters as predictor of response to interferon therapy in chronic hepatitis C. Mod Pathol. 1995; 8: 232-8.