SELÇUK TIP DERGİSİ
SELCUK MEDICAL JOURNAL

Araştırma Makalesi / Research Article

Selcuk Med J 2023;39(4): 178-182
DOI: 10.30733/std.2023.01674

NEÜ YAYINEVİ

Inter Observer Agreement of The Modified Ishak Histological Activity Index in Chronic Viral Hepatitis Among Pathologists Trained in Different Centers

Kronik Viral Hepatitlerde Modifiye Ishak Histolojik Aktivite İndeksinin Farklı Merkezlerde Eğitim Almış Patologlar Arasında Tekrarlanabilirliği

Ilkay Cinar[1]

[1]*Giresun University, Medical Faculty, Department of Pathology, Giresun, Turkiye*

Address correspondence to:
Ilkay Cinar, Giresun University, Medical Faculty, Department of Pathology, Giresun, Turkiye
e-mail: a.ilkaycinar@gmail.com

Geliş Tarihi/Received: *15 September 2023*
Kabul Tarihi/Accepted: *23 November 2023*

Öz

Amaç: Karaciğer biyopsisi kronik hepatitli hastalarda hepatik hasarın gösterilmesinde altın standarttır. Bu çalışmada, kronik hepatitlerin histopatolojik değerlendirmesinde kullanılan Modifiye Ishak Histolojik Aktivite İndeksi (MHAİ)'nin farklı merkezlerde eğitim almış gözlemciler arası tekrarlanabilirliği araştırılmıştır.
Hastalar ve Yöntem: Çalışmaya 2020-2023 yılları arasında toplam 64 karaciğer iğne biyopsisi dahil edildi. Olgulara ait hematoksilen-eozin, gümüş ve trikrom ile boyanmış kesitler araştırmacı (İÇ) tarafından ışık mikroskopunda tekrar değerlendirildi ve MHAİ'ne göre tekrar skorlandı. İstatistiksel değerlendirmede SPSS 21.0 programı kullanılarak Cohen'in kappa istatistiği uygulandı.
Bulgular: İstatistiksel değerlendirmede tüm olgular dikkate alındığında kappa değeri; MAI derece için 0.02(slight), evre için 0.38(fair) bulundu. Araştırmacının diğer patologlar ile uyumu değerlendirildiğinde derecelendirmede kappa değerleri oldukça düşük olup uyumsuz bulundu. Evrede ise dereceye göre daha yüksek değerler(0.12-0.49) olmakla birlikte uyum zayıf ve orta derece idi. Ancak uzun vadeli işbirliği geçmişine sahip patologlar, devam eden bilgi alışverişinin önemini vurgulayarak daha yüksek düzeyde uyum gösterdiler.
Sonuç: Bu çalışma, MHAİ puanlamasının tekrarlanabilirliğini artırmak için yorumlanması ve uygulanmasına yönelik standart bir yaklaşıma duyulan ihtiyacı vurgulamaktadır.

Anahtar Kelimeler: Kronik hepatit, gözlemciler arası uyum, Modifiye Ishak Aktivite İndeksi

Abstract

Aim: Liver biopsy is the gold standard for demonstrating hepatic damage in patients with chronic hepatitis. In this study, the reproducibility of the Modified Ishak Histological Activity Index (MHAI), used in the histopathological evaluation of chronic hepatitis, between observers trained in different centers was investigated.
Patients and Methods: A total of 64 liver needle biopsies were included in the study between 2020 and 2023. Sections of the cases stained with hematoxylin-eosin, silver and trichrome were re-evaluated under a light microscope by the investigator (İÇ) and re-scored according to MHAI. In statistical evaluation, Cohen's kappa statistics were applied using the SPSS 21.0 program.
Results: When all cases are taken into consideration in statistical evaluation, kappa value is; MAI was found to be 0.02 (slight) for degree and 0.38 (fair) for stage. When the researcher's compatibility with other pathologists was evaluated, the kappa values in the grading were quite low and found to be incompatible. Although there were higher values (0.12-0.49) in the stage than in the degree, the fit was poor and moderate. However, pathologists with a long-term history of collaboration showed higher levels of compliance, highlighting the importance of ongoing exchange of information.
Conclusion: This study highlights the need for a standardized approach to interpretation and application of MHAI scoring to increase reproducibility.

Keywords: Chronic hepatitis, interobserver agreement, Modified Ishak Activity Index

Cite this article as: Cinar I. Inter Observer Agreement of The Modified Ishak Histological Activity Index in Chronic Viral Hepatitis Among Pathologists Trained in Different Centers. Selcuk Med J 2023;39(4): 178-182

Disclosure: *Author has not a financial interest in any of the products, devices, or drugs mentioned in this article. The research was not sponsored by an outside organization. Author has agreed to allow full access to the primary data and to allow the journal to review the data if requested.*

INTRODUCTION

Liver biopsy is considered the gold standard for demonstrating hepatic damage in patients with chronic hepatitis. It not only determines the severity of hepatitis during the biopsy but also enables the assessment of treatment effectiveness. Treatment decisions are heavily reliant on both laboratory findings and the biopsy result, which, in turn, dictates the severity of the liver disease. Detecting fibrosis at an early stage can be pivotal in preventing the disease from progressing. Consequently, it is crucial that the reports resulting from the pathological examination of needle biopsies are not only reliable but also repeatable. However, the nature of pathological examination is inherently subjective.

To address this subjectivity and to establish a common language among pathologists for the sake of clinical relevance, various grading and scoring systems have been developed for many diseases. These systems aim to quantify the severity and extent of the disease, thereby providing insights into its course. The initial histological classification of hepatitis dates back to 1968 when De Grote and colleagues introduced it (1). The Histological Activity Index (HAI), proposed by Knodell et al. in 1982, marks the inception of scoring systems for chronic hepatitis (2). The modified Knodell HAI score, as refined by Ishak et al., stands as the most widely employed method in contemporary practice (3). Furthermore, the Scheuer and Metavir systems are routinely adopted by numerous pathologists (4,5). All these methods seek to quantify necroinflammatory activity and fibrosis prevalence in the liver through numerical scores. However, the inherent challenge lies in the fact that these grading and scoring systems rely heavily on subjective descriptions. Pathologists are tasked with interpreting findings, classifying them, and assigning numerical values based on these verbal expressions, which can often lead to discrepancies among pathologists' assessments.

Ideally, a scoring system should confidently serve the purposes of treatment planning and treatment

Table 1. Modified Histological Activity Index Grading

•A Piecemeal necrosis	•None	0
	•Mild (focal, few portal areas)	1
	•Mild/Moderate (focal, most portal areas)	2
	•Moderate (less than 50% continuous tracts or septa)	3
	•Severe (more than 50% continuous in tracts or septa)	4
•B Confleunt necrosis	•None	0
	•Focal confluent necrosis	1
	•Zon 3 necrosis in some areas	2
	•Zon 3 necrosis in most areas	3
	•Zon 3 necrosis + occasional portal-central (P-C) bridging	4
	•Zon 3 necrosis + multiple portal-central (P-C) bridging	5
	•Panacinar or multiacinar necrosis	6
•C Focal necrosis/inflammation	•1 or fewer foci (per x10 magnification)	1
	•2-4 foci (per x10 magnification)	2
	•5-10 foci (per x10 magnification)	3
	•More than 10 foci (per x10 magnification)	4
•D Portal inflammation	•None	0
	•Mild (some or all portal areas)	1
	•Moderate (some or all portal areas)	2
	•Moderate/Marked (all portal areas)	3
	•Marked (all portal areas)	4

Table 2. Modified Histological Activity Index Staging

• No fibrosis	0
• Fibrous expansion of some portal areas and +/- short fibrous septa	1
• Fibrous expansion of most portal areas and +/- short fibrous septa	2
• Fibrous expansion of most portal areas with occasional portal-portal (P-P) bridging	3
• Fibrous expansion of most portal areas with marked P-P bridging as well as portal-central (P-C) bridging	4
• Marked bridging (P-P and/or P-C) with occasional nodules (incomplete cirrhosis)	5
• Cirrhosis (possible or definite)	6

Table 3. Kappa values between researchers and other pathologists in the MHAI assessment (n=64).

	MHAİ grade	MHAİ stage	A Piecemeal necrosis	B Confleunt necrosis inflammation	C Focal necrosis/ inflammation	D Portal
KAPPA	0.02	0.38	0.29	-0.31	0.31	0.41

Table 4. Kappa values between researchers and other pathologists in the MHAI assessment (n=10).

Kappa	P1(n=10)	P2(n=10)	P3(n=10)	P4(n=10)	P5(n=10)	P6(n=10)
MHAI grade	-0.08	0.02	-0.01	0.10	0	0.07
MHAI grade by diagnostic categories *	0	0.40	0.34	0.58	0.20	0.38
MHAİ stage	0.12	0.25	0.49	0.72	0.23	0.32
A Piecemeal necrosis	0.11	0.39	0.48	0.17	0.31	0.23
B Confleunt necrosis	0	-0.15	0.20	0.10	0.03	0.05
C Focal necrosis/ inflammation	0.26	0.40	0.14	0.52	0.26	-0.29
D Portal inflammation	0.50	0.34	0.59	0.16	0.31	0.53

*(Score 1-3; minimal hepatitis, score4-8; mild hepatitis, score9-12; moderate hepatitis, score13-18; severe hepatitis)

effectiveness assessment, while simultaneously exhibiting high interobserver repeatability. This study, therefore, endeavors to explore the interobserver repeatability of the widely used Modified Ishak Histological Activity Index (MHAI) in the histopathological evaluation of chronic hepatitis. Observers trained at different centers are the subjects of this investigation.

PATIENTS AND METHODS

This study is approved by the Local Ethics Committee (Approval number: 4/2023). Between 2020 and 2023, a total of 64 needle liver biopsies were included in this study. Biopsy reports were retrieved from the electronic archive. These biopsies had been reported by seven pathologists working in our department who received specialized training from different centers. Ten reports from each of six pathologists were included in the study. The seventh pathologist, who had recently joined the study, had only four biopsies meeting the criteria and was not evaluated individually. All biopsies had received a diagnosis of chronic hepatitis and were scored according to the Modified Ishak Histological Activity Index (MHAİ) (Table 1).

Sections stained with hematoxylin-eosin, silver, and trichrome were retrieved from the archives and re-evaluated by the researcher (İÇ) under a light microscope. They were then rescored according to the MHAİ. The data were compared with the values from the previously prepared reports. In the initial evaluation, the researcher scores were compared to all other biopsies (n=64) (Table 2). Subsequently, the scores from the researcher and those from the six pathologists were compared individually (n=10) (Table 3).

Statistical analysis was performed using SPSS 21.0, applying Cohen's kappa statistic. The results were evaluated by categorizing them into standard categories as follows (6).

Cohen’s Kappa	İnterpretation
<0	poor
00-0.20	slight
0.21-0.40	fair
0.41-0.60	moderate
0.61-0.80	substantial
0.81- 1.00	almost perfect

RESULTS

In the statistical evaluation, considering all cases, the kappa value for the MAI degree was 0.02 (slight), and for the stage, it was 0.38 (fair). When considering individual parameters, the lowest kappa value was found for B (confluent necrosis), while the highest kappa value was observed in the D (portal inflammation) category (Table 3).

When evaluating the researcher's agreement with other pathologists separately, the kappa values for grading were quite low, indicating poor agreement. For staging, there were relatively higher values (ranging from 0.12 to 0.49), indicating fair to moderate agreement. Substantial agreement (0.72) was observed with one observer (P4). Among the parameters, the highest agreement was found in the assessment of portal inflammation (D), while the weakest agreement was observed in the assessment of confluent necrosis.

The MHAİ grading was also evaluated as diagnostic categories. In this case (score 1-3 minimal hepatitis, score 4-8 mild hepatitis, score 9-12 moderate hepatitis, score 13-18 severe hepatitis), the kappa values were higher but still did not surpass the weak agreement category (Table 4).

DISCUSSION

Studies in the literature investigating the reproducibility and interobserver concordance of scoring methodologies have often reported moderate to low levels of agreement (7–10). This phenomenon can be primarily attributed to variations in individual interpretation when translating verbal descriptions into numerical values in semi-quantitative scoring systems. In our study, we did not observe a high level of agreement among the data assessed by the primary researcher and the other pathologists working within the same institution. The highest level of agreement was achieved with a pathologist (P4) who had a long history of collaboration. This suggests that ongoing consultations and knowledge exchange over the years may have contributed to a shared understanding among these individuals. In a study involving pathologists from the same institution, interobserver agreement in the modified HAI scoring system reportedly reached 95-96% (11). However, in institutions like ours, where pathologists have received training from different institutions and represent diverse schools of thought, interobserver agreement tends to be low, aligning with the findings of certain studies in the literature (7,9,10).

Due to the limited repeatability associated with semi-quantitative scoring systems, there has been a growing interest in the application of computer-assisted morphometry based on image analysis as an alternative approach. In recent years, numerous studies have employed computer-assisted automated algorithms for quantifying fibrosis using biopsy images (12–14). These investigations have convincingly demonstrated that automated fibrosis measurements exhibit greater precision in detecting variations in hepatic fibrosis compared to semi-quantitative histological staging (13,15). Nevertheless, it is imperative to acknowledge that these methods are not without their challenges. Assessing fibrosis in hepatitis necessitates evaluating structural alterations, nodular formations, and changes in microcirculation, in addition to quantifying the increase in fibrous tissue. The automatically measured quantity of fibrous tissue may fail to capture the architectural pattern of fibrosis distribution. Hence, it might be more suitable to combine automatic image analysis with semi-quantitative scoring for histological staging in cases of chronic hepatitis (16).

CONCLUSİON

The Knodell HAI method has been a milestone in liver pathology due to its pioneering role in enabling the comparison of pathology reports for the same sample by different pathologists (17). Although modifications have been made by Ishak, even among globally recognized experts in the field, there are issues with repeatability. Given its direct impact on patient management, there is an evident need for more objective approaches. This underscores the necessity to establish consensus among observers regarding what is meant by the semi-quantitative descriptions and expressions used in the parameters and to create a common approach for implementing scoring.

Furthermore, it is advisable to replace vague expressions such as several or many with more objective numerical values in the descriptions used in scoring systems. Additionally, supplementing the HAI scoring system with schematic images based on it could enhance repeatability, similar to the approach taken in Gleason scoring (18).

Moreover, the use of standardized histometric measurements alongside scoring systems for the assessment of necroinflammation and fibrosis is believed to improve concordance. Further studies on this topic will shed more light on this subject.

Conflict of interest: *Author declares that there is no conflict ofinterest between the authors of the article.*

Financial conflict of interest: *Author declares that he did not receive any financial support in this study.*

Address correspondence to: *Ilkay Cinar, Giresun University, Medical Faculty, Department of Pathology, Giresun, Turkiye*
e-mail: *a.ilkaycinar@gmail.com*

REFERENCES

1. De Grote J, Desmet VJ, Gedigk PKG, et al. A classification of chronic hepatitis. Lancet 1968;2:626-8.
2. Knodell RG, Ishak KG, Black WC CT, et al. Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. Hepatology 1981;1(5):431-5.
3. Ishak K, Baptista A, Bianchi L, et al. Histological grading and staging of chronic hepatitis. J Hepatol 1995;22:696-999.
4. The French METAVIR Cooperative Study Group. Intraobserver and interobserver variations in liver biopsies in patients with chronic hepatitis. Hepatology 1994;20:15-20.
5. Scheuer PJ. Classification of chronic viral hepatitis: A need for reassessment. Hepatology 1991;13:372-4.
6. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-74.
7. Rammeh S, Khadra H, Znaidi N, et al. Inter-observes agreement of Ishak and Metavir scores in histological evaluation of chronic viral hepatitis B and C. Ann Biol Clin 2014;72(1):57-60.
8. Scheuer PJ. Assessment of liver biopsies in chronic hepatitis: How is it best done? J Hepatol 2003;38:240-4.
9. Grønbæk K, Christensen PB, Hamilton -Dutoit S, et al. Interobserver variation in interpretation of serial liver biopsies from patients with chronic hepatitis C. J Viral Hepat 2022;9:443-9.
10. Petz D, Klauck S, Malfertheiner FRP, et al. Feasibility of histological grading and staging of chronic viral hepatitis using specimens obtained by thin-needle biopsy. Virchows Arch 2003;238-44.
11. Westin J, Lagging LM, Wejstål R, et al. Interobserver study of liver histopathology using the Ishak score in patients with chronic hepatitis C virus infection. Liver 1999;19(3):183-7.
12. Zhou Y, Ru GQ, Yan R, et al. An inexpensive digital image analysis technique for liver fibrosis quantification in chronic hepatitis B patients. Ann Hepatol 2017;16(6):881-7.
13. Caballero T, Perez-Milena A, Masseroli M, et al. Liver fibrosis assessment with Sustained-responder, semiquantitative indexes and image analysis quantification in Hepatol, and non-responder interferon-treated patients with chronic hepatitis C. J Hepatol 2001;34(5):740–7.
14. Campos CF, Paiva DD, Perazzo H, et al. JV. An inexpensive and worldwide available Chronic, digital image analysis technique for histological fibrosis quantification in hepatitis C. J Viral Hepat 2014;21(3):216–22.
15. Pavlides M, Birks J, Fryer E et al. Interobserver Variability in Histologic Evaluation of Liver Fibrosis Using Categorical and Quantitative Scores. Am J Clin Path 2017; 147(4):364-9.
16. Masseroli M, Caballero T, O'Valle F, et al. Automatic quantification of liver fibrosis: design and validation of a new image analysis method: Comparison with semiquantitative indexes of fibrosis. J Hepatol 2000;32:453-64.
17. Desmet VJ. Milestones in liver disease Scoring chronic hepatitis. J Hepatol 2003;38:382–6.
18. Gleason DF. Classification of prostatic carcinomas. Cancer Chemother Rep 1966;50:125–8.