Pathologic Criteria for Nonalcoholic Steatohepatitis: Interprotocol Agreement and Ability to Predict Liver-Related Mortality

Zobair M. Younossi,[1,3] Maria Stepanova,[1,3] Nila Rafiq,[1] Hala Makhlouf,[2] Zahra Younoszai,[1] Ritambhara Agrawal,[1] and Zachary Goodman[1]

Since the initial description of nonalcoholic steatohepatitis (NASH), several sets of pathologic criteria for its diagnosis have been proposed. However, their interprotocol agreement and ability to predict long-term liver-related mortality (LRM) have not been demon- strated. In this study, we examined patients with biopsy-proven nonalcoholic fatty liver disease (NAFLD) for whom liver biopsy slides and clinical and mortality data were avail- able. Liver biopsy samples were evaluated for a number of pathologic features and were classified according to the presence or absence of NASH by (1) the original criteria for NAFLD subtypes, (2) the nonalcoholic fatty liver disease activity score (NAS), (3) the Brunt criteria, and (4) the current study's criteria. All NASH diagnostic criteria and individual patho- logic features were tested for agreement and for their independent associations with LRM, which were determined with a Cox proportional hazards model. Two hundred fifty-seven NAFLD patients with complete data were included. The diagnoses of NASH by the original NAFLD subtypes and by the current study's definition of NASH were in almost perfect agreement (j 5 0.896). However, their agreement was moderate with NAS (j 5 0.470 and j 5 0.511, respectively) and only fair to moderate with the Brunt criteria (j 5 0.365 and j 5 0.441, respectively). Furthermore, the agreement of the Brunt criteria with NAS was relatively poor (j 5 0.178). During the follow-up (median 5 146 months), 31% of the patients died (9% were LRM). After we controlled for confounders, a diagnosis of NASH by the original criteria for NAFLD subtypes [adjusted hazard ratio 5 9.94 (95% confidence interval 5 1.28-77.08)] demonstrated the best independent association with LRM. Among the individual pathologic features, advanced fibrosis showed the best independent association with LRM [adjusted hazard ratio 5 5.68 (95% confidence interval 5 1.50-21.45)]. Conclusion: The original criteria for NAFLD subtypes and the current study's criteria for NASH were in almost perfect agreement, but their level of agreement with the NAS and Brunt criteria was lower. A diagnosis of NASH by the original criteria for NAFLD subtypes demonstrated the best predictability for LRM in NAFLD patients. (HEPATOLOGY 2011;53:1874-1882)

See Editorial on page 1792

Nonalcoholic fatty liver disease (NAFLD) is a clinicopathologic spectrum that ranges from simple steatosis to nonalcoholic steatohepatitis (NASH).[1-3] Although the incidence of NAFLD in the US population has been estimated to be 15% to 30%, only 2% to 3% have the potentially progressive sub- type of NAFLD or NASH.[3-5] A number of natural history studies have convincingly shown that among patients on the NAFLD spectrum, only those with NASH are at risk for progression.[1,6-14] Because of this

Abbreviations: aHR, adjusted hazard ratio; CI, confidence interval; HR, hazard ratio; ICD-9, International Classification of Diseases, 9th revision; ICD-10, International Classification of Diseases, 10th revision; LRM, liver-related mortality; NAFLD, nonalcoholic fatty liver disease; NAS, nonalcoholic fatty liver disease activity score; NASH, nonalcoholic steatohepatitis.

From the [1]Center for Liver Diseases, Inova Fairfax Hospital, Falls Church, VA; [2]Armed Forces Institute of Pathology, Washington, DC; and [3]Betty and Guy Beatty Center for Integrated Research, Inova Health System, Falls Church, VA.

Received January 12, 2011; accepted February 16, 2011.

This study was supported in part by the Liver Disease Outcomes Fund of the Center for Liver Diseases at Inova Fairfax Hospital (Falls Church, VA).

Potential conflict of interest: Nothing to report.

Fax:

Pathologic Criteria for Nonalcoholic Steatohepatitis: Interprotocol Agreement and Ability to Predict Liver-Related Mortality

Zobair M. Younossi,[1,3] Maria Stepanova,[1,3] Nila Rafiq,[1] Hala Makhlouf,[2] Zahra Younoszai,[1] Ritambhara Agrawal,[1] and Zachary Goodman[1]

Since the initial description of nonalcoholic steatohepatitis (NASH), several sets of pathologic criteria for its diagnosis have been proposed. However, their interprotocol agreement and ability to predict long-term liver-related mortality (LRM) have not been demonstrated. In this study, we examined patients with biopsy-proven nonalcoholic fatty liver disease (NAFLD) for whom liver biopsy slides and clinical and mortality data were available. Liver biopsy samples were evaluated for a number of pathologic features and were classified according to the presence or absence of NASH by (1) the original criteria for NAFLD subtypes, (2) the nonalcoholic fatty liver disease activity score (NAS), (3) the Brunt criteria, and (4) the current study's criteria. All NASH diagnostic criteria and individual pathologic features were tested for agreement and for their independent associations with LRM, which were determined with a Cox proportional hazards model. Two hundred fifty-seven NAFLD patients with complete data were included. The diagnoses of NASH by the original NAFLD subtypes and by the current study's definition of NASH were in almost perfect agreement (κ = 0.896). However, their agreement was moderate with NAS (κ = 0.470 and κ = 0.511, respectively) and only fair to moderate with the Brunt criteria (κ = 0.365 and κ = 0.441, respectively). Furthermore, the agreement of the Brunt criteria with NAS was relatively poor (κ = 0.178). During the follow-up (median = 146 months), 31% of the patients died (9% were LRM). After we controlled for confounders, a diagnosis of NASH by the original criteria for NAFLD subtypes [adjusted hazard ratio = 9.94 (95% confidence interval = 1.28-77.08)] demonstrated the best independent association with LRM. Among the individual pathologic features, advanced fibrosis showed the best independent association with LRM [adjusted hazard ratio = 5.68 (95% confidence interval = 1.50-21.45)]. Conclusion: The original criteria for NAFLD subtypes and the current study's criteria for NASH were in almost perfect agreement, but their level of agreement with the NAS and Brunt criteria was lower. A diagnosis of NASH by the original criteria for NAFLD subtypes demonstrated the best predictability for LRM in NAFLD patients. (HEPATOLOGY 2011;53:1874-1882)

See Editorial on page 1792

Nonalcoholic fatty liver disease (NAFLD) is a clinicopathologic spectrum that ranges from simple steatosis to nonalcoholic steatohepatitis (NASH).[1-3] Although the incidence of NAFLD in the US population has been estimated to be 15% to 30%, only 2% to 3% have the potentially progressive subtype of NAFLD or NASH.[3-5] A number of natural history studies have convincingly shown that among patients on the NAFLD spectrum, only those with NASH are at risk for progression.[1,6-14] Because of this

Abbreviations: aHR, adjusted hazard ratio; CI, confidence interval; HR, hazard ratio; ICD-9, International Classification of Diseases, 9th revision; ICD-10, International Classification of Diseases, 10th revision; LRM, liver-related mortality; NAFLD, nonalcoholic fatty liver disease; NAS, nonalcoholic fatty liver disease activity score; NASH, nonalcoholic steatohepatitis.

From the [1]Center for Liver Diseases, Inova Fairfax Hospital, Falls Church, VA; [2]Armed Forces Institute of Pathology, Washington, DC; and [3]Betty and Guy Beatty Center for Integrated Research, Inova Health System, Falls Church, VA.

Received January 12, 2011; accepted February 16, 2011.

This study was supported in part by the Liver Disease Outcomes Fund of the Center for Liver Diseases at Inova Fairfax Hospital (Falls Church, VA).

Potential conflict of interest: Nothing to report.

Fax:

differential progression of NAFLD subtypes, establishing the diagnosis of NASH is important both for prognosis and for the identification of potential candidates for future treatment protocols.

In order to establish the diagnosis of NASH, a number of pathologic criteria have been used. Among these, the original criteria for NAFLD subtypes were developed to histologically categorize NAFLD into four subtypes. Specifically, NAFLD subtypes 3 and 4 are now considered to represent NASH.[2,6,15] Subsequently, the Brunt criteria were developed to grade NASH, and they have been used for clinical research in patients with NAFLD.[16] More recently, the nonalcoholic fatty liver disease activity score (NAS) was developed to provide a numerical pathologic score for patients who most likely have NASH.[17]

Over the past decade, these different pathologic criteria have been used to carry out epidemiologic studies or to assess the efficacy of different medications in clinical trials of patients with NASH. Despite their increasing use, the interprotocol agreements of these pathologic criteria have not been assessed. Additionally, the ability of these NASH pathologic criteria to predict adverse outcomes such as liver-related mortality (LRM) has not been assessed. The aims of this study were (1) to assess the agreement of three commonly used pathologic criteria for NASH along with our own pathologic criteria (used in the current study) and (2) to determine their ability to predict long-term LRM in a well-defined cohort of patients with NAFLD for whom liver biopsy slides and mortality data were available.

Patients and Methods

Patient Population. Patients with histologically proven NAFLD, available liver biopsy slides, and adequate clinical information were selected from our fatty liver databases. This NAFLD cohort included patients with available clinical data and liver biopsy slides from the Armed Forces Institute of Pathology (Washington DC) as well as the original NAFLD patients whom we previously reported.[6] For each patient, clinical and demographic data were available (age, sex, race, height, weight, alcohol consumption, medications, presence of diabetes, presence of hyperlipidemia, and results of laboratory tests measuring liver enzymes). The height and the weight were used to calculate the body mass index. To be included in the study, a patient had to have been diagnosed with biopsy-proven NAFLD with a minimum of 5 years of follow-up. Patients were excluded for the following reasons: (1) a daily alcohol intake greater than 20 g in men and greater than 10 g in women; (2) another form of chronic liver disease such as viral hepatitis, autoimmune hepatitis, or medication-induced liver disease; (3) the use of medications associated with fatty liver disease; (4) bariatric surgery or small bowel resection; (5) total parenteral nutrition; and (6) an active or recent malignancy.

The study was approved by the institutional review boards of Inova Health System and the Armed Forces Institute of Pathology.

Pathologic Assessment. For the purpose of this study, all liver biopsy slides were reread at the same time by two hepatopathologists (Z.G. and H.M.) who were blinded to the clinical data. For each liver biopsy, slides stained with hematoxylin-eosin and Masson's trichrome were reviewed in conference by both hepatopathologists (Z.G. and H.M.), and decisions about each pathologic feature and the diagnosis of NASH were made by consensus. Steatosis was scored as an estimate of the percentage of parenchyma replaced by fat: (0) 0%, (1) up to 5%, (2) 6% to 33%, (3) 34% to 66%, or (4) more than 66%. Lobular inflammation, portal inflammation, hepatocellular ballooning, pericellular/perisinusoidal fibrosis, and portal fibrosis were graded on a scale of 0 to 3: (0) none, (1) mild or few, (2) moderate, or (3) marked or many. Bridging fibrosis was scored as (0) none, (1) few bridges, or (2) many bridges. Cirrhosis was scored as (0) absent, (1) incomplete, or (2) established.

Four pathologic protocols or sets of criteria were used to assess each liver biopsy sample. First, we cate- gorized each liver biopsy into one of the following four groups according to the original criteria for NAFLD subtypes[2,15]: (1) steatosis alone (NAFLD type 1), (2) steatosis with lobular inflammation only (NAFLD type 2), (3) steatosis with hepatocellular ballooning (NAFLD type 3), or (4) steatosis with Mallory-Denk bodies or fibrosis (NAFLD type 4). As previously reported, NAFLD types 3 and 4 were con- sidered to be NASH.[6] Furthermore, each liver biopsy sample with at least fat and lobular inflammation was further graded as mild (grade 1), moderate (grade 2), or marked (grade 3) as described by Brunt et al.[16] For the purpose of this study, patients with Brunt grades

Address reprint requests to: Zobair M. Younossi, M.D., M.P.H., Betty and Guy Beatty Center for Integrated Research, Inova Health System, 3300 Gallows Road, Falls Church, VA 22042. E-mail: zobair.younossi@inova.org; Fax: .

Copyright Vc 2011 by the American Association for the Study of Liver Diseases.

View this article online at wil .

DOI 10.1002/hep.24268

differential progression of NAFLD subtypes, establishing the diagnosis of NASH is important both for prognosis and for the identification of potential candidates for future treatment protocols.

In order to establish the diagnosis of NASH, a number of pathologic criteria have been used. Among these, the original criteria for NAFLD subtypes were developed to histologically categorize NAFLD into four subtypes. Specifically, NAFLD subtypes 3 and 4 are now considered to represent NASH.[2,6,15] Subsequently, the Brunt criteria were developed to grade NASH, and they have been used for clinical research in patients with NAFLD.[16] More recently, the nonalcoholic fatty liver disease activity score (NAS) was developed to provide a numerical pathologic score for patients who most likely have NASH.[17]

Over the past decade, these different pathologic criteria have been used to carry out epidemiologic studies or to assess the efficacy of different medications in clinical trials of patients with NASH. Despite their increasing use, the interprotocol agreements of these pathologic criteria have not been assessed. Additionally, the ability of these NASH pathologic criteria to predict adverse outcomes such as liver-related mortality (LRM) has not been assessed. The aims of this study were (1) to assess the agreement of three commonly used pathologic criteria for NASH along with our own pathologic criteria (used in the current study) and (2) to determine their ability to predict long-term LRM in a well-defined cohort of patients with NAFLD for whom liver biopsy slides and mortality data were available.

Patients and Methods

Patient Population. Patients with histologically proven NAFLD, available liver biopsy slides, and adequate clinical information were selected from our fatty liver databases. This NAFLD cohort included patients with available clinical data and liver biopsy slides from the Armed Forces Institute of Pathology (Washington DC) as well as the original NAFLD patients whom we previously reported.[6] For each patient, clinical and demographic data were available (age, sex, race, height, weight, alcohol consumption, medications, presence of diabetes, presence of hyperlipidemia, and results of laboratory tests measuring liver enzymes). The height and the weight were used to calculate the body mass index. To be included in the study, a patient had to have been diagnosed with biopsy-proven NAFLD with a minimum of 5 years of follow-up. Patients were excluded for the following reasons: (1) a daily alcohol intake greater than 20 g in men and greater than 10 g in women; (2) another form of chronic liver disease such as viral hepatitis, autoimmune hepatitis, or medication-induced liver disease; (3) the use of medications associated with fatty liver disease; (4) bariatric surgery or small bowel resection; (5) total parenteral nutrition; and (6) an active or recent malignancy.

The study was approved by the institutional review boards of Inova Health System and the Armed Forces Institute of Pathology.

Pathologic Assessment. For the purpose of this study, all liver biopsy slides were reread at the same time by two hepatopathologists (Z.G. and H.M.) who were blinded to the clinical data. For each liver biopsy, slides stained with hematoxylin-eosin and Masson's trichrome were reviewed in conference by both hepatopathologists (Z.G. and H.M.), and decisions about each pathologic feature and the diagnosis of NASH were made by consensus. Steatosis was scored as an estimate of the percentage of parenchyma replaced by fat: (0) 0%, (1) up to 5%, (2) 6% to 33%, (3) 34% to 66%, or (4) more than 66%. Lobular inflammation, portal inflammation, hepatocellular ballooning, pericellular/perisinusoidal fibrosis, and portal fibrosis were graded on a scale of 0 to 3: (0) none, (1) mild or few, (2) moderate, or (3) marked or many. Bridging fibrosis was scored as (0) none, (1) few bridges, or (2) many bridges. Cirrhosis was scored as (0) absent, (1) incomplete, or (2) established.

Four pathologic protocols or sets of criteria were used to assess each liver biopsy sample. First, we categorized each liver biopsy into one of the following four groups according to the original criteria for NAFLD subtypes[2,15]: (1) steatosis alone (NAFLD type 1), (2) steatosis with lobular inflammation only (NAFLD type 2), (3) steatosis with hepatocellular ballooning (NAFLD type 3), or (4) steatosis with Mallory-Denk bodies or fibrosis (NAFLD type 4). As previously reported, NAFLD types 3 and 4 were considered to be NASH.[6] Furthermore, each liver biopsy sample with at least fat and lobular inflammation was further graded as mild (grade 1), moderate (grade 2), or marked (grade 3) as described by Brunt et al.[16] For the purpose of this study, patients with Brunt grades

Address reprint requests to: Zobair M. Younossi, M.D., M.P.H., Betty and Guy Beatty Center for Integrated Research, Inova Health System, 3300 Gallows Road, Falls Church, VA 22042. E-mail: zobair.younossi@inova.org; Fax:

Copyright © 2011 by the American Association for the Study of Liver Diseases.

View this article online at wil

DOI 10.1002/hep.24268

Table 1. j Scores and Strength of Agreement

j	Interpretation
<0	Poor agreement
0.0-0.20	Slight agreement
0.21-0.40	Fair agreement
0.41-0.60	Moderate agreement
0.61-0.80	Substantial agreement
0.81-1.00	Almost perfect agreement

of 1 to 3 were combined and were considered to have NASH. Next, we used the current study's pathologic criteria for NASH.[18] According to these criteria, NASH was diagnosed for (1) any degree of steatosis along with centrilobular ballooning and/or Mallory-Denk bodies or (2) any degree of steatosis along with centrilobular pericellular/perisinusoidal fibrosis or bridging fibrosis in the absence of another identifiable cause. Finally, for all liver biopsy samples, the elements of NAS and the stage of fibrosis were scored as described by Kleiner et al.[17] with separate scores for steatosis (0-3), hepatocellular ballooning (0-2), lobular inflammation (0-3), and fibrosis (0-4). As recommended, NAS was the sum of the first three features. Fibrosis according to the NAS was scored from 0 to 4 [(0) none, (1) centrilobular/perisinusoidal, (2) centrilobular plus periportal, (3) bridging, and (4) cirrhosis]. Each biopsy sample was examined separately according to these four pathologic criteria, and the readings were recorded into the database.

Mortality Follow-Up. For each patient, the long- term mortality status at the time of the study and the cause of death were obtained from the National Death Index Plus. Maintained by the Center for Disease Control, the National Death Index is a computerized database of all certified deaths in the United States since 1979. In addition to the mortality status, the mortality files contain the dates and causes of death. According to the National Death Index database, peo- ple who died in the United States before 1998 were classified according to the guidelines of the Interna- tional Classification of Diseases, 9th revision (ICD-9), whereas those who died during or after 1998 were classified according to the guidelines of the International Classification of Diseases, 10th revision (ICD-10).[19] In the current study, the causes of death classified as LRM included liver fibrosis and cirrhosis (ICD-10 code K74), chronic liver disease and sequelae of chronic liver disease (ICD-9 code 571-572), liver cell carcinoma (ICD-9 code 155.0 and ICD-10 code C22.0), and hepatic failure (ICD-10 code K72).

Statistical Analysis. The main long-term outcome for this study was LRM. Demographic, clinical, and laboratory characteristics were compared between NAFLD subjects who were dead and those who were alive at the time of follow-up and between LRM and non-LRM groups. The chi-square test or Fisher's exact test was used to compare categorical variables, and nonparametric Mann-Whitney tests were used to compare continuous variables. P values below the level of 0.05 were considered significant.

As described previously, four different sets of pathologic criteria were used to establish the diagnosis of NASH. After the agreement between these different pathologic criteria for NASH was assessed with j statistics[20] (Table 1), each diagnostic criterion for NASH was tested separately for its ability to independently predict LRM (after adjustments for relevant demographic and clinical confounders). Cox proportional hazards models were used to calculate adjusted hazard ratios (aHRs) and to identify independent predictors of LRM, and aHRs with P values 0.05 were considered potentially significant. Furthermore, two different schemes for grading fibrosis (NAS and the current study's criteria) were tested for agreement with Spearman's correlation coefficient.

Next, the individual pathologic features that constituted each of the four pathologic protocols were tested individually for their ability to independently predict LRM. Because in the original criteria for NAFLD these features were described in the form of ordinal variables, series of tests were performed to identify the thresholds for transforming each pathologic feature into a binary classifier. The threshold for each patho- logic feature was selected so that the log-rank test for LRM associated with a transformed binary pathologic feature returned the highest P value in comparison with other possible thresholds for that feature. The transformation of ordinal features into binary features was supposed to eliminate nonequidistant distributions of the degrees of pathologic processes described by the respective pathologic features. The transformed binary pathologic features were further used for building a multivariate survival model for LRM with the aims of (1) identifying those features independently predicting LRM and (2) establishing potential ways of improving existing pathologic protocols.

All analyses were performed with SAS 9.1 (SAS Institute, Inc., Cary, NC).

Results

Liver biopsy slides and clinical data were available for 257 NAFLD patients (67% with NASH and 33% with non-NASH NAFLD); 142 of these patients were from the Armed Forces Institute of Pathology, 72 were from our previously reported NAFLD cohort, and 43

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：https://d.book118.com/907035131052006046