Scoring Systems for Liver Disease, Their Relevance and New Developments in Scoring Biliary Diseases

Stefan Hübscher
Institute of Immunology & Immunotherapy, University of Birmingham
Department of Cellular Pathology, Queen Elizabeth Hospital, Birmingham, UK

Scoring Systems for Liver Disease, Their Relevance and New Developments in Scoring Biliary Diseases

1. **Histological scoring in the assessment of liver disease**
   - **General principles**
   - Clinical relevance
   - Problems with histological scoring
   - Changing role of liver biopsy in assessing disease severity
2. Histological scoring in the assessment of chronic biliary disease (PBC and PSC)

Scoring Systems for Liver Disease, Their Relevance and New Developments in Scoring Biliary Diseases

1. Histological scoring in the assessment of liver disease
   - General principles
   - Clinical relevance
   - Problems with histological scoring
   - Changing role of liver biopsy in assessing disease severity
2. Histological scoring in the assessment of chronic biliary disease (PBC and PSC)

Histological Scoring in Medical Liver Diseases
General Principles

- Reporting of medical liver biopsies was (and still is) subjective, descriptive and opinion-based.
- Scoring systems devised to supplement descriptive reports with semi-quantitative assessments of disease severity.
- Intended to improve objectivity and discriminatory power of liver biopsy, with implications for:
  - Prognosis
  - Treatment
  - Monitoring disease progression/outcomes (clinical trials)

Histological Scoring in Medical Liver Diseases
Grading and Staging

Grading = Disease Activity

- Ongoing damage (e.g. inflammation)
- Potential to progress to chronic damage (e.g. fibrosis)
- Still potentially treatable

Staging = Disease Progression

- Progressive liver injury (e.g. fibrosis)
- May lead to end-stage disease
- Less readily reversible

What Features of Disease Activity can be Graded Histologically?

Histological Feature	Diseases to which grading can be applied
Inflammation (portal/periportal, lobular)	Viral hepatitis (HBV, HCV), autoimmune hepatitis. PBC, PSC, fatty liver disease (ALD, NAFLD), Liver allograft rejection
Hepatocyte death (apoptosis/necrosis)	Diseases associated with lobular inflammation (e.g. acute hepatitis – viral, autoimmune, other)
Steatosis	Fatty liver disease (ALD, NAFLD) Donor biopsies for liver transplantation
Hepatocyte ballooning	Fatty liver disease (ALD, NAFLD)
Bile duct inflammation	PBC, liver allograft rejection

Histological Scoring in Medical Liver Diseases
Grading and Staging

- Grading/staging features used vary according to the disease present
  - Ishak/METAVIR used for chronic hepatitis (e.g. HCV)
  - Brunt/Kleiner used for NAFLD
  - Ludwig system used for PSC
- Grading and staging should be carried out independently

(Details of scoring systems used for non-biliary diseases won't be discussed)

What Features of Disease Progression can be Staged Histologically?

Histological Feature	Diseases to which staging can be applied
Fibrosis (portal/periportal, lobular)	Viral hepatitis (HBV, HCV), autoimmune hepatitis. PBC, PSC, fatty liver disease (ALD, NAFLD),
Bile duct loss	PBC, PSC, other ductopenic liver diseases Liver allograft rejection
Copper/ copper associated protein accumulation	PBC, PSC, other ductopenic liver diseases

Scoring Systems for Liver Disease, Their Relevance and New Developments in Scoring Biliary Diseases

1. **Histological scoring in the assessment of liver disease**
   - General principles
   - **Clinical relevance**
     - ➢ **Prognosis and treatment**
   - Problems with histological scoring
   - Changing role of liver biopsy in assessing disease severity
2. Histological scoring in the assessment of chronic biliary disease (PBC and PSC)

Histological Scoring in Liver Disease - Clinical Relevance of Grading and Staging
2. Therapeutic Implications

Histological Feature	Therapeutic Implications
Interface hepatitis	Presence/ severity of IFH is an indication for: • initiating treatment with immunosuppression in patients with newly diagnosed AIH • Continuing treatment in AIH patients with biochemical resolution after initial treatment with immunosuppression
Fibrosis stage	Early/intermediate stages may be reversible if causative agent is treated Advanced fibrosis/cirrhosis less likely to be reversible. • Patients with cirrhosis require screening for HCC

Histological Scoring in Liver Disease - Clinical Relevance of Grading and Staging
1. Prognostic Value

Histological Feature	Prognostic Significance
Interface hepatitis	Severity predicts progression to fibrosis/cirrhosis in chronic liver disease • HBV, HCV, AIH, PBC
Lobular inflammation/ necrosis	Severity of lobular necrosis predicts progression to liver failure in acute hepatitis
Fibrosis stage	Independently predicts adverse outcomes (cirrhosis, liver failure, HCC) in many chronic liver diseases • HBV, HCV, AIH, PBC, PSC, NAFLD

Scoring Systems for Liver Disease, Their Relevance and New Developments in Scoring Biliary Diseases

1. **Histological scoring in the assessment of liver disease**
   - General principles
   - Clinical relevance
   - **Problems with histological scoring**
   - Changing role of liver biopsy in assessing disease severity
2. Histological scoring in the assessment of chronic biliary disease (PBC and PSC)

Problems With Histological Scoring

1. Sampling variability

Disease heterogeneity intrinsic to all medical liver diseases

- Degree varies according to histological feature and disease process
- Studies of hepatectomy specimens obtained at liver transplantation have shown that fibrosis variability is greater in chronic biliary diseases (PBC, PSC) than in other chronic liver diseases (HBV, HCV, ALD, NAFLD)

Sampling variability influenced by biopsy length and diameter

- Short or narrow biopsies tend to underestimate both disease grade and stage

What is an "adequate biopsy"?

For accurate staging of fibrosis in chronic HCV infection;

- At least 20 - 25 mm long
- At least 1.4 mm diameter
- At least 11 complete portal tracts

Problems With Histological Scoring

3. Interpretation of histological scores

Scores used in grading/staging are categorical (not quantitative) variables:

- e.g. Ishak fibrosis stage F6 does not mean 3x more fibrosis than F2

Adding scores for different histological features may not be appropriate:

- e.g. portal inflammation + interface hepatitis +lobular inflammation + confluent necrosis (Ishak HAI Score)

Statistical methods appropriate for categorical variables should be used

Problems With Histological Scoring

2. Observer variability

Observer agreement (e.g. NAFLD, HCV)

- good for fibrosis, steatosis
- moderate for inflammation
- slight-to-fair for ballooning

Observer reproducibility may be improved by:

- Experienced/expert liver pathologists
- Paired observers reaching consensus (simultaneous or independent)
- Pathologists working together prior to scoring to establish how to implement instructions for scoring ("tuning the violins")

Scoring Systems for Liver Disease, Their Relevance and New Developments in Scoring Biliary Diseases

1. **Histological scoring in the assessment of liver disease**
   - General principles
   - Clinical relevance
   - Problems with histological scoring
   - **Changing role of liver biopsy in assessing disease severity**
     - **e.g. changes in the diagnosis and management of HCV and NAFLD**
2. Histological scoring in the assessment of chronic biliary disease (PBC and PSC)

Assessment of Disease Severity in Chronic Liver Disease
Changing Role of Liver Biopsy

Hepatitis C

1. Non-invasive tools (mainly FibroScan) have reduced the need for liver biopsy to assess disease severity (fibrosis)
2. Novel direct-acting antiviral agents are effective in treating HCV irrespective of disease severity

➢ Liver biopsies now very rarely obtained from patients with uncomplicated HCV infection

Scoring Systems for Liver Disease, Their Relevance and New Developments in Scoring Biliary Diseases

1. Histological scoring in the assessment of liver disease
   - General principles and clinical relevance

2. **Histological scoring in the assessment of chronic biliary disease (PBC and PSC)**
   - **Grading**
   - **Staging**

Assessment of Disease Severity in Chronic Liver Disease
Changing Role of Liver Biopsy

NAFLD

- Non-invasive markers (e.g. NAFLD Fibrosis Score, FibroScan) have reduced the need for liver biopsy to assess disease severity (fibrosis)

Current Indications for Liver Biopsy in NAFLD:

1. Cases where non-invasive investigations have produced an indeterminate / discrepant / unexpected score for fibrosis.
2. Cases where there are concerns about an additional aetiology for liver disease.

Changing Role of Liver Biopsy in PBC and PSC

Liver biopsy no longer required for routine diagnosis:

- PBC diagnosed by cholestatic liver biochemistry (raised Alk Phos) and immunology (anti-mitochondrial antibodies, raised IgM)
- PSC diagnosed by cholestatic liver biochemistry (raised Alk Phos), cholangiography and exclusion of secondary causes of sclerosing cholangitis

Liver biopsy continues to be play a role in atypical cases:

i. Other diagnostic markers lacking – e.g. AMA-negative PBC, small duct PSC
ii. Atypical presentation – e.g. "overlap syndrome" with features of AIH
iii. Suspected additional liver disease e.g. NAFLD
iv. Poor response to treatment with UDCA (PBC)

Scoring Systems for Liver Disease, Their Relevance and New Developments in Scoring Biliary Diseases

1. Histological scoring in the assessment of liver disease
   - General principles and clinical relevance

2. **Histological scoring in the assessment of chronic biliary disease (PBC and PSC)**
   - **Grading**
   - **Staging**

Assessing Inflammatory Activity in PBC and PSC
Hepatitis Activity

PBC

- Presence/severity of inflammatory activity in baseline biopsies predicts progression to fibrosis/cirrhosis & liver failure (Degott 1999, Corpechot 2002, Corpechot 2008, Carbone 2015)

PSC

- Relevance of inflammatory activity for predicting disease progression less clearly established

➢ Histological assessment of inflammatory activity important in the diagnosis of autoimmune "overlap syndromes"

Assessing Inflammatory Activity in PBC and PSC
Cholangitis Activity

- One scoring system proposed for assessing "Cholangitis Activity" in PBC (Hiramatsu 2006, modified in Nakanuma 2010)
  - CA 0 = None
  - CA 1 = One evident chronic cholangitis
  - CA 2 = More than two bile ducts with evident chronic cholangitis
  - CA 3 = At least one destructive duct lesion
- No obvious implications for prognosis or treatment

Autoimmune "Overlap Syndromes"
Problems with Classification / Terminology

- Approximately 5-15% of patients with PBC and PSC have additional features supporting a diagnosis of autoimmune hepatitis (AIH)
- Most cases with "overlap features" (PBC/AIH or PSC/AIH) are best regarded as variants of PBC or PSC, rather than as distinct entities
- Terms such as "PBC/PSC with hepatitic features" or "PBC/PSC with AIH-like features" may be more appropriate

PBC/AIH "overlap syndrome" (PBC with AIH-like features)
Diagnostic Criteria

Diagnostic Criteria for AIH (Chazouilleres 1998, IAHG guidelines 2011, EASL Guidelines 2017)

Any 2 of the following 3 features:

1. ALT ≥5x upper limit of normal (ULN)
2. IgG levels ≥2x ULN OR SMA positive
3. Liver biopsy **with at least moderate interface hepatitis**

- Uncertainties about diagnosing "overlap syndrome" mean that current guidelines recommend histological confirmation of interface hepatitis as "mandatory for diagnosis"
- No clear definition of "moderate interface hepatitis" provided
- No use of formal scoring systems

PSC/AIH Overlap Syndrome
(Beuers 2009, Chapman 2010, Boberg 2011, Czaja 2012, Schulze 2015)

Diagnostic Criteria

- Similar to those used for PBC/AIH overlap (but less clearly defined)

Clinical Significance

- Current guidelines documents (AASLD, EASL, IAIHG) recommend treatment with immunosuppression
- Overall prognosis better than pure PSC (but worse than patients with pure AIH)

PBC/AIH "overlap syndrome" (PBC with "hepatitic features")
Clinical Significance

- PBC with "hepatitic features" - worse outcome than "pure" PBC
- May benefit from treatment with immunosuppression
  - Less fibrosis progression, better short-term outlook
  - No clear effect on long-term outcomes (Zhang 2013, Trivedi 2017)
- Recent EASL PBC Guidelines (J Hepatol 2017) suggest immunosuppression should be stratified according to severity of interface hepatitis:
  - Severe – definitely treat
  - Moderate – consider treatment
  - Mild - uncertain

Scoring Systems for Liver Disease, Their Relevance and New Developments in Scoring Biliary Diseases

1. Histological scoring in the assessment of liver disease
   - General principles and clinical relevance

2. **Histological scoring in the assessment of chronic biliary disease (PBC and PSC)**
   - **Grading**
   - **Staging**

Staging Systems in PBC and PSC

	Rubin (1965) PBC	Scheuer (1967) PBC	Popper (1970) PBC	Ludwig (1978) PBC	Ludwig (1981) PSC	Simplified System
Stage 1	Bile duct damage	Florid duct lesion	Cholangitis	Portal hepatitis or cholangitis	Portal hepatitis or cholangitis	No fibrosis
Stage 2	Ductular proliferation	Ductular proliferation	Ductular proliferation & destruction	Periportal hepatitis or fibrosis	Periportal hepatitis or fibrosis	Periportal fibrosis
Stage 3	Ductular proliferation	Scarring	Pre-cirrhotic fibrosis	Bridging necrosis or fibrosis	Bridging necrosis or fibrosis	Bridging fibrosis
Stage 4	Cirrhosis	Cirrhosis	Cirrhosis	Cirrhosis	Cirrhosis	Cirrhosis

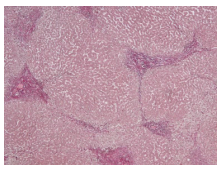
- Above scoring systems all devised at a time when liver biopsies obtained for routine diagnosis
- Incorporate features relating to disease activity (cholangitis, portal hepatitis, periportal hepatitis) as well as disease stage (fibrosis/ cirrhosis)

Staging Systems in PBC and PSC

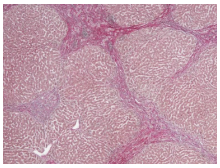
	Rubin (1965) PBC	Scheuer (1967) PBC	Popper (1970) PBC	Ludwig (1978) PBC	Ludwig (1981) PSC	Simplified System
Stage 1	Bile duct damage	Florid duct lesion	Cholangitis	Portal hepatitis or cholangitis	Portal hepatitis or cholangitis	No fibrosis
Stage 2	Ductular proliferation	Ductular proliferation	Ductular proliferation & destruction	Periportal hepatitis or fibrosis	Periportal hepatitis or fibrosis	Periportal fibrosis
Stage 3	Ductular proliferation	Scarring	Pre-cirrhotic fibrosis	Bridging necrosis or fibrosis	Bridging necrosis or fibrosis	Bridging fibrosis
Stage 4	Cirrhosis	Cirrhosis	Cirrhosis	Cirrhosis	Cirrhosis	Cirrhosis

- Scheuer & Ludwig staging systems independently predict adverse outcomes in patients with PBC and PSC
- Prognostic scoring systems incorporating histological stage widely used for many years in the assessment and management of patients with PBC and PSC

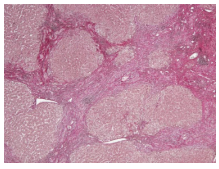
Chronic Biliary Disease – Progression of Periportal Fibrosis



Mild – periportal fibrosis (Ludwig stage 2)



Moderate - bridging fibrosis (Ludwig stage 3)



Severe –cirrhosis (Ludwig stage 4)

Decline of Liver Biopsy for Staging PBC and PSC

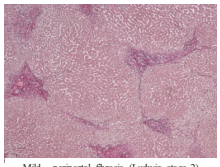
1. Sampling Variability

- Studies of paired liver biopsies and hepatectomy specimens obtained at LT showed variation in stage in 27-52% of cases (Olsson 1995, Garrido & Hubscher 1996)

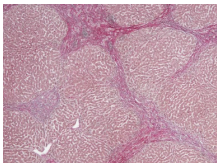
2. Liver biopsy no longer required for routine diagnosis

➢ Subsequent studies have devised prognostic scores not using histological stage

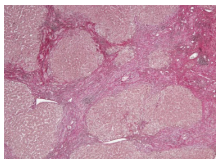
Chronic Biliary Disease – Progression of Periportal Fibrosis



Mild – periportal fibrosis (Ludwig stage 2)



Moderate - bridging fibrosis (Ludwig stage 3)



Severe –cirrhosis (Ludwig stage 4)

All 3 images are from the same block of a PSC liver obtained at transplantation

Proposal of a new staging and grading system for PBC
(Hiramatsu 2006, Nakanuma 2010)

Grading Features
1. Cholangitis Activity (0-3)
2. Hepatitis Activity (0-3)

Staging Features	**Total Staging Score (0-9)**
1.Bile duct loss (0-3)	0 = stage 1 (no progression)
2.Fibrosis (0-3)	1-3 = stage 2 (mild progression)
3.Orcein-positive granules (0-3)	4-6 = stage 3 (moderate progression)
	7-9 = stage 4 (advanced disease)

Staging features correlated with biochemical abnormalities & with Mayo Risk Score (grading features – no correlation)

Staging features could also be applied to PSC

Recent Developments in Grading & Staging PBC

Can these be applied in PSC?

Japanese (Nakanuma) System for Grading & Staging PBC – Further Studies

- Subsequent studies have shown that Nakanuma system for staging PBC correlated better with adverse outcomes than other scoring systems such as Scheuer (1967) and Ludwig (1978) (Harada 2013, Kakuda 2013, Chan 2014)
  - CAP deposition score is most powerful prognostic feature (Kakuda 2013, Chan 2014)
- Nakanuma scoring system also predicts response to treatment with UDCA (Nakamura, Hepatol Res 2015)
  - Higher bile duct loss score = risk factor for worse Alk Phos response
  - Higher hepatitis score = risk factor for worse IgM response

BUT:

- One recent study found that Scheuer stage better than Nakanuma in predicting prognosis and treatment response (Namisaki 2017)

Another Histological Scoring System for PBC

(Wendum, Liver International 2015)

- **Liver biopsies from 33 newly diagnosed PBC patients**
- **Semi-quantitative scoring of features thought to be prognostically important:**

1. Fibrosis (F0 – F4)
2. Interface hepatitis (0-3)
3. Bile duct ratio (number of PTs with BD/total number of PTs)

- **Results (compared with Scheuer & Ludwig systems)**
  - Better inter-observer agreement for each of the 3 features assessed
  - No liver-related events to assess prognostic value

Validation of the Prognostic Value of Histologic Scoring Systems in Primary Sclerosing Cholangitis: An International Cohort Study

Elisabeth M. G. de Vries,[1] Manon de Krijger,[1] Marrit Fiddelia,[2] Johanna Arola,[3] Peter Schirmacher,[4] Daniel Gotthardt,[5] Benjamin Goeppert,[6] Pelin J. Terwil,[7] Gideon M. Hirschfield,[8] Henriette Yeing,[1] Ben Vainer,[9] Henk R. van Buuren,[2] Katharina Beermann,[10] Marrit H. Horns,[1] Olivier Chazouillères,[11] Dominique Wendum,[12] Astrid D. Kengsing,[13] Roger W. Chapman,[13,14] Lai Mun Wang,[15,16] Kate D. Williamson,[15,16] Annette S. H. Gouw,[17] Valerie Paradis,[18] Christine Sempoux,[19] Ulrich Beuers,[1] Stefan G. Hübscher,[3,20] Joanne Verheij,[17] and Cyriel Y. Ponsioen[1,*]

(HEPATOLOGY 2017;65:907-919)

119 patients from 7 European centres, median follow-up 11.3 years

- Baseline liver biopsies assessed by two observers in tandem
  - Ludwig stage, Ishak grade & stage, Nakanuma grade & stage
- Staging components of all 3 scoring systems associated with adverse outcomes:

Scoring System	Hazard Ratios		
	PSC-related death/ liver transplantation	Liver transplantation	Liver-related events
Ludwig	NS	2.62	2.16
Ishak	NS	1.55	1.48
Nakanuma	2.14	3.16	2.49

Prognostic value of individual components (Nakanuma system)

- CAP > fibrosis > duct loss (NS)

Applicability and prognostic value of histologic scoring systems in primary sclerosing cholangitis

Elisabeth M.G. de Vries[1], Joanne Verheij[2], Stefan G. Hubscher[3], Mariska M.G. Leeflang[4], Kirsten Boonstra[1], Ulrich Beuers[1], Cyriel Y. Ponsioen[1,*]

Journal of Hepatology 2015: 63:1212-1219

64 patients with PSC (Dutch cohort), median follow-up 9.3 years

- Baseline liver biopsies assessed by two observers in tandem
  - Ludwig stage, Ishak grade & stage, Nakanuma grade & stage
- Staging components of all 3 scoring systems associated with adverse outcomes:

Scoring System	Hazard Ratios	
	Transplant-free Survival	Time to LT
Ludwig	1.94	3.13
Ishak	2.56	4.18
Nakanuma	6.53	7.05

Prognostic value of individual components (Nakanuma system)

- Fibrosis and CAP scores predict transplant-free survival and time to LT
- Bile duct loss not predictive

Validation of the Prognostic Value of Histologic Scoring Systems in Primary Sclerosing Cholangitis: An International Cohort Study

Elisabeth M. G. de Vries,[1] Manon de Krijger,[1] Marrit Fiddelia,[2] Johanna Arola,[3] Peter Schirmacher,[4] Daniel Gotthardt,[5] Benjamin Goeppert,[6] Pelin J. Terwil,[7] Gideon M. Hirschfield,[8] Henriette Yeing,[1] Ben Vainer,[9] Henk R. van Buuren,[2] Katharina Beermann,[10] Marrit H. Horns,[1] Olivier Chazouillères,[11] Dominique Wendum,[12] Astrid D. Kengsing,[13] Roger W. Chapman,[13,14] Lai Mun Wang,[15,16] Kate D. Williamson,[15,16] Annette S. H. Gouw,[17] Valerie Paradis,[18] Christine Sempoux,[19] Ulrich Beuers,[1] Stefan G. Hübscher,[3,20] Joanne Verheij,[17] and Cyriel Y. Ponsioen[1,*]

(HEPATOLOGY 2017;65:907-919)

Interobserver Reproducibility

- 73/119 biopsies scored by 6 other liver pathologists, following training session

TABLE 5. Interobserver Agreement for Nakanuma Grade and Stage, Ishak Grade and Stage, and Ludwig Stage

	Interobserver Agreement, κ (Range)*
Nakanuma stage	0.56 (0.41-0.75)
Bile duct loss	0.47 (0.31-0.71)
Fibrosis	0.67 (0.57-0.81)
Orcein deposition	0.58 (0.42-0.77)
Ishak stage	0.64 (0.53-0.73)
Ludwig stage	0.62 (0.53-0.71)

Kappa scores for agreement

0–0.20= slight, 0.21–0.40 = fair, 0.41–0.60 = moderate, 0.61–0.80 = substantial, 0.81–1 = almost perfect.

Role of Liver Biopsy in Monitoring Outcomes in PBC/PSC Clinical Trials

- PBC and PSC are both relatively uncommon diseases with long natural histories
- Need to have surrogate endpoints to determine treatment outcomes

Liver biopsy still viewed as "gold standard" for assessing disease severity

- Diagnostic/prognostic utility of non-invasive markers less clearly established in chronic biliary diseases (PBC and PSC) than in non-biliary diseases (e.g. HCV, NAFLD) (Poupon 2015, Corpechot 2016, Stasi 2016)
- Non-invasive markers of fibrosis unable to detect other histological staging features relevant in disease progression (e.g. ductopenia, CAP deposits)

Role of Liver Biopsy in Scoring Disease Severity
"Beyond Grading and Staging"

Use of other methods for assessing histological disease severity

i. Morphometry for collagen proportionate area (CPA)
ii. Other computer-assisted approaches for assessing fibrosis
   - e.g. Second harmonic generation/two photon excitation fluorescence to assess texture features of collagen using unstained sections (qFibrosis algorithm)

Potential advantages over conventional histological staging scores:

- Numerical data
- More objective/reproducible
- Better at predicting outcomes
  - e.g. CPA predicts adverse outcomes better than Ishak staging in chronic HCV infection
- May allow dynamic assessment of fibrosis progression/regression
  - e.g. qFibrosis scores identify different patterns of fibrosis predicting likely fibrosis progression or regression in HBV positive patients with "stable" Ishak fibrosis scores following treatment (Sun, Nature Scientific Reports 2018)


BARCLAYS
Manager of the Month
DM