Diagnostic and prognostic scoring systems for autoimmune hepatitis: a review

O. Ducazu[1], H. Degroote[1,2], A. Geerts[1,2], A. Hoorens[3], J. Schouten[2,4], H. Van Vlierberghe[1,2], X. Verhelst[1,2]

(1) Department of Gastroenterology and Hepatology, Ghent University Hospital, Ghent, Belgium ; (2) Euopean Reference Network for Rare Liver Disease (ERN RARE LIVER) ; (3) Department of Anatomical Pathology, Ghent University Hospital, Ghent, Belgium ; (4) Department of Gastroenterology, AZ Nikolaas, Sint-Niklaas, Belgium.

Abstract

Introduction: **Auto-immune hepatitis (AIH) is a rare condition which primarily affects young women. Several diagnostic scoring systems exist based on clinical, biochemical, immunologic and histologic characteristics of AIH. Additionally, prognostic parameters can be identified. The purpose of this literature review is to compare the clinical value, strengths and limitations of these diagnostic and prognostic scoring systems.**

Methods: **A literature search was performed in two databases and selected based on diagnostic and prognostic criteria. Only studies concerning AIH in adults were included.**

Results: **The backbone of scoring systems remains the revised AIH criteria published in 1999 and the simplified from 2008. The revised system shows a higher sensitivity, lower specificity and lower diagnostic accuracy compared to the simplified. Limitations to these scoring systems include limited diagnostic accuracy in acute or fulminant liver failure, insufficient inclusion of atypical auto-antibodies and lacking diagnostic power in presence of overlap syndromes. Concerning these overlap syndromes, the Paris criteria show a higher diagnostic accuracy compared to the scoring systems for AIH. Presently, no clinical prognostic scoring systems are available. However, a first system based on response to treatment accurately predicts long-term survival in AIH.**

Conclusion: **Diagnostic scoring systems are useful in diagnosing AIH and have complementary value. However, they are no substitute for the gold standard of appropriate clinical assessment and are mostly useful in defining cohorts for research purposes. An evolution towards a more dynamic scoring system, using prognostic parameters and the progression of typical features, seems more valuable than the current diagnostic systems.** (Acta gastroenterol. belg., **2021**, 84, **487-495).**

Keywords: autoimmune hepatitis, diagnosis, scoring system, prognosis.

Introduction

Auto-immune hepatitis (AIH) is an immune-mediated inflammatory disease affecting mostly young women (71-95 %). Age of onset is seen mostly at 10-30 and 40-60 years age. It is relatively rare, with estimated global incidence of 0.67-2 per 100.000 person years and prevalence ranging from 160-180 per 1.000.000 person years (1,2,3).

Presentation is heterogeneous with multiple clinical, biochemical and histologic features (3,4). AIH can remain asymptomatic or progress into chronic hepatic inflammation, fibrosis, cirrhosis and chronic liver failure with need for transplantation or death. Additionally, first presentation with acute liver failure is possible (1,3).

Diagnosis of AIH is based on the appearance of multiple typical yet aspecific features, which are also present in other liver diseases. No clinical signs or symptoms are pathognomonic for AIH. Presentation is mostly asymptomatic with symptoms such as fatigue, fluctuating jaundice or arthralgia. In contrary, acute liver failure is also possible at first presentation (1).

Approximately 25-34 % of patients are asymptomatic at presentation; symptoms arise within two to 120 months in 26-70 % of patients. In the absence of symptoms, aspecific hepatic biochemical and histologic abnormalities can be found during this period, including features of chronic hepatic failure (3,5). Acute presentation is seen in 25-40 % of cases, mostly children and adolescents. (6).

Acute jaundice is seen in 25 % and hepatic encephalopathy in 3-6 % of North-American and European populations (5,7,8). Clinical signs are hepato-splenomegaly, ascites, epigastric pain and abdominal mass (9).

Autoimmune comorbidity is frequent: thyroiditis (10-23 %), ulcerative colitis (2-8 %), celiac disease (1-2 %), diabetes mellitus type 1 (7-9 %), rheumatoid arthritis (2-5 %), systemic lupus erythematosus (1-2%), vitiligo and psoriasis have been reported, with strongest association with Hashimoto thyroiditis (3,9).

Serological assessment typically (but not necessarily) shows hypergammaglobulinemia with elevated IgG and normal IgM and IgA (1,3). Immunohistochemistry is useful in differentiating with overlap syndromes; IgM is mostly seen in primary biliary cirrhosis (PBC) (92.3 % of cases) and IgG in AIH (90 %), primary sclerosing cholangitis (PSC) (75 %) and chronic hepatitis (100 %). Intrahepatic IgM/IgG ratio of 1 is able to differentiate between PBC and AIH with 92.3 % sensitivity and 90 % specificity (10). Some patients with PBC show a IgM/IgG ratio lower than 1, with the absence of overlap syndrome (11).

Biochemical signs of hepatitis cannot be used to assess severity of the disease; both bilirubin and transaminase levels fluctuate over time and can spontaneously normalize in histologically proven active disease (7). Exclusion of viral hepatitis in diagnosing AIH has been documented for hepatitis B virus (HBV) and hepatitis C virus (HCV) but not for other hepatotropic viruses (12).

Assessment of autoantibodies is crucial in diagnosing AIH. Type 1 AIH comprises 80 % of all cases and

Correspondence to : Prof. Dr. Xavier Verhelst, UZ Gent, Corneel Heymanslaan 10, 9000 Gent, Belgium.
E-mail: xavier.verhelst@uzgent.be

Submission date : 06/02/2020
Acceptance date : 15/10/2020

includes mostly antinuclear antibodies (ANA), smooth muscle antibodies (SMA) and anti-actin antibodies; type 2 mostly comprises anti-liver-kidney microsomal antibody 1 (anti-LKM1) and seronegativity for ANA and SMA. Type 2 is considered to have a more aggressive course (4,8). Classification in subtypes is questioned as no significant differences in clinical presentation or outcome were found (5).

ANA, SMA and anti-LKM1 serology is a first step in diagnosing AIH. ANA is found in 80 %, SMA in 63 % and anti-LKM1 in 3 % of white North-Americans suggestive for AIH at first presentation; multiple autoantibodies are found in 51 % (1). Anti-LKM1 assessment is recommended subsequent to seronegativity to ANA or SMA, due to low diagnostic sensitivity and frequent presence in patients seronegative for ANA or SMA (1,13).

Seronegativity for ANA, SMA or anti-LKM1 occurs in up to 20 % of clinical pictures suggestive for AIH. (1). Seronegativity for ANA, SMA and anti-LKM1 at first presentation with acute hepatitis should be reassessed after three to six months, due to variable serology during the course of disease (13).

Atypical autoantibodies can be assessed in seronegative AIH. These include anti-soluble liver antigen/liver-pancreas antigen (anti-SLA/LP), atypical perinuclear anti-neutrophil cytoplasmic antibodies (p-ANCA), anti-actin antibodies, anti-α-actinin antibodies, anti-liver cytosol 1 (anti-LC1) antibodies, anti-asialoglycoprotein receptor antibodies (anti-ASGPR) and double-strand DNA (ds-DNA) antibodies (1,13,14,15). Seropositivity is aspecific for AIH with important comorbidity and manifestation in healthy controls; concomitant autoantibodies may increase diagnostic accuracy (4,5). A stepwise assessment of autoantibodies is recommended, with uncommon antibodies following seronegativity for ANA, SMA and anti-LKM1 (1,2). Antimitochondrial antibodies (AMA) assessment is recommended simultaneous to common antibodies as to uncover possible overlap syndromes (2). A case report describes multiple myeloma oncogene 1 (MUM1) immunostaining to be useful when clinical picture is unclear; this could prove a contribution of immunohistochemistry in diagnosing AIH (16).

Histologic assessment through biopsy is crucial in diagnosing AIH. The most common features are interface hepatitis (Fig. 1) (70-80 % of patients) plasma cell infiltration (66 %), emperipolesis (65 %), lobular hepatitis (47 %), hepatocytic rosettes (33 %) and centrilobular necrosis (29 %). (1,17,18). Although none of the above histologic features are pathognomonic for AIH, appearance of interface hepatitis with emperipolesis, rosettes and lobular lymphoplasmacytic invasion are considered 'typical' histologic pictures of AIH (1,2,18). Histology is useful in diagnosing differential or concomitant diseases (2).

Differential diagnosis of AIH mainly consists of acute or chronic viral hepatitis, alcoholic liver disease, drug-induced liver injury (DILI), PBC, Wilson's disease,

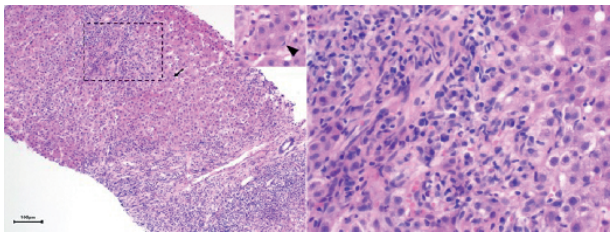


Figure 1. — Anatomopathological findings in autoimmune hepatitis. Left: Liver biopsy with dense portal inflammation with diffuse interface hepatitis, moderate lobular inflammation and hepatic rosette formation (indicated with an arrow). Inset: higher magnification of the hepatic rosette showing in addition emperipolesis (indicated with an arrowhead). Right: Higher magnification of area marked with a rectangle shows profusion of plasma cells in the inflammatory infiltrate.

hereditary hemochromatosis, non-alcoholic fatty liver disease (NAFLD) or non-alcoholic steatohepatitis (NASH), celiac disease and α1-antitrypsin deficiency. IgG4-associated cholangitis, systemic lupus erythematosus (SLE) or HIV-induced cholangiopathy show similar clinical features (2). Liver disease caused by SLE can be wrongly diagnosed as AIH and autoantibody serology is similar to both diseases; histology remains the most important assessment in differentiating SLE from other autoimmune diseases (19).

Overlap syndromes of AIH with both PBC and PSC exist and complicate differentiation and diagnosis. AIH-PBC is estimated to occur in 10 % of adults with AIH or PBC and is most commonly diagnosed using the Paris criteria; overlap syndrome should be suspected in refractory AIH as response to conventional immunomodulatory therapy is lacking (3,5). Seropositivity for AMA, elevated IgM and biliary histologic abnormalities are suggestive but aspecific for PBC; features can develop years after diagnosing AIH. Reassessment of serology for ANA and AMA is recommended in patients with persistent or new cholestatic signs or symptoms suggestive for PBC (13).

Diagnosis of AIH-PBC overlap syndrome occurs mostly using the Paris criteria, with sensitivity of 58.46 % and specificity of 99.52 % (20,21). Using these criteria, diagnosis of PBC is made when two of following criteria are met: alkaline phosphatase (ALP) levels two times above upper limit or gamma-glutamyltransferase (γ-GT) levels five times above upper limit, AMA seropositivity, florid histologic duct lesions. AIH is found when two of following criteria are met: alanine aminotransferase (ALT) levels five times above upper limit, IgG levels two times above upper limit or SMA seropositivity, histologic presence of interface hepatitis. Diagnosis of AIH-PBC overlap syndrome is made when these features are present for each disease in absence of biliary obstruction, hepatitis C, alcohol abuse and use of hepatotoxic medication (20).

Differentiation with DILI is difficult as histologic features are similar to AIH. On average, 9-12 % of patients with clinical signs of AIH will receive DILI diagnosis (2). Discontinuing immunomodulatory therapy

causes relapse in AIH; this is not the case in DILI (3). Autoimmune DILI (AI-DILI) has recently been defined as a separate disease; markers suggestive for AIH are concomitant with clinical signs of DILI. Although histopathology is similar and overlap is possible, significant chronic fibrosis is more commonly seen in AIH (22).

Diagnosis of AIH relies upon the syndromal concomitant appearance of multiple typical clinical, biochemical, immunohistochemical and histologic features. Multiple diagnostic scoring systems were developed to compile these features into a diagnostic tool useful in clinical practice (2,7,12,23).

In this article, we will study and compare different available diagnostic and prognostic scoring systems for AIH in adults. We will discuss the strengths, limitations and relative advantages of these systems, as well as their usefulness in clinical practice. AIH in children or adolescents is not included in this article.

Methods

Results were searched from two databases, Pubmed and Embase. Mainly studies between 1998 and 2019 were included. In Pubmed, articles were found using following search terms:

('autoimmune hepatitis' [Mesh]) AND ('diagnosis' [Mesh] OR 'diagnosis criteria' OR ('diagnosis' [Mesh] AND 'criteria' [Mesh]) OR 'scoring system' OR ('diagnosis' [Mesh] AND 'scoring system') OR ('diagnosis' [Mesh] AND ('scoring system' OR 'criteria')) OR ('prognosis' [Mesh] OR 'outcome') OR ('autoimmune hepatitis' [Mesh]) AND ('comparing' OR 'comparison' OR 'versus') AND ('diagnosis' [Mesh] OR 'diagnosis criteria' OR ('diagnosis' [Mesh] AND 'criteria' [Mesh])) OR 'scoring system' OR ('diagnosis' [Mesh] AND 'scoring system') OR ('diagnosis' [Mesh] AND ('scoring system' OR 'criteria'))).

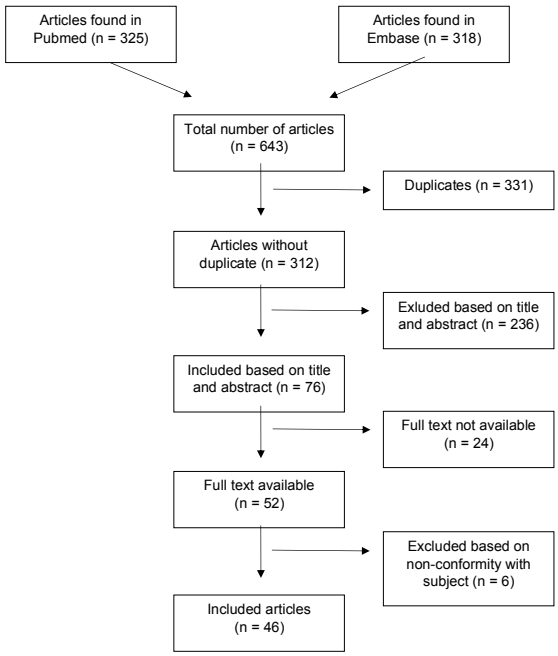
This search generated 325 results. The search terms from Pubmed were translated into applicable terms for research in Embase. Following search terms were used:

('autoimmune hepatitis'/exp OR 'autoimmune hepatitis') AND ('diagnosis'/exp OR 'diagnosis') AND ('scoring system'/exp OR 'scoring system') AND [2010-2020]/py

This search generated 318 results. Further selection was made based on duplicates, abstract and availability of full text. Articles concerning overlap syndromes were included, but only when found meaningful regarding AIH and scoring systems. Results were excluded which primarily treated overlap syndromes as such, rather than the differential diagnostic problems of overlap syndromes for the scoring systems used for AIH. Results regarding AIH in children were excluded.

A detailed flowchart of the used search strategy with the selection and number of results based on inclusion- and exclusion criteria is available as addendum A.

ADDENDUM



Results

1. Diagnosis

1.1 Scoring systems

Multiple scoring systems were developed to organise the diverse features found in AIH into a structured assessment. The International Autoimmune Hepatitis Group (IAIHG) created a first scoring system in 1993; this system was subsequently revised in 1999 and simplified in 2008 (7,12,25).

Original scoring system

The original system (Table 1) was developed by an international panel based upon expert consensus. The primary goal of this system was to provide a standardised tool for choosing adult patient populations in future research, rather than to be used in clinical practice (25).

No separate criteria were defined regarding the diagnosis or exclusion of overlap syndromes. However, it was decided that patients with histologic or cholangiographic biliary changes should not be diagnosed with AIH. The criteria above were developed as 'minimally required'; additional criteria were included regarding histologic features in AIH (Table 2) (25).

Criteria should be reassessed when the clinical presentation changes over time. As such, a patient diagnosed with 'probable' AIH can be diagnosed with 'definite' AIH later on (12). A review found a sensitivity of 54.1-81.5 % in diagnosing AIH for this scoring system. In excluding hepatitis C, a specificity of 66.1-92 % was found; the specificity in excluding AIH in patients with biliary diseases was found to be 44.8-64.9 % (7).

Table 1. — **Original diagnostic scoring system for AIH as developed in 1993**

Criterium	Score
Sex	
Female	+ 2
Male	0
AF/AST (or ALT) ratio	
> 3.0	- 2
< 3.0	+ 2
Serum globulin, γ-globulin or IgG above upper limit	
> 2.0	+ 3
1.5-2.0	+ 2
1.0-1.5	+ 1
< 1.0	0
ANA, SMA or anti-LKM1	
> 1:80	+ 3
1:80	+ 2
1:40	+ 1
< 1:40	0
AMA	
Positive	- 2
Negative	0
Viral markers	
Positive IgM HAV, HBsAg or IgM anti-HBc	- 3
Positive anti-HCV on ELISA	- 2
Positive anti-HCV on PCR	- 3
Positive test indicating active infection by any other virus	- 3
Seronegativity for all of the above	+ 3
Recent use of hepatotoxic medication or parenteral blood products	
Positive	- 2
Negative	+ 1
Mean alcohol intake in men	
< 35 g daily	+ 2
35-50 g daily	0
50-80 g daily	- 2
> 80 g daily	- 1
Mean alcohol intake in women	
< 25 g daily	+ 2
25-40 g daily	0
40-60 g daily	- 2
> 60 g daily	- 1
Other autoimmune disease in patient or first-grade relative	+ 1

Interpretation: diagnosis of 'definite' AIH with score of > 15 before and > 17 after treatment; 'probable' AIH with score of 10 - 12 before and 12 -17 after treatment. AIH = autoimmune hepatitis, AF = alkaline phosphatase, AST = aspartate transaminase, ALT = alanine transaminase, ANA = antinuclear antibodies, SMA = smooth muscle antibodies, anti-LKM = anti liver/kidney microsome antibodies, AMA = antimitochondrial antibodies, HBsAg = hepatitis B surface antigen, anti-HBc = hepatitis B core-antibody.

Revised scoring system

The revised scoring system (Table 3) was developed mainly due to the limited ability of the original system to differentiate between AIH and cholestatic diseases, resulting in unclear diagnosis (24).

Additional parameters can be assessed. Two extra points are given for seropositivity to other auto-antibodies such as p-ANCA, anti-LC1, anti-SLA, anti-ASGPR, anti-LP and anti-sulfatide. 1 point is given to positivity to HLA DR3 or DR4, 2 points to complete response to therapy and 3 points to relapse after therapy. The cut-off value post-therapy is set to > 17 points for 'definite' AIH and 12-17 points for 'probable' AIH. It should be noted that the values of these criteria were arbitrarily chosen.

Table 2. — **Additional criteria in diagnosing AIH as defined in 1993**

Criterium	Score
Histology	+ 3
Chronic active hepatitis with lobular participation and bridging necrosis	+ 2
Chronic active hepatitis without the above features	
Rosettes	+ 1
Extensive plasma cell infiltration	+ 1
Biliary changes	- 1
Other features suggesting different etiology	- 3
Antibodies (anti-SLA, anti-ASGPR, etc.) with sero-negativity for ANA, SMA, anti-LKM1	
Positive	+ 2
Negative	0
HLA B8-DR3 haplotype or DR4 allotype	+ 1
Response to treatment	
Full response	+ 2
Partial response	0
Failure of treatment	0
No response in disease activity	- 2
Relapse after initial response	+ 3

Interpretation: diagnosis of 'definite' AIH with a score of > 15 before and > 17 after treatment; 'probable' AIH with a score of 10 - 15 before and 12 - 17 after treatment. AIH = autoimmune hepatitis, anti-SLA = anti soluble liver antigen antibodies, anti-ASGPR = anti asialoglycoprotein receptor antibodies, ANA = antinuclear antibodies, SMA = smooth muscle antibodies, anti-LKM = anti liver/kidney microsome antibodies.

Table 3. — **Revised scoring system for diagnosing AIH as developed in 1999**

Criterium	Score
Female sex	+ 2
AF/AST (or ALT) ratio	
< 1.5	+ 2
1.5-3.0	0
> 3.0	- 2
Serum globulin or IgG above upper limit	
> 2.0	+ 3
1.5-2.0	+ 2
1.0-1.5	+ 1
< 1.0	0
ANA, SMA or anti-LKM1	
> 1:80	+ 3
1:80	+ 2
1:40	+ 1
< 1:40	0
AMA positive	- 4
Viral hepatitis markers	
Positive	- 3
Negative	+ 3
Medication use	
Positive	- 4
Negative	+ 1
Alcohol intake	
< 25 g daily	+ 2
> 60 g daily	- 2
Histology	
Interface hepatitis	+ 3
Lymphoplasmacytic infiltration	+ 1
Rosettes	+ 1
None of the above	- 5
Biliary changes	- 3
Other features	- 3
Other autoimmune disease in patient or first-grade relative	+ 2

Interpretation: diagnosis of 'definite' AIH with a score of > 15, 'probable' AIH with 10 - 15. AIH = autoimmune hepatitis, AF = alkaline phosphatase, AST = aspartate transaminase, ALT = alanine transaminase, ANA = antinuclear antibodies, SMA = smooth muscle antibodies, anti-LKM = anti liver/kidney microsome antibodies, AMA = antimitochondrial antibodies.

Table 4. — **Simplified scoring system for diagnosing AIH as developed in 2008**

Criterium	Score
ANA or SMA 1:40 or more 1:80 or more	 1 1
'or' anti-LKM 1:40 or more	2 (*)
'or' anti-SLA positive	2 (*)
IgG > upper limit > 1.10x upper limit	 1 2
Liver histology (**) compatible with AIH typical for AIH	 1 2
No viral features	2

Interpretation: diagnosis of 'definite' AIH with a score of 7 or more, 'probable' AIH with 6 or more. (*) When multiple auto-antibodies are present, the respective points can be accumulated up to a maximum of two additional points. (**) In liver histology, the presence of features of hepatitis is necessary; it does not get any score. The presence of interface hepatitis, emperipolesis and rosettes gains two points. One point is added when these typical features of AIH are not present, but there is evidence for histological chronic hepatitis with lymphocytic infiltration. AIH = autoimmune hepatitis, ANA = antinuclear antibodies, SMA = smooth muscle antibodies, anti-LKM = anti liver/kidney microsome antibodies, anti-SLA = anti soluble liver antigen antibodies.

As such, they do not provide insight into the severity or extent of AIH and cannot be included in statistical models (7).

Re-evaluation of a population of 114 patients diagnosed with PSC was used to implement the revised scoring system into clinical practice. Specificity of 89.5 % for the exclusion of AIH was found, opposite 64.9 % using the original system (7). A study found that 34 % of patients with chronic hepatitis with unclear etiology could be diagnosed with 'probable' or 'definite' AIH using the revised criteria; other studies had results of 19-22 % of patients (5).

Simplified scoring system

Both the original and revised scoring system were developed using mainly expert consensus. Since these criteria were found to be impractical, it was decided in 2008 to develop a simplified system based on data from 359 patients already diagnosed with AIH to provide a useful clinical instrument (Table 4) (12).

Cut-off values were defined using statistical analysis with logistic regression. When excluding patients with viral hepatitis, sensitivity of 88 % and specificity of 97 % for the diagnosis of AIH was found at cut-off of 6 points. At 7 points cut-off, sensitivity and specificity were found to be respectively 81 % and 99 % (12).

The simplified system relies heavily on histologic features. However, both emperipolesis and rosettes are difficult to recognize and have a sensitivity of 27 % in adults with AIH. The relatively low sensitivity of the simplified system as well as the low diagnostic accuracy in acute presentation could be explained by these histologic criteria, as all three criteria of interface hepatitis, emperipolesis and rosettes are needed for a score of 2 points. The necessity of the presence of all three criteria as 'typical' histologic feature in AIH could as such result in an underdiagnosis of AIH using this scoring system (17).

In developing the simplified system, priority was given to diagnosing AIH rather than excluding overlap syndromes. Nonetheless, excluding performance was found to be similar to the revised system, which aimed to more adequately exclude patients with primary biliary liver pathology (12).

1.2 Comparison

Comparing the revised and the simplified scoring system is difficult, as patient data used in developing the simplified system are based upon diagnosis of AIH using the revised system. This could generate certain bias towards the revised system (25). It was found that patients with hepatitis of unclear etiology were more likely to be diagnosed with AIH using the revised system than the simplified; respectively 95 % against 24 %. Exclusion of AIH in patients with liver disease and comorbid autoimmune pathology was more likely to occur using the simplified system than the revised; respectively 83 % against 64 % (26).

The American Association for the Study of Liver Diseases (AASLD) 2019 Guidelines describe higher sensitivity for diagnosing AIH for the revised system than the simplified (respectively 100 % against 95 %), in addition to a lower specificity (respectively 73 % against 90 %). Diagnostic accuracy is higher in the simplified system (92 % versus 82 % in the revised) (1). In addition to the histologic criteria, the exclusion of response to therapy could give another explanation for the lower sensitivity of the simplified system (5). As the revised scoring system includes the criterion of response to immunomodulatory therapy, it could be useful in diagnosing AIH in patients with a more unclear clinical presentation (3).

A comparative study found that 64. 9 % of a population of Chinese patients diagnosed with AIH using the revised system was given a decreased score using the simplified system. It was found that using 1:80 dilution or more in ANA assessment significantly reduced the variation in diagnostic accuracy, as well as positive anti-SLA or p-ANCA serology. Comorbid autoimmune diseases increased the variation in diagnostic accuracy; no significant difference was found regarding IgG serology (27,28).

Diagnostic accuracy regarding 'definite' AIH was found to be similar between the revised and simplified system in a Mexican cohort. Higher sensitivity and lower specificity was found in the revised system. The simplified system was found to be more practical and able to exclude patients with a possible diagnosis other than AIH (29). A Korean study found a sensitivity of 69.9 % and positive predictive value of 86.4 % for the simplified

system (30). A Chinese study found similar performance between revised and simplified criteria. Additionally, the simplified system was able to more specifically distinguish between 'definite' AIH or reduced diagnosis of 'probable' AIH or hepatitis with unclear etiology (31). It should be noted that AIH is ethnically diverse: disease is more progressive in Afro-Americans and cirrhosis more prevalent in patients of Latin-American heritage. This last group also reports the best prognosis and overall survival; survival is worst in Asian patients (4,32).

Seropositivity of AMA or biliary histologic features resulted in a negative score in the revised criteria; the simplified system does not include these (33). A review concluded that the revised system could not sufficiently distinguish between AIH and cholestatic syndromes (2). The simplified system was developed to include overlap syndromes. A comparative study found that five in six patients with AIH-PBC overlap syndrome were correctly diagnosed with the simplified criteria, versus one in six with the revised (25). Higher sensitivity in diagnosing overlap syndromes was found for the simplified system compared to the revised, with similar specificity. It should however be noted that neither of these criteria were developed to diagnose overlap syndromes (33). Another study found higher specificity in recognizing PBC for the simplified system than the revised, in addition to a higher liver-related mortality and poorer outcome. It was concluded that the simplified criteria could replace the revised in diagnosing PBC-AIH overlap syndrome (20).

Diagnosis of AIH following the first presentation with fulminant liver failure was made in 40 % with the revised system, against 24 % with the simplified. In another cohort, diagnosis was made in 91 % using the revised system and 40 % using the simplified. However, prospective studies regarding this subgroup are lacking (2).

Regarding diagnostic performance, AASLD considers the simplified system to be more applicable for patients with typical clinical presentation; the revised system is recommended for patients with more unclear features of AIH (1). A comparative multinational study found that 95 % of patients were correctly diagnosed with AIH using the simplified criteria (9). Another study concluded that the revised criteria remain the gold standard regarding patients with comorbid autoimmune disease or patients without typical autoantibody serology. However, it was concluded that neither the simplified nor the revised system should replace the other; combining both systems according to individual presentation and clinical features could adequately diagnose AIH (27).

1.3 Limitations

These scoring systems yield several limitations. The validity of each of these systems is subpar when used in prospective studies, as well as the diagnostic accuracy in presence of comorbid PSC, PBC, NAFLD, NASH or fulminating liver failure (1). Differentiation between AIH and DILI remains a challenge in all three scoring systems since clinical knowledge was subpar at the time (25). The simplified system does not include response to therapy as a criterion, which could be useful due to the variable nature of AIH and the importance of adequate therapy (2).

As all diagnostic scoring systems were developed using patient data from multiple specialised centres, different methods were used in detecting auto-antibodies. The lack of a standardised procedure for assessing serology with enzyme-linked immuno sorbent assay (ELISA) or immunofluorescence could generate suboptimal values. IAIHG recommends maintaining a uniform standard in serologic assessment (1,12). Furthermore, patients were mostly selected from tertiary or specialised centres resulting in significant selection bias (12).

Finally, the scoring systems are inadequately validated in prospective studies, with limited data regarding performance and utility in common clinical practice. The performance of both the revised and the simplified scoring system is measured against the gold standard of adequate clinical assessment. As such, it could be stated that these systems could never replace this clinical judgment; interpreting these scores and their implication could prove a challenge in correctly diagnosing AIH (5).

2. Prognosis

Untreated AIH has a 50 % 5-year mortality (12). Early diagnosis and adequate therapy prolong overall survival. Biochemical remission defined as normalised serum transaminases and IgG occurs in 80-90 % of patients treated with immunomodulators (34,35). Liver transplantation remains the only therapeutic alternative for non-responders to immunomodulatory therapy (36). When appropriately treated, 10-year survival of AIH is thought to be 85-95 %; two studies found no difference in overall survival with healthy controls. The 20-year survival has been unsufficiently studied (34). Spontaneous progress of the clinical picture is seen in 12 % of untreated patients with mild or asymptomatic AIH. 10-year survival of these patients was found to be lower than that of treated patients with severe symptomatology: respectively 67 % to 98 % (5).

Cirrhosis, anti-SLA/LP seropositivity and diagnosis at a young age (mostly pediatric) are all predictors of worse outcome. Diagnosis of AIH before the age of 18 is associated with higher risk of relapse after initial therapy, need for liver transplantation and lower overall survival. However, it was found that immunomodulatory treatment is discontinued more and at a younger age in children versus adults; the possible effect on patient outcome is not well known (36).

Assessment of ANA and SMA is of diagnostic rather than prognostic value and does not significantly predict outcome (4). Correlation between anti-SLA/LP seropositivity and poor outcome has generated

contradictory results; its inclusion in diagnostic scoring systems is nonetheless recommended (36). Anti-actin antibodies seropositivity is associated with higher liver failure-related mortality and need for transplantation. However, prognostic value depends on the assay used; with no standardised version available, predicting outcome could prove difficult (5).

Higher risk of liver transplantation was seen in patients who did not show at least a 50 % reduction of transaminase levels after six months of adequate therapy. Moreover, it was found that reduction of at least 80 % of transaminase levels during the first eight weeks of therapy was correlated with lower risk of liver related mortality and transplantation, as well as higher chance of remission and normalisation of biochemical AIH markers after six and 12 months therapy. In contrast, risk of liver related mortality and transplantation was highest in patients with absence of normalisation of biochemical markers after 12 months therapy. No prognostic advantage towards risk of relapse after therapy could be found for normalisation of serum transaminases versus reduction (36,37). Discontinuing immunomodulatory of treatment in patients with long-term remission of biochemical AIH markers resulted in relapse of disease or loss in biochemical remission in practically all patients (38).

Liver transplantation can be essential for patients with acute liver failure. Need for transplantation is determined using MELD-score, with transplantation needed at a score of 16 or more. Biochemical markers such as creatinine and coagulation factors are included in the model for end-stage liver disease (MELD) criteria; other markers as IgG, gamma globulins or aminotransferase were not found to have prognostic relevance (36). Post-transplantation survival is estimated to be 70 %. Relapse varies from 8-68 %; it is seen in 8-12 % of patients one year post-transplantation and 36-68 % after five years (5). Relapse of overlap syndromes post-transplantation also occurs and further complicaties differentiation with AIH (39).

'De novo' AIH occurs when features of AIH develop in donor tissue after non-AIH liver transplantation (40). It is seen in 1-7 % of patients one month to nine years after procedure (5). Correlation between persistently high titers of SMA or ANA post-transplantation and relapse of disease is reported. As antibody seropositivity can occur without failure of donor tissue, histologic assessment through biopsy is recommended (39). Cirrhosis occurs in 60 % of 'de novo' AIH cases and re-transplantation is needed in 8-50 % (5).

Our group developed a treatment-response prognostic score for AIH, based on mean survival after 5.5 years follow-up. A strong correlation was found between biochemical markers γ-GT, bilirubin, aspartate amino transaminase (AST) or ALP and decompensation, development of hepatocellular carcinoma, liver transplantation and death. ROC-curve shows a sensitivity of 75 % and specificity of 81 % for prediction of poor outcome for a patient with AIH, providing evidence that treatment-response score is an excellent predictor of long-term overall survival in AIH (41).

Discussion

Diagnosing AIH is based upon a combination of typical clinical, biochemical and histological features. The diagnostic scoring systems were developed to organize these features in a structured and standardized manner. The revised and simplified systems in particular can be applied in daily clinical practice (7,25,26).

The original system applied 'minimally required' criteria, in which no histologic features were included. The revised system included these criteria; histology is currently considered to be crucial in diagnosing AIH (2,18). The simplified system was developed to account for the cumbersome application of the revised system. Whereas the original and revised scoring system were both developed through consensus by experts with a primarily academic goal, the simplified system was developed using statistical methods and mathematical models. Its main purpose was to provide a useful instrument in daily clinical practice (7,12).

In literature, comparison is made mostly between the revised and simplified scoring systems. The revised system was found to have a higher sensitivity in diagnosing AIH compared to the simplified, respectively 100 % versus 95 %), as well as a lower specificity (respectively 73 % versus 90 %). The diagnostic accuracy was found to be higher in the simplified system as well: 92 % of patients with a clinical presentation suggestive of AIH were accurately diagnosed using this system, versus 82 % using the revised system (1,25,27,28)

Because of these differences in statistical properties, the simplified system is generally considered in clinical practice to be more applicable to patients with typical presentation suggesting AIH. The revised system on the other hand is suggested to be more usable in patients with more unclear presentation (1,2,26). Additionally, the revised system is useful in reassessing patients who were given a low diagnostic score with the simplified system. The revised system remains the most desirable for diagnosing AIH in patients with comorbid autoimmune diseases or atypical serology (27).

A limitation found in all diagnostic scoring systems is the insufficient ability to distinguish between AIH, primarily biliary diseases and overlap syndromes (1). Neither the original nor revised system include separate criteria to diagnose or exclude overlap syndromes. In developing the simplified system, overlap syndromes were included (7,12). The simplified system was found to have a higher sensitivity towards diagnosing overlap syndromes than the revised systems, with similar specificity. It was concluded that the simplified scoring system could be superior in assessing overlap syndromes (20).

It should be noted that neither the revised nor the simplified system was developed to diagnose overlap

syndromes. The most common instrument in assessing AIH-PBC overlap syndrome is given by the Paris criteria, with a sensitivity of 92 % and a specificity of 97%. These criteria could be superior to both the revised and the simplified scoring system (5,33).

Diagnostic accuracy of the scoring systems is affected by first presentation with acute or fulminant liver disease: results varied from 40-91 % with the revised to 24-40 % with the simplified scoring system (2). Evidence is lacking regarding this clinical subgroup; further research is required regarding correct and early diagnosis (1,5).

The population of patients with AIH used in developing the simplified scoring system was diagnosed with the revised system; this could mean a certain form of bias towards the revised system (25). Regarding the rare prevalence of AIH, patients are mostly selected in specialised or tertiary centres; this selection bias could provide difficulties in diagnosing AIH in common clinical practice (2). Both the revised and the simplified system are inadequately validated in prospective studies; more research is needed towards performance of these systems (1).

Clinical features such as serology, cirrhosis or transaminase levels are all variable over the course of the AIH and are useful in assessing the severity, extent or therapeutic response of the disease over time. Additionally, overlap syndromes could be versatile and arise only after diagnosing AIH. As such, regular reassessment of clinical features is recommended according to the clinical presentation (1,13). Histologic assessment is considered to be crucial in diagnosing AIH (17,18).

Dynamic systems regarding the evolution of the disease could be useful in managing AIH or overlap syndromes. In both the original and revised scoring system, response to therapy was included as an 'additional' criterion; the simplified system does not include this (12). For now, no dynamic tools predicting the evolution of a suggestive clinical picture are available; this could prove an important subject for future research.

Because of the need to adequately diagnose and treat AIH at an early stage to safeguard prognosis and reduce evolution towards liver failure and transplantation, prognostic scoring systems will probably gain importance (35,37). Previous prognostic scoring systems such as a treatment-response score appear to be useful (41).

Conclusion

Diagnostic scoring systems have their value as an auxiliary tool in diagnosing AIH. However, they will never replace the gold standard of clinical assessment and are primarily useful in selecting patient populations in research. The revised and simplified scoring system can be used in a complementary manner based upon their differences in performance and emphasis on diagnostic criteria, with neither system replacing the other. Due to its higher specificity, the simplified scoring system is more useful in assessing a clinical presentation of typical features seen in AIH, whereas the revised system can be applied to a more unclear clinical presentation due to its higher sensitivity. Nevertheless, the validity of diagnostic scoring systems in AIH has been insufficiently investigated; more research is recommended, particularly prospective studies.

Because of the variable nature of AIH, in which certain diagnostic features change over the course of the disease and the evolution of a clinical parameter can be more useful than its value upon assessment at first presentation, it could be stated that static diagnostic methods such as scoring systems will never be sufficient to adequately diagnose AIH. A transition towards a more dynamic system in which the evolution of certain clinical parameters and response to therapy are more prominent, could offer a more valuable solution. However, such prognostic systems are not available for the time being. Future research and potential development of such a dynamic diagnostic tool could be useful, both in assessing a clinical presentation suggestive for AIH as well as starting an adequate therapeutic policy for this incapacitating disease.

Conflict of interest

None.

References

1. MACK CL, ADAMS D, ASSIS DN, KERKAR N, MANNS MP, MAYO MJ, *et al.* Diagnosis and Management of Autoimmune Hepatitis in Adults and Children: 2019 Practice Guidance and Guidelines From the American Association for the Study of Liver Diseases. *Hepatology* (Baltimore, Md). 2020, **72**(2): 671-722.
2. European Association for the Study of the L. EASL Clinical Practice Guidelines: Autoimmune hepatitis. *Journal of hepatology*. 2015, **63**(4): 971-1004.
3. DALEKOS GN, KOSKINAS J, PAPATHEODORIDIS GV. Hellenic Association for the Study of the Liver Clinical Practice Guidelines: Autoimmune hepatitis. *Annals of gastroenterology*. 2019, **32**(1): 1-23.
4. GATSELIS NK, ZACHOU K, KOUKOULIS GK, DALEKOS GN. Autoimmune hepatitis, one disease with many faces: etiopathogenetic, clinico-laboratory and histological characteristics. *World journal of gastroenterology*. 2015, **21**(1): 60-83.
5. CZAJA AJ. Diagnosis and Management of Autoimmune Hepatitis: Current Status and Future Directions. *Gut and liver*. 2016, **10**(2): 177-203.
6. VERGANI D, MIELI-VERGANI G. Autoimmune hepatitis: Diagnostic criteria and serological testing. *Clinical liver disease*. 2014, **3**(2): 38-41.
7. ALVAREZ F, BERG PA, BIANCHI FB, BIANCHI L, BURROUGHS AK, CANCADO EL, *et al.* International Autoimmune Hepatitis Group Report: review of criteria for diagnosis of autoimmune hepatitis. *Journal of hepatology*. 1999, **31**(5): 929-38.
8. KRIESE S, HENEGHAN MA. Current concepts in the diagnosis and management of autoimmune hepatitis. *Frontline gastroenterology*. 2013, **4**(1): 2-11.
9. ABDOLLAHI MR, SOMI MH, FARAJI E. Role of international criteria in the diagnosis of autoimmune hepatitis. *World journal of gastroenterology*. 2013, **19**(23): 3629-33.
10. MOREIRA RK, REVETTA F, KOEHLER E, WASHINGTON MK. Diagnostic utility of IgG and IgM immunohistochemistry in autoimmune liver disease. *World journal of gastroenterology*. 2010, **16**(4): 453-7.
11. CABIBI D, TARANTINO G, BARBARIA F, CAMPIONE M, CRAXI A, DI MARCO V. Intrahepatic IgG/IgM plasma cells ratio helps in classifying autoimmune liver diseases. *Digestive and liver disease*: official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2010, **42**(8): 585-92.

12. HENNES EM, ZENIYA M, CZAJA AJ, PARES A, DALEKOS GN, KRAWITT EL, *et al.* Simplified criteria for the diagnosis of autoimmune hepatitis. *Hepatology* (Baltimore, Md). 2008, **48**(1): 169-76.
13. SEBODE M, WEILER-NORMANN C, LIWINSKI T, SCHRAMM C. Autoantibodies in Autoimmune Liver Disease-Clinical and Diagnostic Relevance. *Frontiers in immunology*. 2018, **9**: 609.
14. TERJUNG B, SPENGLER U, SAUERBRUCH T, WORMAN HJ. "Atypical p-ANCA" in IBD and hepatobiliary disorders react with a 50-kilo-dalton nuclear envelope protein of neutrophils and myeloid cell lines. *Gastroenterology*. 2000, **119**(2): 310-22.
15. HAUSDORF G, ROGGENBUCK D, FEIST E, BUTTNER T, JUNGBLUT PR, CONRAD K, *et al.* Autoantibodies to asialoglycoprotein receptor (ASGPR) measured by a novel ELISA – revival of a disease-activity marker in autoimmune hepatitis. *Clinica chimica acta*, international journal of clinical chemistry. 2009, **408**(1-2): 19-24.
16. ULPIANO TRILLIG A, RAMOISIAUX L, KOMUTA M, LANTHIER N. Autoimmune hepatitis in a patient with thymectomy: the benefit of histology and MUM-1 immunostaining. *Acta Gastroenterol Belg.* 2019 Oct-Dec, **82**(4): 550-551.
17. BALITZER D, SHAFIZADEH N, PETERS MG, FERRELL LD, ALSHAK N, KAKAR S. Autoimmune hepatitis: review of histologic features included in the simplified criteria proposed by the international autoimmune hepatitis group and proposal for new histologic criteria. *Modern pathology*: an official journal of the United States and Canadian Academy of Pathology, Inc. 2017, **30**(5): 773-83.
18. JABLONSKA J, CIELECKA-KUSZYK J, MIKULA T, KOZLOWSKA J, WIERCINSKA-DRAPALO A. Hepatopathy of unknown etiology-is liver biopsy a good tool in differential diagnosis? *Archives of medical science: AMS*. 2019, **15**(6): 1462-7.
19. GAO Y, LI W, MENG QH, LIU HX, ZHU YK. Analysis of 15 patients with abnormal liver function as the first systemic lupus erythematosus symptom. *Acta Gastroenterol Belg.* 2016 Sep-Dec, **79**(2) : 441-446.
20. CHAZOUILLERES O, WENDUM D, SERFATY L, MONTEMBAULT S, ROSMORDUC O, POUPON R. Primary biliary cirrhosis-autoimmune hepatitis overlap syndrome: clinical features and response to therapy. *Hepatology* (Baltimore, Md). 1998, **28**(2): 296-301.
21. ZHANG W, DE D, MOHAMMED KA, MUNIGALA S, CHEN G, LAI JP, *et al.* New scoring classification for primary biliary cholangitis-autoimmune hepatitis overlap syndrome. *Hepatology communications*. 2018, **2**(3): 245-53.
22. FEBRES-ALDANA CA, ALGHAMDI S, KRISHNAMURTHY K, POPPITI RJ. Liver Fibrosis Helps to Distinguish Autoimmune Hepatitis from DILI with Autoimmune Features: A Review of Twenty Cases. *Journal of clinical and translational hepatology*. 2019, **7**(1): 21-6.
23. JOHNSON PJ, MCFARLANE IG. Meeting report: International Autoimmune Hepatitis Group. *Hepatology* (Baltimore, Md). 1993, **18**(4): 998-1005.
24. CZAJA A, CARPENTER HA. Validation of scoring system for diagnosis of autoimmune hepatitis. *Digestive diseases and sciences*. 1996, 41(2): 305-14.
25. LOHSE AW. Recognizing autoimmune hepatitis: scores help, but no more. *Journal of hepatology*. 2011, **54**(2): 193-4.
26. FUJIWARA K, YASUI S, TAWADA A, FUKUDA Y, NAKANO M, YOKOSUKA O. Diagnostic value and utility of the simplified International Autoimmune Hepatitis Group criteria in acute-onset autoimmune hepatitis. *Liver international*: official journal of the International Association for the Study of the Liver. 2011, **31**(7): 1013-20.
27. LI Y, PENG M, GONG G. Evaluation of the revised versus the simplified scoring system in patients with autoimmune hepatitis. *Experimental and therapeutic medicine*. 2014, **7**(1): 131-6.
28. GATSELIS NK, ZACHOU K, PAPAMICHALIS P, KOUKOULIS GK, GABETA S, DALEKOS GN, *et al.* Comparison of simplified score with the revised original score for the diagnosis of autoimmune hepatitis: a new or a complementary diagnostic score? *Digestive and liver disease*: official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2010, **42**(11): 807-12.
29. MUNOZ-ESPINOSA L, ALARCON G, MERCADO-MOREIRA A, CORDERO P, CABALLERO E, AVALOS V, *et al.* Performance of the international classifications criteria for autoimmune hepatitis diagnosis in Mexican patients. *Autoimmunity*. 2011, **44**(7): 543-8.
30. KIM BH, KIM YJ, JEONG SH, TAK WY, AHN SH, LEE YJ, *et al.* Clinical features of autoimmune hepatitis and comparison of two diagnostic criteria in Korea: a nationwide, multicenter study. *Journal of gastroenterology and hepatology*. 2013, **28**(1): 128-34.
31. QIU D, WANG Q, WANG H, XIE Q, ZANG G, JIANG H, *et al*. Validation of the simplified criteria for diagnosis of autoimmune hepatitis in Chinese patients. *Journal of hepatology*. 2011, **54**(2): 340-7.
32. LIBERAL R, VERGANI D. Effect of ethnicity on the clinical presentation and outcome of autoimmune hepatitis. *Expert review of gastroenterology & hepatology*. 2012, **6**(3): 267-9.
33. KUIPER EM, ZONDERVAN PE, VAN BUUREN HR. Paris criteria are effective in diagnosis of primary biliary cirrhosis and autoimmune hepatitis overlap syndrome. *Clinical gastroenterology and hepatology*: the official clinical practice journal of the American Gastroenterological Association. 2010, **8**(6): 530-4.
34. HOEROLDT B, MCFARLANE E, DUBE A, BASUMANI P, KARAJEH M, CAMPBELL MJ, *et al.* Long-term outcomes of patients with autoimmune hepatitis managed at a nontransplant center. *Gastroenterology*. 2011, **140**(7): 1980-9.
35. PAPE S, GEVERS TJG, BELIAS M, MUSTAFAJEV IF, VROLIJK JM, VAN HOEK B, *et al.* Predniso(lo)ne Dosage and Chance of Remission in Patients With Autoimmune Hepatitis. *Clinical gastroenterology and hepatology*: the official clinical practice journal of the American Gastroenterological Association. 2019, **17**(10): 2068-75 e2.
36. KIRSTEIN MM, METZLER F, GEIGER E, HEINRICH E, HALLENSLEBEN M, MANNS MP, *et al.* Prediction of short- and long-term outcome in patients with autoimmune hepatitis. *Hepatology* (Baltimore, Md). 2015, **62**(5): 1524-35.
37. PAPE S, GEVERS TJG, VROLIJK JM, VAN HOEK B, BOUMA G, VAN NIEUWKERK CMJ, et al. Rapid Response to Treatment of Autoimmune Hepatitis Associated With Remission at 6 and 12 Months. *Clinical gastroenterology and hepatology*: the official clinical practice journal of the American Gastroenterological Association. 2020, **18**(7): 1609-17 e4.
38. VAN GERVEN NM, VERWER BJ, WITTE BI, VAN HOEK B, COENRAAD MJ, VAN ERPECUM KJ, *et al.* Relapse is almost universal after withdrawal of immunosuppressive medication in patients with autoimmune hepatitis in remission. *Journal of hepatology*. 2013, **58**(1): 141-7.
39. EDMUNDS C, EKONG UD. Autoimmune Liver Disease Post-Liver Transplantation: A Summary and Proposed Areas for Future Research. *Transplantation*. 2016, **100**(3): 515-24.
40. KERKAR N, YANNI G. 'De novo' and 'recurrent' autoimmune hepatitis after liver transplantation: A comprehensive review. *Journal of autoimmunity*. 2016, **66**: 17-24.
41. SIMOEN C, SCHOUTEN J, COLLE I, *et al.* A New Treatment-Response Score for AUTO Immune Hepatitis Is Predictive of LONG-TERM Event-Free Survival and Overall Survival. A Multicenter Belgian Cohort Study. AASLD The Liver Meeting, San Francisco, *Hepatology* 2018, **68**: S1; 1977B.