Digestive and Liver Disease 51 (2019) 1308–1313

Contents lists available at ScienceDirect

Digestive and Liver Disease

journal homepage: www.elsevier.com/locate/dld

Liver, Pancreas and Biliary Tract

Development and validation of a new simplified diagnostic scoring system for pediatric autoimmune hepatitis

José Vicente Arcos-Machancoses[a,*], Cristina Molera Busoms[a,c], Ecaterina Julio Tatis[a], María Victoria Bovo[a], Jesús Quintero Bernabeu[b,c], Javier Juampérez Goñi[b,c], Vanessa Crujeiras Martínez[d], Javier Martín de Carpi[a,c]

[a] Sant Joan de Déu Hospital (HSJD), Department of Pediatric Gastroenterology, Hepatology and Nutrition, Barcelona, Spain
[b] Vall d'Hebron Hospital (HVH), Department of Pediatric Gastroenterology, Hepatology and Nutrition, Barcelona, Spain
[c] HSJD-HVH Comprehensive Unit of Complex Hepatology and Pediatric Liver Transplantation, Barcelona, Spain
[d] University Hospital Complex of Santiago de Compostela, Department of Pediatric Gastroenterology, Hepatology and Nutrition, Spain

ARTICLE INFO

Article history:
Received 6 November 2018
Accepted 25 February 2019
Available online 28 March 2019

Keywords:
Diagnostic criteria
Liver diseases
Sensitivity
Specificity
Autoimmune hepatitis

ABSTRACT

Background: Children with autoimmune hepatitis (AIH) often exhibit particular features. Accordingly, seven pediatric-specific criteria have been proposed.
Aim: To develop a prediction model based on them, transform it into a scoring system and study its accuracy.
Methods: A cohort of children under study for liver disease was consecutively selected. AIH diagnosis was based on classical criteria. Already proposed pediatric criteria were recorded. The best possible regression model was selected, and the beta coefficient of each criterion was translated into a whole number (points). Total scores were obtained following the points system and the best cut-off was calculated. Subsequently, accuracy of the diagnostic score was studied in the validation set.
Results: Among 212 included patients, 100 had AIH. The score included 5 criteria: autoantibodies (0–2 points), hypergammaglobulinemia, exclusion of viral hepatitis, exclusion of Wilson's disease (1 point each) and liver histology (3 points). In addition, a normal cholangiogram is mandatory. The validation set was formed of 70 patients (24 with AIH). In this subsample, a score of $\geq$6 renders a sensitivity/specificity of 95.8%/100%. The area under the receiver operating characteristic curve was 97.1%.
Conclusion: Pediatric-specific criteria for the diagnosis of AIH can be reliably used as a scoring system.

1. Introduction

Autoimmune hepatitis (AIH) is an inflammatory liver disorder with a wide clinical spectrum, that ranges from isolated hypertransaminasemia to acute liver failure (ALF) [1,2]. It is characterized by several analytical features, especially the presence of autoantibodies, high levels of immunoglobulin G (IgG) and, histologically, by liver lymphoplasmacytic chronic infiltration that is typically displayed as interface hepatitis [3].

In the absence of a single diagnostic test, the International Autoimmune Hepatitis Group (IAIHG) proposed, in 1993, some classification criteria to facilitate reproducibility of clinical studies [4]. Criteria were revised in 1999 and, as a result, their specificity was improved [5]. IAIHG revised diagnostic criteria are considered to be the best diagnostic reference despite not being a truly gold standard [6,7]. However, their practical application remains challenging for clinical use due to its complexity, including 13 categories, some of them impractical in children. To overcome these difficulties, the IAIHG proposed a simplified scoring system in 2008, that takes into account only the presence of autoantibodies, IgG levels, histopathology and absence of viral markers [8,9]. Some validation studies have been carried out on the accuracy of simplified criteria, showing that they are not perfectly suitable to the juvenile form of the disease, even further in cases with ALF onset [10–18]. Furthermore, children with AIH often exhibit lower autoantibodies

* Corresponding author at: Sant Joan de Déu Hospital, Pediatric Gastroenterology and Hepatology Department, Passeig de Sant Joan de Déu, 2, Esplugues de Llobregat, Barcelona, 08950, Spain.
E-mail addresses: jvicentearcos@gmail.com (J.V. Arcos-Machancoses), cristinamolera@gmail.com (C. Molera Busoms), ecaterina02@hotmail.com (E. Julio Tatis), mavibovo@gmail.com (M.V. Bovo), jquintero@vhebron.net (J. Quintero Bernabeu), jjuamperez@vhebron.net (J. Juampérez Goñi), vanecrujeiras@hotmail.com (V. Crujeiras Martínez), javiermartin@sjdhospitalbarcelona.org (J. Martín de Carpi).

https://doi.org/10.1016/j.dld.2019.02.018
1590-8658/

titers than adults, are less prone to hypergammaglobulinemia, and autoimmune sclerosing cholangitis (AISC) require exclusion as it can easily be mistaken for AIH [17,19,20]. Taking into account these particularities, seven pediatric-specific criteria were proposed in 2009: hypertransaminasemia, elevated IgG, presence of autoantibodies even at low levels, liver biopsy with interface hepatitis or multilobular collapse, exclusion of Wilson's disease, exclusion of viral hepatitis and normal cholangiogram [21]. We hypothesized that 2009 criteria can function adequately as a diagnostic tool. Our aim was to develop a prediction model based on them, transform it into a simple scoring system and study its accuracy.

2. Methods

2.1. Design

The study was conceived in two phases. The first one was the modelling through logistic regression of a predictive formula for the diagnostic probability of AIH. The second step consisted in the assessment of the validity of the new scoring system, which was conducted under a cross-sectional diagnostic study design. Data were collected mainly in a retrospective way since January 2005, but a prospective wing was included from January 2016 to January 2017 to increase the sample size, verify already retrieved information, correct missing data and reclassify patients' first diagnosis if needed. Needed sample size was estimated according to the first phase requirements and considering the need for extra cases to carry out the external validation [22]. Patients were randomly assigned to either the training or the validation set with a probability 2/3 and 1/3, respectively.

2.2. Study population

2.2.1. Inclusion criteria

Patients were all under 18 years old and were enrolled from two tertiary care hospitals that share a common Pediatric Hepatology unit. They were intended to be representative of a population in clinical or analytical situation potentially attributable to AIH. Inclusion criteria were: (1) patients with liver biopsy performed because of signs of acute or chronic hepatocellular damage, regardless of the coexistence of cholestasis and the technique for sample collection (percutaneous, laparoscopic or transjugular), and (2) ALF defined according to the consensus reached by the Pediatric Acute Liver Failure Study Group [23].

2.2.2. Exclusion criteria

Patients with the following conditions were excluded: (1) children who underwent liver or hepatocyte transplantation. (2) Previous diagnosis of any congenital or acquired liver disease. (3) Previous diagnosis of any systemic or metabolic disorder with liver involvement. (4) Hepatic space-occupying lesion. (5) Liver biopsy indicated for sample collection in the setting of fever of undetermined origin investigation. (6) Infants less than six month of age (to rule out alloimmune hepatitis and other causes of young infant ALF). Participants formed a consecutive series and were identified through International Classification of Diseases codes, 9th revision (ICD-9), cross referencing with Pediatric Hepatology database and radiology reports

2.2.3. Patient classification

AIH was defined according to the 1999 classical revised criteria [5]. As it is not a completely accurate gold standard, the study was designed to minimize misestimation of validity indexes[7]. In this regard, post-treatment scores were taken into account if possible. Therefore, we established a threshold of ≥12 to be a positive case, regardless its classification in probable (12–17) or definite AIH (>17). In a pre-treatment basis, a 10 to 15 score meant probable AIH, and >15 meant definite AIH. Labelling between these two categories is based on variations in clinical manifestations and does not reflect differences in the reliability of the diagnosis [24]. Auto-antibodies titers under 1:40 were given 1 point [3,5] and, unless specified in another way in the medical reports, alcohol intake was judged not significant. To ensure diagnostic robustness and reduce false positives, two conditions were considered necessary: (1) Treatment response defined as symptoms relief and transaminases normalization (complete or with posterior relapse) and (2) liver histology describing, at least, features compatibles with AIH (chronic hepatitis with lymphocytic infiltration) [25]. During the prospective wing, in order to reduce false negatives, those cases initially labelled as non AIH by classical criteria were reviewed. They were reclassified to AIH if there was an explicit AIH diagnosis in medical reports and the two necessary conditions described above were simultaneously fulfilled.

The protocol was approved by the ethics committee of Hospital Sant Joan de Déu, Barcelona.

Table 1
Criteria for the diagnosis of autoimmune hepatitis in childhood.

Parameter	Description
Elevated transaminases	
Positive autoantibodies	ANA and/or anti-SMA titers ≥1:20 Anti-LKM1 tires ≥1:10 Anti-LC1 Anti-SLA
Elevated IgG	
Liver biopsy	Interface hepatitis Multilobular collapse
Exclusion of viral hepatitis	
Exclusion of Wilson's disease	
Normal cholangiogram (nuclear magnetic resonance or retrograde cholangiography)	

Source: Mieli-Vergani G, Heller S, Jara P, et al. Autoimmune hepatitis. *J Pediatr Gastroenterol Nutr*, 2009;49(2):159.
ANA: anti-nuclear antibody; anti-SMA: anti-smooth muscle antibody; anti-LKM1; anti-liver/kidney microsomal antibody; anti-LC1: anti-liver cytosol 1 antibody; anti-SLA: anti-soluble liver antibody; IgG: immunoglobulin G.

2.3. Development and validation of the diagnostic scoring system

Pediatric 2009 AIH criteria (Table 1) were recorded as independent binary variables, except the autoantibodies criterion, which was categorized into three levels (negative, low or high) by convention according to current knowledge [20,21]. Autoantibodies were classified as *high* when either anti-soluble liver antigen (SLA) or anti-liver cytosol type 1 (LC1) were positive, or antinuclear antibodies (ANA), anti-smooth muscle antibodies (SMA) or anti-liver kidney microsomal antibodies type 1 (LKM1) titers were above 1:80. The threshold between *negative* and *low* was established at the level suggested by the pediatric 2009 criteria for positive ANA, anti-SMA and anti-LKM1. The planning consisted of excluding those patients from analyses if missing data recovery was not possible in the prospective wing. Only missing values regarding the cholangiogram variable were solved by multiple imputation [26].

The diagnostic equation was modelled through logistic regression and transformed into a scoring system following the Framingham study example [27]. For this purpose, we only employed the training set. All possible models were studied by an iterative approach. Regarding the finally selected model, the standardized β coefficient of each variable was translated into a whole number (points) maintaining their proportionality. Total scores were obtained for each patient according to the point system and the best cut-off was calculated following the Zweig and Campbell strategy with a cost ratio of 1 [28].

Table 2
Patients' clinical, analytical and demographic characteristics according to final diagnosis.

	AIH patients (n = 100)	Non-AIH patients (n = 112)	*P* Value
Specific diagnosis	Association with other autoimmune related disease (23) Acute liver failure (4) Type 2 AIH (17)	Acute cryptogenic hepatitis without cholestasis (27) Wilson's disease (22) Viral hepatitis (15) Alagille syndrome (7) Toxic hepatitis (6) Primary sclerosing cholangitis (5) Congestive hepatopathy (5) Acute cryptogenic hepatitis with cholestasis (4) Congenital hepatic fibrosis (4) Deposit disease (4) Non-alcoholic steatohepatitis (3) Progressive familial intrahepatic cholestasis (3) Mitochondrial disease (3) Cryptogenic chronic liver disease (2) Giant cell hepatitis (1) Hereditary fructose intolerance (1)	
Proportion of females	72.0% (62.5%–79.9%)	42.0% (33.2%–51.2%)	<0.001
Age at diagnosis (years)[a]	7.9 (3.9–11.6)	8.1 (3.8–12.5)	0.665
Personal of family history of other autoimmune disease	29.0% (21.0%–38.5%)	9.8% (5.6%–16.7%)	<0.001
Hyper-IgG levels	67.0% (57.3%–75.4%)	18.8% (12.6%–27.0%)	<0.001
Immunoglobulin G (mg/dL)	1598 (1130.5–2373.5)	960 (776–1143)	<0.001
AST (U/L)	788.5 (141.5–1730.5)	96 (53.5–209.5)	<0.001
ALT(U/L)	678 (174–1833)	105 (52–387.5)	<0.001
AP (U/L)	288 (215–407.5)	290 (201–380.5)	0.935
GGT (U/L)	76 (36–145.5)	42 (22.5–80)	0.001
Proportion of patients with positive ANA and/or anti-SM titers <1:40	0	0	–
Proportion of patients with ANA ≥1:40	66.0% (56.3%–74.5%)	31.3% (23.4%–40.3%)	<0.001
Proportion of patients with ANA ≥1:80	60.0% (50.2%–69.1%)	25.9% (18.7%–34.7%)	<0.001
Proportion of patients with anti-SM ≥1:40	63.0% (53.2%–71.8%)	52.7% (43.5%–61.7%)	0.129
Proportion of patients with anti-SM ≥1:80	58.0% (48.2%–67.2%)	24.1% (17.1%–32.8%)	<0.001
ANA titers (1:X)	240 (80–640)	80 (80–320)	0.001
Anti-SM titers (1:X)	160 (140–640)	40 (40–80)	<0.001
Proportion of patients with positive anti-LKM1 titers <1:40	0	0	–
Proportion of patients with anti-LKM1 ≥1:40	15.0% (9.3%–23.3%)	1.8% (0.5%–6.3%)	<0.001
Anti-LKM1 titers (1:X)	160 (160–240)	80[b]	0.059
Revised 1999 criteria scoring	14 (12–17)	2 (0–4)	<0.001
Simplified 2008 criteria scoring	6 (5–7)	3 (2–4)	<0.001

Source: Arcos-Machancoses JV et al. *Pediatr Gastroenterol Hepatol Nutr*. 2018 April;21(2):121.

AIH: Autoimmune hepatitis. AST: Aspartate aminotransferase. ALT: Alanine aminotransferase. AP: Alkaline phosphatase. GGT: gamma glutamyl transferase. ANA: Antinuclear antibodies. Anti-SM: Anti-smooth muscle antibodies. Anti-LKM1: Anti-liver/kidney microsomal antibodies. Data between parenthesis represent interquartile range for quantitative variables and 95% confidence interval for proportions.

[a] Or at first biopsy date if diagnosis initially unclear.

[b] Interquartile range not informative due to few patients.

Binary results were reported as percentages with 95%CI, by Wilson method, and continuous variables as median and interquartile range (IQR). The Mann–Whitney *U* test was used to evaluate differences in continuous variables between groups and X^2 test for dichotomous variables. Fisher's exact test was used when appropriate. A *p*-value of <0.05 was considered statistically significant. The script *AllSetsReg* was used for the mathematical modelling [29].

Sensitivity, specificity, likelihood ratios (LR) and predictive values (PV) were calculated with their 95% confidence interval also by Wilson method (95%CI). A receiver operating characteristic (ROC) curve was plotted, and its area under the curve calculated, to assess the discriminative power of the new diagnostic score. These analyses were performed within the training set and the validation set separately, for exploratory purposes. However, only the validity indicators obtained from the validation set were considered to be truly representative of the accuracy of the new scoring system. All statistical analyses were performed on IBM SPSS® version 21.0 (IBM, Armonk, NY, USA).

3. Results

Initially, a total of 425 patients showing a variety of liver diseases were identified. Out of them, 207 met exclusion criteria. We finally intended to conduct the analysis in 218 patients, but 6 of them were not taken into account because data recovery was impossible due to loss to follow-up. In the remaining 212 children, the prevalence of AIH was 47.2% (95%CI 40.6%–53.9%). Seventeen of these children were included in the prospective wing. Clinical and demographic characteristics according to final diagnosis are shown in Table 2. Proportion between AIH type 1 and 2 in our sample was nearly 5:1. None of the patients had an explicit diagnosis of overlap syndrome with primary sclerosing cholangitis (PSC). Nevertheless, not all patients enrolled at the first years of the inclusion period had a cholangiographic evaluation performed.

3.1. Diagnostic scoring system: training set, modelling and internal validity

The training set included data from 76 children with AIH and 66 controls with several liver diseases, among which 21 were acute cryptogenic hepatitis with or without cholestasis, 14 Wilson's disease, 9 viral hepatitis, and a miscellany of Alagille syndrome, PSC and others. Clinical and analytical characteristics of patients in the training set did not statistically differed from those in the entire sample. Twelve children with AIH could be classified as type 2. ALF was the clinical onset of four patients finally diagnosed with AIH. The proportion of females within this subset was of 71.1% in the AIH group and 37.9% in the non-AIH group. Up to 69.7% of AIH cases

Table 3
Predictive regression model for the diagnosis of pediatric autoimmune hepatitis with the criteria proposed by Mieli-Vergani et al. (J Pediatr Gastroenterol Nutr. 2009;49:159).

Variable	Type	Meaning	β Coefficient	Multivariate analysis
Constant	–	–	–57.595	NS
Hypertransaminasemia	Binary (Y/N)	Elevated transaminases	2.453	NS
Hypergammaglobulinemia	Binary (Y/N)	Elevated IgG	18.176	NS
Wilson	Binary (Y/N)	Rule out Wilson's disease	17.997	NS
Virus	Binary (Y/N)	Rule out viral hepatitis	16.644	NS
Histopathology	Binary (Y/N)	Interface hepatitis or multilobular collapse	52.090	NS
Cholangiogram	Binary (Y/N)	Normal cholangio-MR or retrograde cholangiography	1.267	NS
Autoantibodies (reference category = negative)	Categorical (negative, low or high)	• *Negative*: ANA and anti-SMA <1:20, and anti-LKM1 <1:10. • *Low*: ANA or anti-SMA ≥1:20 and <1:80, or anti-LKM1 tires ≥1:10 and <1:80. • *High*: ANA, anti-SMA or anti-LKM1 ≥1:80, or positive anti-LC1 or anti-SLA.	18.150	NS

Y/N: Yes/No. IgG: immunoglobulin G. MR: magnetic resonance. ANA: antinuclear antibody. SMA: smooth muscle antibody. LKM1: liver kidney microsomal antibody type 1. LC1: liver cytosol type 1. SLA: soluble liver antigen. NS: non-significant.

Table 4
New proposal for pediatric autoimmune hepatitis diagnostic score.

Points[a]	Parameter
	Autoantibodies[b]
+1	ANA or anti-SMA ≥1:20 and <1:80, or anti-LKM1 tires ≥1:10 and <1:80
+2	ANA, anti-SMA or anti-LKM1 ≥ 1:80, or positive anti-LC1 or anti-SLA
+1	Hypergammaglobulinemia
+3	Liver biopsy (histopathology)
	Interface hepatitis
	Multilobular collapse
+1	Exclusion of viral hepatitis
+1	Exclusion of Wilson's disease

ANA: anti-nuclear antibody; anti-SMA: anti-smooth muscle antibody; anti-LKM1; anti-liver/kidney microsomal antibody; anti-LC1: anti-liver cytosol 1 antibody; anti-SLA: anti-soluble liver antibody.

[a] Autoimmune hepatitis can be diagnosed in children with hypertransaminasemia, 6 or more points in the scoring system and normal cholangiogram.

[b] Maximum number of points for all autoantibodies is 2. Maximum total points are 8.

had hyper-IgG, while only 21.2% of children with other liver disorders presented with this condition. Autoantibodies titers of 1:80 or above were found in 75.0% of AIH patients, and only in 25.0% of alternative diagnoses. Of all AIH patients with anti-SLA/LC1 tested (40/76), 12.5% gave a positive result. Anti-SLA was not found to be positive in any of the controls. Regarding the score reached in the classical criteria, AIH patients had a median of 14 points (IQR 12–17 points) and controls obtained 2 points (0–5 points). Simplified criteria median scoring was also closed to that from the whole sample: 6 points (2–7 points) and 4 points (2–4 points) for the AIH group and the non-AIH group respectively. Only 37 AIH cases had a magnetic resonance cholangiogram performed.

In relation to the development of the diagnostic model by logistic regression, the recommendation of including at least 10 cases per each candidate variable (7) was met [22]. The maximum model with the seven pediatric-specific criteria achieved an area under the ROC curve of 99.8%. We found an equation with only 5 variables that reached an area under the ROC curve of 99.6%. However, it did not include neither the normal cholangiogram variable nor the presence of high transaminases, which we considered mandatory in the AIH diagnostic model. Consequently, their presence was forced in the score system and the maximum regression model was finally selected (Table 3).

On the basis of previous results, points for every criterion of the final selected prediction model were assigned by calculating the quotient between their beta coefficient and the lowest one, rounded to the nearest whole number. The new proposed score for pediatric AIH ended as set out in Table 4. Even though no points could be assigned to the hypertransaminasemia and the cholangiogram criteria, their presence has to be met to establish the suspicion of AIH and to rule out AISC. Optimal cut-off was stablished at 6 points, that gave a sensitivity of 96.1% (95%CI 89.0%–98.7%) and a specificity of 100% (95%CI 94.5%–100%) in the training set. Positive and negative LR were calculated at 64.51 and 0.04 respectively. The median of points reached in the new scoring system was 8 (IQR 7–8) for children with AIH, and 4 (IQR 2–4) for the non-AIH group. Finally, the area under the ROC curve within the training set was 98.9% (95.5%–99.9%). Only three cases were misclassified by the new scoring system as non-AIH.

3.2. Accuracy of the new pediatric AIH diagnostic score: external validity

Provision of a validation set gave us the possibility to estimate a non-biased body of validity indicators. This subsample was composed of 24 AIH cases and 46 cases with alternative diagnoses. The determined cut-off value (6 or more points leading to the diagnosis of AIH) delivered a sensitivity of 95.8% (95%CI 79.8% to 99.3%) and a specificity of 100% (95%CI 92.3%–100%). The positive and negative LR were 24.9 and 0.04. For the prevalence estimated in our clinical scenario, the positive and negative PV were, consequently, 100% (95%CI 89.2%–100%) and 96.4% (95%CI 85.2%–99.2%). The median of points, and its IQR, that AIH and non-AIH patients got in the validation set were the same than in the training set. Nevertheless, the area under the ROC curve was slightly lower in the first one (Fig. 1): 97.1% (95%CI 91.2%–99.9%). An exhaustive display of the accuracy measures of the new scoring system (both in the training and validation set), is shown in Table 5. Optimal cut-offs following the predefined strategy were simulated for several prevalences (from 5%–50%) and cost ratio assumptions (from 1/8–8). For every studied combination, the best cut-off for diagnosing AIH accordingly to the score results was established at 6 points.

4. Discussion

The diagnosis of AIH is made under a combination of clinical, biochemical, immunological and histological characteristics, and after excluding other causes of liver disease that can mimic AIH. To streamline the cognitive process of diagnosis and facilitate the reproducibility of clinical studies, several diagnostic criteria for AIH have been proposed and validated [30]. Some previous efforts have been made to assess the accuracy of the IAIHG 2008 simplified score in children, as a clinically suitable alternative to classical criteria. It has been found that, despite showing a good specificity, their sensitivity is not enough to make them totally useful in the pediatric form of the disease [17,18,31]. In this respect, we designed this study in order to develop a pediatric-specific simplified scoring system to diagnose AIH. The list of criteria already suggested by Mieli-Vergani

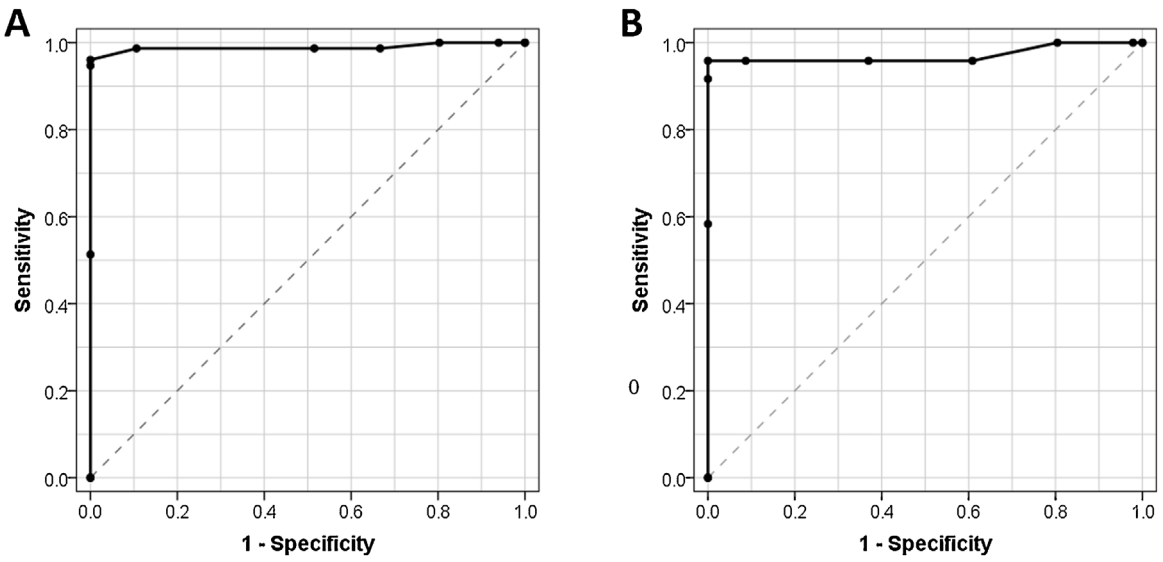

Fig. 1. Panel A: receiver operating characteristic curve in the training set for the new pediatric autoimmune hepatitis diagnostic score. Panel B: receiver operating characteristic curve in the validation set for the new pediatric autoimmune hepatitis diagnostic score.

Table 5
Internal and external validity indicators of the new diagnostic score for pediatric autoimmune hepatitis based on the 2009 criteria.

Cut-off	Internal validation					External validation				
	Sensitivity	Specificity	Correct classifications	LR +	LR −	Sensitivity	Specificity	Correct classifications[b]	LR +	LR −
≥1	100%	6.1%	50.4%	1.1	–	100%	2.2%	47.2%	1.0	–
≥2	100%	19.7%	57.6%	1.3	–	100%	19.6%	57.5%	1.2	–
≥3	98.7%	33.3%	64.2%	1.5	0.04	95.8%	39.1%	68.9%	1.6	0.11
≥4	98.7%	48.5%	72.2%	1.9	0.03	95.8%	63.0%	78.5%	2.6	0.07
≥5	98.7%	89.4%	93.8%	9.3	0.01	95.8%	91.3%	93.4%	11.0	0.05
≥6	96.1%	100%	98.2%	64.6[a]	0.04	95.8%	100%	98.0%	24.9[c]	0.04
≥7	94.7%	100%	97.5%	–	0.05	91.7%	100%	96.1%	–	0.08
8	51.3%	100%	77.0%	–	0.49	58.3%	100%	80.3%	–	0.42

LR: Likelihood ratio.
[a] Approximate value after correction of an empty cell in the contingency table.
[b] Calculated with an assumed autoimmune hepatitis prevalence of 47.2% (that of the entire sample).
[c] Calculated for the average specificity between the limits of the real value's confidence intervals.

et al. in 2009 were considered as far as they are sustained by clinical evidence [21]. Several facts support the hypothesis of its better appropriateness for pediatric AIH, specially the lower autoantibodies titers that children tend to show, and the consideration of Wilson's disease as a major alternative diagnosis. The 2009 criteria also propose the presence of hypertransaminasemia, do not mention any threshold for a significative hypergammaglobulinemia and relax the definition of a positive histopathology.

In our opinion, we have managed to build representative samples of AIH and its differentials to create a diagnostic score and explore its goodness and discriminative capacity. However, several limitations must be pointed out.

There are three forms of pediatric liver disease with an autoimmune component in their pathogenesis: AIH, AISC and de novo AIH after liver transplant [30,32]. The present study lacks external evidence to infer the results to transplanted children, as they were excluded to study specifically *pure* AIH. Therefore, under this circumstance, the application of the score should be avoided, regardless the similarity between AIH in the native liver and the recurrent disease in the allograft. A future specifically designed accuracy study may amend the lack of external validity for this concrete target population.

Additionally, distinguishing between AIH and patients with autoimmune cholestatic liver disease is considered a major problem. However, the role of the new score to solve this clinical challenge could not be validated neither. Indeed, in comparison with other settings routine clinical practice, patients with AISC, were underrepresented in our sample. This might lead to overestimate the validity indicators of the score. To compensate, we decided to force the inclusion of the cholangiogram criterion. Linking with that, it is noted that the IAIHG simplified criteria do not solve this problem, because they are not always able to properly differentiate between AIH and some diseases involving biliary tract impairment such as PSC and AISC [31,33]. A number of points could not be assigned to the cholangiogram variable but, nevertheless, a normal biliary tract imaging test should be mandatory to establish AIH diagnosis. As a result, positive scoring in our new criteria can reliably rule out AISC but is not helpful to confirm this diagnosis. In this respect, a new scoring diagnostic system have recently been proposed for both pediatric AIH and AISC. They have not been validated yet but, interestingly, the number of points assigned to the autoantibodies criterion, to the liver histology or to the absence of alternative diagnoses are equal or very closed to that calculated in this study [30].

Classical and simplified criteria's utility in adult and pediatric patients have proven to be poor in ALF [18,34]. However, the four cases in our cohort matching the definition of ALF were correctly classified by the new score. All of them had mild signs of encephalopathy that recovered after treatment. It is possible that clinical severity of ALF could have an impact in AIH diagnostic criteria reliability and, as a result, our results could not be strictly inferred to more severe cases.

Finally, our new proposal allows seronegative AIH to be diagnosed if the other criteria are met. However, over half of the children with no autoantibodies do not show high IgG levels, which has been the case of our false negatives. Even when presenting with other conditions suggestive of seronegative AIH, such as aplastic anemia or peripheral thrombocytopenia, these criteria still have a limited role in this particular phenotype [35].

It would be desirable to validate this diagnostic scoring system in a larger cohort of patients presenting with hypertransaminasemia. Notionally, the resulting sample of a multicenter collaborating study would include the whole clinical spectrum of AIH. Thus, it would be possible to assess its accuracy, even in the abovementioned non-typical profiles of children with AIH.

As a conclusion, in this non-selected population, the new scoring diagnostic system based on 2009 pediatric AIH criteria have shown an overall appropriate performance (sensitivity of 95.8% and specificity of 100%). Additional external validation is required in samples with a wider variety of presentations concerning AIH, liver transplanted patients and a complete list of differential diagnoses. This is highly advisable as our validation set belongs to the same population than the training set. We suggest to carry out a further comparative assessment with the scoring proposed by the ESPGHAN Hepatology Committee, that includes additional criteria such the absence of NASH and toxic hepatitis, the presence of extrahepatic autoimmunity, family history of autoimmune disease and the finding of peripheral anti-nuclear neutrophil antibodies [30]. To evaluate the practical usefulness of both scoring criteria, clinical decision analysis with data gathered from diagnostic studies should be designed.

Conflict of interest

None declared.

References

[1] Krawitt E-L. Clinical features and management of autoimmune hepatitis. World J Gastroenterol 2008;14:3301–5.
[2] Vergani D, Mieli-Vergani G. Cutting edge issues in autoimmune hepatitis. Clin Rev Allergy Immunol 2012;42:309–21.
[3] Mieli-Vergani G, Vergani D. Autoimmune hepatitis in children: what is different from adult AIH? Semin Liver Dis 2009;29:297–306.
[4] Johnson PJ, McFarlane IG. Meeting report: International Autoimmune Hepatitis Group. Hepatology 1993;18:998–1005.
[5] 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. J Hepatol 1999;31:929–38.
[6] Ebbeson RL, Schreiber RA. Diagnosing autoimmune hepatitis in children: is the International Autoimmune Hepatitis Group scoring system useful? Clin Gastroenterol Hepatol 2004;2:935–40.
[7] Kochar R, Fallon M. Diagnostic criteria for autoimmune hepatitis: what is the gold standard? Hepatology 2010;51:350–1, author reply 351.
[8] Hennes EM, Zeniya M, Czaja AJ, Parés A, Dalekos GN, Krawitt EL, et al. Simplified criteria for the diagnosis of autoimmune hepatitis. Hepatology 2008;48:169–76.
[9] Choi G, Peters MG. The challenge of diagnosing autoimmune hepatitis: less is more. Hepatology 2008;48:10–2.
[10] 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. J Hepatol 2011;54:340–7.
[11] Li Y, Peng M, Gong G. Evaluation of the revised versus the simplified scoring system in patients with autoimmune hepatitis. Exp Ther Med 2014;7:131–6.
[12] 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 Int 2011;31:1013–20.
[13] Liu F, Pan ZG, Ye J, Xu D, Guo H, Li GP, et al. Primary biliary cirrhosis-autoimmune hepatitis overlap syndrome: simplified criteria may be effective in the diagnosis in Chinese patients. J Dig Dis 2014;15:660–8.
[14] 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? Dig Liver Dis 2010;42:807–12.
[15] Yeoman AD, Westbrook RH, Al-Chalabi T, Carey I, Heaton ND, Portmann BC, et al. Diagnostic value and utility of the simplified International Autoimmune Hepatitis Group (IAIHG) criteria in acute and chronic liver disease. Hepatology 2009;50:538–45.
[16] Muratori P, Granito A, Pappas G, Muratori L. Validation of simplified diagnostic criteria for autoimmune hepatitis in Italian patients. Hepatology 2009;49:1782–3, author reply 1783.
[17] Mileti E, Rosenthal P, Peters MG. Validation and modification of simplified diagnostic criteria for autoimmune hepatitis in children. Clin Gastroenterol Hepatol 2012;10:417–21, e1-2.
[18] Hiejima E, Komatsu H, Sogo T, Inui A, Fujisawa T. Utility of simplified criteria for the diagnosis of autoimmune hepatitis in children. J Pediatr Gastroenterol Nutr 2011;52:470–3.
[19] Ferri PM, Ferreira AR, Miranda DM, Simões E Silva AC. Diagnostic criteria for autoimmune hepatitis in children: a challenge for pediatric hepatologists. World J Gastroenterol 2012;18:4470–3.
[20] Vergani D, Alvarez F, Bianchi FB, Cançado ELR, Mackay IR, Manns MP, et al. Liver autoimmune serology: a consensus statement from the committee for autoimmune serology of the International Autoimmune Hepatitis Group. J Hepatol 2004;41:677–83.
[21] Mieli-Vergani G, Heller S, Jara P, Vergani D, Chang M-H, Fujisawa T, et al. Autoimmune hepatitis. J Pediatr Gastroenterol Nutr 2009;49:158–64.
[22] Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9.
[23] Squires RH, Shneider BL, Bucuvalas J, Alonso E, Sokol RJ, Narkewicz MR, et al. Acute liver failure in children: the first 348 patients in the pediatric acute liver failure study group. J Pediatr 2006;148:652–8.
[24] Czaja AJ. Comparability of probable and definite autoimmune hepatitis by international diagnostic scoring criteria. Gastroenterology 2011;140:1472–80.
[25] Tiniakos DG, Brain JG, Bury YA. Role of Histopathology in Autoimmune Hepatitis. Dig Dis 2015;33(Suppl 2):53–64.
[26] Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med 1991;10:585–98.
[27] Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med 2004;23:1631–60.
[28] Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993;39:561–77.
[29] Sesma R. Extension Command UAB AllSetsReg v0.0.7 (c) JM Doménech & JB Navarro. Laboratori d'Estadística Aplicada — Universitat Autònoma de Barcelona; 2012.
[30] Mieli-Vergani G, Vergani D, Baumann U, Czubkowski P, Debray D, Dezsofi A, et al. Diagnosis and management of pediatric autoimmune liver disease: ESPGHAN hepatology committee position statement. J Pediatr Gastroenterol Nutr 2018;66:345–60.
[31] Arcos-Machancoses JV, Molera Busoms C, Julio Tatis E, Bovo MV, Quintero Bernabeu J, Juampérez Goñi J, et al. Accuracy of the 2008 simplified criteria for the diagnosis of autoimmune hepatitis in children. Pediatr Gastroenterol Hepatol Nutr 2018;21:118–26.
[32] Mieli-Vergani G, Vergani D. Paediatric autoimmune liver disease. Arch Dis Child 2013;98:1012–7.
[33] Arcos-Machancoses JV, Molera Busoms C, Julio Tatis E, et al. Accuracy of the Simplified Criteria for Autoimmune Hepatitis in Children: Systematic Review and Decision Analysis. J Clin Exp Hepatol 2019, http://dx.doi.org/10.1016/j.jceh.2018.10.006 [forthcoming].
[34] Yeoman AD, Westbrook RH, Al-Chalabi T, Carey I, Heaton ND, Portmann BC, et al. Diagnostic value and utility of the simplified International Autoimmune Hepatitis Group (IAIHG) criteria in acute and chronic liver disease. Hepatology 2009;50:538–45.
[35] Maggiore G, Socie G, Sciveres M, Roque-Afonso A-M, Nastasio S, Johanet C, et al. Seronegative autoimmune hepatitis in children: Spectrum of disorders. Dig Liver Dis 2016;48:785–91.