Validation of the Simplified Criteria for the Diagnosis of Autoimmune Hepatitis in Chilean-Hispanic Patients

Roberto Candia,* Blanca Norero,* Carlos Agüero,* Luis Díaz,* Juan Pablo Ortega,* Rodrigo Wolff,* Cristian Hernández-Rocha,* Ignacio Duarte,** Alejandro Soza,* Carlos Benítez,* Marco Arrese*

* Departamento de Gastroenterología, Facultad de Medicina, Pontificia Universidad Católica de Chile. Santiago, Chile.
** Departamento de Anatomía Patológica, Facultad de Medicina, Pontificia Universidad Católica de Chile. Santiago, Chile.

INTRODUCTION

Autoimmune hepatitis (AIH) is a progressive inflammatory liver disease of unknown origin and is a frequent cause of liver transplantation.1-6 This disease is characterized by the presence of autoantibodies, high levels of gammaglobulins, interphase hepatitis on biopsy examination and being more frequent in women (sex ratio 3.6:1).1,7 AIH has been described in all ethnic groups around the world. Epidemiologic data from Caucasian population estimate a prevalence ranging from 11 to 24/100,000 inhabitants,5-10 while in Asiatic population the prevalence appears to be lower (4-5/100,000 inhabitants).11,12 In Latin America the information is scarce.

The diagnosis of AIH is challenging for many reasons. First, there is not a specific laboratory test to make the diagnosis with certainty.1 Second, clinical presentation is variable and includes a wide spectrum of conditions: asymptomatic patients, acute and chronic hepatitis, fulminating liver failure and cirrosis.1,7 Finally, its clinical behavior varies among different races. For instance, AIH is more frequently diagnosed at cirrhosis stage in Afro-American and Hispanic patients than in Caucasian population; similarly a worst long-term survival has been described in Asian patients.13,14

In this context, an international panel published a descriptive diagnostic scoring system for AIH in 1993 and then a revised version in 1999.15,16 This criteria is composed by 13 items scored between -5 to +3 points and has been validated in many populations,17-20 however, its application in clinical practice is complicated. To resolve this issue the International Autoimmune Hepatitis Group pro-

ABSTRACT

Introduction and aim. In 2008 the International autoimmune hepatitis (AIH) Group proposed the simplified diagnostic criteria for this disease. The original cohort study was performed in 11 international centers, but validation studies are scarce in Latin-America. The aim of this study is validate these criteria in Hispanic patients. Material and methods. A retrospective cohort of patients undergoing percutaneous liver biopsy and follow-up of at least 12 months was recruited from a Chilean University hospital. Patients with previous immunosuppressive therapy and liver transplant recipients were excluded. The diagnostic accuracy was analyzed using as gold standard the clinical course during long-term follow-up. Sensitivity, specificity, positive and negative predictive values (PPV and NPV) and area under the ROC curve (AUROC) were calculated. Results. Four hundred eighty one patients were evaluated, 294 were included. 218 (74.15%) were female, mean age 48.5 (± 12.3) years, mean follow-up 34 (± 16) months. 66 patients had AIH or overlap syndrome (22.45%), 96 (32.65%) non-alcoholic steatohepatitis, 40 (13.61%) primary biliary cholangitis, 31 (10.54%) hepatitis C, 8 (2.72%) hepatitis B, 53 (18.02%) other etiologies. The AUROC for AIH simplified criteria was 0.976. Using a cutoff ≥ 6 and ≥ 7 points, the sensitivity was 86.4% and 54.6%; specificity, 98.7% and 99.6%; and NPV, 96.2% and 88.6%, respectively. Conclusion. Simplified criteria for the diagnosis of AIH have a high accuracy in our Chilean-Hispanic cohort. The female gender is strongly associated to AIH and could help in difficult cases. Further studies with a prospective design are necessary to confirm these observations.

Key words. Diagnostic criteria. AUROC. AIH. Hepatitis. Hispanic population.
posed a simplified scoring system. These new criteria include only 4 variables and the score varies between 0 and 8 points, with sensitivity and specificity of 88% and 97% at a cutoff of 6 points, respectively. Considering its advantages in clinical application, this simplified score has been validated in Caucasian, Chinese and Japanese population, showing a high performance. In Hispanic Latin-American population the available evidence is limited. Considering this, the main aim of our study is to validate the simplified criteria for the diagnosis of AIH in a cohort of adult Hispanic patients from Chile.

MATERIAL AND METHODS

Patients and diagnostic criteria

All patients undergoing percutaneous liver biopsy in a Chilean university hospital from January 2007 to January 2013 were evaluated. In our center (Hospital Clínico Pontificia Universidad Católica de Chile; Santiago, Chile) the liver biopsy is obtained using the Menghini technique after abdominal ultrasound visualization. We selected a retrospective cohort according to the following inclusion criteria:

- Adults (≥ 18 years old).
- Complete clinical follow-up and laboratory data for at least 12 months after biopsy.
- Patient’s last name or mother’s maiden name from Hispanic origin.

We excluded patients with the following characteristics:

- Previous immunosuppressive therapy.
- Being a liver transplant recipient.
- Incomplete laboratory data (inability to calculate the simplified criteria of AIH or to establish definitive diagnosis).
- Tissue sample that does not allow a suitable interpretation by pathologist (insufficient sample, advanced cirrhosis).

In patients with 2 or more liver biopsies (i.e. patients under histological follow-up to evaluate the progression of a known liver disease) we only included the first biopsy. Clinical records were reviewed to obtain medical history, age, gender, comorbidities, autoimmune serology, levels of immunoglobulin, hepatitis B surface antigen (HBsAg), hepatitis B anti-core antibodies (antibodies HBV), hepatitis C serology (anti-HCV), histology, therapy and clinical evolution. If data were incomplete or confusing, we contacted the treating gastroenterologist in order to obtain further details.

The study was approved by the Institutional Review Board Ethics Committee for Human Studies of the Pontificia Universidad Católica de Chile (Project number: 14-449).

Determination of definitive diagnosis (gold standard)

AIH does not have a diagnostic gold standard, for that reason in our study the definitive diagnosis was established according to clinical follow-up, as it has been proposed by previous studies and QUADAS-2. We considered a follow-up of at least 12 months as an appropriate minimum time to determine the final diagnosis of the liver disease. The definitive diagnosis was established by the treating gastroenterologist, as it was described in the medical record. Two authors (BN and RC) retrospectively reviewed all medical records and confirmed the diagnosis. We contacted the treating gastroenterologist in case of discrepancies, which were resolved by consensus. The definitive diagnostic was established according to criteria described below:

- For the diagnosis of AIH we took in consideration clinical characteristics and laboratory tests similar to the revised criteria proposed by the International Autoimmune Hepatitis Group (IAIHG): female gender; acute hepatitis or persistent abnormal liver tests for more than 2 months and negative viral serology: anti-HAV (hepatitis A virus) IgM, anti-HCV, antico-re-HBV IgM and HBsAg (-); serum IgG levels over the upper limit of normal level; a positive test for antinuclear antibody (ANA) and/or anti-smooth muscle antibody (SMA); daily alcoholic intake < 20 g; a liver biopsy showing moderate or severe periporal or periseptal hepatitis with lymphocytic piecemeal necrosis. The most important factor to establish the gold standard was the clinical course of the disease during the follow-up (at least 12 months). We did not calculate the revised diagnostic score proposed by IAIHG in 1999, because the most important criterion was clinical follow-up. Similarly to original studies, patients with overlap syndrome were included as AIH.
- The diagnosis of non-alcoholic steatohepatitis (NASH) was established according to the following criteria:
  a. Persistent abnormal liver tests for more than 2 months.
  b. Liver biopsy compatible with NASH.
  c. Daily alcohol consumption of < 20 g.
  d. Appropriate exclusion of other liver diseases.
Chronic hepatitis C was considered the definitive diagnosis when a positive HCV antibody test and quantifiable HCV RNA in serum were positive and liver biopsy was compatible with the diagnosis of chronic hepatitis C. Overlap syndrome was considered the definitive diagnosis when the patient simultaneously met the criteria for PBC or primary sclerosing cholangitis and AIH. An adequate response to double therapy (immunosuppressive treatment plus ursodeoxycholic acid) was considered a supporting factor for this diagnosis.

Simplified diagnostic criteria of AIH

We applied the simplified criteria for the diagnosis of AIH according to the original study to all included patients.21 These criteria include 4 variables and the score varies between 0 and 8 points. The variables used were the following:

- Autoimmune serology. ANA, SMA, liver-kidney microsomal antibodies (LKM). If ANA/SMA \( \geq 1:40 \), we assigned 1 point; if ANA/SMA \( \geq 1:80 \) or LKM \( \geq 1:40 \), we assigned 2 points. For this variable the maximum score was 2 points. In our center soluble liver/liver-pancreas antibodies is not available.
- Immunoglobulin G levels. If levels > upper normal limit, we assigned 1 point; if levels > 1.1 times the upper normal limit, we assigned 2 points. For this variable the maximum score was 2 points.
- Liver histology (evidence of hepatitis was a necessary condition). If the findings were compatible, we assigned 1 point; if the findings were typical, we assigned 2 points. For this variable the maximum score was 2 points. Liver histology was evaluated by the local pathologists, who were not blinded to the patient’s history.
- Viral hepatitis status. If there was not viral hepatitis evidence, we assigned 2 points.

Each medical record was evaluated by two authors to calculate the simplified criteria. The discrepancies were resolved by consensus.

Statistical analysis

Continuous data were described using mean and standard deviation (SD) or median and interquartile range (25 and 75 percentiles) (IQR), according to data distribution. Nominal data were described using percentages. A multivariate logistic regression model was conducted to establish if age or gender are independently associated to AIH after adjustment by the simplified diagnostic criteria. Odds ratio (OR) and 95% confidence interval (CI) were used to report these results. Additionally, we calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and likelihood ratio (LR) of the simplified diagnostic criteria of AIH with diagnos-
tic cutoffs 5, 6 and 7 points. The global diagnostic accuracy was assessed by the area under the receiver operating characteristic curve (AUROC). Statistical analysis was performed with SAS software version 9.4 (SAS Institute, Cary, NC). A p value of < 0.05 was considered statistically significant.

RESULTS

General characteristics

A total of 481 patients were evaluated, 294 met the inclusion criteria. The main causes of exclusion were cirrhosis and liver transplanted patients. Two hundred eighteen patients (74.15%) were female, the mean age was 48.5 years (SD: 12.3), the median follow-up 33 months (IQR: 18-51) (Table 1). The distribution of patients according to the score obtained in the simplified diagnostic criteria of AIH and the definitive diagnosis is detailed in figure 1.

The most frequent liver disease detected was NASH (96 patients, 32.65%). AIH was the second diagnosis most frequent (66 patients, 22.45%). Other liver diseases detected were PBC (40 patients, 13.61%), hepatitis C (31 patients, 10.54%), DILI (9 patients, 3.06%), hepatitis B (8 patients, 2.72%), hemochromatosis (6 patients, 2.04%) and self-limited hepatitis or cholestasis of unknown origin (16 patients, 5.44%). Sixteen female patients (5.44%) had overlap syndrome. General characteristics of patients and the simplified score for AIH according to the main causes of liver disease are described in tables 1 and 2.

In the multivariable analysis female gender was strongly associated to the diagnosis of AIH (OR: 13.9, 95% CI: 1.9-100.6, p = 0.009), after adjustment by age (OR: 0.98, 95% CI: 0.94-1.03, p = 0.5) and the simplified diagnostic criteria (OR: 8.6, 95% CI: 4.5-16.6, p < 0.001).

Table 1. General characteristics of included patients.

<table>
<thead>
<tr>
<th>Total, n (%)</th>
<th>294 (100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, n (%)</td>
<td>218 (74.15%)</td>
</tr>
<tr>
<td>Age (years), mean, SD.</td>
<td>48.5 ± 12.3</td>
</tr>
<tr>
<td>Follow up (months) median, (IQR).</td>
<td>33 (18-51)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis, n (%)</th>
<th>96 (32.65%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NASH</td>
<td>66 (22.45%)</td>
</tr>
<tr>
<td>Autoimmune hepatitis</td>
<td>40 (13.61%)</td>
</tr>
<tr>
<td>Primary biliary cholangitis</td>
<td>31 (10.54%)</td>
</tr>
<tr>
<td>Hepatitis C virus</td>
<td>9 (3.06%)</td>
</tr>
<tr>
<td>DILI</td>
<td>8 (2.72%)</td>
</tr>
<tr>
<td>Hepatitis B virus</td>
<td>6 (2.04%)</td>
</tr>
<tr>
<td>Hemochromatosis</td>
<td>3 (1.02%)</td>
</tr>
<tr>
<td>Sarcoidosis/Granulomatous hepatitis</td>
<td>3 (1.02%)</td>
</tr>
<tr>
<td>Coeliac disease</td>
<td>2 (0.68%)</td>
</tr>
<tr>
<td>Amyloidosis</td>
<td>2 (0.68%)</td>
</tr>
<tr>
<td>Primary sclerosing cholangitis</td>
<td>1 (0.34%)</td>
</tr>
<tr>
<td>Alcoholic hepatitis</td>
<td>16 (5.44%)</td>
</tr>
<tr>
<td>Self-limited hepatitis and/or cholestasis</td>
<td>11 (3.74%)</td>
</tr>
</tbody>
</table>

Simplified score for diagnosis of AIH, median (IQR). 3 (2-5)

Diagnostic accuracy of simplified criteria of AIH

In our cohort the AUROC of the simplified criteria for the diagnosis of AIH was 0.976 (Figure 2). The sensitivity, specificity, PPV, NPV and LR were calculated using diagnostic criteria of $\geq 5$, $\geq 6$ and $\geq 7$ points (Table 3).

When a cutoff $\geq 6$ points is considered diagnostic of AIH, the performance of the simplified criteria looks suitable, with sensitivity and specificity of 86.4% and 98.7%, respectively, and LR (+) of 65.6. Using $\geq 7$ points as diagnostic criteria, the sensitivity decreases to 54.6%, but specificity and LR (+) increase to 99.6% and 124.3, respectively. Using $\geq 5$ points as diagnostic criteria, the sensitivity was 93.9%, but specificity, PPV and LR (+) decrease to 90.4%, 73.8% and 9.7, respectively. For more details, PPV, NPV and LR (-) (Table 3).

Description of false negative and false positive patients

Using a cutoff $\geq 6$ as diagnostic criteria, 12 patients were misclassified:

- **False negatives.** Nine patients had score $\leq 6$ and the clinical follow-up was compatible with AIH. Five patients had 5 points in the simplified diagnostic score, 6 had seronegative AIH and 5 had normal IgG levels. Despite this, all false negative were female, had successful response to corticosteroids or immunosuppressive therapy and biopsy was suggestive or compatible with AIH.

- **False positives.** Three patients had score $\geq 6$ but the clinical follow-up was not compatible with AIH. One patient (male gender) obtained 6 points in the simplified score, but he had acquired immune deficiency syndrome (AIDS) and the biopsy was compatible with a granulomatous hepatitis. The other male patient had 7 points in the simplified score and the histological evaluation suggested an overlap syndrome (PBC plus...

Table 2. Characteristics of included patients according to main diagnoses.

<table>
<thead>
<tr>
<th></th>
<th>Age, years Mean (SD)</th>
<th>Sex, female n (%)</th>
<th>Follow-up, months Median (IQR)</th>
<th>Simplified score of AIH, points Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autoimmune hepatitis</td>
<td>46.2 (12.9)</td>
<td>62 (93.9%)</td>
<td>30 (20-56)</td>
<td>7 (6-8)</td>
</tr>
<tr>
<td>NASH</td>
<td>48.8 (12.3)</td>
<td>68 (70.8%)</td>
<td>34 (18-53)</td>
<td>2 (2-4)</td>
</tr>
<tr>
<td>Primary biliary cirrhosis</td>
<td>52.2 (10.5)</td>
<td>37 (92.5%)</td>
<td>30 (17-45)</td>
<td>4 (2-5)</td>
</tr>
<tr>
<td>Hepatitis C virus</td>
<td>50.8 (11.4)</td>
<td>17 (54.8%)</td>
<td>35 (15-49)</td>
<td>0 (0-1)</td>
</tr>
<tr>
<td>DILI</td>
<td>49 (13.5)</td>
<td>9 (100%)</td>
<td>31 (23-50)</td>
<td>3 (2-4)</td>
</tr>
<tr>
<td>Hepatitis B virus</td>
<td>48.5 (8.5)</td>
<td>1 (12.5%)</td>
<td>21 (14-49)</td>
<td>0 (0-1)</td>
</tr>
<tr>
<td>Hemochromatosis</td>
<td>51.5 (11.1)</td>
<td>0 (0%)</td>
<td>37 (30-47)</td>
<td>2 (2-2)</td>
</tr>
<tr>
<td>Self-limited hepatitis or cholestasis</td>
<td>41.7 (15.3)</td>
<td>12 (75%)</td>
<td>31 (19-52)</td>
<td>3 (2-4)</td>
</tr>
</tbody>
</table>

AIH), however, he had an excellent response to ursodeoxycholic acid (UDCA). Considering that this patient never received immunosuppressive therapy, the definitive diagnosis was established as PBC. The third patient (female gender) had 6 points in the simplified score, however, her hepatitis was detected in relation to use of nitrofurantoin. The patient suspended the drug and received a short course of corticoids with an excellent response. After 5 years of follow up she is asymptomatic and has normal liver tests, for that reason the definitive diagnosis was DILI.

DISCUSSION

In our retrospective cohort of Chilean-Hispanic patients, the simplified score of AIH had a high global diagnostic accuracy with an AUROC of 0.976. According to our data, the most suitable cutoff is ≥ 6 points, with sensitivity 86.4%, specificity 98.7% and predictive values over 95%. These findings are similar to the original study and other validation articles.21-28 Despite this, our data had a lower performance than previous studies using cutoff ≥ 7 points.21-28 The global accuracy of these criteria could be put in practice using a dynamic approach: Considering the LR (+) of 65 and specificity over 98% at cutoff ≥ 6 points and LR (+) of 124 and specificity over 99% at cutoff ≥ 7 points, patients with 6, 7 and 8 points could confidently start with immunosuppressive therapy. On the other hand, considering the LR (-) < 0.1 and sensitivity of 93.9% at cutoff ≥ 5, AIH could be safely ruled out in patients with less than 5 points, especially if they are male. A challenging scenario is a patient with 5 points: in our cohort 20% of patients with 5 points had AIH (Figure 1). Considering that at the same age and simplified diagnostic score the OR of having AIH is almost 14 times higher in female than male patients, our suggestion is to consider immunosuppressive therapy as a suitable strategy if the patient is a woman with no other obvious diagnostic alternative.

According to our literature review, this is the first validation study of the simplified criteria for the diagnosis of AIH focused in Hispanic patients. This point matters, because this ethnic group has been under-represented in previous studies.21-28 Although our study was only conducted on Chilean patients, the genetic background of this population is similar to other Hispanic populations, especially those with low African heritage.39-44 Considering this, we feel that our data could be extrapolated to other Latin-American countries, except Brazil and some countries from Central-America, where the African genetic background is > 5%.45

The main limitation of our study is related to its retrospective design. On one hand, there is risk of measure-bias secondary to mistakes of data collection; on the other hand, the calculation of simplified score of AIH was not blinded to our gold standard, because the definitive diagnosis was always available in the medical record. To address these limitations and reduce the risk of bias, all data and the calculation of the simplified score of AIH were obtained by one author and afterwards reviewed independently by a second author. All discrepancies were resolved by consensus.

Our research study has strengths. The previous validation studies have used as gold standard the diagnostic criteria revised by IAIHG in 1999.16 Considering that both diagnostic criteria have similar items (auto-antibodies, viral serology, immunoglobulins levels), it is highly probable a biased overestimation of the diagnostic accuracy of simplified criteria. In our opinion the definitive diagnosis of AIH established with long-term follow-up is a more suitable gold standard. This situation could explain the differences observed among our data and previous studies, specifically our lower sensitivity and NPV at cutoff ≥ 7 points. Other strength of our study is the spectrum of patients included in our cohort. Some validation studies mainly included as controls patients with PBC, NASH or viral hepatitis,21,23,26 all of them with a clinical presentation easily differentiable from AIH without sophisticated diagnostic tools. This situation usually overestimates the diagnostic accuracy of the new test through spectrum bias.33 A diagnostic tool for AIH is useful in clinical practice when it is able to resolve more challenging scenarios, as DILI, self-limited hepatitis/cholestasis of unknown origin and other immunological conditions. In our cohort approximately 10% of patients met these conditions, decreasing the risk of spectrum bias.

In summary, this study supports that simplified criteria for the diagnosis of AIH have a high diagnostic accuracy in Chilean-Hispanic population. The female gender is strongly associated to AIH and could help in difficult cases. Further studies with a prospective design and using as gold-standard long-term follow-up are necessary to confirm these observations.

ABBREVIATIONS

- AIDS: acquired immune deficiency syndrome.
- AIH: autoimmune hepatitis.
- ANA: antinuclear antibody.
- Anti-HCV: hepatitis C serology.
- AUROC: area under the receiver operating characteristic curve.
- DILI: Drug-induced liver injury.
- HAV: hepatitis A virus.
- HBsAg: hepatitis B surface antigen.
HBV: hepatitis B virus.
HCV: hepatitis C virus.
IAIHG: International Autoimmune Hepatitis Group.
IQR: interquartile range.
LKM: liver-kidney microsomal antibodies.
LR: likelihood ratio.
NASH: non-alcoholic steatohepatitis.
NPV: negative predictive value.
PBC: primary biliary cholangitis.
PPV: positive predictive value.
SD: standard deviation.
SMA: anti-smooth muscle antibody.
UDCA: ursodeoxycholic acid.

CONFLICTS OF INTEREST
The authors of this article do not have conflicts of interest.
This article was not funded.

REFERENCES
28. Gatselis NK, Zachou K, Papamichalis P, Koukoulis GK, Gabe ta S, Dalekos GN, Rigopoulou EI. 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.


Correspondence and reprint request:
Roberto Candia Balboa, M.D. MSc.
Departamento de Gastroenterología. Facultad de Medicina.
Pontificia Universidad Católica de Chile.
Diagonal Paraguay, No. 362, 4th Floor, Office 424.
Postal code: 833-0077. Santiago, Región Metropolitana, Chile.
Tel./Fax: 56-2-23543820
E-mail: rcandiab@med.puc.cl