International multicenter validation of GES score for HCC risk stratification in chronic hepatitis C patients

Gamal Shiha, Reham Soliman, Nabiel Mikhail, Fabrice Carrat, Jessica Azzi,
Ganne-Carrié Nathalie, Hidenori Toyoda, Haruki Uojima, Akito Nozaki,
Koichi Takaguchi, et al.

HAL Id: hal-03785890

https://hal.sorbonne-universite.fr/hal-03785890v1

Submitted on 23 Sep 2022

Shiha Gamal (Orcid ID: 0000-0002-9338-8854)
Toyoda Hidenori (Orcid ID: 0000-0002-1652-6168)
Atsukawa Masanori (Orcid ID: 0000-0003-3374-7111)
Sarin Shiv Kumar (Orcid ID: 0000-0002-0544-5610)
mousa Nasser hamed (Orcid ID: 0000-0001-8329-2587)

International Multicenter Validation of GES score for HCC risk stratification in chronic hepatitis C patients

Shiha G[1,2] · Soliman R[1,3], Mikhail NNH[1,4], Carrat F [5,6], Azzi J [5], Nathalie Ganne-Carrié [7,8,9], Toyoda H [10], Uojima H [11], Nozaki A [12], Takaguchi K [13], Hiraoka A [14], Atsukawa M [15], Abe H [16], Matsuura K [17], Mikami S [18], Watanabe T [19], Tsuji K [20], Ishikawa T [21], Suri V [22], Osinusi A [22], Ni L [22], Zou J [22], Sarin S [23], Kumar M [23], Jalal PK [24], Hashim MA [24], Hassan M [24,25], Lopez SA [26], Bañares R [26], Ahumada AM [27], Mousa N [28], Eslam M [29], Waked I [30]

[1] Egyptian Liver Research Institute and Hospital (ELRIAH), Sherbin, El Mansoura, Egypt
[2] Hepatology and Gastroenterology Unit, Internal Medicine Department, Faculty of Medicine, Mansoura University, Egypt

[3] Tropical Medicine Department, Faculty of Medicine, Port Said University, Egypt
[4] Biostatistics and Cancer Epidemiology Department, South Egypt Cancer Institute, Assiut University, Egypt
[5] Sorbonne Université, Institut National de la Santé et de la Recherche Médicale (INSERM), Institut Pierre Louis d'Epidémiologie et de Santé Publique, Paris, France
[6] AP-HP, Sorbonne Université, Hôpital Saint-Antoine, Santé Publique, Paris, France.
[7] AP-HP, Hôpitaux Universitaires Paris Seine Saint-Denis, APHP, Liver Unit, Bobigny, France.
[8] Université Sorbonne Paris Nord, Bobigny, France
[9] Inserm, UMR-1138 « Functional Genomics of solid tumors », Centre de Recherche des Cordeliers, Université de Paris, France.
[10] Department of Gastroenterology, Ogaki Municipal Hospital, Ogaki, Japan

[11] Department of Gastroenterology, Internal Medicine, Kitasato University School of Medicine, Sagamihara, Japan
[12] Gastroenterology Center, Yokohama City University Medical Center, Yokohama, Japan
[13] Department of Hepatology, Kagawa Prefectural Central Hospital, Takamatsu, Japan
[14] Gastroenterology Center, Ehime Prefectural Central Hospital, Matsuyama, Japan
[15] Department of Internal Medicine, Division of Gastroenterology and Hepatology, Nippon Medical School, Tokyo, Japan

[16] Department of Internal Medicine, Division of Gastroenterology and Hepatology, Shinmatusdo Central General Hospital, Matsudo, Japan

[17] Department of Virology & Liver Unit, Nagoya City University, Graduate School of Medical Sciences, Nagoya, Japan

[18] Department of Internal Medicine, Division of Gastroenterology, Kikkoman General Hospital, Noda, Japan

[19] Department of Internal Medicine, St. Marianna University School of Medicine, Kawasaki, Japan

[20] Center for Gastroenterology, Teine Keijinkai Hospital, Sapporo, Japan

[21] Department of Hepatology, Saiseikai Niigata Hospital, Niigata, Japan

[22] Gilead Sciences, Inc, Foster City, California, USA

[23] Department of Hepatology, Institute of Liver and BiliarySciences, New Delhi, India

[24] Division of Abdominal Transplantation Baylor College of Medicine, USA

[25] Department of Epidemiology, Division of Cancer Prevention and Population Sciences, MD Anderson, USA

[26] Liver Unit, Hospital General Universitario Gregorio Marañon, Madrid, Spain; Instituto De Investigación SanitariaGregorio Marañón (IiSGM), Centro de Investigación Biomédica En Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), Madrid, Spain

[27] Liver Unit, Hospital General Universitario Gregorio Marañon, Madrid, Spain

[28] Tropical Medicine Department, Faculty of Medicine, Mansoura University, Egypt

[29] Storr Liver Centre, Westmead Institute for Medical Research, Westmead Hospital and University of Sydney, NSW, Australia

[30] Hepatology Department, National Liver Institute, Menoufia University, Shebeen, Elkom, Egypt

Corresponding author:

Prof. Gamal Shiha, Internal Medicine Department, Faculty of Medicine, Mansoura University, Egypt. CEO, Egyptian Liver Research Institute and Hospital (ELRIAH), Mansoura, Egypt

Email: g_shiha@hotmail.com

Mob: +2 01212222765

Running title: International validation of GES for HCC risk

Number of tables: 2, supplementary tables: 1

Number of figures: 1

Words: 3298

ANRS-AFEF Hepather Study group

Investigators

Laurent Alric, Delphine Bonnet, Virginie Payssan-Sicart, Chloe Pomes (CHU Purpan, Toulouse, France), Fabien Zoulim, Marianne Maynard, Roxane Bai, Lucie Hucault, François Bailly (Hospices Civils de Lyon, Lyon, France), François Raffi, Eric Billaud, David Boutoille, Maeva Lefebvre, Elisabeth André-Garnier (Hôpital Hôtel-Dieu, Nantes, France), Paul Cales, Isabelle Hubert, Adrien Lannes, Françoise Lunel, Jérôme Boursier (CHU Angers, Angers, France), Tarik Asselah, Nathalie Boyer, Nathalie Giuily, Corinne Castelnau, Giovanna Scoazec (Hôpital Beaujon, Clichy, France), Stanislas Pol, Hélène Fontaine, Emilie Rousseaud, Anaïs Vallet-Pichard, Philippe Sogni (Hôpital Cochin, Paris, France), Victor de Ledinghen, Juliette Foucher, Jean-Baptiste Hiriart, Jancell M'Bouyou, Marie Irlès-Depé (Hôpital Haut-Lévêque, Pessac, Bordeaux, France), Marc Bourlière, Si Nafa Si Ahmed, Valérie Oules (Hôpital Saint Joseph, Marseille, France), Albert Tran, Rodolphe Anty, Eve Gelsi, Régine Truchi (CHU de Nice, Nice, France), Dominique Thabut, Saloua Hammeche, Joseph Moussali (Hôpital de la Pitié Salptétrière, Paris, France), Xavier Causse, Barbara De Dieuleveult, Brahim Ouarani, Damien Labarrière (CHR La Source, Orléans, France), Nathalie Ganne, Véronique Grando-Lemaire, Pierre Nahon, Séverine Brulé, Betul ULKER (Hôpital Jean Verdier, Bondy, France), Dominique Guyader, Caroline Jezequel, Audrey Brener, Anne Laligant, Aline Rabot, Isabelle Renard (CHU Rennes, Rennes, France), François Habersetzer, Thomas F. Baumert, Michel Doffoel, Catherine Mutter, Pauline Simo-Noumbissie, Esma Razi (Hôpitaux Universitaires de Strasbourg, Strasbourg, France), Jean-Pierre Bronowicki, Hélène Barraud, Mouni Bensenane, Abdelbasset Nani, Sarah Hassani-Nani, Marie-Albertine Bernard (CHU de Nancy, Nancy, France), Georges-Philippe Pageaux, Dominique Larrey, Magda Meszaros (Hôpital Saint Eloi, Montpellier, France), Sophie Metivier, Christophe Bureau, Thibault Morales, Jean Marie Peron, Marie Angèle Robic (CHU Purpan, Toulouse, France),

Thomas Decaens, Marine Faure, Bruno Froissart, Marie-Noelle Hilleret, Jean-Pierre Zarski (CHU de Grenoble, Grenoble, France), Ghassan Riachi, Odile Goria, Fatima Paris, Hélène Montialoux (CHU Charles Nicolle, Rouen, France), Vincent Leroy, Giuliana Amaddeo, Anne Varaut, Mélanie Simoes, Rachida Amzal (Hôpital Henri Mondor, Créteil, France), Olivier Chazouillières, Tony Andreani, Bénédicte Angoulevant, Azeline Chevance, Lawrence Serfaty (Hôpital Saint-Antoine, Paris, France), Didier Samuel, Teresa Antonini, Audrey Coilly, Jean-Charles Duclos Vallée, Mariagrazia Tateo (Hôpital Paul Brousse, Villejuif, France), Armand Abergel, Maud Reymond, Chanteranne Brigitte, Buchard Benjamin, Léon Muti (Hôpital Estaing, Clermont-Ferrand, France), Claire Geist, Guillaume Conroy, Raphaëlle Riffault (Centre Hospitalier Régional, Metz, France), Isabelle Rosa, Camille Barrault, Laurent Costes, Hervé Hagège (Centre Hospitalier Intercommunal, Créteil, France), Véronique Loustaud-Ratti, Paul Carrier, Maryline Debette-Gratien, (CHU Limoges, Limoges, France), Philippe Mathurin, Guillaume Lassailly, Elise Lemaitre, Valérie Canva, Sébastien Dharancy, Alexandre Louvet (CHRU Claude Huriez, Lille, France), Anne Minello, Marianne Latournerie, Marc Bardou, Thomas Mouillot (Dijon University Hospital, Dijon, France), Louis D'Alteroche, Didier Barbereau, Charlotte Nicolas, Laure Elkrief, Anaïs Jaillais (CHU Trousseau, 37044 Tours, France), Jérôme Gournay, Caroline Chevalier, Isabelle Archambeaud, Sarah Habes (CHU de Nantes, Nantes, France), Isabelle Portal (CHU Timone, Marseille, France), Moana Gelu-Simeon, Eric Saillard, Marie-Josée Lafrance, Lucie Catherine (CHU de Pointe-à-Pitre, Pointe-à-Pitre, Guadeloupe).

Methodology and Coordinating Centre.

Fabrice Carrat (coordinator), Frederic Chau, Céline Dorival, Isabelle Goderel, Clovis Lusivika-Nzinga, Marc-Antoine Bellance, Jonathan Bellet, Priscilla Monfalet, Jessica Chane-Teng, Sephora Bijaoui, Grégory Pannetier, François Téoulé, Jérôme Nicol, Florian Sebal, Rafika Bekhti (Sorbonne University & INSERM U1136 - IPLESP, Paris, France).

Abstract

Background and Aims:

We have recently demonstrated the ability of a simple predictive model (GES) score to determine the risk of hepatocellular carcinoma (HCC) after using direct-acting antivirals. However, our results were restricted to Egyptian patients with hepatitis C virus (HCV) genotype 4. Therefore, we studied a large, independent cohort of multiethnic populations through our international collaborative activity.

Methods:

Depending on their GES scores, patients are stratified into low risk (≤6/12.5), intermediate risk (>6–7.5/12.5), and high risk (>7.5/12.5) for HCC. A total of 12038 patients with chronic HCV were analyzed in this study, of whom 11202 were recruited from 54 centers in France, Japan, India, the U.S., and Spain, and the remaining 836 were selected from the Gilead-sponsored randomized controlled trial conducted across the U.S., Europe, Canada, and Australia. Descriptive statistics and log-rank tests. The performance of the GES score was evaluated using Harrell's C-index (HCI).

Results:

The GES score proved successful at stratifying all patients into 3 risk groups, namely low-risk, intermediate-risk, and high-risk. It also displayed significant predictive value for HCC development in all participants (P< 0.0001), with HCI ranging from 0.55 to 0.76 among all cohorts after adjusting for HCV genotypes and patient ethnicities.

Conclusion:

The GES score can be used to stratify HCV patients into 3 categories of risk for HCC, namely low-risk, intermediate-risk, and high-risk, irrespective of their ethnicities or HCV genotypes. This international multicenter validation may allow the use of GES score in individualized HCC risk-based surveillance programs.

Word count: 250

Summary Box:

What is already known about this subject?

Many HCC prediction models have been recently proposed for HCC risk stratification which is vital for HCC surveillance programs. These scores either rely on readily available clinical and laboratory parameters or depend on molecular and genetic factors or complicated mathematical methods. Almost all these scores have not been validated neither outside the country of origin nor in all HCV genotypes. Hence, thier use for HCC risk stratification cannot be generalized.

What are the new findings?

Using readily available clinical and laboratory parameters (Age, Sex, S. albumin, S. AFP and fibrosis stage) we developed HCC risk prediction score (GES) which was internally and externally validated in HCV genotype 4 only. Herein, GES score has been validated in 12038 CHC patients with all HCV genotypes from 5 countries in 54 centers in addition to a randomized clinical trial in diverse countries. Log-rank test demonstrated highly significant predictive value of GES score in the overall validation cohorts (P< 0.0001) with Harrell's C ranging from 0.55 to 0.76 among all cohorts after adjusting for HCV genotypes and patient ethnicities.

How might it impact on clinical practice in the foreseeable future?

This international multicenter validation may pave the way for using the GES score in individualized HCC risk-based surveillance programs. This simple score can be used to stratify HCV patients into low, intermediate, and high-risk groups, irrespective of their ethnicities or HCV genotypes.

Introduction:

It is estimated that 70 million people are infected with chronic hepatitis C virus (HCV), which remains a major cause of cirrhosis and hepatocellular carcinoma (HCC) worldwide [1]. The availability of highly effective, well-tolerated direct-acting antivirals (DAAs) has allowed for increased HCV testing and improved access to HCV care and treatment. Eradication of HCV by antiviral therapy reduces but does not eliminate the risk of HCC. In particular, patients with liver cirrhosis (stage F4) and advanced fibrosis (stage F3) have a substantial residual risk of HCC that remains after viral eradication [2,3].

Current guidelines recommend biannual screening of patients with cirrhosis for HCC by using ultrasound with or without alpha-fetoprotein (AFP) measurement [4,5]. These recommendations have been validated by data suggesting improved survival, higher rates of early tumor detection, and curative treatments among patients screened for HCC [4]. However, the "one-size-fits-all" strategy increases the burden on the health care system—particularly in low- to middle-income countries, where there is a high prevalence of HCV infection; moreover, a limited number of patients with cirrhosis undergo surveillance consistent with the guidelines, emphasizing the urgent unmet clinical need to develop a better prediction model to guide recommendations regarding HCC surveillance among patients with advanced fibrosis who have sustained virological response (SVR) [5].

Recently, there has been a growing interest in the use of HCC prediction scores as an important prognostic tool to enable individualized surveillance for HCC. Several HCC prediction scores have been proposed in the last 5 years, with some relying on molecular and genetic risk factors [6-8] or complicated mathematical methods [9] and others depending on readily available clinical and laboratory parameters [10-11]. However, most of these scores are either too expensive or have not been validated outside the country of origin.

We have recently developed a predictive model called General Evaluation Score (GES) as a simple tool for HCC risk stratification in patients with chronic HCV who achieved SVR following treatment with DAAs [12]. This score, which incorporates readily available clinical and laboratory parameters (including albumin and AFP levels, stage of liver fibrosis plus age, and gender),

could successfully stratify HCV patients into 3 different categories of risk for HCC with a rational predictive ability (log-rank, P< 0.001; Harrell's C-index, 0.801). However, our previous results were restricted to patients with HCV genotype 4 only.

In a previous study, Bergna et al. externally validated the GES score in a single-center cohort of European cirrhotic patients (mostly with HCV genotype 1) in Italy and reported that the score was able to stratify 577 patients into low-risk (n = 188, 32.5 %), intermediate-risk (n = 243, 42.1%), and high-risk (n = 146, 25.3%) groups, with the 5-year cumulative incidence of HCC being 4.7%, 10%, and 13.8%, respectively (log-rank, P = 0.01); nevertheless, they did not present Harrell's C-index [13,14]. Very recently, a Japanese study examining 689 patients with chronic HCV (the FLAG cohort) over an observation period of 35.25 ± 13.24 (range, 0–55) months validated the GES score, demonstrating successful use of this score for stratifying their patients into 3 risk groups with significant log-rank test results (P < 0.001) and a Harrell's C-index of 0.6919.

Here, we aimed to evaluate the performance and clinical utility of the GES score as a tool for HCC risk stratification in a large, independent cohort of patients with different ethnicities and various HCV genotypes from multiple centers through our international collaborative activity.

Patients and methods:

Cohorts

This study was conducted in 5 countries (i.e., France, Japan, India, the U.S., and Spain) and included cohorts from 54 centers. Patients were also recruited from the Gilead-sponsored randomized controlled trial (RCT) conducted across the U.S., Europe, Canada, and Australia.

Patients were included if they met the following criteria: age≥18 years, infection with HCV, treatment with DAAs, and no history of or current HCC. Patients with either hepatitis B virus or human immunodeficiency virus coinfection as well as those with a previous history of interferon therapy, liver transplantation, renal impairment, or other malignancies were excluded. All participants underwent a course of treatment with one of several DAAs for chronic HCV infection according to local guidelines.

This study was conducted in accordance with the principles of the Declaration of Helsinki and its amendments in 2008. The protocol was approved by the Institutional Research Board of each participating center, as per local regulations. The need to obtain informed consent from the participants was waived due to the retrospective nature of the study.

Patient evaluation

Clinical and laboratory data were collected at 6-month intervals, from the initiation of antiviral treatment until the last visit, according to the treatment protocol. All patients underwent virological, hematological, and biochemical laboratory testing, abdominal ultrasound examination, and triphasic multislice spiral computed tomography (CT) if indicated. The follow-up duration was calculated as the time between treatment termination and the last follow-up session or the date of event development (HCC occurrence), whichever occurred first.

Biochemical parameters included alanine aminotransferase, aspartate aminotransferase, prothrombin time, international normalized ratio, total bilirubin, albumin, platelets, and AFP.

HCC diagnosis

The diagnosis of HCC was made in accordance with the European Association for the Study of the Liver and American Association for the Study of Liver Diseases guidelines. Multiphase CT or magnetic resonance imaging (MRI) was performed if the patients had any focal hepatic lesions diagnosed by abdominal ultrasound and/or AFP values >20 ng/mL, with the hallmark diagnostic features of HCC being arterial enhancement and early washout in the delayed phase.

Statistical analysis

Statistical analyses were performed using SPSS version 26 (IBM Corp., USA). Continuous variables were reported as median (interquartile range). Categorical variables were expressed as frequency (percentage). Cox regression analysis, and log-rank test were used to evaluate the effect of the GES score on the cumulative hazard of HCC.

The performance of the GES score was evaluated using:

- Overall performance by Brier score. The lower the Brier score is for a set of predictions, the better the predictions are calibrated. [15]
- Discrimination by Harrell's C-statistic. A rough rule for interpretation is that values above 0.80 indicate very good models; between 0.70 and 0.80, good models; and between 0.50 and 0.70, fair models [16]
- Calibration using Hosmer-Lemeshow test. The output returns a chi-square value (a Hosmer-Lemeshow chi-squared) and a p-value. Small p-values mean that the model is a poor fit. Small p-values (usually under 5%) mean that model is not a good fit [17]
- Evaluating the performance of the risk stratification as a screening procedure against HCC development as the gold standard. Using the risk stratification results, patients are classified into risky group (intermediate and high risk score) and less-risky group (low risk score) and then performance statistics (sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy) are calculated.

Statistical analysis of all data was carried out at ELRIAH, except for the French (HEPATHER) and Gilead RCT cohorts, whose data were analyzed in-house due to internal regulations preventing original data transfer.

Results:

The present study included a total of 12038 HCV patients, of whom 11202 were enrolled from 54 centers in France, Japan, India, the U.S., and Spain, and the remaining 836 were selected from the Gilead-sponsored RCT conducted across the U.S., Europe, Canada, and Australia. Cases with incomplete data were removed from the analysis. The observation period ranged from 0 month to 76 months (Table 1).

France cohorts

These cohorts included 7752 patients from 32 centers. The patients had a median age of 56.2 (50.4–64.3) years, and 54.0% were male. The most common HCV genotype was genotype 1 (62.2%), followed by genotype 3 (11.8%) and genotype 4 (11.8%). Patients with chronic HCV were classified under F0, F1, and F2 (n= 3291), F3 (n= 1097), and F4 (n= 2840). The observation period was 31.85±16.26 (range, 19–45) months after the end of DAA therapy (Table 1 and supplement table).

Based on the GES score, 5098 (65.76%), 1884 (24.30%), and 770 (9.93%) of the studied patients were found to be at low, intermediate, and high risk for HCC, respectively. In total, 227 cases of HCC were observed during the study period, of which 37 occurred in the low-risk group (37/5098, 0.73%), 91 in the intermediate-risk group (91/1884, 4.83%), and 99 in the high-risk group (99/770, 12.86%).

The 5-year cumulative incidence of HCC was 1.23% in the low-risk group with a 95% confidence interval (CI) of 0.84%–1.75%, 7.95% (95% CI, 6.19%–9.98%) in the intermediate-risk group, and 17.69% (95% CI, 14.25%–21.45%) in the high-risk group. Analysis of the cumulative incidence of HCC showed a highly significant difference between the 3 risk groups (P< 0.0001; Fig. 1-A). Harrell's C-statistic for this external validation group was 0.7688. The area under the receiver operating characteristic curve was 0.74 at 36 months and 0.72 at 48 months after the end of DAA treatment.

Japan cohorts

These cohorts consisted of 2331 patients from 12 centers. The median age was 71.0 (63.0–77.0) years, and males represented 44.4% of all patients. The study included 1003 (43.0%) cirrhotic patients, who were followed up for 29.0±14.39 months (Table 1 and supplement table).

According to the GES score, 1430 (61.3%), 443 (19.0%), and 458 (19.6%) of the studied patients turned out to be at low, intermediate, and high risk for HCC, respectively. As shown in Table 2, HCC developed in 212 patients during the study period, of whom 87 belonged to the low-risk group (87/1430, 6.1%), 49 to the intermediate-risk group (49/443, 11.1%), and 76 to the high-risk group (76/458, 16.6%).

The cumulative incidence of HCC at 3 years from EOT was 2.31% in the low-risk group, 4.08% in the intermediate-risk group, and 6.68% in the high-risk group (Table 2). Analysis of the cumulative incidence of HCC revealed a significant difference between the 3 risk groups ($P<0.001$). Harrell's C-statistic for this cohort was 0.622. Brier score was 0.247 and Hosmer-Lemeshow test p-value was 0.582 (Fig. 1-B). NPV comparing risky patients with less risky was 93.9% (95% CI = 92.6-95.0).

Gilead cohort

This cohort was comprised of 836 patients from different medical centers included in an RCT conducted by Gilead. Of these patients, 592 (70.8%) were from the U.S. and 244 (29.2%) were non-Americans. The most frequently observed HCV genotype was genotype 1 (68.9%), followed by genotype 3 (21.5%) and genotype 4 (4.2%). Chronic HCV was diagnosed as F3 in 271 patients and as F4 in 565 patients. The observation period was 46.9±13.28 (range, 10.9–78.5) months after the end of DAA treatment (Table 1 and supplement table).

The GES score identified 374 (44.7%) of patients as low-risk, 384 (45.9%) as intermediate-risk, and 78 (9.3%) as high-risk (Table 2).

During the study period, HCC occurred in 54 cases (excluding patients with follow-up periods less than 6 months and those with HCC before study enrollment) as follows: 23 cases (6.1%) in the low-risk group, 20 cases (5.2%) in the intermediate-risk group, and 11 cases (14.1%) in the high-risk group (Table 1).

The 3-year cumulative incidence of HCC was 7.46% (95% CI, 4.93%–11.2%) in the low-risk group, 5.88% (95% CI, 3.82%–9.0%) in the intermediate-risk group, and 19.61% (95% CI, 11.14%–33.21%) in the high-risk group. Analysis of the cumulative incidence of HCC displayed a significant difference between the 3 risk groups (P = 0.0055; Fig. 1-B). Harrell's C-statistic for this external validation group was 0.55 (Table 2 and Fig. 1-C).

India cohorts

This cohort was composed of 662 patients from the Institute of Liver and Biliary Science, who had a median age of 50.0 (42.0–59.0) years and were predominantly male (67.5%). The study contained 574 (86.7%) cirrhotic patients (Table 1), who were followed up for 30.7±14.7 months (Table 1 and supplement table).

According to the GES score, 227 (34.3%), 62 (9.4%), and 373 (56.36%) of the studied patients were at low, intermediate, and high risk for HCC. During the study period, 9 patients in the low-risk group (9/227, 4.0%), 5 patients in the intermediate-risk group (5/62, 8.1%), and 59 patients in the high-risk group (59/373, 15.8%) developed HCC (Table 2).

The cumulative incidence of HCC at 5 years from EOT was found to be 1.50% in the low-risk group, 3.06% in the intermediate-risk group, and 6.35% in the high-risk group (Table 2). Analysis of the cumulative incidence of HCC indicated a significant difference between the 3 risk groups (P<0.001). Harrell's C-statistic for this cohort was 0.643. Brier score was 0.328 and Hosmer-Lemeshow test p-value was 0.278 (Fig. 1-D). NPV comparing risky patients with less risky was 96.0% (95% CI = 92.6-97.9).

USA cohort

This cohort encompassed 276 patients from MD Anderson Cancer Center-Baylor college of Medicine. The median age was 63.0 (59.0–68.0) years, and males accounted for 59.1% of all participants. The study included 162 (58.7%) cirrhotic patients that were followed up for 33.9 ± 16.0 months (Table 1 and supplement table).

The GES score stratified patients into low-risk (148, 53.6%), intermediate-risk (49, 17.8%), and high-risk (79, 28.6%) groups in terms of the risk for HCC. During the study period, 3 low-risk (3/148, 2.0%), 2 intermediate-risk (2/49, 4.1%), and 9 high-risk (9/79, 11.4%) patients developed HCC (Table 2).

The cumulative incidence of HCC after 5 years from EOT was 1.07% in the low-risk group, 0.91% in the intermediate-risk group, and 5.17% in the high-risk group (Table 2). Analysis of the cumulative incidence of HCC suggested a significant difference between the 3 risk groups (P=0.046). Harrell's C-statistic for this cohort was 0.666. Brier score was 0.301 and Hosmer-Lemeshow test p-value was 0.736 (Fig. 1-E). NPV comparing risky patients with less risky was 98.0% (95% CI = 94.2-99.3).

Spain cohorts

This cohort included 181 patients from 8 medical centers. They had a median age of 63.7 (56.5–71.1) years, and 53.0% were male. In this study, 125 (69.1%) cirrhotic patients were followed up for 45.5±7.6 months (Table 1 and supplement table).

Of the studied patients, 102 (56.4%), 62 (34.3%), and 17 (9.4%) had low, intermediate, and high GES scores, respectively. As presented in Table 2, HCC developed in 6 cases during the study period—i.e., 2 in the low-risk group (2/102, 2.0%), 3 in the intermediate-risk group (3/62, 4.8%), and 1 in the high-risk group (1/17, 5.9%).

The cumulative incidence rates of HCC after 4 years from EOT were 0.51%, 1.14%, and 1.56% for the low-risk, intermediate-risk, and high-risk groups, respectively (Table 2). Analysis of the cumulative incidence of HCC demonstrated no significant difference between the 3 risk groups (P=0.509). Harrell's C-statistic for this cohort was 0.615. Brier score was 0.240 and Hosmer-Lemeshow test p-value was 0.672 (Fig. 1-F). NPV comparing risky patients with less risky was 98.0% (95% CI = 93.1-99.5).

Discussion:

One of the hallmarks of validating the clinical utility of an HCC risk stratification score is to establish its generalizability to populations that are different from the original model derivation cohort. Compared with our previous single-center study of 4400 patients (derivation cohort) that introduced the GES score, this work represents one of the largest, multicenter, external validation studies of this HCC risk stratification score, where a total of 12038 HCV patients with different clinical characteristics, ethnicities, and HCV genotypes were recruited from the Gilead RCT cohort and 54 centers across diverse countries.

Notably, despite significant differences in patient characteristics between the derivation and validation cohorts, the GES score maintained very good accuracy in predicting the risk of HCC development after DAA therapy. Log-rank test results demonstrated that the GES score had significant predictive value for HCC development in the overall validation cohorts (P< 0.0001), which is consistent with the original cohort.

The GES score proved successful at stratifying all patients into low-, intermediate-, and high-risk groups. Specifically, the GES score identified most of the patients as being at low risk for HCC (34.3%–65.7%) with a very low Cumulative incidence (1.07–1.23) at 5 years, suggesting that these patients could either have a longer follow-up period or be safely excluded from the surveillance program to avoid the potential harms of screening and subsequent unnecessary confirmatory imaging (CT/MRI) [18,19]. On the other hand, some patients (9.4%–56.3%) were categorized under the high-risk group with a high cumulative incidence (1.56–19.6) at 5 years, which necessitates more intense surveillance to establish the diagnosis of HCC at early stages. Curative treatment is possible in many HCC patients with BCLC stages 0 and A, which could ultimately increase their survival [20,21].

The best performance of the GES score was observed in the French cohort, consisting of 7752 patients with different ethnicities from 32 centers in North Africa (12.9%), Sub-Saharan Africa (5.2%), Asia (2.1%), and Europe (74.9%). Almost all HCV genotypes were detected in this cohort, with the most frequent one being genotype 1 (66.7%). The GES score stratified the HEPATHER cohort into low-risk (65.7%), intermediate-risk (24.3%), and high-risk groups (9.93%). Interestingly, the 5-year cumulative incidence of HCC was 1.23% (very low) in the low-risk group, 7.95% in the intermediate-risk group, and 17.69% in the high-risk group, with highly significant log-rank test results (P<0.0001; Fig. 1). Harrell's C-statistic for this cohort was very good

(0.7688). Remarkably, these results are almost the same as those of the original study, despite considerable difference in patients' characteristics, ethnicities, and HCV genotypes.

About 65% of the HEPATHER cohort were considered at low risk for HCC with a very low 5-year cumulative incidence (1.23%), implying that they could be either followed up at a longer period or safely excluded from the surveillance program. By contrast, more intense screening should be directed to the high-risk group because of having a high 5year cumulative incidence rate (17.6%) in only 9.9% of the patients. The intermediate-risk group with a relatively high 5-year cumulative incidence rate (7.9%), which constituted 25% of the patients, may need to continue screening according to the current guidelines. This obvious applicability of the GES score for HCC risk stratification not only highlights its clinical utility in real-life cohorts with different clinical characteristics, ethnicities, and HCV genotypes, but also may add to its cost-effectiveness. A similar pattern was also noted in the Japanese, American, and Spanish cohorts. However, in the Indian cohort, a different pattern was noticed, that is, 56.3% of the patients were stratified into the high-risk group. This could be explained by the different clinical characteristics in that cohort compared with others, as the Indian cohort had the highest percentages of both cirrhosis (86.7%) and male gender (67.5%). Furthermore, 65.0% of the patients had HCV genotype 3, which is known to be associated with a higher incidence of HCC in patients with CHC and cirrhosis [22,23].

We previously compared the performance of different HCC risk prediction scores that were developed to guide HCC risk stratification and identify CHC patients who either need intensified surveillance or may not require screening [24]. Scores were evaluated in 3075 CHC patients who achieved SVR following DAAs. ADRES [25], GES (pre- and post-treatment) [12], GES algorithm [26] and Watanabe (post-treatment) scores [27] offered acceptable HCC-risk predictability and clinical utility in CHC patients.

Our study had several limitations. First, the performance of the GES score in the Gilead cohort was not optimal; in other words, although the GES score was able to stratify patients into low-risk (44.7%), intermediate-risk (45.9%), and high-risk (9.3%) groups with 3-year cumulative incidence rates of 7.46%, 5.88%, and 19.61%, respectively (P<0.005), the cumulative hazard of HCC in the low-risk group was higher than that in the intermediate-risk group. Moreover, Harrell's C-statistic was only 0.55, which is the

lowest among all cohorts. Second, analysis of the cumulative incidence of HCC for the Spanish Cohort showed no significant difference between the 3 risk groups (P=0.509), which could probably be attributed to the small sample size (n= 181) and the very small number of events (n=6) in this cohort. Third, the dynamics of the GES score after achieving SVR were not reported. These dynamic changes could be more important than the baseline calculation [26,28]. Finally, the GES score was limited to only HCV patients, so further work is needed to validate the GES score in HCC of different etiologies.

Our current analysis, which revealed far greater accuracy and generalizability of GES in determining HCC risk after DAA therapy in a real-world context, provides a strong rationale for the widespread routine use of this score for clinical outcome prediction as well as in clinical research. This finding suggests that the GES score can be incorporated in an individualized HCC risk-based surveillance strategy, where patients stratified as low-risk may have a longer follow-up period or even be omitted from the surveillance program; conversely, those stratified as high-risk may undergo intensified surveillance.

In conclusion, the present validation study of the GES score represents one of the largest, multicenter, external validation studies to date of an HCC risk score in a real-world context involving international patients. The GES score, which is able to incorporate 5 well-established and easy-to-implement clinical risk factors for HCC, was used herein. The derivation and external validation cohorts revealed good discrimination and calibration for both the overall GES risk model and the identification of at-risk patient groups, which could improve the assessment of at-risk HCV patients by avoiding unnecessary harm from the overuse of surveillance procedures in the low-risk population. Yet, additional prospective management studies together with economic evaluations and cost-effectiveness analyses of the GES score remain paramount.

Data availability statement:

GS & NM had full access to all the data in the study and takes responsibility for the integrity and accuracy of the data analysis. Data is available upon request.

References:

1. WHO? Global health sector strategy on viral hepatitis 2016-2021. World Health Organization, 2016. https://www.who.int/hepatitis/strategy2016-2021/ghss-hep/en/.

2. Carrat F, Fontaine H, Dorival C, et al. Clinical outcomes in patients with chronic hepatitis C after direct-acting antiviral treatment: A prospective cohort study. Lancet. 2019;393(10179):1453-1464.

3. Shiha G, Mousa N, Soliman R, et al. Incidence of HCC in chronic hepatitis C patients with advanced hepatic fibrosis who achieved SVR following DAAs: A prospective study. J Viral Hepat. 2020 Jul;27(7):671-679.

4. European Association for the study of the liver. EASL recommendations on treatment of hepatitis C 2018. J Hepatol 2018;69:461-511.

5. Ghany MG, Morgan TR, AASLD-IDSA. Hepatitis C guidance 2019 update: American Association for the Study of Liver Diseases-Infectious Diseases Society of America recommendations for testing, managing, and treating hepatitis C virus infection. Hepatology 2020;71(2):686-721.

6. Degasperi E, Galmozzi E, Pelusi S, et al. Hepatic fat - Genetic risk score predicts hepatocellular carcinoma in HCV cirrhotic patients treated with DAAs. Hepatology. 2020 Aug 6;72(6):1912-1923.

7. Matsuura K, Sawai H, Ikeo K, et al. Genome-wide association study identifies TLL1 variant associated with development of hepatocellular carcinoma after eradication of hepatitis C virus infection. Gastroenterology. 2017 May;152(6):1383-1394.

8. Liu W, Ma N, Zhao D, et al. Correlation between the DEPDC5 rs1012068 polymorphism and the risk of HBV-related hepatocellular carcinoma. Clin Res Hepatol Gastroenterol. 2019 Aug;43(4):446-450.

9. El-Serag HB, Kanwal F, Davila JA, et al. A new laboratory-based algorithm to predict development of hepatocellular carcinoma in patients with hepatitis C and cirrhosis. Gastroenterology. 2014 May;146(5):1249-55. e1.

10. Watanabe T, Tokumoto Y, Joko K, et al. Predictors of hepatocellular carcinoma occurrence after direct-acting antiviral therapy in patients with hepatitis C virus infection. Hepatol Res. 2019 Feb;49(2):136-146.

11. Fan R, Papatheodoridis G, Sun J, et al. aMAP risk score predicts hepatocellular carcinoma development in patients with chronic hepatitis. J Hepatol. 2020 Jul 21;S0168-8278(20):30478-30475.

12. Shiha G, Waked I, Soliman R, et al. GES: A validated simple score to predict the risk of HCC in patients with HCV-GT4-associated advanced liver fibrosis after oral antivirals. Liver Int. 2020;40(11):2828-2833.

13. Bergna I, Degasperi E, D'Ambrosio R. Suboptimal accuracy of GES score to stratify post-SVR HCC risk in a single center cohort of European cirrhotics infected with any HCV genotype. Liver Int. 2021 Oct 11;41(5):1152-1153.

14. Shiha G, Soliman R, Mikhail N, Eslam M. Reply to: Suboptimal accuracy of GES score to stratify post SVR HCC risk in a single-centre cohort of European cirrhotics. Liver Int. 2020;00:1-2.

15. Brier (1950). Verification of Forecasts Expressed in Terms of Probability. Monthly Weather Review 1950;78:1–3.

16. Harrell FE, Jr., Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. JAMA 1982;247:2543-2546.

17. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied logistic regression: John Wiley & Sons, 2013.

18. Ioannou G. HCC surveillance after SVR in patients with F3/F4 fibrosis. J Hepatol. 2021 Feb;74(2):458-465.

19. Singal AG, Patibandla S, Obi J, et al. Benefits and harms of hepatocellular carcinoma surveillance in a prospective cohort of patients with cirrhosis. Clin Gastroenterol Hepatol. 2020 Sep 10; S1542;3565(20):31270-31272.

20. van Meer S, de Man RA, Coenraad MJ, et al. Surveillance for hepatocellular carcinoma is associated with increased survival: Results from a large cohort in the Netherlands. J Hepatol. 2015 Nov;63(5):1156-1163.

21. Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening forhepatocellular carcinoma. J Cancer Res Clin Oncol. 2004 Jul;130(7):417-422.

22. Nkontchou G, Ziol M, Aout M, et al. HCV genotype 3 is associated with a higher hepatocellular carcinoma incidence in patients with ongoing viral C cirrhosis. J Viral Hepat. Oct 2011;18(10):e516-e522.

23. Kanwal F, Kramer JR, Ilyas J, Duan Z, El-Serag HB. HCV genotype 3 is associated with an increased risk of cirrhosis and hepatocellular cancer in a national sample of U.S. Veterans with HCV. Hepatology. 2014;60(1):98-105.

24. Shiha G, Mikhail NNH, Soliman R, et al. Predictive performance and clinical utility of HCC risk scores in chronic hepatitis C: a comparative study. Hepatol Int. 2022 Feb;16(1):159-170.

25. Hiraoka A, Kumada T, Ogawa C, Kariyama K, Morita M, Nouso K, et al. Proposed a simple score for recommendation of scheduled ultrasonography surveillance for hepatocellular carcinoma after direct acting antivirals: multicenter analysis. J Gastroenterol Hepatol 2019;34(2):436–441.

26. Shiha G, Soliman R, Mikhail NNH, et al. Development of a simple dynamic algorithm for individualized HCC risk-based surveillance using pre- and post-treatment GES score. Liver Int. 2021 Jun 26; 41:2768-2776.

27. Watanabe T, Tokumoto Y, Joko K, Michitaka K, Horiike N, Tanaka Y, et al. Predictors of hepatocellular carcinoma occurrence after direct-acting antiviral therapy in patients with hepatitis C virus infection. Hepatol Res 2019;49(2):136–146.

28. Bird T, Dimitropoulou P, Turner R, et al. Alpha-fetoprotein detection of hepatocellular carcinoma leads to A standardized analysis of dynamic AFP to improve screening based detection. PLOS ONE. 2016;11(6):e0156801.

Figure legends:

Fig. 1: Cumulative hazard (%) of HCC in patients with HCV after the end of DAA therapy, shown by curves comparing different risk groups

Table (1): Baseline characteristics of the studied cohorts

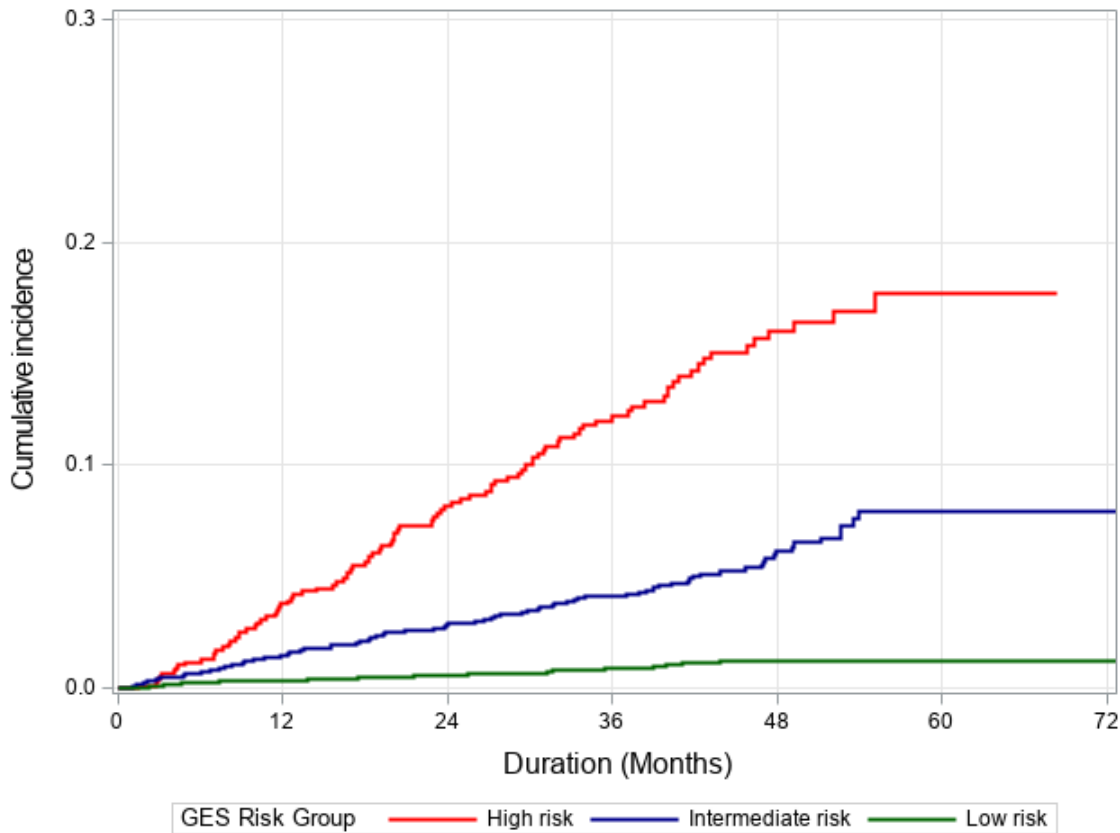|	France	Japan	GILEAD	India	USA	Spain
No. of Centers	32	12	RCT	1	1	8
All patients	7752	2331	836	662	276	181
Age	56.2 (50.4–64.3)	71.0 (63.0–77.0)	60.0 (56.0–64.0)	50.0 (42.0–59.0)	63.0 (59.0–68.0)	63.7 (56.5–71.1)
Gender						
- Male	4185 (54.0%)	1036 (44.4%)	642 (76.8%)	447 (67.5%)	163 (59.1%)	96 (53.0%)
- Female	3567 (46.0%)	1295 (55.6%)	194 (23.2%)	215 (32.5%)	113 (40.9%)	85 (47.0%)
HCV genotypes						
- Genotype 1	4818 (66.7%)	576 (72.3%)	573 (68.9%)	217 (32.8%)	226 (81.8%)	
- Genotype 2	420 (5.8%)	221 (27.7%)	43 (5.2%)	5 (0.8%)	21 (7.6%)	
- Genotype 3	918 (12.7%)	0 (0.0%)	179 (21.5%)	430 (65.0%)	25 (9.05%)	
- Genotype 4	918 (12.7%)	0 (0.0%)	35 (4.2%)	8 (1.2%)	3 (1.08%)	
- Genotypes 5, 6, and 7	153 (2.1%)	0 (0.0%)	2 (0.2%)	2 (0.3%)	1 (0.36%)	
- Unknown	525	1534	4	0	0	
AFP (ng/mL)	5.0 (3.0–9.0)	5.70 (3.00–12.60)	4.5(3.2-6.3)	4.3 (3.2–5.69)	7.3 (3.8–14.2)	4.3 (4.0–4.5)
Albumin (g/dL)	4.2 (3.9–4.4)	4.00 (3.70–4.30)	4.4(4.0-4.6)	2.9 (2.4–3.4)	3.9 (3.5–4.1)	6.1 (3.5–10.8)
Fibrosis stage						
- Non-cirrhotic	4912 (63.4%)	1328 (57.0%)	271 (32.4%)	88 (13.3%)	114 (41.3%)	56 (30.9%)
- Cirrhotic	2840 (36.6%)	1003 (43.0%)	565 (67.6%)	574 (86.7%)	162 (58.7%)	125 (69.1%)

HCV, hepatitis C virus; APF, alpha-fetoprotein.

Table (2): External evaluation of the GES score in different cohorts

	France	Japan	GILEAD	India	USA	Spain
All patients	7752	2331	836	662	276	181
Follow-up period, month (range)	31.85 ± 16.26 (19–45)	29.02 ± 14.39 (0–76)	46.9 ± 13.28 (10.9–78.5)	30.69 ± 14.69 (0–64)	33.91 ± 15.97 (0–62)	45.5 ± 7.6 (7–51)

HCC cases	227	212	54	73	14	6
GES risk groups						
- Low	5098 (65.76%)	1430 (61.3%)	374 (44.7%)	227 (34.3%)	148 (53.6%)	102 (56.4%)
- Intermediate	1884 (24.30%)	443 (19.0%)	384 (45.9%)	62 (9.4%)	49 (17.8%)	62 (34.3%)
- High	770 (9.93%)	458 (19.6%)	78 (9.3%)	373 (56.3%)	79 (28.6%)	17 (9.4%)
HCC in the 3 risk groups						
- Low	37/5098 (0.73%)	87/1430 (6.1%)	23/374 (6.1%)	9/227 (4.0%)	3/148 (2.0%)	2/102 (2.0%)
- Intermediate	91/1884 (4.83%)	49/443 (11.1%)	20/384 (5.2%)	5/62 (8.1%)	2/49 (4.1%)	3/62 (4.8%)
- High	99/770 (12.86%)	76/458 (16.6%)	11/78 (14.1%)	59/373 (15.8)	9/79 (11.4%)	1/17 (5.9%)
Cumulative Incidence, %						
- Low	1.23 (5 years)	2.31 (3 years)	7.46 (3 years)	1.50 (5 years)	1.07 (5 years)	0.51 (4 years)
- Intermediate	7.95 (5 years)	4.08 (3 years)	5.88 (3 years)	3.06 (5 years)	0.91 (5 years)	1.14 (4 years)
- High	17.69 (5 years)	6.68 (3 years)	19.61 (3 years)	6.35 (5 years)	5.17 (5 years)	1.56 (4 years)
Log-rank test	<0.0001	<0.001	0.0055	<0.001	0.046	0.509
Harrell's C-statistic	0.7688	0.622	0.55	0.643	0.666	0.615
Brier score	0.027	0.247	NA	0.328	0.301	0.240
Hosmer-Lemeshow test	NA	0.582	NA	0.278	0.736	0.672
Performance statistics[#]						
- Sensitivity	83.7 (78.9-88.5)	59.0 (52.2-695.4)	NA	87.7 (78.2-93.4)	78.6 (52.4-92.4)	66.7 (30.0-90.3)
- Specificity	67.3 (66.2-68.3)	63.4 (61.3-65.4)		37.0 (33.2-41.0)	55.3 (49.3-61.2)	57.1 (49.7-64.2)
- PPV	NA	13.9 (11.8-16.3)		17.7 (11.7-18.4)	8.6 (4.9-14.7)	5.1 (2.0-12.3)
- NPV	NA	93.9 (92.6-95.0)		96.0 (92.6-97.9)	98.0 (94.2-99.3)	98.0 (93.1-99.5)
- Accuracy	67.7 (66.7-68.8)	63.0 (61.0-64.9)		42.6 (38.9-46.4)	56.5 (50.6-62.24)	57.5 (50.2-64.4)

Comparing risky patients (high + intermediate risk groups) with less risky patients (low risk group)
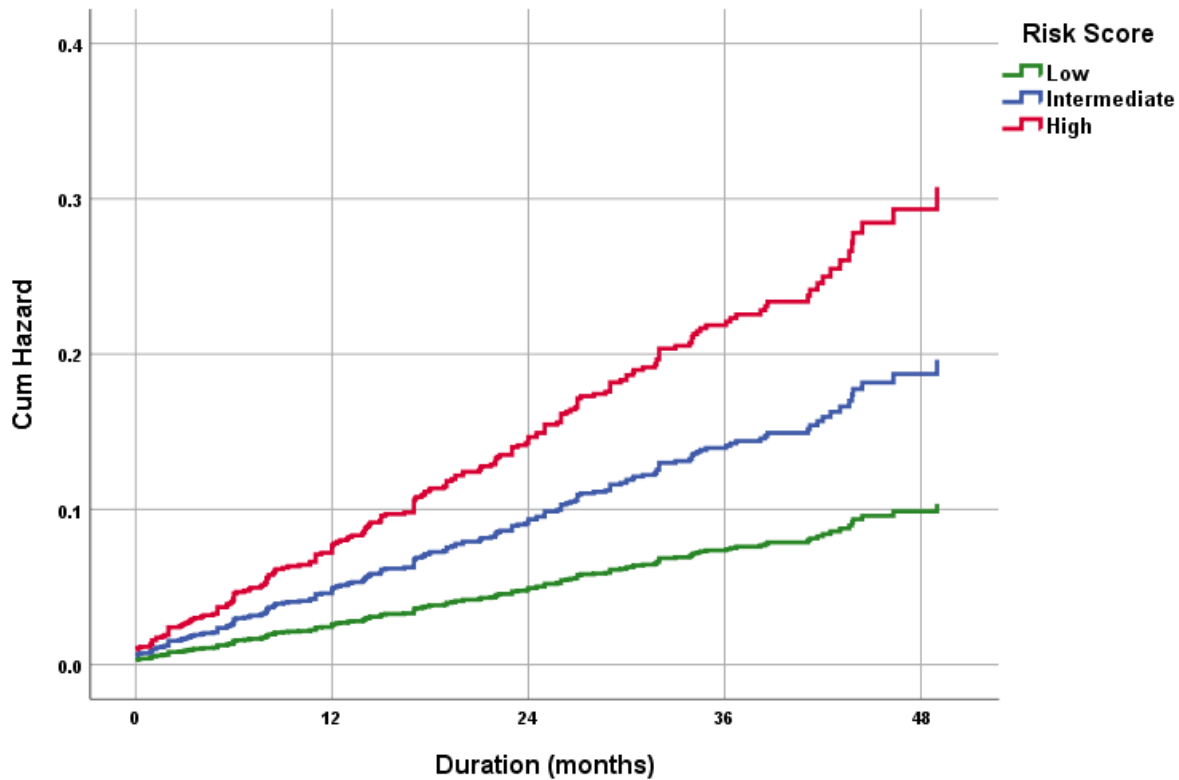
GES, predictive model; HCC, hepatocellular carcinoma.

Fig. 1: Cumulative hazard (%) of HCC in patients with HCV after the end of DAA therapy, shown by Kaplan-Meier curves comparing different risk groups (**France** cohort)
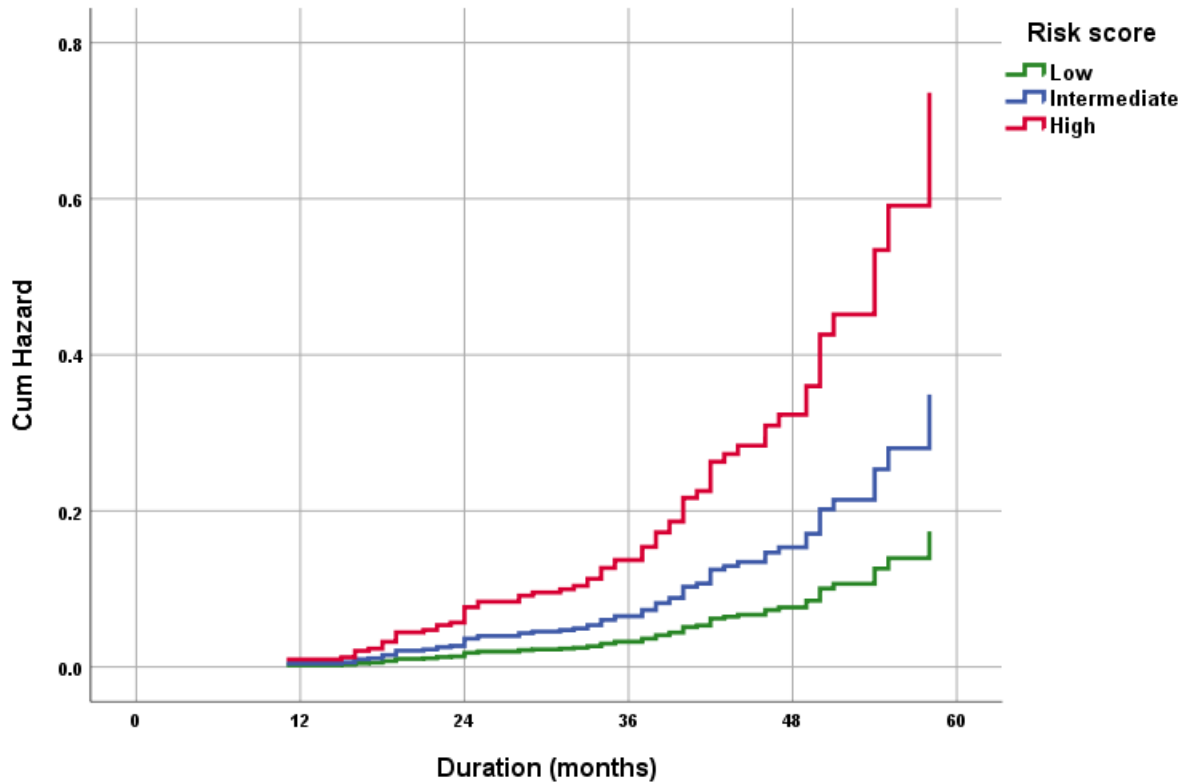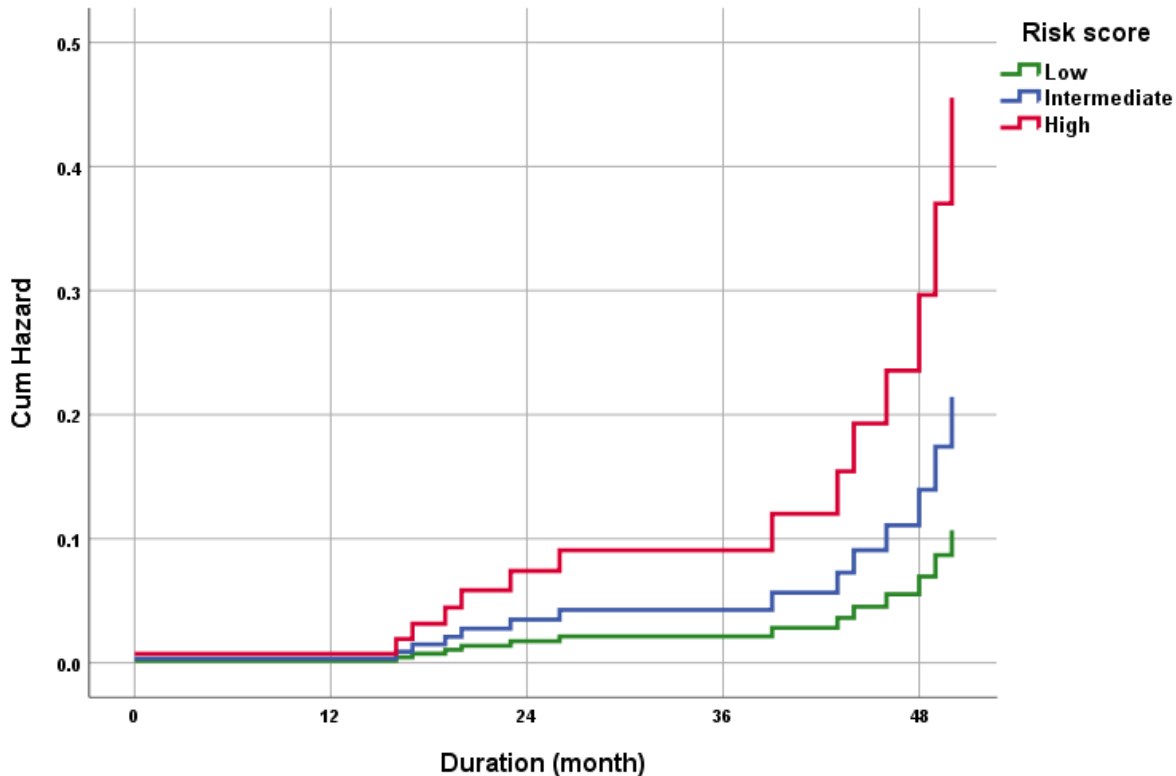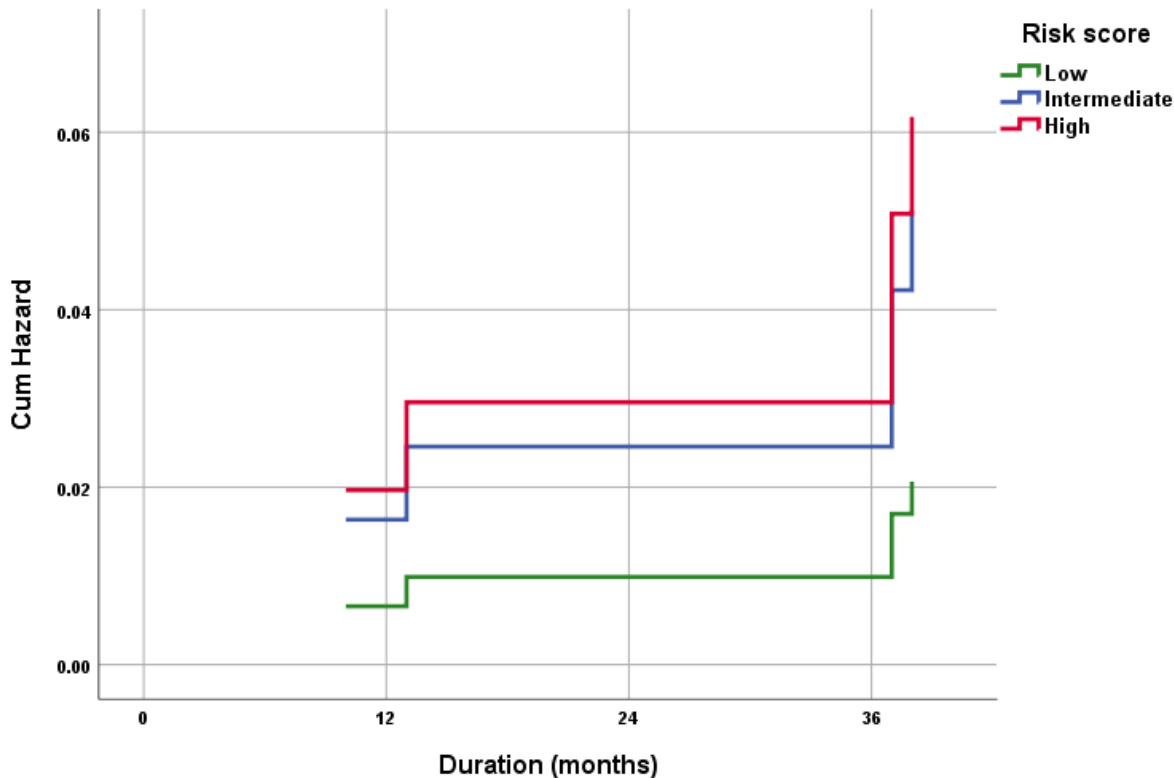
GES, predictive model; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; DAA, direct-acting antiviral.

	Number at risk				
High	458	358	256	167	39
Intermed.	443	368	294	167	23
Low	1430	1206	964	584	90

Fig. 2: Cumulative hazard (%) of HCC in patients with HCV after the end of DAA therapy, shown by Kaplan-Meier curves comparing different risk groups (**Japan** cohort)
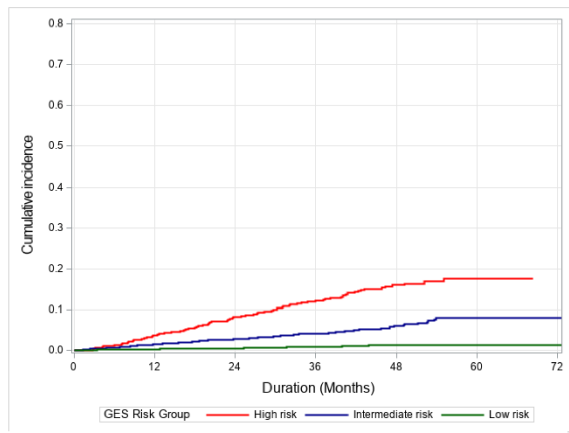
HCC, hepatocellular carcinoma; HCV, hepatitis C virus; DAA, direct-acting antiviral.
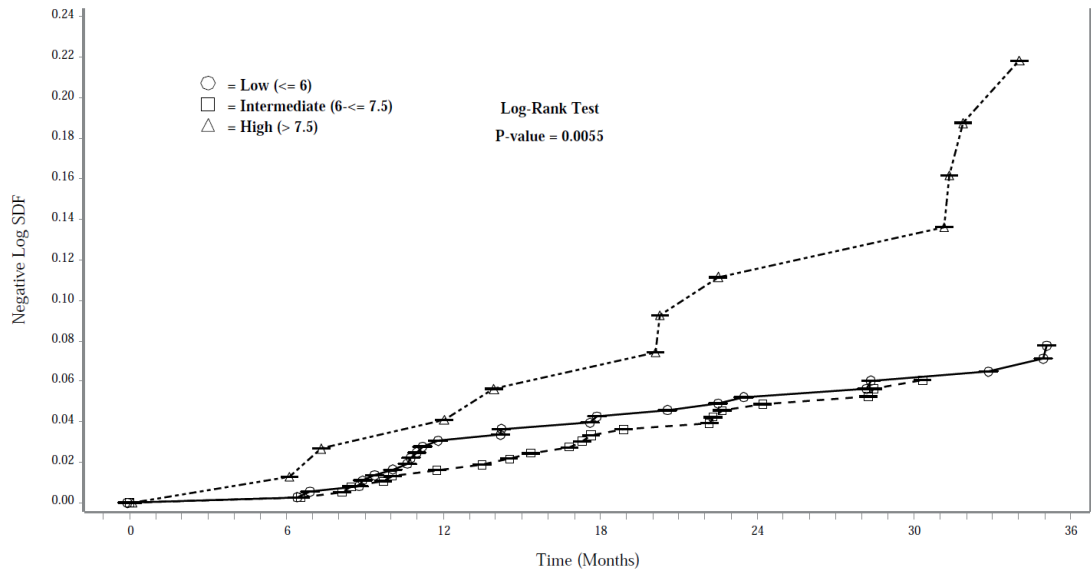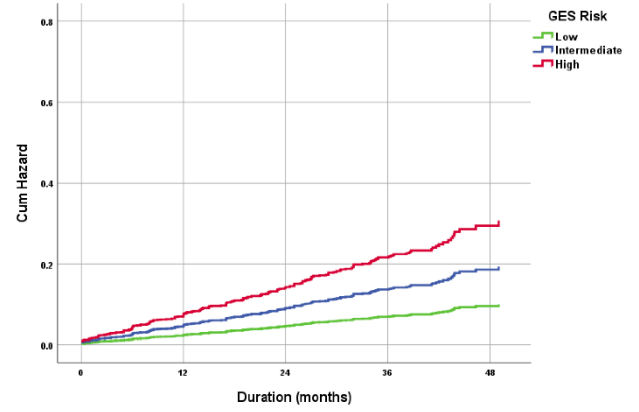
Fig. 3: Cumulative hazard (%) of HCC in patients with HCV after the end of DAA therapy, shown by Kaplan-Meier curves comparing different risk groups (**Gilead** cohort)

HCC, hepatocellular carcinoma; HCV, hepatitis C virus; DAA, direct-acting antiviral.
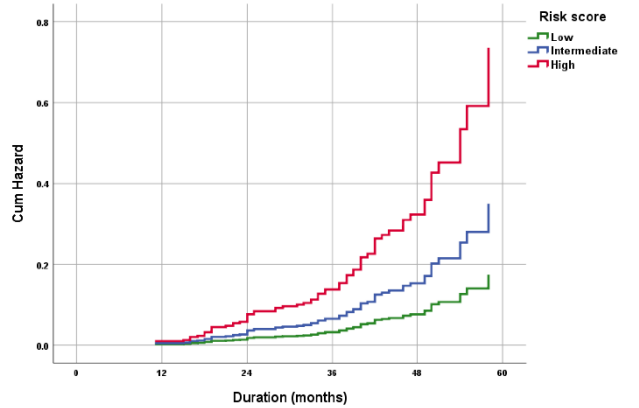
	Number at risk					
High	373	296	248	149	54	5
Intermed.	62	53	46	29	8	1
Low	227	189	160	104	27	2

Fig. 4: Cumulative hazard (%) of HCC in patients with HCV after the end of DAA therapy, shown by Kaplan-Meier curves comparing different risk groups (**India** cohort)

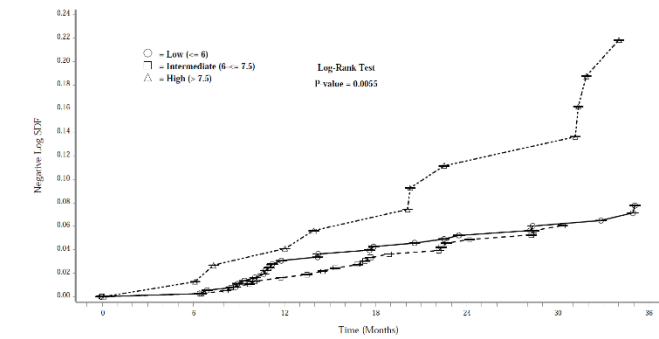HCC, hepatocellular carcinoma; HCV, hepatitis C virus; DAA, direct-acting antiviral.
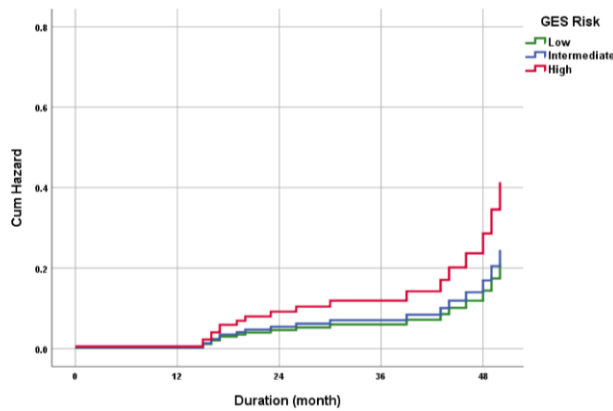
	Number at risk				
High	79	59	37	24	11
Intermed.	49	35	51	15	5
Low	148	93	61	37	14

Fig. 5: Cumulative hazard (%) of HCC in patients with HCV after the end of DAA therapy, shown by Kaplan-Meier curves comparing different risk groups (**USA** cohort)

HCC, hepatocellular carcinoma; HCV, hepatitis C virus; DAA, direct-acting antiviral.
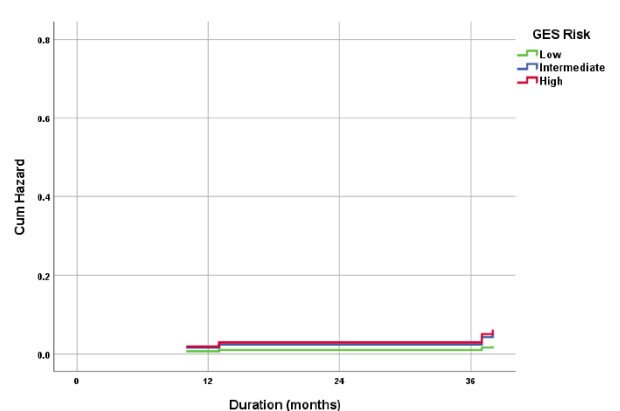
	Number at risk			
High	17	17	17	16
Intermed.	62	61	60	57
Low	102	100	99	96

Fig. 6: Cumulative hazard (%) of HCC in patients with HCV after the end of DAA therapy, shown by Kaplan-Meier curves comparing different risk groups (**Spain** cohort)

HCC, hepatocellular carcinoma; HCV, hepatitis C virus; DAA, direct-acting antiviral.

A- France

B- Japan

C- Gilead

D- India

E- USA

F- Spain

Fig. 1: Cumulative hazard (%) of HCC in patients with HCV after the end of DAA therapy, shown by curves comparing different risk groups

GES, predictive model; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; DAA, direct-acting antiviral