

Article

Machine Learning-Based Assessment of Survival and Risk Factors in Non-Alcoholic Fatty Liver Disease-Related Hepatocellular Carcinoma for Optimized Patient Management

Miguel Suárez^{1,2,3}, Sergio Gil-Rojas^{1,2,3} , Pablo Martínez-Blanco^{1,2,3}, Ana M. Torres^{2,3}, Antonio Ramón⁴ , Pilar Blasco-Segura⁴, Miguel Torralba^{5,6,7}  and Jorge Mateo^{2,3,*}

¹ Gastroenterology Department, Virgen de la Luz Hospital, 16002 Cuenca, Spain; msuarezmatias@sescam.jccm.es (M.S.)

² Medical Analysis Expert Group, Institute of Technology, Universidad de Castilla-La Mancha, 16071 Cuenca, Spain

³ Medical Analysis Expert Group, Instituto de Investigación Sanitaria de Castilla-La Mancha (IDISCAM), 45071 Toledo, Spain

⁴ Department of Pharmacy, General University Hospital, 46014 Valencia, Spain

⁵ Internal Medicine Unit, University Hospital of Guadalajara, 19002 Guadalajara, Spain

⁶ Faculty of Medicine, Universidad de Alcalá de Henares, 28801 Alcalá de Henares, Spain

⁷ Translational Research Group in Cellular Immunology (GITIC), Instituto de Investigación Sanitaria de Castilla-La Mancha (IDISCAM), 45071 Toledo, Spain

* Correspondence: jorge.mateo@uclm.es

Simple Summary: Non-alcoholic fatty liver disease (NAFLD) is the most prevalent chronic liver condition globally. The increasing incidence of NAFLD suggests that in the upcoming years, NAFLD-related hepatocellular carcinoma (HCC) is poised to become the leading cause of this type of tumor. The aim of this study is to evaluate the survival rates of these patients and identify the primary risk factors contributing to a less favorable prognosis. To accomplish this, we have employed machine learning techniques. This introduces a novel approach for identifying these factors that can be targeted to enhance the life expectancy of these patients, offering a more personalized and effective management strategy. This enhanced management approach not only aids in the optimization of patient care but also facilitates the delivery of the most effective available treatments.

Abstract: Non-alcoholic fatty liver disease (NAFLD) is the most common chronic liver disease worldwide, with an incidence that is exponentially increasing. Hepatocellular carcinoma (HCC) is the most frequent primary tumor. There is an increasing relationship between these entities due to the potential risk of developing NAFLD-related HCC and the prevalence of NAFLD. There is limited evidence regarding prognostic factors at the diagnosis of HCC. This study compares the prognosis of HCC in patients with NAFLD against other etiologies. It also evaluates the prognostic factors at the diagnosis of these patients. For this purpose, a multicenter retrospective study was conducted involving a total of 191 patients. Out of the total, 29 presented NAFLD-related HCC. The extreme gradient boosting (XGB) method was employed to develop the reference predictive model. Patients with NAFLD-related HCC showed a worse prognosis compared to other potential etiologies of HCC. Among the variables with the worst prognosis, alcohol consumption in NAFLD patients had the greatest weight within the developed predictive model. In comparison with other studied methods, XGB obtained the highest values for the analyzed metrics. In conclusion, patients with NAFLD-related HCC and alcohol consumption, obesity, cirrhosis, and clinically significant portal hypertension (CSPH) exhibited a worse prognosis than other patients. XGB developed a highly efficient predictive model for the assessment of these patients.

Keywords: non-alcoholic fatty liver disease; hepatocellular carcinoma; NAFLD-related HCC; machine learning; mortality; extreme gradient boosting; alcohol



Citation: Suárez, M.; Gil-Rojas, S.; Martínez-Blanco, P.; Torres, A.M.; Ramón, A.; Blasco-Segura, P.; Torralba, M.; Mateo, J. Machine Learning-Based Assessment of Survival and Risk Factors in Non-Alcoholic Fatty Liver Disease-Related Hepatocellular Carcinoma for Optimized Patient Management. *Cancers* **2024**, *16*, 1114. <https://doi.org/10.3390/cancers16061114>

Academic Editors: Dania Cioni and Kathryn Effendi

Received: 2 February 2024

Revised: 27 February 2024

Accepted: 7 March 2024

Published: 10 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Non-alcoholic fatty liver disease (NAFLD) is currently the most prevalent chronic liver disease worldwide. It is known to affect 25–30% of the global population, but is estimated to actually impact around 40% [1]. There are geographical differences within this prevalence [2]. It is defined as the presence of macrovesicular steatosis in $\geq 5\%$ of hepatocytes in the absence of other diseases, such as hepatitis B or C, or excessive alcohol consumption (20 g/day in women and 30 g/day in men) [3]. It is closely correlated with increasing rates of obesity and metabolic comorbidities. Its significance today is such that there are various intercountry projects aimed at addressing this growing issue [4,5]. The definition has also been modified and updated to better characterize patients [6], evolving from the initial definition of NAFLD, transitioning through metabolic-associated fatty liver disease (MAFLD) [7], to the currently proposed definition as metabolic-associated steatotic liver disease (MASLD) [8,9].

The disease progression is erratic, characterized by numerous steps forward and backward between a non-alcoholic fatty liver (NAFL) and non-alcoholic steatohepatitis (NASH) [10]. These changes are influenced by multiple factors driving disease progression, primarily comorbid conditions (diabetes mellitus, insulin resistance, dyslipidemia, obesity, etc.) and lifestyle factors, such as alcohol, physical exercise, coffee consumption, or dietary patterns [11,12]. Additionally, genetic predisposition in each individual plays a significant role, being one of the primary contributors to the heterogeneity among patients [13,14].

It is estimated that only 10% of patients with NAFLD will eventually develop complications stemming from liver disease and cirrhosis [15]. Considering the total number of patients in this situation, this figure is not negligible at all. One of these potential consequences is the development of hepatocellular carcinoma (HCC). HCC is the most frequent primary liver cancer, being the third most common cause of death due to cancer and the sixth in terms of diagnosis [16].

Most HCC cases will develop on a cirrhotic liver irrespective of the etiology. However, when referring to HCC related to NAFLD, this scenario varies. There is a possibility of its development on a non-cirrhotic liver, which poses a significant challenge, particularly concerning its detection [17–19].

There is limited published data on the difference in survival between patients with HCC related to NAFLD compared to other causes. The aim of this study is to compare the survival rates among patients with NAFLD-related HCC versus other etiologies. It is also intended to identify prognostic factors that may impact on patient survival. The identification of these factors will facilitate targeted interventions, designed to enhance life expectancy outcomes. For this purpose, machine learning (ML) techniques will be employed. ML presents remarkable superiority by enabling the identification of complex patterns and precise result prediction. These methodologies not only handle large volumes of data efficiently but also have the capability to adapt and learn from the data, unveiling relationships that might go unnoticed using traditional methods [20,21]. The algorithm extreme gradient boosting (XGB) was chosen as the reference method, and it was compared to other systems widely used in the scientific literature [22]. XGB was selected for its fast execution, high scalability, and superior accuracy in results obtained in other medical fields, including hepatology [23,24].

2. Materials and Methods

A multicenter retrospective cohort study was conducted at the Virgen de la Luz Hospital in Cuenca and the University Hospital of Guadalajara in Spain. Data collection occurred between January 2008 and December 2022. Inclusion criteria encompassed patients aged 18 and above with a confirmed HCC diagnosis via the presence of an imaging test compatible with its vascular behavior in patients with liver cirrhosis or via liver biopsy for those without cirrhosis or with diagnostic uncertainty [17,25]. Exclusion criteria involved patients diagnosed in a different facility and those lacking available prognostic variables at the time of diagnosis. This study obtained approval from the ethics committee of the University

Hospital of Guadalajara. Due to the type of study, an exemption from obtaining informed consent from the patients was granted.

The demographic variables and comorbidities collected were age, sex, date of HCC diagnosis, date of death or censoring date, active alcohol consumption, active smoking, diagnosis of diabetes mellitus (DM) [26] and dyslipidemia (DL) [27], and obesity (considered as a body mass index ≥ 30 kg/m² [28]). The censoring date was considered as the last visit date in the clinic for those patients who were alive. Alcohol consumption was categorized as follows: for patients diagnosed with NAFLD, any amount of alcohol was considered active consumption, taking into account the diagnostic criteria for NAFLD (<20 g/day in women, <30 g/day in men); for the rest of the patients and etiologies, active alcohol consumption was considered as consumption at risk (>20 g/day in women, >30 g/day in men).

Regarding variables related to the diagnosis of HCC, the following were analyzed: etiology (NAFLD vs. any other cause), presence of cirrhosis [29], Child–Pugh score [30], diagnosis within an HCC screening program, Eastern Cooperative Oncology Group (ECOG) stage [31], method of diagnosis (imaging technique or biopsy), model for end-stage liver disease (MELD) [32], presence of clinically significant portal hypertension (CSPH) [33], Barcelona Clinic Liver Cancer (BCLC) stage [34], meeting Milan criteria [35], number of lesions, size of the largest lesion (cms), presence of portal thrombosis [36], metastasis, and lymphadenopathy.

Analytical variables collected included the following: neutrophils (cells/mm³), lymphocytes (cells/mm³), platelets (10³/dL), International Normalized Ratio (INR), creatinine (mg/dL), albumin (g/dL), sodium (Na) (mEq/L), bilirubin (mg/dL), Aspartate Aminotransferase (AST) (U/L), Alanine Aminotransferase (ALT) (U/L), and alpha-fetoprotein (AFP) (ng/mL). These variables were collected at the time of diagnosis or within the first month of diagnosis.

To conduct the data analysis, the XGB algorithm was proposed as the reference method. As mentioned earlier, this algorithm was selected for its scalability, execution speed, and excellent results in terms of accuracy. Furthermore, it is a versatile system that allows for parallel computing [37]. Additionally, other ML methods employed in the scientific literature were implemented. These were performed to assess the utility and performance of this system. Those that achieved better performance were support vector machine (SVM) [38], Bayesian linear discriminant analysis (BLDA) [39], decision tree (DT) [40], Gaussian naïve Bayes (GNB) [41], and K-nearest neighbors (KNN) [42]. The resulting models were developed using MATLAB (The MathWorks, Natick, MA, USA; MATLAB R2023a).

To enhance the ML algorithms performance, hyperparameters of each method were adjusted during training utilizing Bayesian techniques in this study. Bayesian optimization, a sequential model-based algorithm, optimized hyperparameter values by leveraging outcomes from previous iterations, reducing the number of model tests, and focusing on parameters likely to improve validation scores [43]. This approach significantly improved the developed model performance. The simulations involved 100 iterations for mean and standard deviation values, reducing noise impact and ensuring statistically valid results [44].

The representation of the steps carried out for the implementation of the ML algorithms is shown in Figure 1. Cross-validation with 5 folds was performed. The analysis was conducted in this manner to prevent overfitting. The developed database was divided into two groups: 70% of it was used in the training phase and the remaining 30% in the testing phase. This approach ensured that patients were not used in both phases simultaneously. After completing this process, the analysis was conducted.

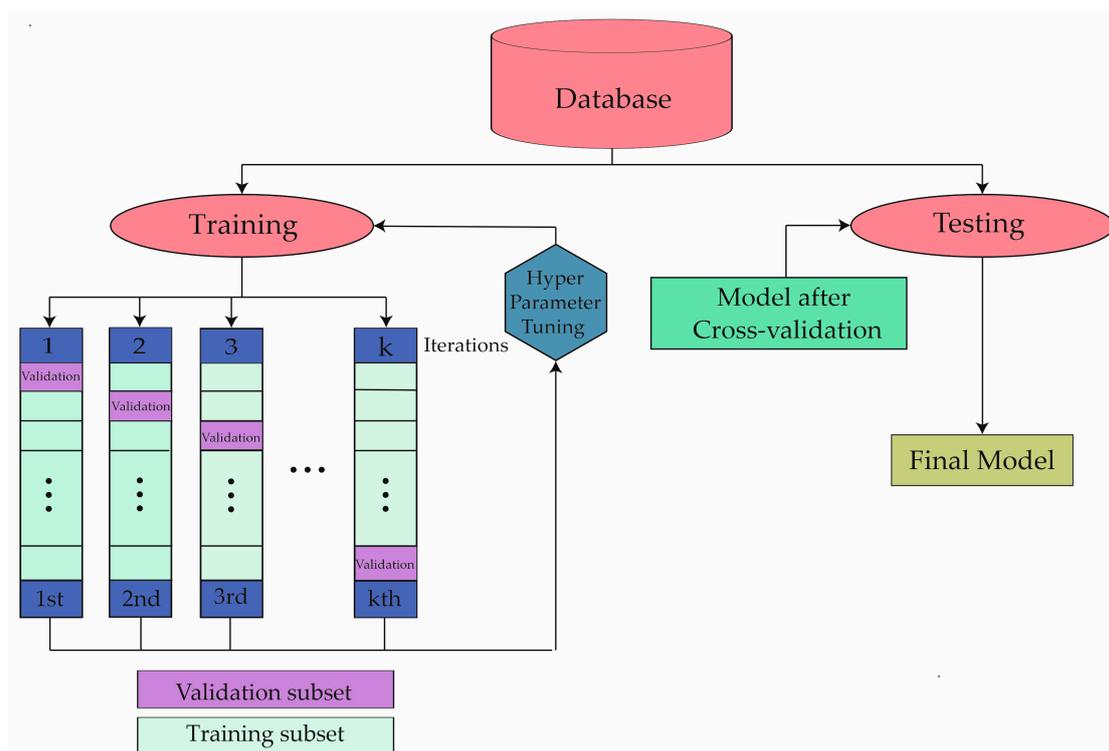


Figure 1. Description of the machine learning methodology development process.

3. Results

This section presents the results for the training and validation phases for identifying the main prognostic factors for mortality in patients diagnosed with HCC and NAFLD. It also demonstrates the comparison between (the proposed method) and the rest of the analyzed ML algorithms.

A total of 191 patients were included in the study based on the inclusion and exclusion criteria. Among them, 29 patients developed NAFLD-related HCC, with 24.2% being women. Within this group, low-level alcohol consumption was present in 48.3%. Only 31% of these patients had obesity, and 55.2% were diagnosed with diabetes mellitus (DM). The majority of them (>85%) were incidentally diagnosed outside the HCC screening program, but 41.4% had cirrhosis. Nearly 69% presented with an ECOG score of 0 at diagnosis, and 79.3% had a BCLC score between 0 and A.

In the control group ($n = 162$), the main causes of HCC were alcohol (38.3%) and hepatitis C (34.6%). Except for patients whose primary cause was alcohol consumption, most of them were either undergoing treatment or had received treatment for their underlying condition causing liver damage. In this group, 61.73% of patients reported abusive alcohol consumption, with DM prevalence at 40.1% and obesity at 31.5%. The percentage of patients with cirrhosis was higher than in the NAFLD group, with 93.2% of patients being cirrhotic. A total of 53% of cases were diagnosed by the screening program. Up to 66.5% of them presented with an ECOG score of 0 at diagnosis, and 42.6% had a BCLC score of 0 or A. More data are available in Table 1.

The average survival within the NAFLD patient group was 9.65 months, while in the rest of etiologies, it approached 12.4 months. The difference between both groups was statistically significant, with a p -value of 0.003. Patients with NAFLD-related HCC had portal hypertension (CSPH) diagnosed in 41.38% of cases, compared to 66.66% in the control group. The MELD score was 9 in the NAFLD group, while the other group had a score of 11. Summary of these data is found in Table 2.

Table 1. Summary of descriptive characteristics of included patients. NAFLD: non-alcoholic fatty liver disease. BMI: body mass index. ECOG: Eastern Cooperative Oncology Group. BCLC: Barcelona Clinic Liver Cancer stage.

		Other Etiologies (%)	NAFLD
N		162	29
Sex	Female	18 (11%)	7 (24.2%)
	Male	144 (89%)	22 (75.8%)
Alcohol	None/low risk	62 (38.27%)	None: 15 (51.7%) Low risk: 14 (48.3%)
	Risk	100 (61.73%)	0
Smoker	Never	53 (32.71%)	19 (65.51%)
	Active/ex-smoker	109 (67.28%)	10 (34.48%)
Diabetes mellitus	No	97 (59.87%)	13 (44.82%)
	Yes	65 (40.12%)	16 (55.17%)
Obesity	BMI < 30 kg/m ²	111 (68.51%)	20 (68.96%)
	≥30 kg/m ²	51 (34.48%)	9 (31.03%)
Dyslipidemia	No	127 (78.39%)	18 (62.06%)
	Yes	35 (21.6%)	11 (37.93%)
ECOG	0	124 (76.54%)	20 (68.96%)
	1	16 (9.87%)	3 (10.34%)
	2	14 (8.64%)	3 (10.34%)
	3	7 (4.32%)	1 (3.44%)
	4	1 (0.61%)	2 (6.89%)
Diagnostic method	Biopsy	54 (33.33%)	16 (55.17%)
	Imaging test	108 (66.66%)	13 (44.87%)
Surveillance	No	76 (46.91%)	25 (86.2%)
	Yes	86 (53.08%)	4 (13.79%)
Cirrhosis	No	11 (6.79%)	12 (41.37%)
	Yes	151 (93.2%)	17 (58.62%)
Etiology	Alcohol	62 (38.27%)	0
	HCV	56 (34.56%)	0
	NAFLD	0	29 (100%)
	Other etiologies	42 (25.92%)	0
CSPH	No	54 (33.33%)	17 (58.62%)
	Yes	108 (66.66%)	12 (41.37%)
Ascites	No	101 (62.34%)	19 (65.51%)
	Yes	61 (37.65%)	10 (34.48%)
Encephalopathy	No	142 (87.65%)	27 (93.1%)
	Yes	20 (12.34%)	2 (6.89%)
Portal thrombosis	No	130 (80.24%)	23 (79.31%)
	Yes	30 (19.75%)	6 (20.69%)
Metastasis	No	146 (90.12%)	21 (72.41%)
	Yes	15 (9.87%)	8 (27.89%)

Table 1. *Cont.*

		Other Etiologies (%)	NAFLD
Lymphadenopathy	No	140 (86.41%)	22 (75.86%)
	Yes	21 (13.58%)	7 (24.13%)
Milan criteria	No	100 (61.73%)	20 (69%)
	Yes	62 (38.27%)	9 (31%)
BCLC	0	9 (5.56%)	1 (3.45%)
	A	60 (37.03%)	11 (37.93%)
	B	27 (16.67%)	2 (6.9%)
	C	48 (29.63%)	10 (34.48%)
	D	18 (11.11%)	5 (17.24%)

Table 2. Descriptive values of survival, MELD, and laboratory parameters used. NAFLD: non-alcoholic fatty liver disease. MELD: model for end-stage liver disease. INR: international normalized ratio.

		Mean Value ± Standard Deviation
Survival (months)	Other etiologies	12.4 ± 23.9
	NAFLD	9.65 ± 22.64
MELD	Other etiologies	11
	NAFLD	9
Albumin (g/dL)	Other etiologies	3.70 ± 0.67
	NAFLD	3.58 ± 0.75
INR	Other etiologies	1.27 ± 0.58
	NAFLD	1.20 ± 0.40
Na (mEq/L)	Other etiologies	138.61 ± 3.42
	NAFLD	139.14 ± 3.49
Lymphocytes (cells/mm ³)	Other etiologies	1436.69 ± 762.29
	NAFLD	1638.21 ± 825.72
Neutrophils (cells/mm ³)	Other etiologies	3772.69 ± 1999.79
	NAFLD	4663.93 ± 2233.46
Platelets (10 ³ /dL)	Other etiologies	140.13 ± 81.60
	NAFLD	178.80 ± 99.15
Creatinine (mg/dL)	Other etiologies	1.02 ± 0.65
	NAFLD	0.95 ± 0.22

In Figure 2, the results of the developed predictive model are presented. Alcohol consumption emerged as the most important variable, followed at some distance by the second variable, obesity. The presence of cirrhosis and the presence of CSPH data were the subsequent variables concerning the mortality of these patients. Both variables showed a similar weight. The prognostic differences among ECOG, MELD, and Child–Pugh stage were not remarkable, as all three presented a similar value. The most significant factor was an advanced ECOG stage. It is noteworthy that alpha-fetoprotein (AFP) levels are insignificant for the prognosis of these patients.

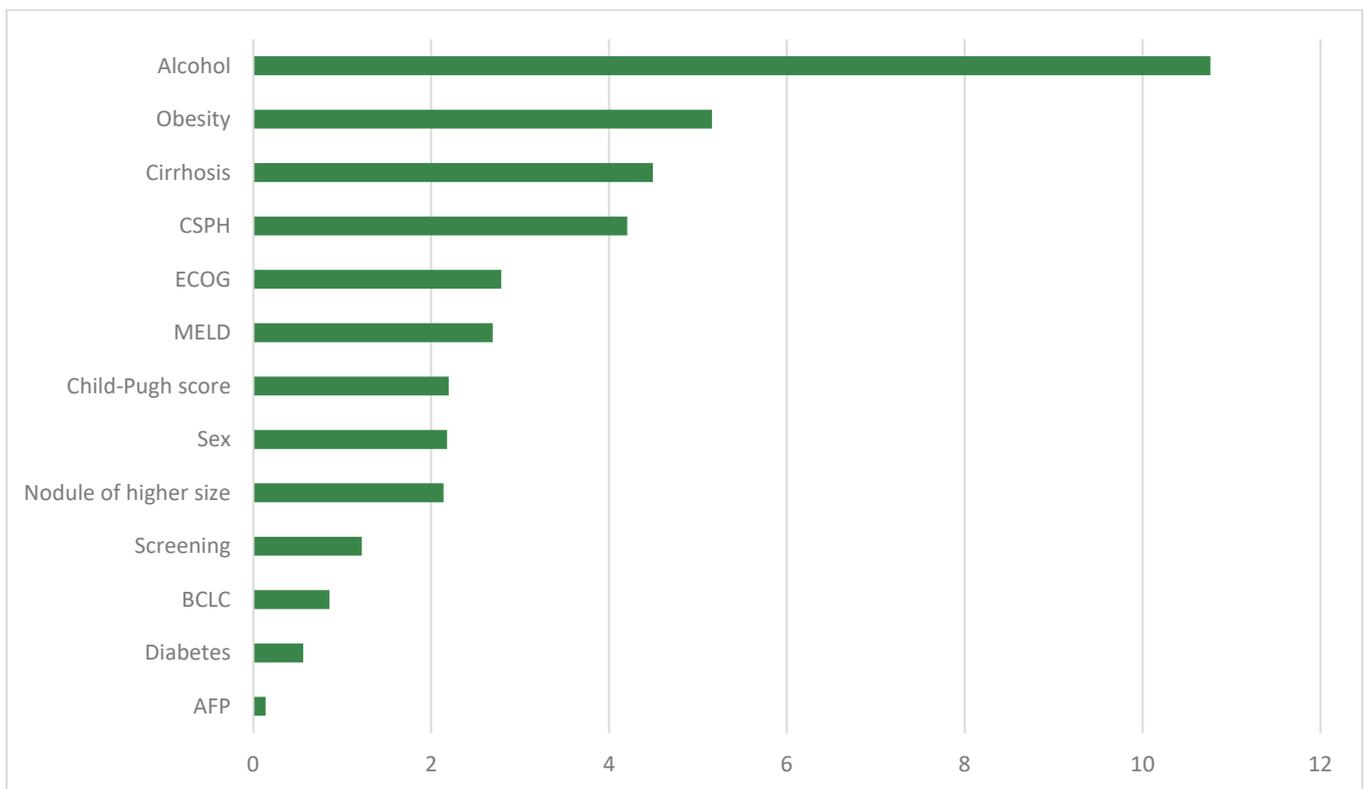


Figure 2. Weight of the most representative variables within the developed predictive model. CSPH: Clinically significant portal hypertension. ECOG: Eastern Cooperative Oncology Group. MELD: model for end-stage liver disease. BCLC: Barcelona Clinic Liver Cancer stage. AFP: alpha-fetoprotein.

In the next table, the values obtained for various metrics analyzed for the developed models are presented. On one hand, the values for balanced accuracy, recall, specificity, and precision were analyzed. On the other hand, to assess the performance of the methods, area under the curve (AUC), F1 score, Matthews correlation coefficient (MCC), Youden's dependent index (DYI), and Kappa score were employed. The latter are commonly used methods in the scientific literature for this validation purpose.

As can be observed in Table 3, XGB presents values higher than 94% for balanced accuracy, recall, and specificity, and very close to this value for precision. This implies a significant difference compared to the closest method, KNN, with differences of around 8% for these values. The differences are more substantial for the rest of the algorithms, especially with GNB. In this case, the differences are around 12%. This superiority also translates into the rest of the metrics.

Table 3. Compilation of the analyzed metrics for all algorithms. SVM: support vector machine. BLDA: Bayesian linear discriminant analysis. DT: decision tree. GNB: Gaussian naïve Bayes. KNN: K-nearest neighbors. XGB: extreme gradient boosting. AUC: area under the curve.

Methods	Accuracy	Recall	Specificity	Precision	AUC
SVM	86.96	87.06	86.85	86.34	0.87
BLDA	84.32	84.42	84.23	83.72	0.84
DT	86.11	86.51	86.41	85.69	0.86
GNB	82.18	82.27	82.08	81.59	0.82
KNN	88.93	89.03	88.82	88.29	0.89
XGB	94.29	94.40	94.18	93.61	0.94

When MCC results are observed, there is a difference of 4.75% between XGB and KNN in favor of the proposed model. This is one of the most reliable statistical indices, yielding high values only when correctly performed across all four categories of the confusion matrix [45]. The differences are significantly higher when comparing the values of F1 score, Kappa, and DYI. Clearly, the differences are also more pronounced for the other proposed algorithms compared to XGB, ranging favorably between 8–12% for XGB. All the aforementioned data are detailed in Table 4.

Table 4. Summary of metrics collected and analyzed for all methods. SVM: support vector machine. BLDA: Bayesian linear discriminant analysis. DT: decision tree. GNB: Gaussian naïve Bayes. KNN: K-nearest neighbors. XGB: extreme gradient boosting. MCC: Matthews correlation coefficient. DYI: degenerated Youden index.

Methods	MCC	DYI	F1 Score	Kappa
SVM	77.16	86.96	86.70	77.41
BLDA	74.82	84.32	84.07	75.07
DT	76.54	86.11	86.02	76.89
GNB	72.92	82.18	81.93	73.16
KNN	78.91	88.93	88.66	79.17
XGB	83.66	94.29	94.00	83.94

Regarding the receiver operating characteristic (ROC) curves, XGB achieves an AUC superior to the rest of the systems. These curves represent sensitivity and specificity for the study's purpose (Figure 3). The XGB algorithm obtained a value of 0.94, the largest curve among the proposed methods. This higher AUC translates to it being the best method for predicting mortality in patients diagnosed with HCC-NAFLD and for identifying the most influential variables affecting their mortality.

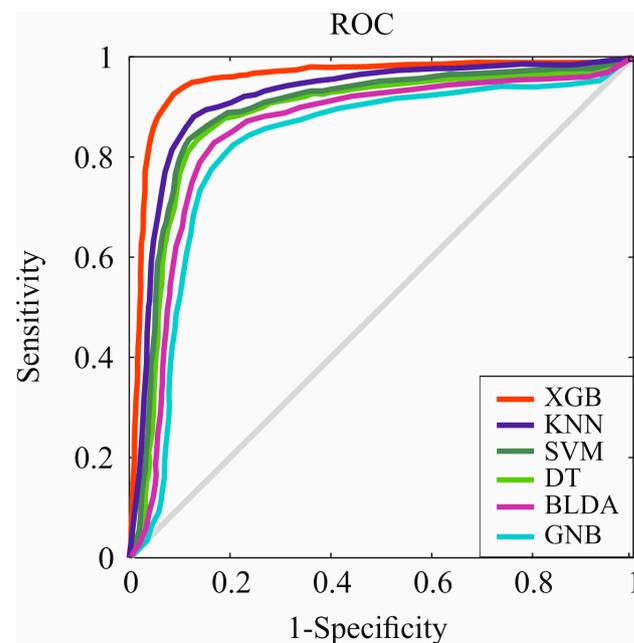


Figure 3. Representation of ROC curves of the analyzed algorithms. XGB: extreme gradient boosting. KNN: K-nearest neighbors. SVM: support vector machine. DT: decision tree. BLDA: Bayesian linear discriminant analysis. GNB: Gaussian naïve Bayes.

Finally, to depict all this data collectively, a radar plot was created. It showcases the training phase data (above of Figure 4) and the test phase data (below of Figure 4). As

observed, the obtained XGB algorithm presents similar data in both phases. This indicates there is no overfitting, implying that the resulting model generates a good predictive model with the capacity for generalizability. A smaller area obtained in this representation implies lower reliability for the study's objective.

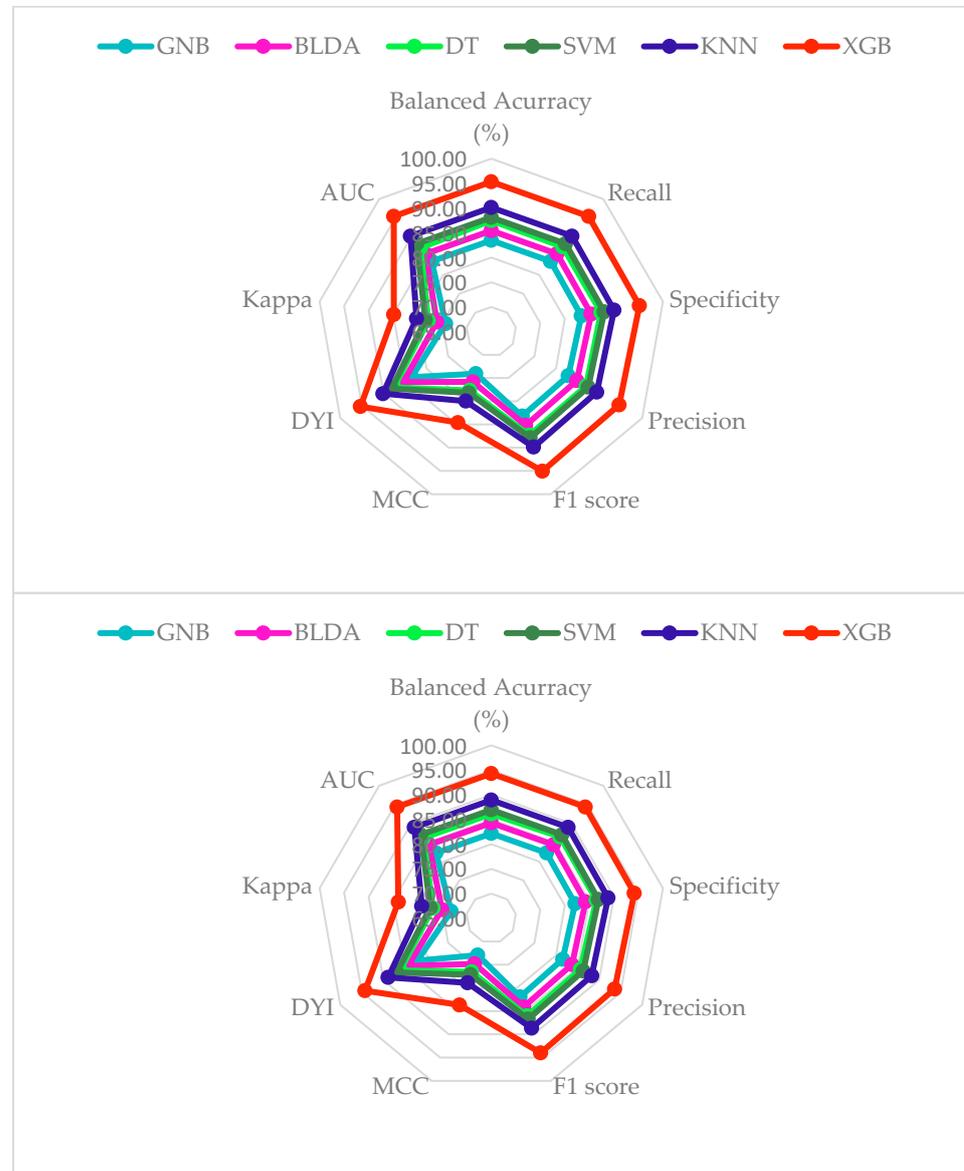
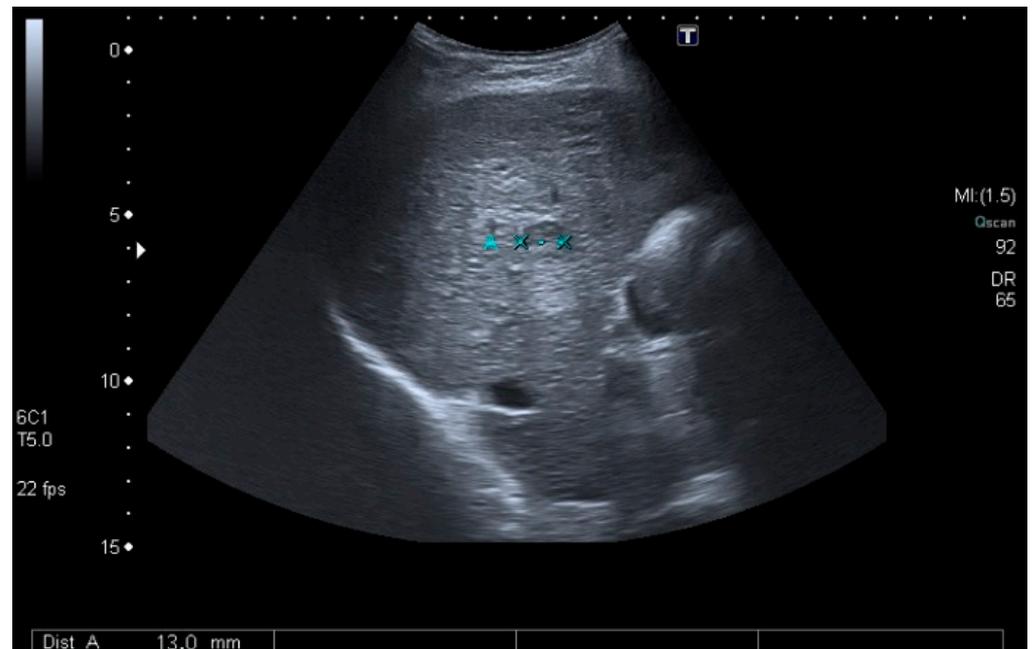


Figure 4. Radar plot comparing all the analyzed methods. The image above is the training phase and the image below is the test phase.

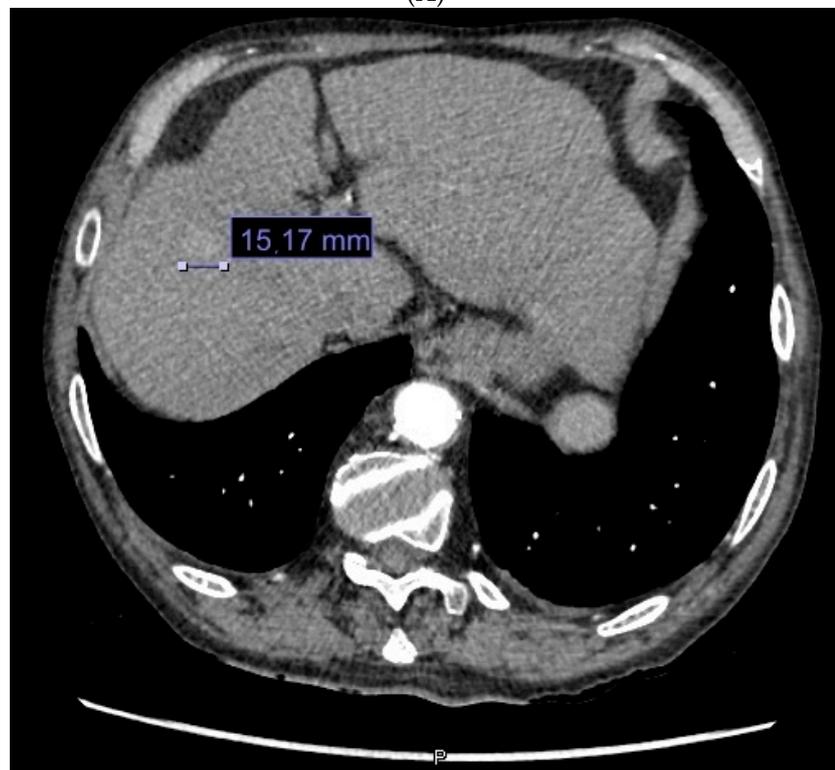
4. Discussion

NAFLD is one of the most relevant chronic diseases today and is clearly on the rise. Since 1990, it is estimated that the prevalence of this disease has risen by 50% [46]. This disease is becoming a significant challenge, especially from the perspective of gastroenterology, particularly hepatology. So much so that different scientific societies and governments of multiple countries are attempting to implement programs and public policies for awareness and to collectively confront this new epidemic [5,47]. The large number of NAFLD patients makes it one of the leading causes of HCC, and in the future, it will be the main one [48,49]. In addition, detecting fibrosis in these patients is complex. This point is crucial because the primary risk factor for this progression is the degree of hepatic fibrosis present,

with patients in stages F3-F4 being highlighted, along with the presence of non-alcoholic steatohepatitis (NASH) [50]. To this fact, the possibility of developing NAFL-associated HCC must be added [49]. That is, these patients can develop HCC without significant fibrosis or cirrhosis. The number of patients, its silent nature, and the lack of adequate screening make it impossible for healthcare systems to manage. Moreover, the diagnosis of HCC remains challenging despite the available advances in imaging tests, as depicted in Figure 5.



(A)



(B)

Figure 5. Cont.

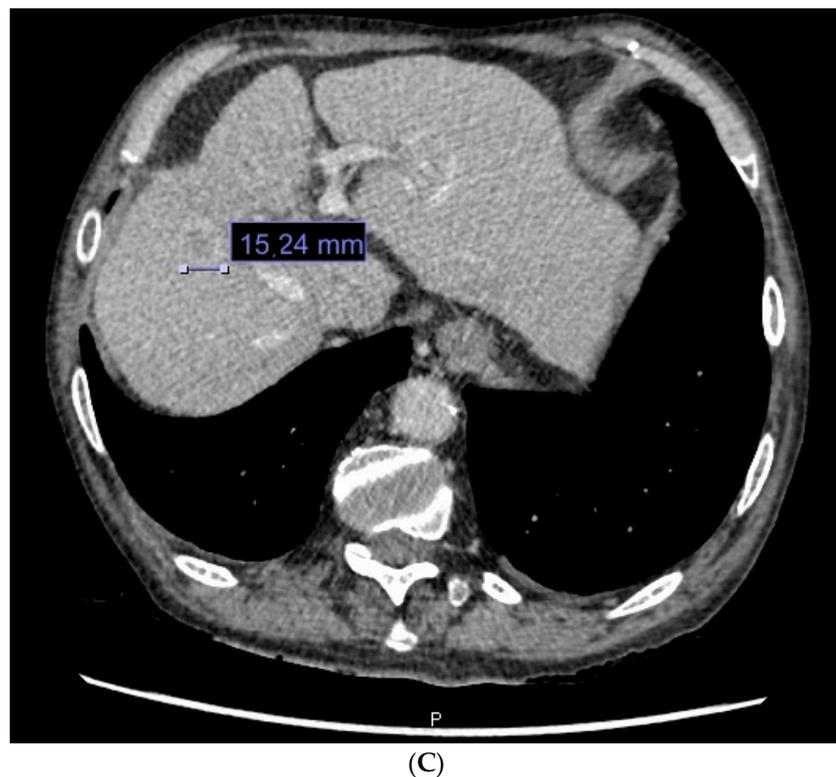


Figure 5. Diagnostic images of HCC. In image (A), a hypoechogenic lesion of 13 mm is observed on cirrhotic liver using ultrasound. Image (B) displays the arterial phase of the subsequent triphasic CT scan, revealing an image slightly larger than previously described. Finally, image (C) shows the venous phase with the characteristic washout, confirming the diagnosis of HCC. HCC: hepatocellular carcinoma. CT: computed tomography.

In line with all the aforementioned information, the trends in HCC etiology are changing. While years ago the primary causes of HCC development were viral hepatitis (chronic hepatitis B and C) and alcohol consumption, NAFLD is now becoming one of the main causes of HCC, already being the leading cause when referring to non-cirrhotic patients [51,52]. Simultaneously, it is also becoming one of the main reasons for liver transplantation. The importance of early detection is crucial since it is a potentially curable tumor. Adequate assessment and monitoring of these patients are essential to detect the disease in time. It is crucial to identify potential risk factors that may contribute to the development of HCC, aiming to act upon them and prevent their occurrence. The list is extensive, including smoking, alcohol consumption, obesity, lifestyle, detection of other underlying liver pathologies, and exposure to certain substances such as aflatoxin [25].

In this study, initially, the differences in mortality between patients with NAFLD-related HCC and other causes are compared. As can be observed, the difference in terms of survival between both groups was statistically significant. Once this was assessed, it was decided to investigate the main prognostic factors at the diagnosis of HCC associated with a poorer prognosis and higher mortality.

It is concluded that the primary risk factor for mortality in these patients is alcohol consumption. Alcohol has been shown to be a potential carcinogen not only at the hepatic level but also in other locations, such as the pancreas or colorectal area, among many others [53]. In this case, alcohol consumption emerges as the worst prognostic factor at the diagnosis of HCC in patients with NAFLD. This can be explained because alcohol acts as an additional incentive for liver damage. Even in low amounts, alcohol consumption enhances the progression of hepatic fibrosis, the degree of inflammation, and the development of HCC [54]. Additionally, although there are no significant histological differences between NASH and alcohol-induced steatohepatitis, alcohol induces characteristic epigenetic changes and

alteration in the intestinal microbiome, leading to increased intestinal permeability that may pose a higher risk of HCC [55–57].

The second factor with a worse prognosis is obesity. Obesity stands as one of the principal risk factors for the development of NAFLD. These patients also present a higher degree of fibrosis and the possibility of progressing to cirrhosis in proportion to BMI and abdominal circumference. This results from hepatic immune activation, leading to secondary inflammation and fibrosis, heightening the risk of HCC development [58,59]. Additionally, obesity also shows an association and a worse prognosis among patients who consume alcohol. This is explained by the direct relationship between alcohol intake and an increased risk of overweight and obesity [60]. The higher cardio-metabolic risk of these patients is likely an explanation for the worse prognosis.

Other factors associated with a worse prognosis are the presence of cirrhosis and CSPH, both equally significant. It is noteworthy that, despite these results, the Child–Pugh score exhibits lower significance within the predictive model. Unlike the risk factors for HCC development and the published literature, smoking and the presence of DM are not decisive regarding the prognosis of these patients [25]. In fact, being or having been a smoker hardly showed any significance within the predictive model. The low importance of AFP in the prognosis of patients is remarkable. This aligns with the latest scientific evidence published. It is also noteworthy that among the variables related to the tumor, only the largest nodule attains significant value within the predictive model for survival prognosis upon the diagnosis of these patients.

A bibliographic search was conducted on prognostic factors at the diagnosis of HCC focusing on ML techniques, but no results were obtained. Studies conducted previously have focused more on imaging techniques, genomics, and molecular biology [61,62]. For this reason, it was decided to perform the analysis by evaluating various methods. The XGB system demonstrated the best results in all parameters analyzed compared to others. Except for two metrics, the results obtained are around 94%. These figures confirm the utility of XGB for classifying these patients. Moreover, the similarity between the training and test phases shown in the radar plot indicates the absence of overfitting and overtraining. This implies high generalizability, so that when introducing new data, the results obtained are consistent with those obtained at the current time [63]. The method also exhibits high scalability and execution speed, allowing its usefulness in daily clinical practice to assist in decision making.

There are two limitations that need to be addressed. Apart from those inherent in a retrospective study, the primary variable being alcohol consumption raises doubts about whether it was accurately recorded. After analyzing the database, the data were corroborated by at least two researchers from each center, reviewing available medical records of the patients. On the other hand, there might be a question about whether the number of included patients was sufficient to draw these conclusions. To address this, ML techniques can mitigate this issue by optimizing hyperparameters to achieve the highest possible accuracy [64]. These methods more efficiently surpass traditional logistic regressions. The study is replicated 100 times to obtain primary values and standard deviations. This way, statistically significant results are obtained with small samples, also avoiding the potential noise present within [65].

5. Conclusions

In conclusion, patients with NAFLD-related HCC exhibit an unfavorable prognosis in terms of survival. Even low alcohol consumption in patients with NAFLD was associated with a poorer prognosis. Obesity, cirrhosis at any stage, and CSPH emerged as additional risk factors conditioning increased mortality at the time of HCC diagnosis in these patients.

XGB proved to be the algorithm that developed a more efficient predictive model in identifying prognostic factors for mortality at the diagnosis of HCC in patients with NAFLD. This model can serve as a valuable tool in the daily management of these patients. Thanks to these results, a more personalized management approach can be offered for

these individuals. In addition to the established therapeutic approach, optimizing these conditions may contribute to an improvement in the survival of these patients.

Author Contributions: Conceptualization, M.S., S.G.-R., P.M.-B., A.M.T., M.T. and J.M.; methodology, M.S., S.G.-R., A.M.T., M.T. and J.M.; software, A.M.T. and J.M.; validation, A.M.T. and J.M.; formal analysis, A.M.T. and J.M.; investigation, M.S., S.G.-R., P.M.-B. and M.T.; resources, M.S., S.G.-R., P.M.-B., A.M.T., A.R., M.T. and J.M.; data curation, M.S., S.G.-R., P.M.-B., A.M.T., M.T. and J.M.; writing—original draft preparation, M.S. and J.M.; writing—review and editing, M.S., S.G.-R., P.M.-B., A.M.T., A.R., P.B.-S., M.T. and J.M.; visualization, M.S., S.G.-R., P.M.-B., A.M.T., A.R., P.B.-S., M.T. and J.M.; supervision, M.T. and J.M.; project administration, M.S., S.G.-R., M.T. and J.M.; and funding acquisition, A.R., P.B.-S. and J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fundación Investigación Hospital General Universitario de Valencia and University of Castilla-La Mancha (Spain).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of University Hospital of Guadalajara.

Informed Consent Statement: Patient consent was waived due to the number of patients, study design (retrospective), absence of medical prescription, and the number of deceased patients.

Data Availability Statement: The datasets used and/or analyzed during the present study are available from the corresponding author on reasonable request.

Acknowledgments: This study was sponsored by Virgen de la Luz Hospital of Cuenca (Spain), Fundación Investigación Hospital General Universitario de Valencia and the Institute of Technology of the University of Castilla-La Mancha (Spain).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Le, M.H.; Yeo, Y.H.; Li, X.; Li, J.; Zou, B.; Wu, Y.; Ye, Q.; Huang, D.Q.; Zhao, C.; Zhang, J. 2019 Global NAFLD prevalence: A systematic review and meta-analysis. *Clin. Gastroenterol. Hepatol.* **2022**, *20*, 2809–2817.e28. [[CrossRef](#)]
2. Younossi, Z.M.; Golabi, P.; Paik, J.M.; Henry, A.; Van Dongen, C.; Henry, L. The global epidemiology of nonalcoholic fatty liver disease (NAFLD) and nonalcoholic steatohepatitis (NASH): A systematic review. *Hepatology* **2023**, *77*, 1335–1347. [[CrossRef](#)] [[PubMed](#)]
3. Rinella, M.E.; Neuschwander-Tetri, B.A.; Siddiqui, M.S.; Abdelmalek, M.F.; Caldwell, S.; Barb, D.; Kleiner, D.E.; Loomba, R. AASLD practice guidance on the clinical assessment and management of nonalcoholic fatty liver disease. *Hepatology* **2023**, *77*, 1797–1835. [[CrossRef](#)]
4. Lazarus, J.V.; Ekstedt, M.; Marchesini, G.; Mullen, J.; Novak, K.; Pericàs, J.M.; Roel, E.; Romero-Gómez, M.; Ratziu, V.; Tacke, F. A cross-sectional study of the public health response to non-alcoholic fatty liver disease in Europe. *J. Hepatol.* **2020**, *72*, 14–24. [[CrossRef](#)] [[PubMed](#)]
5. Lazarus, J.V.; Mark, H.E.; Allen, A.M.; Arab, J.P.; Carrieri, P.; Noureddin, M.; Alazawi, W.; Alkhouri, N.; Alqahtani, S.A.; Arrese, M. A global research priority agenda to advance public health responses to fatty liver disease. *J. Hepatol.* **2023**, *79*, 618–634. [[CrossRef](#)]
6. Lonardo, A.; Leoni, S.; Alswat, K.A.; Fouad, Y. History of nonalcoholic fatty liver disease. *Int. J. Mol. Sci.* **2020**, *21*, 5888. [[CrossRef](#)]
7. Eslam, M.; Sanyal, A.J.; George, J.; Sanyal, A.; Neuschwander-Tetri, B.; Tiribelli, C.; Kleiner, D.E.; Brunt, E.; Bugianesi, E.; Yki-Järvinen, H. MAFLD: A consensus-driven proposed nomenclature for metabolic associated fatty liver disease. *Gastroenterology* **2020**, *158*, 1999–2014.e1991. [[CrossRef](#)]
8. Staufer, K.; Stauber, R.E. Steatotic Liver Disease: Metabolic Dysfunction, Alcohol, or Both? *Biomedicines* **2023**, *11*, 2108. [[CrossRef](#)] [[PubMed](#)]
9. Rinella, M.E.; Lazarus, J.V.; Ratziu, V.; Francque, S.M.; Sanyal, A.J.; Kanwal, F.; Romero, D.; Abdelmalek, M.F.; Anstee, Q.M.; Arab, J.P. A multi-society Delphi consensus statement on new fatty liver disease nomenclature. *Ann. Hepatol.* **2023**, *29*, 101133. [[CrossRef](#)]
10. McPherson, S.; Hardy, T.; Henderson, E.; Burt, A.D.; Day, C.P.; Anstee, Q.M. Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: Implications for prognosis and clinical management. *J. Hepatol.* **2015**, *62*, 1148–1155. [[CrossRef](#)]
11. Shaikh, A.; Pedra, G.; Ruiz-Casas, L.; Franks, B.; Dhillon, H.; Fernandes, J.D.d.R.; Mangla, K.K.; Augusto, M.; Romero-Gómez, M.; Schattenberg, J.M. Risk factors for fibrosis progression in non-alcoholic steatohepatitis: Analysis of the European cohort in the real-world GAIN study. *Gastroenterol. Y Hepatol.* **2023**, *in press*. [[CrossRef](#)]

12. Estes, C.; Anstee, Q.M.; Arias-Loste, M.T.; Bantel, H.; Bellentani, S.; Caballeria, J.; Colombo, M.; Craxi, A.; Crespo, J.; Day, C.P. Modeling nafld disease burden in china, france, germany, italy, japan, spain, united kingdom, and united states for the period 2016–2030. *J. Hepatol.* **2018**, *69*, 896–904. [[CrossRef](#)] [[PubMed](#)]
13. Zhu, Y.; Zhang, H.; Jiang, P.; Xie, C.; Luo, Y.; Chen, J. Transcriptional and Epigenetic Alterations in the Progression of Non-Alcoholic Fatty Liver Disease and Biomarkers Helping to Diagnose Non-Alcoholic Steatohepatitis. *Biomedicines* **2023**, *11*, 970. [[CrossRef](#)] [[PubMed](#)]
14. Huang, G.; Wallace, D.F.; Powell, E.E.; Rahman, T.; Clark, P.J.; Subramaniam, V.N. Gene Variants Implicated in Steatotic Liver Disease: Opportunities for Diagnostics and Therapeutics. *Biomedicines* **2023**, *11*, 2809. [[CrossRef](#)] [[PubMed](#)]
15. Powell, E.E.; Wong, V.W.-S.; Rinella, M. Non-alcoholic fatty liver disease. *Lancet* **2021**, *397*, 2212–2224. [[CrossRef](#)] [[PubMed](#)]
16. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
17. European Association for the Study of the Liver. EASL clinical practice guidelines: Management of hepatocellular carcinoma. *J. Hepatol.* **2018**, *69*, 182–236. [[CrossRef](#)] [[PubMed](#)]
18. Heimbach, J.K.; Kulik, L.M.; Finn, R.S.; Sirlin, C.B.; Abecassis, M.M.; Roberts, L.R.; Zhu, A.X.; Murad, M.H.; Marrero, J.A. AASLD guidelines for the treatment of hepatocellular carcinoma. *Hepatology* **2018**, *67*, 358–380. [[CrossRef](#)]
19. Reig, M.; Forner, A.; Ávila, M.A.; Ayuso, C.; Mínguez, B.; Varela, M.; Bilbao, I.; Bilbao, J.I.; Burrel, M.; Bustamante, J. Diagnosis and treatment of hepatocellular carcinoma. Update of the consensus document of the AEEH, AEC, SEOM, SERAM, SERVEI, and SETH. *Med. Clín. Engl. Ed.* **2021**, *156*, 463.e1–463.e31. [[CrossRef](#)]
20. Beam, A.L.; Kohane, I.S. Big data and machine learning in health care. *JAMA* **2018**, *319*, 1317–1318. [[CrossRef](#)]
21. Deo, R.C. Machine learning in medicine. *Circulation* **2015**, *132*, 1920–1930. [[CrossRef](#)]
22. Rajkomar, A.; Dean, J.; Kohane, I. Machine learning in medicine. *N. Engl. J. Med.* **2019**, *380*, 1347–1358. [[CrossRef](#)] [[PubMed](#)]
23. Zhang, D.; Gong, Y. The comparison of LightGBM and XGBoost coupling factor analysis and prediagnosis of acute liver failure. *IEEE Access* **2020**, *8*, 220990–221003. [[CrossRef](#)]
24. Suárez, M.; Martínez, R.; Torres, A.M.; Torres, B.; Mateo, J. A Machine Learning Method to Identify the Risk Factors for Liver Fibrosis Progression in Nonalcoholic Steatohepatitis. *Dig. Dis. Sci.* **2023**, *68*, 3801–3809. [[CrossRef](#)] [[PubMed](#)]
25. Singal, A.G.; Llovet, J.M.; Yarrow, M.; Mehta, N.; Heimbach, J.K.; Dawson, L.A.; Jou, J.H.; Kulik, L.M.; Agopian, V.G.; Marrero, J.A. AASLD practice guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. *Hepatology* **2023**, *78*, 1922–1965. [[CrossRef](#)] [[PubMed](#)]
26. Blonde, L.; Umpierrez, G.E.; Reddy, S.S.; McGill, J.B.; Berga, S.L.; Bush, M.; Chandrasekaran, S.; DeFronzo, R.A.; Einhorn, D.; Galindo, R.J. American Association of Clinical Endocrinology clinical practice guideline: Developing a diabetes mellitus comprehensive care plan—2022 update. *Endocr. Pract.* **2022**, *28*, 923–1049. [[CrossRef](#)] [[PubMed](#)]
27. Jellinger, P.S.; Handelsman, Y.; Rosenblit, P.D.; Bloomgarden, Z.T.; Fonseca, V.A.; Garber, A.J.; Grunberger, G.; Guerin, C.K.; Bell, D.S.; Mechanick, J.I. American Association of Clinical Endocrinologists and American College of Endocrinology guidelines for management of dyslipidemia and prevention of cardiovascular disease. *Endocr. Pract.* **2017**, *23*, 1–87. [[CrossRef](#)] [[PubMed](#)]
28. Apovian, C.M. Obesity: Definition, comorbidities, causes, and burden. *Am. J. Manag. Care* **2016**, *22*, S176–S185.
29. Ginès, P.; Krag, A.; Abraldes, J.G.; Solà, E.; Fabrellas, N.; Kamath, P.S. Liver cirrhosis. *Lancet* **2021**, *398*, 1359–1376. [[CrossRef](#)]
30. Kok, B.; Abraldes, J.G. Child–Pugh Classification: Time to Abandon? In *Seminars in Liver Disease*; Thieme Medical Publishers: Leipzig, Germany, 2019; pp. 096–103.
31. Sok, M.; Zavrl, M.; Greif, B.; Srpčić, M. Objective assessment of WHO/ECOG performance status. *Support. Care Cancer* **2019**, *27*, 3793–3798. [[CrossRef](#)]
32. Peng, Y.; Qi, X.; Guo, X. Child–Pugh versus MELD score for the assessment of prognosis in liver cirrhosis: A systematic review and meta-analysis of observational studies. *Medicine* **2016**, *95*, e2877. [[CrossRef](#)]
33. De Franchis, R.; Bosch, J.; Garcia-Tsao, G.; Reiberger, T.; Ripoll, C.; Abraldes, J.G.; Albillos, A.; Baiges, A.; Bajaj, J.; Bañares, R. Baveno VII—renewing consensus in portal hypertension. *J. Hepatol.* **2022**, *76*, 959–974. [[CrossRef](#)]
34. Vogel, A.; Cervantes, A.; Chau, I.; Daniele, B.; Llovet, J.M.; Meyer, T.; Nault, J.-C.; Neumann, U.; Rieke, J.; Sangro, B. Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **2018**, *29*, iv238–iv255. [[CrossRef](#)]
35. Toniutto, P.; Fumolo, E.; Fornasiero, E.; Bitetto, D. Liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: A comprehensive review. *J. Clin. Med.* **2021**, *10*, 3932. [[CrossRef](#)] [[PubMed](#)]
36. Reig, M.; Forner, A.; Rimola, J.; Ferrer-Fàbrega, J.; Burrel, M.; Garcia-Criado, Á.; Kelley, R.K.; Galle, P.R.; Mazzaferro, V.; Salem, R. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *J. Hepatol.* **2022**, *76*, 681–693. [[CrossRef](#)] [[PubMed](#)]
37. Sagi, O.; Rokach, L. Approximating XGBoost with an interpretable decision tree. *Inf. Sci.* **2021**, *572*, 522–542. [[CrossRef](#)]
38. Sahmadi, B.; Boughaci, D.; Rahmani, R.; Sissani, N. A modified firefly algorithm with support vector machine for medical data classification. In Proceedings of the 6th IFIP TC 5 International Conference—Computational Intelligence and Its Applications, CIIA 2018, Oran, Algeria, 8–10 May 2018; Proceedings 6. Springer: Berlin/Heidelberg, Germany, 2018; pp. 232–243.

39. Suárez, M.; Martínez, R.; Torres, A.M.; Ramón, A.; Blasco, P.; Mateo, J. Personalized Risk Assessment of Hepatic Fibrosis after Cholecystectomy in Metabolic-Associated Steatotic Liver Disease: A Machine Learning Approach. *J. Clin. Med.* **2023**, *12*, 6489. [[CrossRef](#)] [[PubMed](#)]
40. Liang, J.; Qin, Z.; Xiao, S.; Ou, L.; Lin, X. Efficient and secure decision tree classification for cloud-assisted online diagnosis services. *IEEE Trans. Dependable Secur. Comput.* **2019**, *18*, 1632–1644. [[CrossRef](#)]
41. Bafjaish, S.S. Comparative analysis of Naive Bayesian techniques in health-related for classification task. *J. Soft Comput. Data Min.* **2020**, *1*, 1–10.
42. Xing, W.; Bei, Y. Medical health big data classification based on KNN classification algorithm. *IEEE Access* **2019**, *8*, 28808–28819. [[CrossRef](#)]
43. Turner, R.; Eriksson, D.; McCourt, M.; Kiili, J.; Laaksonen, E.; Xu, Z.; Guyon, I. Bayesian optimization is superior to random search for machine learning hyperparameter tuning: Analysis of the black-box optimization challenge 2020. In Proceedings of the NeurIPS 2020 Competition and Demonstration Track, Virtual, 6–12 December 2020; pp. 3–26.
44. Yang, J.; Li, Y.; Liu, Q.; Li, L.; Feng, A.; Wang, T.; Zheng, S.; Xu, A.; Lyu, J. Brief introduction of medical database and data mining technology in big data era. *J. Evid.-Based Med.* **2020**, *13*, 57–69. [[CrossRef](#)]
45. Chicco, D.; Warrens, M.J.; Jurman, G. The Matthews correlation coefficient (MCC) is more informative than Cohen’s Kappa and Brier score in binary classification assessment. *IEEE Access* **2021**, *9*, 78368–78381. [[CrossRef](#)]
46. Estes, C.; Razavi, H.; Loomba, R.; Younossi, Z.; Sanyal, A.J. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology* **2018**, *67*, 123–133. [[CrossRef](#)]
47. Karlsen, T.H.; Sheron, N.; Zelber-Sagi, S.; Carrieri, P.; Dusheiko, G.; Bugianesi, E.; Pryke, R.; Hutchinson, S.J.; Sangro, B.; Martin, N.K. The EASL–Lancet Liver Commission: Protecting the next generation of Europeans against liver disease complications and premature mortality. *Lancet* **2022**, *399*, 61–116. [[CrossRef](#)] [[PubMed](#)]
48. Foerster, F.; Gairing, S.J.; Müller, L.; Galle, P.R. NAFLD-driven HCC: Safety and efficacy of current and emerging treatment options. *J. Hepatol.* **2022**, *76*, 446–457. [[CrossRef](#)] [[PubMed](#)]
49. Shah, P.A.; Patil, R.; Harrison, S.A. NAFLD-related hepatocellular carcinoma: The growing challenge. *Hepatology* **2023**, *77*, 323–338. [[CrossRef](#)] [[PubMed](#)]
50. Taylor, R.S.; Taylor, R.J.; Bayliss, S.; Hagström, H.; Nasr, P.; Schattenberg, J.M.; Ishigami, M.; Toyoda, H.; Wong, V.W.-S.; Peleg, N. Association between fibrosis stage and outcomes of patients with nonalcoholic fatty liver disease: A systematic review and meta-analysis. *Gastroenterology* **2020**, *158*, 1611–1625.e12. [[CrossRef](#)] [[PubMed](#)]
51. Kanwal, F.; Kramer, J.R.; Mapakshi, S.; Natarajan, Y.; Chayanupatkul, M.; Richardson, P.A.; Li, L.; Desiderio, R.; Thrift, A.P.; Asch, S.M. Risk of hepatocellular cancer in patients with non-alcoholic fatty liver disease. *Gastroenterology* **2018**, *155*, 1828–1837.e2. [[CrossRef](#)] [[PubMed](#)]
52. Ioannou, G.N. Epidemiology and risk-stratification of NAFLD-associated HCC. *J. Hepatol.* **2021**, *75*, 1476–1484. [[CrossRef](#)] [[PubMed](#)]
53. Rumgay, H.; Shield, K.; Charvat, H.; Ferrari, P.; Sornpaisarn, B.; Obot, I.; Islami, F.; Lemmens, V.E.; Rehm, J.; Soerjomataram, I. Global burden of cancer in 2020 attributable to alcohol consumption: A population-based study. *Lancet Oncol.* **2021**, *22*, 1071–1080. [[CrossRef](#)]
54. Barbería-Latasa, M.; Gea, A.; Martínez-González, M.A. Alcohol, drinking pattern, and chronic disease. *Nutrients* **2022**, *14*, 1954. [[CrossRef](#)] [[PubMed](#)]
55. Takahashi, Y.; Dungubat, E.; Kusano, H.; Fukusato, T. Pathology and Pathogenesis of Metabolic Dysfunction-Associated Steatotic Liver Disease-Associated Hepatic Tumors. *Biomedicines* **2023**, *11*, 2761. [[CrossRef](#)] [[PubMed](#)]
56. Shen, H.; French, B.A.; Tillman, B.C.; Li, J.; French, S.W. Increased DNA methylation in the livers of patients with alcoholic hepatitis. *Exp. Mol. Pathol.* **2015**, *99*, 326–329. [[CrossRef](#)]
57. Zhu, R.; Baker, S.S.; Moylan, C.A.; Abdelmalek, M.F.; Guy, C.D.; Zamboni, F.; Wu, D.; Lin, W.; Liu, W.; Baker, R.D. Systematic transcriptome analysis reveals elevated expression of alcohol-metabolizing genes in NAFLD livers. *J. Pathol.* **2016**, *238*, 531–542. [[CrossRef](#)] [[PubMed](#)]
58. Motta, B.M.; Masarone, M.; Torre, P.; Persico, M. From Non-Alcoholic Steatohepatitis (NASH) to Hepatocellular Carcinoma (HCC): Epidemiology, Incidence, Predictions, Risk Factors, and Prevention. *Cancers* **2023**, *15*, 5458. [[CrossRef](#)]
59. Mota, M.; Banini, B.A.; Cazanave, S.C.; Sanyal, A.J. Molecular mechanisms of lipotoxicity and glucotoxicity in nonalcoholic fatty liver disease. *Metabolism* **2016**, *65*, 1049–1061. [[CrossRef](#)]
60. Bataller, R.; Cabezas, J.; Aller, R.; Ventura-Cots, M.; Abad, J.; Albillos, A.; Altamirano, J.; Arias-Loste, M.T.; Bañares, R.; Caballería, J. Enfermedad hepática por alcohol. Guías de práctica clínica. Documento de consenso auspiciado por la AEEH. *Gastroenterol. Hepatol.* **2019**, *42*, 657–676. [[CrossRef](#)]
61. Xia, T.; Zhao, B.; Li, B.; Lei, Y.; Song, Y.; Wang, Y.; Tang, T.; Ju, S. MRI-Based Radiomics and Deep Learning in Biological Characteristics and Prognosis of Hepatocellular Carcinoma: Opportunities and Challenges. *J. Magn. Reson. Imaging* **2023**, *59*, 767–783. [[CrossRef](#)]
62. Feng, Q.; Huang, Z.; Song, L.; Wang, L.; Lu, H.; Wu, L. Combining bulk and single-cell RNA-sequencing data to develop an NK cell-related prognostic signature for hepatocellular carcinoma based on an integrated machine learning framework. *Eur. J. Med. Res.* **2023**, *28*, 306. [[CrossRef](#)] [[PubMed](#)]

63. Mateo-Sotos, J.; Torres, A.; Santos, J.; Quevedo, O.; Basar, C. A machine learning-based method to identify bipolar disorder patients. *Circuits Syst. Signal Process.* **2022**, *41*, 2244–2265. [[CrossRef](#)]
64. Sammut, C.; Webb, G.I. *Encyclopedia of Machine Learning and Data Mining*; Springer Publishing Company, Incorporated: Berlin/Heidelberg, Germany, 2017.
65. Han, J.; Pei, J.; Tong, H. *Data Mining: Concepts and Techniques*; Morgan Kaufmann: Burlington, MA, USA, 2022.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.