

RESPONSE TO REVIEWERS

Journal title: World Journal of Gastroenterology

Manuscript NO: 79914

Reviewer #1

1. Authors could mention papers using image segmentation in Chronic Liver Disease.

R: Thank you for the comment. CT images and X-ray scans were two variables considered as inputs to the ML algorithms in several research papers already mentioned in the manuscript. To make it clearer, the following sentences were added:

-Line 321-3: "For example, to determine the stage of liver fibrosis, some authors have used CT images processed by segmentation algorithms." -Line 445-7: "These authors pre-processed by normalization the images and then segmented them by fuzzy c-means clustering."

Furthermore, the following sentence was added in line 284-7 and the paragraph rephrased as follows: *"Therefore, there is a need for ML approaches to help in image segmentation and some authors have already implemented this technique to improve clinical practice [60,61]. Moreover, ML can help with the integration of more complex information beyond imaging to study and diagnose liver diseases, since patients with CLD..."*

2. In conclusion, you should summarize interpret the most significant factors outlined for this disease after analysing so many papers.

R: Thank you for the suggestion, as we believe it improves the reading of the article. We added a table (Table 2) that summarizes the most repeated entries from all compiled ML models along with the most repeated predictive results for the top four inflammation-related liver conditions. Please find it below.

Table 2. Summary of the most repeated inputs of the machine learning models with the most repeated predictor outcomes for the four main inflammatory-related liver conditions. Classifier (SRC), a able d nicable and viral diseases.

Inflammatory-related liver condition	Inputs	Most repeated predictors
FLD	Age, sex, blood biomarkers, demographic, anthropometric and clinical data	BMI, uric acid, TG, and ALT levels
Liver fibrosis	Age, sex, and CT images	Better diagnosis compared to classical methods like APRI and FIB-4 indexes.
Virus-induced hepatitis	Age, sex, blood biomarkers, demographic, anthropometric and clinical data	AST, platelet levels, APRI index, and age
COVID-19	Age, sex, blood biomarkers, CT images, and demographic, anthropometric and clinical data	Age, BMI, CT images, oxygen rate, AST, and ALT levels

FLD: fatty liver disease; CT: computed tomography.

3. You should briefly explain the machine learning performance measures and what they refer to.

R: Thank you for the comment. The following clarification was added in lines 250-53: *“ML algorithms are generally assessed by simple methodologies like sensitivity,*

specificity, and accuracy. While sensitivity evaluates the proportion of true positives correctly identified, specificity evaluates the proportion of true negatives. Meanwhile, the accuracy value indicates the number of times the model is correct [54]."

Reviewer #2

This is an interesting paper but there are some issues that must be addressed before it can be published:

1. The paper' contribution needs to be stated clearly.

R: Thank you for your suggestion. To clarify the paper's contribution, the following statements were added:

-Line 46: "...and enlighten the factors involved in CLD development."

-Line 125-8: "Its goal is to shed light on the factors involved in CLD to help health professionals in clinical management with the support of ML, and identify new targets that can define therapeutic care lines in viral infections and non-communicable diseases (NCD), with an impact on liver functions with an inflammatory component."

-Line 561-2: "This review clarifies and compiles the importance of the different factors involved in CLD and analysed by ML algorithms..."

2. It may be helpful to identify the target audience for the paper, given the rather technical nature of research topic.

R: Thank you for the comment. To clarify the target, the following phrase was added in conclusion lines 563-5: *"...which can be useful information for clinicians, like endocrinologists and gastroenterologists, and other healthcare professionals with a focus on hepatology and bioinformatics."*

3 and 5. Specify what makes this article different from the rest of studies that are available in the literature. The contribution of your study needs to be clearly articulated in this section. The contribution is not clear and need to be stated.

R: This article is valuable to the scientific community because it brings together data from traditional interventions with ML algorithms in the context of liver damage. Thus, it compiles information on diseases that affect liver function from an inflammatory perspective, including what has been done so far regarding new

diseases such as COVID-19. Along with the amends made in point 1 based on this reviewer’s comment, the following sentence was also added in the conclusion section in line 553-5: *“The objective of this article was to collect the information derived from ML techniques in liver damage induced by inflammatory conditions, including the new disease COVID-19”*.

4. Please summarize the influencing factors of hepatitis and these can be used as a summary of the AI input variables.

R: Thank you for the suggestion. In the wake of a previous comment from another reviewer, we have now added a table summarizing these factors (Table 2). Please find it below.

Table 2. Summary of the most repeated inputs of the machine learning models with the most repeated predictor outcomes for the four main inflammatory-related liver conditions. and viral diseases

Inflammatory-related liver condition	Inputs	Most repeated predictors
FLD	Age, sex, blood biomarkers, demographic, anthropometric and clinical data	BMI, uric acid, TG, and ALT levels
Liver fibrosis	Age, sex, and CT images	Better diagnosis compared to classical methods like APRI and FIB-4 indexes.
Virus-induced hepatitis	Age, sex, blood biomarkers, demographic, anthropometric and clinical data	AST, platelet levels, APRI index, and age

Age, sex, blood biomarkers, CT images, Age, BMI, CT images, COVID-19 and demographic, oxygen rate, AST, and anthropometric and ALT levels clinical data

FLD: fatty liver disease; CT: computed tomography.

6. The section needs to include some recommendations for practitioners based on the findings, if appropriate.

R: Regarding this comment, the following paragraph was added in lines 544-50:

“Personalized and precision medicine aims to harmonize the greatest number of factors so that diagnosis, prognosis, and treatment are based on the greatest number of decision elements. Much remains to be investigated to establish guidelines in the context of personalized medicine. However, it is safe to say that precision medicine will drive modern medicine, combining the most classic variables with the newest digital ones. Health professionals must be prepared to understand and implement these new technologies in the near future.”

Reviewer #3

This manuscript is well and good at innovation and clears the clarity of the reader. It is well structured and well written. The author does a good job of presenting a highly technical and complicated process in an easy-to-understand manner.

R: Thank you for the nice words about our work.

Authors need to cross check the reference section by addressing the cited contents in the introduction and related work part.

R: Thank you for the comment. The reference section has been cross checked and the problem has been solved regarding this part of the manuscript.

The introduction must be an extended version of the abstract. The authors must elaborate on the points highlighted on the abstract and give supportive ideas and references.

R: As this reviewer indicates, the abstract was rephrased as follows:

“The liver is a key organ involved in a wide range of functions, whose damage can lead to chronic liver disease (CLD). CLD accounts for more than two million deaths worldwide, becoming a social and economic burden for most countries. Among the different factors that can cause CLD, alcohol abuse, viruses, drug treatments, and unhealthy dietary patterns top the list. These conditions prompt to perpetuate an inflammatory environment and oxidative stress imbalance that favour the development of hepatic fibrogenesis. High stages of fibrosis can eventually lead to cirrhosis or hepatocellular carcinoma (HCC). Despite the advances achieved in this field, new approaches are needed for the prevention, diagnosis, treatment, and prognosis of CLD. In this context, the scientific community is using machine learning (ML) algorithms to integrate and process vast amounts of data with unprecedented performance. ML techniques allow the integration of anthropometric, genetic, clinical, biochemical, dietary, lifestyle and omics data, giving new insights to tackle CLD and bringing personalized medicine a step closer. This review summarises the investigations, where ML techniques have been applied to study new approaches that could be used in inflammatory-related, hepatitis viruses-induced, and COVID-19-induced liver damage, and enlighten the factors involved in CLD development.”

The conclusions in this manuscript are primitive. Rewrite your conclusions.

R: Following this comment and others made by the previous reviewer, the conclusions were rewritten as follows in lines 552-65:

“In summary, ML science can process and integrate a vast amount of different data with unprecedented outstanding performance. The objective of this article was to collect the information derived from ML techniques in liver damage induced by inflammatory conditions, including the new disease COVID-19. The main role of ML in liver pathologies is to help identify high risk patients for referral to specialized centres. Results show that the use of ML models have brought new insights into biology and medicine questions that can be very useful in determining the next directions towards research in diagnosis, prognosis, and treatment of inflammatory and virus-related liver diseases, leading the way to personalized medicine. Also biomarkers concerning inflammation/insulin resistance related to liver disease can be boosted by ML strategies.

This review clarifies and compiles the importance of the different factors involved in CLD and analysed by ML algorithms, which can be useful information for clinicians, like endocrinologists and gastroenterologists, and other health science professionals with a focus on hepatology and bioinformatics."

References aren't formatted according to rules.

R: Thank you for the note. We have checked the reference format and believe they are now cited according to the journal guidelines.

Additional References: The following articles could be useful: • Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning. <https://doi.org/10.52866/ijcsm.2022.01.01.013> • A diagnostic testing for people with appendicitis using machine learning techniques. <https://doi.org/10.1007/s11042-022-11939-8> • A Scoping Review of Machine Learning Techniques and Their Utilisation in Predicting Heart Diseases. <https://doi.org/10.30526/35.3.2813>

R: Thank you for these suggestions. We have added them with a little introduction in the following lines, respectively:

Line 240-1: *"AI seeks to mimic human behaviour, and within this science, ML is the most common approach [52]."*

Line 273-5: *"For instance, Mijwil and Aggarwal [57] analysed and compared 7 ML algorithms to predict appendix illness in the same dataset, revealing that certain models performed better than others, allowing for higher accuracy and results."*

Line 497-9: *"This methodology has been used in a wide range of diseases in the search of more efficient and effective approaches, like heart and liver diseases [109,110]."*