



Applications of artificial intelligence (AI) in researches on non-alcoholic fatty liver disease (NAFLD) : A systematic review

Yifang Li¹ · Xuetao Wang¹ · Jun Zhang¹ · Shanshan Zhang¹ · Jian Jiao¹

Accepted: 7 August 2021 / Published online: 15 August 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Non-alcoholic fatty liver disease (NAFLD) is one of the most important causes of chronic liver disease in the world, it has been found that cardiovascular and renal risks and diseases are also highly prevalent in adults with NAFLD. Diagnosis and treatment of NAFLD face many challenges, although the medical science has been very developed. Efficiency, accuracy and individualization are the main goals to be solved. Evaluation of the severity of NAFLD involves a variety of clinical parameters, how to optimize non-invasive evaluation methods is a necessary issue that needs to be discussed in this field. Artificial intelligence (AI) has become increasingly widespread in healthcare applications, and it has been also brought many new insights into better analyzing chronic liver disease, including NAFLD. This paper reviewed AI related researches in NAFLD field published recently, summarized diagnostic models based on electronic health record and lab test, ultrasound and radio imaging, and liver histopathological data, described the application of therapeutic models in personalized lifestyle guidance and the development of drugs for NAFLD. In addition, we also analyzed present AI models in distinguishing healthy VS NAFLD/NASH, and fibrosis VS non-fibrosis in the evaluation of NAFLD progression. We hope to provide alternative directions for the future research.

Keywords Non-alcoholic fatty liver disease (NAFLD) · Non-alcoholic steatohepatitis (NASH) · Metabolic associated fatty liver disease (MAFLD) · Artificial intelligence (AI) · Machine learning (ML)

Abbreviations

AI	artificial intelligence	NASH	nonalcoholic steatohepatitis
ML	machine learning	ALT	alanine aminotransferase
DL	deep learning	AST	aspartate aminotransferase
NAFLD	non-alcoholic fatty liver disease	GGT	gamma-glutamyl transpeptidase
MAFLD	metabolic associated fatty liver disease	TG	triglycerides
NAFL	nonalcoholic fatty liver	WBC	white blood cell count
		PLT	platelet
		HCC	hepatocellular carcinoma
		BP	blood pressure
		SS	simple steatosis
		BN	bayesian network
		HNB	hidden naïve Bayes
		CNN	convolutional neural networks
		kNN	k-nearest neighbors
		GLCM	gray level cooccurrence matrix
		SVM	support vector machine
		FDM	fuzzy data-mining
		CPG	clinical practice guidelines
		AdaBoost	adaptive boosting
		RF	random forest
		GB	gradient boosting

✉ Jian Jiao
jjian@jlu.edu.cn

Yifang Li
1441479175@qq.com

Xuetao Wang
wxt6518@163.com

Jun Zhang
zjzdfsyfzh@163.com

Shanshan Zhang
zhangss20@mails.jlu.edu.cn

¹ Department of Gastroenterology & Hepatology, China-Japan Union Hospital, Jilin University, Changchun 130033, China

EHR	electronic health records
NLP	natural language processing
ALARM	automatic liver attenuation ROI-based measurement
AUROC	area under the receiver-operator characteristics
FXR	Farnesoid X Receptor
hFXR	human FXR
ASCVD	arteriosclerotic cardiovascular disease
SPPARMs	selective PPAR- γ modulators
SGLT-2is	sodium-glucose cotransporter 2 inhibitors
DPP-4is	dipeptidyl peptidase-4 inhibitors
HSI	hepatic steatosis index
BMI	body mass index
FLI	fatty liver index
WC	waist circumference
FPG	fasting plasma glucose
F(0-4)	Fibrosis stage (0-4)
NAS	non-alcoholic fatty liver disease activity score
HDL-C	high-density lipoprotein cholesterol
LDL-C	low-density lipoprotein cholesterol
MCDD	methionine choline deficient diet
HbA1c	glycated hemoglobin
HCD	high cholesterol and cholate diet
HFD	high-fat diet
aP	alkaline phosphatase
HOMA-IR	homeostasis model assessment-estimated insulin resistance
LR	logistic regression
DT	decision trees
AODE	aggregating one-dependence estimators
CDSS	clinical decision support system
ICD	international classification of diseases
EMR	electronic medical record
QUS	quantitative ultrasound
SWE	shear wave elastography
RFE	recursive feature elimination
HT	hypertension
ANN	artificial neural networks
CAD	computer-aided diagnosis
ROI	region of interest
CT	computed tomography
DCNN	deep convolutional neural network
DM	diabetes mellitus
SBVS	Structure-based Virtual Screening
MetS	Metabolic Syndrome
CVD	cardiovascular disease
GLP-1RA	glucagon-like peptide-1 receptor agonist
MI	myocardial infarction
XAI	Explainable Artificial Intelligence

1 Introduction

Non-alcoholic Fatty Liver Disease (NAFLD) has become one of the most important causes of chronic liver diseases in the world. Because the pathogenesis of NAFLD is closely related to obesity, hypertension, hyperlipidemia, hyperglycemia and the other metabolic factors, it has been proposed that NAFLD may be regarded as a hepatic manifestation of Metabolic Syndrome (MetS), and should be renamed G as Metabolic Associated Fatty Liver Disease (MAFLD) [1, 2]. It was reported that from 1999 to 2016, cardiovascular and renal risks and diseases have become highly prevalent in adults with MAFLD [3]. Patients with NASH or advanced fibrosis are at an increased risk of developing arteriosclerotic cardiovascular disease (ASCVD) compared with non-NAFLD controls or subjects with NAFL, independent of conventional metabolic risk factors for cardiovascular disease (CVD). Histological information on NAFLD may be helpful to promote our understanding of extrahepatic complications, such as ASCVD, resulting from NAFLD progression [4, 5]. The correlation between NAFLD and MetS showed that both NAFLD and NASH increased the incidence of cardiovascular events. Therefore, NASH should be considered in cardiovascular risk assessment.

The course of NAFLD includes Simple Steatosis (SS), Nonalcoholic Steatohepatitis (NASH), fatty liver cirrhosis and Hepatocellular Carcinoma (HCC) that may develop further [6]. Up to now, liver biopsy is still the “gold standard” for the diagnosis of NAFLD, but biopsy is invasive and may result in serious complications, such as pain, infection, and bleeding, and it cannot be carried out in all suspected patients. An ultrasound examination is noninvasive, convenient, and widely used in clinical screening of NAFLD, but the diagnosis results are affected by too many factors, especially the subjectivity of the examiner.

Genetic factors, dyslipidaemia and alteration of glucose metabolism have been proven to be associated with NAFLD [7, 8]. Some studies identify sets of lipids as predictive biomarkers of NAFLD progression by using machine learning (ML) [9]. A variety of comprehensive evaluation indexes have been used in the assessment of NAFLD. For example, Hepatic Steatosis Index (HSI) [10] is calculated according to alanine aminotransferase (ALT), aspartate aminotransferase (AST), Body Mass Index (BMI), gender and diabetes history. Fatty Liver Index (FLI) [11] based on triglyceride (TG), BMI, gamma-glutamyl transpeptidase (GGT) and Waist Circumference (WC). Chinese ZJU index [12] calculated by BMI, Fasting Plasma Glucose (FPG), TG, ALT and AST, etc. Broadly speaking, these indexes are models belong to the early

elementary artificial intelligence. Recently, some studies have included biological markers such as adiponectin and Caspase-cleaved cytokeratin 18 fragment (M30) into evaluation models to predict NAFLD activity score (NAS). The model was not only able to monitor disease progression and weight change but also can distinguish NASH from NAFLD with an area under the receiver-operator characteristics (AUROC) of 0.70-0.73 [13]. However, different algorithms and data sources from particular populations certainly affected the applicability and reliability of the model. Any simple formula or limited index is limited in evaluating the disease comprehensively, the predicting parameters and optimal algorithms to predict NAFLD are still heterogenous and varied across different studies [14]. The development of “big data” network and the application of AI algorithms provide new methods and possibilities for the better understanding of NAFLD. The advantage of AI over traditional statistical modelling is that it can recognize unique patterns and incorporate multiple factors to create predictive models, risk stratification [15] and outcomes, it is particularly suited for use in chronic diseases given the heterogeneity, complex nature, and overlapping confounding factors.

2 Overview of AI and ML in NAFLD research

AI is poised to influence nearly every aspect of the human conditions, and they are the collection of data-analytical techniques, aiming at building predictive models from multi-dimensional data sets. AI is a general term in which computer programs could think and behave like humans. After the introduction of ML, AI is the most searched term. Some describe ML as the primary AI application, while others describe it as a subset of AI. ML generates a mathematical algorithm from the training data set and uses it to predict results or make decisions. Later, with the development of neural networks, machines could classify and organize the input data like the human brain. The term “deep learning (DL)” is proposed based on multi-layer neural networks that can be applied to large datasets, which is suitable for the detection, classification and segmentation of biomedical images. DL is a subset of ML and ML is a subset of AI, the order of evolution of AI-ML-DL is shown in Fig. 1 [16, 17].

According to the model of the training methods, ML can be divided into supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning (Fig. 2). Supervised learning is suitable for data with labels and it deals with annotated data with input–output pairs and common techniques including linear regression, logistic regression (LR), decision trees (DT), k-nearest neighbors (kNN), support vector machine (SVM), random forest (RF), Naive Bayes classification, and gradient

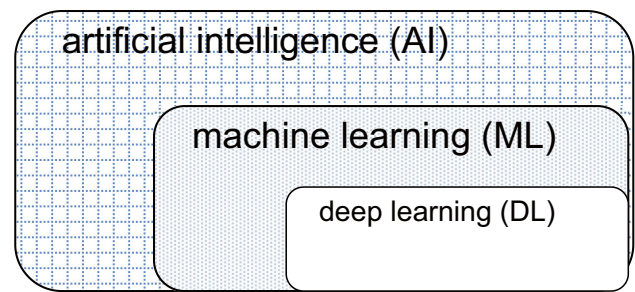


Fig. 1 Evolution of artificial intelligence (AI)-machine learning (ML)-deep learning (DL)

boosting (GB) [18, 19]. Unsupervised learning works with data without a label and needs to classify the data according to its own structural characteristics. Semi-supervised learning is an amalgamation of supervised and unsupervised ML that can analyse a large amount of unclassified data whilst augmenting its pattern recognition abilities with a small amount of preclassified data. Moreover, it can increase the speed and accuracy of information extraction from large data sets [12]. What makes reinforcement learning different from other algorithms is that there is no supervisor and only one reward signal, it can learn interactive situations with the purpose of maximizing the reward signal. At present, supervised learning [20–29] and semi-supervised learning [30] are the most widely applied in medicine, this is also determined by the source and characteristics of medical data.

According to the purpose of modeling, the models can be divided into diagnostic models, therapeutic models and prognostic models. The diagnostic models, also can be called as assessment models. These assessment models mainly complete the diagnosis tasks on the basis of large data set, and compare with the evidence-based medicine data, continuously optimize their functions according to the degree of task completion, so as to achieve more accurate diagnosis effects, and they have important values for non-invasive assessment of NAFLD. It has been also increasingly applied in the detection and prediction of NAFLD and fibrosis outcomes in recent years.

However, the choice of algorithms may affect the final judgment. A study evaluated different algorithms including LR, ridge regression, AdaBoost and DT models in detecting NAFLD from the general population based on 23 routine clinical and laboratory parameters. The result showed that Ridge score was the best performing algorithm with AUROC of 0.87 (95% CI 0.83-0.90) and 0.88 (0.84-0.91) in the training and validation groups respectively [31]. With the use of large-scale AI training cohorts for model development, the prediction on validation cohorts is probably more accurate compared with traditional biostatistical methods [22].

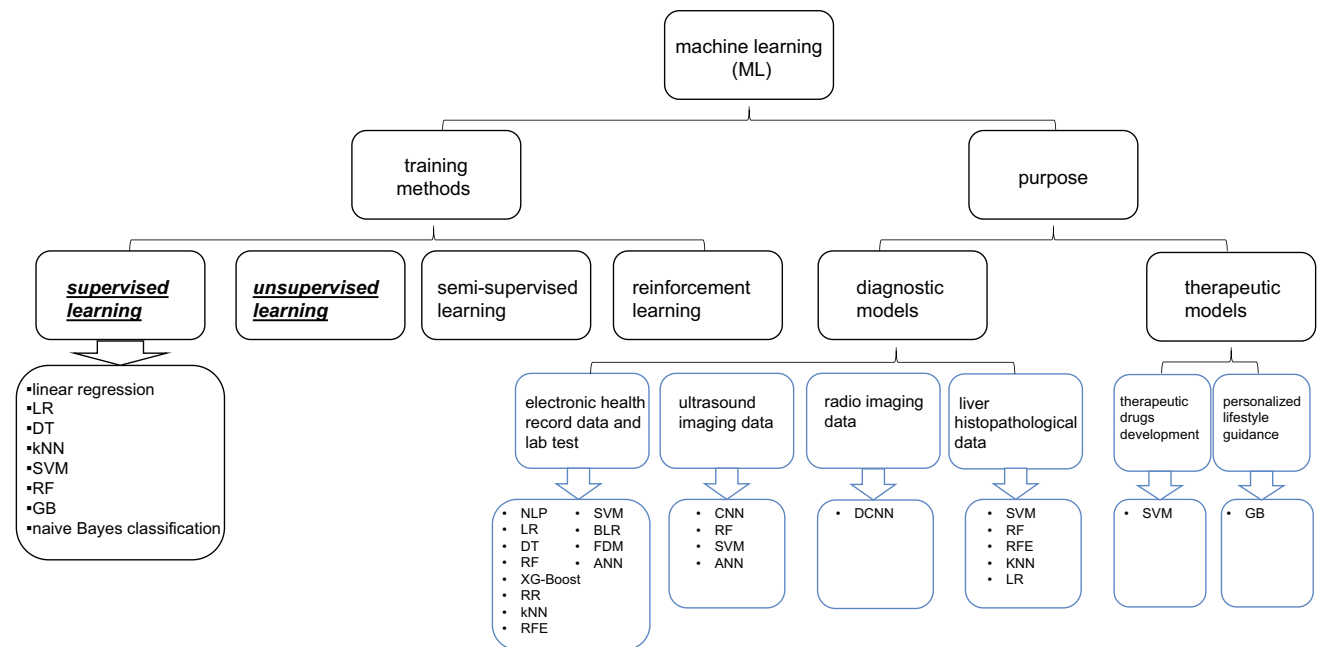


Fig. 2 Classification and common algorithms of ML

The therapeutic model focuses on new drug development, personalized lifestyle guidance and follow-up of therapeutic effects. It helps to develop personalized diets, to accelerate new drug researches and to improve the therapeutic schedules. At present, clinical medicine, basic biomedical research, and new health-related data sources (such as wearable devices with biosensors) generate a large amount of biomedical and health-related data, these data go beyond the limits of human analysis and need machines to help. Therefore, AI is providing NAFLD patients with a paradigm shift to precision medical practice.

In addition, the development of omics techniques in recent years may be used to further investigate the pathophysiological mechanisms of NAFLD. genomics, epigenomics, transcriptomics, metabolomics/lipidomics and glycomics in relation to the pathophysiology provided invaluable potential for the diagnosis and treatment of NAFLD. AI is a data-driven and hypothesis-free approach, which better incorporates clinical factors to detect hidden patterns for disease detection/prediction. The accurate analysis of large data sets from omics is an important advantage of artificial intelligence in NAFLD field.

3 AI application in NAFLD diagnosis

According to the data sources analyzed by AI, studies applied to NAFLD diagnosis can be classified as follows: electronic health record (EHR) data and lab test based,

ultrasound imaging data based, radio imaging data based and liver histopathological data based.

3.1 Diagnostic models based on electronic health record data and lab test

A common goal of AI algorithm is to develop a predictive model based on statistical associations among features from a given data set. With the development of medical diagnosis technology, the EHR system with intelligent diagnostic function has become one of the important scientific issues of science for medical information. The EHR system contains structured data (such as diagnostic codes) and unstructured data (clinical documents including lab test), by analyzing these data with AI algorithms, the patients with NAFLD/NASH can be screened and the risks of disease progression also can be derived. Presently, AI algorithms used in EHR analysis include Natural Language Processing (NLP)-based approaches [60], text search-based approaches and International Classification of Diseases (ICD)-based approaches [63].

NAFLD progresses insidiously, and the presence of fibrosis is hard to spot, high-throughput identification of NAFLD with “electronic” follow-up through the EHR, could aid in understanding the risk factors for progression to cirrhosis. Van Vleck et al. [32] assessed above approaches to identify NAFLD within the EHR data compared to manual validation by clinicians in a large, multiethnic cohort. Results suggested that NLP approaches had the best overall performance

compared to ICD and text search-based approaches, although there were numerous patients identified by ICD that were missed by NLP.

While with medical practice becomes more specialized and patient care is provided by more physicians, the opportunities for information loss at patient handoffs increase, and the analysis of progression to NASH was complicated by the fact that due to the complexity of diagnosing NASH, a firm diagnosis is often not made, the application of NLP still needs to be further improved. Koleck et al. [33] reviewed the literatures on the use of NLP to process or analyze symptom information documented in EHR free-text narratives. They found that NLP tools, classification methods, and manually curated rule-based processing are being used to extract information from EHR free-text narratives written by a variety of healthcare providers on a wide range of symptoms across diverse clinical specialties. Clear statement of the symptoms being evaluated as part of the study, a detailed description of the clinical population from which symptom information was extracted and analyzed, open sharing of user-developed symptom-related NLP algorithms or pipelines and vocabularies, and the establishment of formal reporting standards for investigations using NLP methodologies would be of help.

3.2 Assessment models based on ultrasound imaging data

Ultrasonography is a well-established and cost-effective imaging technique for the diagnosis of hepatic steatosis, especially for screening a large population at the risk of NAFLD, while it has shortcomings, including lower accuracy in mild steatosis confirmation, operator-dependent, and rather qualitative. Quantitative Ultrasound (QUS) imaging methods, including elastography, echogenicity analysis, and speckle statistical modeling can be obtained from single ultrasound radio-frequency data acquisition. Because these ultrasound imaging methods provide complementary quantitative tissue information, the characteristics of the liver can be obtained from their combination. The ML model is the main method to realize this function, up to now, ML models derived from ultrasound examination have been used in cardiovascular, nervous system diseases, and chronic hepatitis B [34], but there are very few reports on the application of ultrasound-based ML models in diagnosis of NAFLD. Tang et al. [35] established a ML model based on QUS parameters by using histopathology scoring as the reference standard to improve the classification of steatohepatitis with shear wave elastography (SWE) in rats. Their results proven that QUS parameters improved the classification accuracy of steatohepatitis, liver steatosis, inflammation, and fibrosis compared with the shear wave elastography alone. This attempt

provided a basis for relevant researches in human beings. Wu et al. developed a RF model to predict ultrasonographic fatty liver disease by using age, gender, systolic and diastolic blood pressure, abdominal girth, glucose, TG, HDL, AST, and ALT, which outperformed the Naïve Bayes, artificial neural networks (ANN), and LR model with an AUROC of 0.93 [36]. It is worth noting that the ultrasound data have racial differences, although many ultrasound images have been accumulated and are now available and ready for the preparation of a database for the development of computer-aided ultrasound diagnosis with DL technology [37]. The establishment of the model should also consider individual factors, such as races and geographic regions.

3.3 Assessment model based on radio imaging data

The use of ML has been increasing rapidly in the medical imaging field, including Computer-Aided Diagnosis (CAD), radiomics, and medical image analysis. The combination of imaging and artificial intelligence improves the accuracy of liver fibrosis staging. The deep learning radiology of shear wave elastography significantly improved the diagnostic performance of evaluating liver fibrosis [38]. In addition, as a tool of radiology, derivative parameters of image-based texture analysis combined with ML of non-contrast-enhanced T1-weighted magnetic resonance images can be as accurate as magnetic resonance elastography in the quantification of liver fibrosis (82%) [39]. Manual tracking of the Region of Interest (ROI) of the liver is a standard method to measure liver attenuation on Computed Tomography (CT) to diagnose NAFLD. However, manual tracking is resource intensive. To address these limitations and expand the effectiveness of quantitative CT measurement of hepatic steatosis, Huo et al. [40] proposed an Automatic Liver Attenuation ROI-based Measurement (ALARM) method for automated liver attenuation estimation. It consists of two major stages: (a) deep convolutional neural network (DCNN)-based liver segmentation and (b) automated ROI extraction. The combination of DCNN and morphological operation can achieve "excellent" consistency with manual estimation of fatty liver detection. The whole pipeline is implemented as a docker container, which enables users to complete liver attenuation assessments within 5 minutes of each CT scan. Graffy et al. [41] developed a deep learning-based automated liver fat quantification tool with non-enhanced CT for determining the prevalence of steatosis in a large screening cohort. By using three-dimensional convolutional neural networks, including a sub cohort with follow-up scans, the automated volume-based liver attenuation, including conversion to CT fat fraction was analyzed and compared with manual measurement in a large subset of scans. The results showed that this fully automated CT-based liver fat quantification tool could

be used for population-based assessment of liver steatosis and NAFLD, and the objective data matched well with the manual measurement results. By utilizing NLP, many studies developed algorithms which were capable of "reading" full-text radiology reports to accurately identify the presence of fatty liver disease [42]. Abdominal ultrasound, computerized tomography, and magnetic resonance imaging reports retrieved from random samples were analyzed by physicians and expert radiologists, the radiographic fatty liver disease could be determined by manual review. These algorithms could be used to rapidly screen patient records to establish a large cohort to facilitate epidemiological and clinical studies.

It can be seen from the above that, supervised learning played an important role in collecting and analyzing radio imaging data, therefore, radiologists have to face with the difficult tasks of annotating these large data sets. In addition, data sources and governance policies need to be developed to address the concerns of institutional review boards, as well as broader ethical issues around the use of large patient data sets.

3.4 Assessment models based on liver histopathological data

AI software may identify histological features of NAFLD with high levels of interobserver and intraobserver agreement (0.95 to 0.99) [43]. An automatic assessment of histological characteristics of NAFLD may reduce human variability and provide continuous, rather than semi-quantitative assessment of liver damages. However, the limited availability of experienced liver pathologists, the variability in pathologists' agreement on detecting and quantifying various histological features of liver diseases, and the limited use of semi-quantitative manual grading scores hindered the application of AI in histopathological evaluation.

Macro-steatosis is the cardinal lesion of NAFLD, and it is commonly used as a major endpoint in therapeutic clinical trials in human NAFLD [32, 34]. The pathological assessment of NAFLD consists of four key features: steatosis, lobular inflammation, fibrosis, and hepatocyte ballooning. Several studies aimed to automatically quantify the NAFLD score by ML algorithms [45]. Some aim to predict the severity of liver fibrosis [46]. In view of this, many attempts have been made in the automatic histopathological classification of fatty liver in rodents and humans. Deepak Sethunath et al. [33] created classifiers to detect macro- and micro-steatosis by using experts' annotations, supervised ML and image processing techniques. For macro-steatosis prediction, the model's precision, sensitivity, and AUROC were 94.2%, 95%, 99.1% respectively. By using similar approaches, Scott Vanderbeck et al. [20, 47] developed an automated classifier, including naïve Bayes, LR, DT, and neural networks, which

can detect and quantify steatosis of the liver in humans. The classification algorithm performs with 89 % overall accuracy. It identified macro-steatosis, bile ducts, portal veins and sinusoids with high precision and recall ($\geq 82\%$). The above tentative work demonstrated that the automatic quantification of cardinal NAFLD histologic lesions is feasible and offers promise for further development of automatic quantification as a potential aid to pathologists to evaluate NAFLD biopsies in clinical practices and clinical trials.

3.5 Assessment models on evaluation of NAFLD, NASH and fibrosis

From what has been discussed above, the diagnosis of NAFLD focused on distinguishing healthy VS NAFLD/NASH [20, 22–26, 31, 35, 38, 48–51], NAFL VS NASH [32, 52, 53] and fibrosis VS non-fibrosis [21, 27, 54]. We summarized literatures according to "what is being predicted" and listed them in detail. (See Table 1)

In fact, it is not easy to accurately distinguish between NAFLD and healthy by using regular imaging and laboratory examination, the presence of NAFLD may change dynamically in one case. Those suspected patients with NAFLD on imaging are often not acknowledged in subsequent clinical documentation. Many of such patients are later found to have more advanced liver diseases. SVM, LR and DT algorithms are mainly used in ML research for NAFLD diagnosis, most of them were retrospective studies, and no biopsy was included. The parameters involved were mainly derived from EMR data, morphology features and lab tests. The specificity ranges from 60% to 92%. Ridge score was reported to be more advantageous.

NASH indicates the progression and deterioration of NAFLD, which can only be confirmed by biopsy, so it is of great value to explore the role of non-invasive model of ML in the assessment of NASH. This kind of research included mice and human studies, almost always use biopsy as the standard, the main purpose is to investigate the predictive accuracy of NASH, NAFL or health status and quantitatively interpret biopsy specimens of patients with NAFLD. The highest specificity was 99% [51]. It should be mentioned that NLP-based approaches facilitated analyses of knowledge flow among physicians and enabled the identification of breakdowns where the key information was lost that could have slowed and prevented the progression of early NAFLD to NASH or cirrhosis. Koleck et al. [33] synthesized literatures on the use of NLP in order to process or analyze symptom information documented in EHR free-text narratives. They suggested that future NLP studies should focus on the investigation of symptoms, and symptom documentation in EHR free-text narratives should strive to examine the patients' characteristics and make symptom-related NLP algorithms or pipelines and vocabularies openly available.

Table 1 ML models in NAFLD diagnosis on evaluation of NAFLD,NASH and fibrosis

Categories	Contribution	Author	Country	Years	AI algorithms	Parameters in the model	Biopsy (yes/no)	Study cohort	subjects	Sensitivity	Specificity
NAFLD healthy VS NAFLD	Identify the four histological features of NAFLD in liver biopsy images	Vanderbeck et al. [20]	United states	2014	SVM	morphological features, Texture, GLCM Statistics, N_jet, Nuclei Density	yes	Retrospective	Patients with normal liver histology and NAFLD	/	/
	Identify NAFLD in EMR	Corey et al. [38]	United states	2015	LR	number of natural language mentions of NAFLD in the EMR, lifetime number of ICD-9 codes for NAFLD and triglyceride level	yes	Longitudinal cohorts	Patients was randomly selected from the Research Patient Data Registry at Partners Healthcare	60% (RF)	77% (RF)
	Ridge score is the most predictive ML model for detection of NAFLD	Yip et al. [31]	China	2017	RR, AdaBoost, DT	ALT, HDL-C, TG, HbA1C, WBC, HT	no	Cross-sectional study	Populations from the Chinese government census database	training group: LR:90% RR:92% AdaBoost:91% DT:95%	training group: LR:90% RR:90% AdaBoost:90% DT:92%
	Using MetS risk factors to investigate the risk of NAFLD in Canadian population	Perveen et al. [22]	Pakistan	2018	DT	systolic blood pressure, diastolic blood pressure, HDL, TG, BMI, FG, age, sex	no	Retrospective	Canadian population	/	/
healthy VS mild/moderate/severe NAFLD	Evaluate the optimal predictive clinical model of NAFLD	Ma et al. [49]	China	2018	KNN, SVM, LR, BN, DT, AdaBoost, bootstrap aggregating, RF, HNB, AODE	weight, TG, ALT, GGT, and serum uric acid	no	Cross-sectional study	Participants in physical examination	67.5%	87.8%
	Objective to verify the value of DL in the diagnosis of NAFLD	Cao et al. [24]	China	2020	CNN	Ultrasound	no	Prospective	Adults undergoing routine abdominal examination	/	/

Table 1 (continued)

Categories	Contribution	Author	Country	Years	AI algorithms	Parameters in the model	Biopsy (yes/no)	Study cohort	subjects	Sensitivity	Specificity
NASH Non-NASH VS NASH	Improve the classification of steatohepatitis over SWE in rats and identify QUS provide the highest classification accuracy for steatosis, inflammation and fibrosis	Tang et al. [35]	Canada	2018	RF	QUS parameters	yes	/	11-week-old male Sprague-Dawley rats fed with MCDD	/	/
	The accuracy of NASH diagnosis by CDSS is high	Donali et al. [48]	France	2013	LR and FDM	scientific literature and CPG	no	Prospective	Morbidly obese subjects	/	/
	Predicting the actual health status of patients with benign fatty liver disease	Fialoke et al. [23]	United states	2018	LR,DT,RF,XG-Boost	age, gender,race,ALT, AST, PLT,Type 2 diabetes status	no	/	Adult patients from Optum's Integrated Claims-Clinical dataset	92%	/
NAFLD VS NASH	The accuracy of NAFLD recognition based on NLP is better than that based on ICD /text search in EHR	Van Vleck et al. [32]	United states	2019	NLP	EHR systems	yes	Retrospective	Ethnically diverse patients recruited from primary care and subspecialty clinics within the Mount Sinai Health System	93%	89%
healthy VS NAFL VS NASH	Objective to describe the characteristics of specific lipids in liver and serum in early NASH	Chiappini et al. [52]	France	2016	Microarray Data Analysis, RF	Histology	yes	/	Mouse models that developed NAFL when HFD or NASH when fed using MCDD	/	/
	To establish a non-invasive model for the diagnosis of liver fibrosis	Perakakis et al. [25]	United states	2019	RFE,SVM,kNN,RF	lipids, hormones and glycans	yes	Prospective	Healthy people and patients with NAFLD	/	/

Table 1 (continued)

Categories	Contribution	Author	Country	Years	AI algorithms	Parameters in the model	Biopsy (yes/no)	Study cohort	subjects	Sensitivity	Specificity
	NASH can progress to HCC without cirrhosis	Hwang et al. [28]	United states	2018	Supervised learning	osteopontin alpha-fetoprotein Dickkopf-1	yes	/	Adult male C57BL/6 mice	82%	86%
	Evaluated the predictive accuracy of NASH, NAFL or healthy status	Katsiki et al. [51]	Greece	2019	RFE,SVM,kNN,RF	lipids, glycans and hormones	yes	Retrospective	Healthy individuals and patients with biopsy-proven with NAFL	95%	99%
	Quantification of steatosis, inflammation, ballooning and fibrosis	Forlano et al. [30]	England	2019	LR	age,gender,BMI,ethnic,type 2 DM,arterial,hyper tension,dyslipidemia ,PLT,ALT,AST,total cholesterol,TG,HDL-C, LDL-C,HbA1c, Ferritin	yes	Retrospective	Patients with biopsy-proven NAFLD	80%	62%
Fibrosis											
NAFLD F1 VS F2	Differentiation of stage 1 and stage 2 fibrosis in a cohort of NAFLD patients	Sowa et al. [21]	Germany	2013	LR,kNN,SVM,DT,RF	ALT, AST, M30, M65,HA	yes	Retrospective	Patients undergoing bariatric surgery for morbid obesity	RF:>60% kNN:30.8% DT:53.9%	RF:77% kNN:91.3% DT:94.2%
F0 VS F1 VS F2 VS F3 VS F4	Detection and quantification of liver fibrosis	Gawrieh et al. [27]	India	2020	SVM	Histology	yes	/	A set of digitalimages of the trichrome stained slides of 18 unique liver biopsy	/	/
Others											
healthy VS liver injury	Identify risk factors for elevated ALT,NAFLD may be the main cause of liver injury	Qu et al. [39]	United states	2012	SVM,BLR	gender, age,rs738409 genotype,BMI, body height, WC, hip circumference, waist/hip ratio, pulse rate, BP, physical activity, alcohol consumption, smoking, education levels, status of diabetes, history of hepatitis, medications,PPG, fasting plasma insulin level, HOMA-IR,TG, HDL-C, LDL-C	no	Prospective	Populations with high rates of obesity and diabetes	/	/

Table 1 (continued)

Categories	Contribution	Author	Country	Years	AI algorithms	Parameters in the model	Biopsy (yes/no)	Study cohort	subjects	Sensitivity	Specificity
DM VS DM with CVD	A prospective correlation between baseline ALT level and incidence of CVD in patients with type 2 diabetes mellitus	Corey et al. [38]	Iran	2016	ANN	age, gender, BMI, WC, lipid-lowering and antihyperglycemic medication use, family history of CVD, smoking status, HT, HbA1C, ALT, AST, aP, GGT, LDL-C, HDL-C, TG	no	Prospective	Patients with type 2 diabetes and community-dwellers without diabetes	/	/
steatosis VS inflammation VS ballooning VS fibrosis	To determine the prevalence of steatosis	Graffy et al. [41]	United states	2019	three-dimensional CNN	CT, age, gender, BMI	no	Retrospective	Patients who underwent initial nonenhanced abdominal CT at a single academic medical center for the purpose of colonography screening	normal82.6% mild steatosis88.7% moderate steatosis87.5% severe steatosis:74.6%	normal96.3% mild steatosis83.5% moderate steatosis98.3% Severe steatosis99.9%
HCC-free VS HCC	Relationship between metabolomics and NAFLD	Qu et al. [39]	Germany	2017	LR	branched-chain amino acids ,derived gamma-glutamyl dipeptide,BMI,WC,age, sex,alcohol consumption,diabetes status	no	Retrospective	Population-based sample from Northern Germany	/	/
AI VS pathologist	To compare the quantitative results of AI application of fatty acid accumulation with the semi quantitative evaluation results by committee certified toxicology pathologists using the full microscopic method	Ramot et al. [29]	Israel	2020	CNN	Quantification of fatty vacuoles	yes	/	1. Male C57BL/6 J mice with HFD acquired from Harlan Laboratories 2. Male C57BL/6 J mice with HCD acquired from Harlan Laboratories	/	/

There are not many studies on fibrosis. Assessment models on evaluation of NAFLD, NASH and fibrosis were listed in the Table 1. Model studies related to NAFLD and liver injury, Diabetes Mellitus (DM), CVD and hepatocellular carcinoma (HCC) are also listed, maybe future research should focus more on these aspects.

4 AI application in NAFLD treatment

4.1 Therapeutic models for personalized lifestyle guidance

It has been suggested that NAFLD should be listed as one of the complications of DM since NAFLD has been renamed as MAFLD, and more attention has been paid to the therapies related to diet and glycemic control in the patients with MAFLD. However, the composition of healthy diet and personalized glycemic control have not been clearly established yet.

A new approach called "nutritional geometry" considers the importance of integrating nutrition, animals, and the environment. The goal of this approach is to combine nutrients and foods in a model to understand how food ingredients interact to regulate dietary characteristics that affect health and diseases. The AI algorithms may create a personalized diet for patients, which can provide personalized nutrition consultation to prevent and treat NAFLD [55]. Zeevi et al. [56] measured the postprandial glucose responses to 46,898 foods in 800 individuals and found a great variability in responses to the same foods. Based on these data, they designed a ML algorithm-GB regression, a data-driven unbiased approach, which integrates blood parameters, dietary habits, anthropometrics, physical activity, and gut microbiota to predict personalized postprandial glycemic response in real-life meals. Then, they validated these predictions in an independent 100-person cohort. The results showed that the blinded randomized controlled diet intervention based on this algorithm resulted in a significant decrease in postprandial response and continuous alterations in intestinal flora configuration, which meant that personalized diet could successfully improve postprandial hyperglycemia and its metabolic consequences.

There are also ideas that collect dietary and microbiome data from many individuals [57], derive AI models of how diet affects the composition of the microbiome, and then validate the models with controlled dietary interventions.

The contribution of the above-mentioned AI algorithms in personalized diet guidance is helpful to control the development of NAFLD. A few studies are using smart phones to collect clinical images to assist in the diagnosis of hepatic steatosis [58]. In the future, patients may wear devices to record what they eat, and then the information will be

processed through DL and integrated by AI, combined with a variety of data (physical activity, stress level, sleep, microbiome, physiological constants, medications and genome) to provide individualized dietary recommendation and nutritional counseling to prevent and treat NAFLD.

4.2 AI models related to the development of therapeutic drugs for NAFLD

One way to reduce costs is to use genetic data to inform drug design. Genetic information helps researchers to demonstrate that drug targets are relevant to the disease from the start, and drugs with this evidence are twice as likely to be approved as those without [59]. By using AI techniques, we can further optimize drug discovery. If we start with a 'deep' molecular profile that includes data about the microbiome and genome, as well as information about metabolites and proteins (the metabolome and proteome), coupled with physiological measurements, we may be able, in some cases, to skip animal testing and move straight to human trials [57]. Various ML systems and AI platforms had been used to search for immuno-oncology drugs and metabolic-disease therapies. It is predicted that AI and ML will create an era of faster, cheaper, and more effective drug discovery.

While, there is no effective drug to control the progression of MAFLD in the clinic up to now. Farnesoid X Receptor (FXR) agonists can reverse deregulated bile acid metabolism, and thus, they are potential therapeutics to prevent and treat NAFLD [44]. Unfortunately, Structure-based Virtual Screening (SBVS) that can speed up drug discovery has rarely been reported with success for FXR, which was likely hindered by the failure in addressing protein flexibility. Xia et al. [61] devised human FXR (hFXR)-specific ensemble learning models based on pose filters from 24 agonist-bound hFXR crystal structures and coupled them to traditional SBVS approaches of the FRED docking plus Chemgauss 4 scoring function, as a result, they provided a promising lead compound for further development.

Newer antidiabetic drugs, such as selective PPAR- γ modulators (SPPARMs), glucagon-like peptide-1 receptor agonist (GLP-1RA) and sodium-glucose cotransporter 2 inhibitors (SGLT-2is) were recommended in T2DM, might contribute substantially to NAFLD/NASH amelioration, possibly reducing not only liver-specific but also cardiovascular morbidity. Tomohide Yamada et al. compared the risk of myocardial infarction (MI) among SGLT-2is, GLP-1RAs and dipeptidyl peptidase-4 inhibitors (DPP-4is) and developed a machine learning model for predicting MI in patients without prior heart disease. After developing a machine learning model to predict MI, proportional hazards analysis of MI incidence was conducted using the risk obtained with this model and the drug classes as explanatory variables.

Receiver operating characteristics analysis showed higher precision of machine learning over logistic regression analysis. Finally, the machine learning analysis suggested the risk of MI was 37% lower in type 2 diabetes patients without prior MI using GLP-1RAs versus DPP-4is, while the risk was 19% lower for SGLT-2is versus DPP-4is. While there have been no studies using machine learning methods to evaluate the efficacy of these drugs in NAFLD to date [62].

5 Summary

Diagnosis and treatment of NAFLD face many challenges although the medical science has been very developed. Limitations of AI technology include the lack of high-quality data sets for ML development. Most of the evidence used to develop ML algorithm comes from preclinical research. Efficiency, accuracy, and individualization are the main goals to be solved.

In addition, the ambiguity of AI made it problematic for machine learning systems to be adopted in a sensitive yet critical domains, such as healthcare. As a result, scientific interest in the field of Explainable Artificial Intelligence (XAI), a field that is concerned with the development of new methods that explain and interpret machine learning models, has been tremendously reignited over recent years [64]. However, AI provides a new way for disease understanding by extracting the characters of complex data and combining them with the mode of “automatic learning”, which will contribute to an increase in diagnostic quality, facilitate the development of remote medicine, and reduce the costs in the national health care. From this point of view, it may have a large enough potential to induce a paradigm shift in the handling of NAFLD. Certainly, ML itself is far from fulfilling its potential in NAFLD researches, and we have a long way to go to uncover the networked intricacies and complexities of living systems. Accumulation of subjective and objective data and long-term follow-up verification are still the most basic, individual factors, which should be considered in the application of AI model, XAI for NAFLD studies is also need to be explored.

Acknowledgments We thank Dr. Steven Limbara for revising the language of the manuscript.

Author contributions All authors contributed to the study conception and design, and All authors read and approved the final manuscript. Yifang Li _Organized information and wrote the article. Xuetao Wang _Collected and organized information. Jun Zhang _Collected and organized information. Shanshan Zhang _Collected and organized information. Jian Jiao _Supervisor, gave the idea and revised the article

Funding No funding was received to assist with the preparation of this manuscript.

Declarations

Ethics approval The authors have no ethical conflicts to disclose.

Consent to participate Not applicable.

Consent for publication Not applicable

Conflict of interest The authors declare that they have no competing interests.

References

1. Eslam M, Sanyal AJ, George J. MAFLD: A Consensus-Driven Proposed Nomenclature for Metabolic Associated Fatty Liver Disease. *Gastroenterology* 2020;158(7):1999–2014.e1991 [PMID: 32044314. <https://doi.org/10.1053/j.gastro.2019.11.312>].
2. Fouad Y, Waked I, Bollipo S, Gomaa A, Ajlouni Y, Attia D. What's in a name? Renaming 'NAFLD' to 'MAFLD'. *Liver Int.* 2020 [PMID: 32301554. <https://doi.org/10.1111/liv.14478>].
3. Zhang HJ, Wang YY, Chen C, Lu YL, Wang NJ. Cardiovascular and renal burdens of metabolic associated fatty liver disease from serial US national surveys, 1999–2016. *Chin Med J (Engl)*. 2021 [PMID: 34091530. <https://doi.org/10.1097/CM9.0000000000001513>].
4. Sanyal AJ, Brunt EM, Kleiner DE, et al. Endpoints and clinical trial design for nonalcoholic steatohepatitis. *Hepatology* 2011;54(1):344–353 [PMID: 21520200 PMCID: PMC4014460. <https://doi.org/10.1002/hep.24376>].
5. Park JH, Koo BK, Kim W, Kim WH; Innovative Target Exploration of NAFLD (ITEN) Consortium. Histological severity of nonalcoholic fatty liver disease is associated with 10-year risk for atherosclerotic cardiovascular disease. *Hepatology Int.* 2021 [PMID: 34081289. <https://doi.org/10.1007/s12072-021-10209-3>].
6. Estes C, Anstee QM, Arias-Loste MT, et al. 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(4):896–904 [PMID: 29886156. <https://doi.org/10.1016/j.jhep.2018.05.036>].
7. Zhu J, Xu D, Yang R, Liu M, Liu Y. The triglyceride glucose index (TyG) and CDKAL1 gene rs10946398 SNP are associated with NAFLD in Chinese adults. *Minerva Endocrinol.* 2020 [PMID: 33269568. <https://doi.org/10.23736/S0391-1977.20.03273-3>].
8. Wang S, Sun W, Cheng Y, Wang L, Ma S, Jing F, Zhang X, Zhou X. Relationship between plasma 12,13-diHOME level and nonalcoholic fatty liver disease in patients with type 2 diabetes and obesity. *Minerva Endocrinol (Torino)*. 2021 [PMID: 33855386. <https://doi.org/10.23736/S2724-6507.21.03424-6>].
9. Castañé H, Baiges-Gaya G, Hernández-Aguilera A, Rodríguez-Tomás E, Fernández-Arroyo S, Herrero P, Delpino-Rius A, Canela N, Menendez JA, Camps J, Joven J. Coupling Machine Learning and Lipidomics as a Tool to Investigate Metabolic Dysfunction-Associated Fatty Liver Disease. A General Overview. *Biomolecules*. 2021;11(3):473 [PMID: 33810079; PMCID: PMC8004861. <https://doi.org/10.3390/biom11030473>].
10. Deo RC. Machine Learning in Medicine. *Circulation* 2015;132(20):1920–1930 [PMID: 26572668 PMCID: PMC5831252. <https://doi.org/10.1161/circulationaha.115.001593>].
11. Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine Learning and Data Mining Methods in Diabetes Research. *Comput Struct Biotechnol J.* 2017;15:104–116 [PMID: 28138367 PMCID: PMC5257026. <https://doi.org/10.1016/j.csbj.2016.12.005>].

12. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med.* 2018;284(6):603–619 [PMID: 30102808 DOI: <https://doi.org/10.1111/joim.12822>].
13. Canbay A, Kälsch J, Neumann U, Rau M, Hohenester S, Baba HA, Rust C, Geier A, Heider D, Sowa JP. Non-invasive assessment of NAFLD as systemic disease-A machine learning perspective. *PLoS One.* 2019;14(3):e0214436 [PMID: 30913263; PMCID: PMC6435145. <https://doi.org/10.1371/journal.pone.0214436>].
14. Wong GL, Yuen PC, Ma AJ, Chan AW, Leung HH, Wong VW. Artificial intelligence in prediction of non-alcoholic fatty liver disease and fibrosis. *J Gastroenterol Hepatol.* 2021;36(3):543–550. [PMID: 33709607. <https://doi.org/10.1111/jgh.15385>].
15. Dinani AM, Kowdley KV, Noureddin M. Application of Artificial Intelligence for Diagnosis and Risk Stratification in NAFLD and NASH- the State of the Art. *Hepatology.* 2021 [PMID: 33928671. <https://doi.org/10.1002/hep.31869>].
16. Badillo S, Banfai B, Birzele F, Davydov II, Hutchinson L, Kam-Thong T, Siebourg-Polster J, Steiert B, Zhang JD. An Introduction to Machine Learning. *Clin Pharmacol Ther.* 2020;107(4):871–885 [PMID: 32128792; PMCID: PMC7189875. <https://doi.org/10.1002/cpt.1796>].
17. Rizwan I, Haque I, Neubert J. Deep learning approaches to biomedical image segmentation. *Inf Med Unlocked.* 2020;18:100297.
18. Su TH, Wu CH, Kao JH. Artificial intelligence in precision medicine in hepatology. *J Gastroenterol Hepatol.* 2021;36(3):569–580 [PMID: 33709606. <https://doi.org/10.1111/jgh.15415>].
19. Ahn JC, Connell A, Simonetto DA, Hughes C, Shah VH. The application of artificial intelligence for the diagnosis and treatment of liver diseases. *Hepatology.* 2020 [PMID: 33098140.DOI: <https://doi.org/10.1002/hep.31603>].
20. Vanderbeck S, Bockhorst J, Komorowski R, Kleiner DE, Gawrieh S. Automatic classification of white regions in liver biopsies by supervised machine learning. *Hum Pathol.* 2014;45(4):785–792 [PMID: 24565203 <https://doi.org/10.1016/j.humpath.2013.11.011>].
21. Sowa JP, Heider D, Bechmann LP, Gerken G, Hoffmann D, Canbay A. Novel algorithm for non-invasive assessment of fibrosis in NAFLD. *PLoS One* 2013;8(4):e62439 [PMID: 23638085 PMCID: PMC3640062 <https://doi.org/10.1371/journal.pone.0062439>].
22. Perveen S, Shahbaz M, Keshavjee K, Guergachi A. A Systematic Machine Learning Based Approach for the Diagnosis of Non-Alcoholic Fatty Liver Disease Risk and Progression. *Sci Rep.* 2018;8(1):2112 [PMID: 29391513 PMCID: PMC5794753. <https://doi.org/10.1038/s41598-018-20166-x>].
23. Fialoke S, Malarstig A, Miller MR, Dumitriu A. Application of Machine Learning Methods to Predict Non-Alcoholic Steatohepatitis (NASH) in Non-Alcoholic Fatty Liver (NAFL) Patients. *AMIA Annu Symp Proc.* 2018;2018:430–439 [PMID: 30815083 PMCID: PMC6371264].
24. Cao W, An X, Cong L, Lyu C, Zhou Q, Guo R. Application of Deep Learning in Quantitative Analysis of 2-Dimensional Ultrasound Imaging of Nonalcoholic Fatty Liver Disease. *J Ultrasound Med.* 2020;39(1):51–59 [PMID: 31222786. <https://doi.org/10.1002/jum.15070>].
25. Perakakis N, Polyzos SA, Yazdani A, et al. Non-invasive diagnosis of non-alcoholic steatohepatitis and fibrosis with the use of omics and supervised learning: A proof of concept study. *Metabolism* 2019;101:154005 [PMID: 31711876. <https://doi.org/10.1016/j.metabol.2019.154005>].
26. Garcia-Carretero R, Vigil-Medina L, Barquero-Perez O, Ramos-Lopez J. Relevant Features in Nonalcoholic Steatohepatitis Determined Using Machine Learning for Feature Selection. *Metab Syndr Relat Disord.* 2019;17(9):444–451 [PMID: 31675274. <https://doi.org/10.1089/met.2019.0052>].
27. Gawrieh S, Sethunath D, Cummings OW, et al. Automated quantification and architectural pattern detection of hepatic fibrosis in NAFLD. *Ann Diagn Pathol.* 2020;47:151518 [PMID: 32531442. <https://doi.org/10.1016/j.anndiagpath.2020.151518>].
28. Hwang A, Shi C, Zhu E, et al. Supervised learning reveals circulating biomarker levels diagnostic of hepatocellular carcinoma in a clinically relevant model of non-alcoholic steatohepatitis; An OAD to NASH. *PLoS One* 2018;13(6):e0198937 [PMID: 29944670 PMCID: PMC6019748 KJ, LP, SVG, DD, QA, MAY, IDG, PN. The specific roles of these authors are articulated in the 'author contributions' section. There are no patents, products in development or marketed products to declare. This does not alter our adherence to PLoS One policies on sharing data and materials. <https://doi.org/10.1371/journal.pone.0198937>].
29. Ramot Y, Zandani G, Madar Z, Deshmukh S, Nyska A. Utilization of a Deep Learning Algorithm for Microscope-Based Fatty Vacuole Quantification in a Fatty Liver Model in Mice. *Toxicol Pathol.* 2020;48(5):702–707 [PMID: 32508268. <https://doi.org/10.1177/0192623320926478>].
30. Forlano R, Mullish BH, Giannakeas N, et al. High-Throughput, Machine Learning-Based Quantification of Steatosis, Inflammation, Ballooning, and Fibrosis in Biopsies From Patients With Non-alcoholic Fatty Liver Disease. *Clin Gastroenterol Hepatol.* 2019 [PMID: 31887451. <https://doi.org/10.1016/j.cgh.2019.12.025>].
31. Yip TC, Ma AJ, Wong VW, et al. Laboratory parameter-based machine learning model for excluding non-alcoholic fatty liver disease (NAFLD) in the general population. *Aliment Pharmacol Ther.* 2017;46(4):447–456 [PMID: 28585725. <https://doi.org/10.1111/apt.14172>].
32. Van Vleck TT, Chan L, Coca SG, et al. Augmented intelligence with natural language processing applied to electronic health records for identifying patients with non-alcoholic fatty liver disease at risk for disease progression. *Int J Med Inform.* 2019;129:334–341 [PMID: 31445275 PMCID: PMC6717556. <https://doi.org/10.1016/j.ijmedinf.2019.06.028>].
33. Koleck TA, Dreisbach C, Bourne PE, Bakken S. Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. *J Am Med Inform Assoc.* 2019;26(4):364–379 [PMID: 30726935 PMCID: PMC6657282. <https://doi.org/10.1093/jamia/ocy173>].
34. Wang K, Lu X, Zhou H, Gao Y, Zheng J, Tong M, Wu C, Liu C, Huang L, Jiang T, Meng F, Lu Y, Ai H, Xie XY, Yin LP, Liang P, Tian J, Zheng R. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. *Gut.* 2019;68(4):729–741 [PMID: 29730602; PMCID: PMC6580779. <https://doi.org/10.1136/gutjnl-2018-316204>].
35. Tang A, Destremes F, Kazemirad S, Garcia-Duitama J, Nguyen BN, Cloutier G. Quantitative ultrasound and machine learning for assessment of steatohepatitis in a rat model. *Eur Radiol.* 2019;29(5):2175–2184 [PMID: 30560362. <https://doi.org/10.1007/s00330-018-5915-z>].
36. Wu CC, Yeh WC, Hsu WD, Islam MM, Nguyen PAA, Poly TN, Wang YC, Yang HC, Jack Li YC. Prediction of fatty liver disease using machine learning algorithms. *Comput Methods Programs Biomed.* 2019;170:23–29 [PMID: 30712601. <https://doi.org/10.1016/j.cmpb.2018.12.032>].
37. Che H, Brown LG, Foran DJ, Noshier JL, Hacihaliloglu I. Liver disease classification from ultrasound using multi-scale CNN. *Int J Comput Assist Radiol Surg.* 2021 [PMID: 34097226. <https://doi.org/10.1007/s11548-021-02414-0>].
38. Corey KE, Kartoun U, Zheng H, Shaw SY. Development and Validation of an Algorithm to Identify Nonalcoholic Fatty Liver Disease in the Electronic Medical Record. *Dig Dis Sci.* 2016;61(3):913–919 [PMID: 26537487 PMCID: PMC4761309. <https://doi.org/10.1007/s10620-015-3952-x>].

39. Qu HQ, Li Q, Grove ML, et al. Population-based risk factors for elevated alanine aminotransferase in a South Texas Mexican-American population. *Arch Med Res.* 2012;43(6):482–488 [PMID: 22959976 PMID: PMC3590902. <https://doi.org/10.1016/j.arcmed.2012.08.005>].
40. Huo Y, Terry JG, Wang J, et al. Fully automatic liver attenuation estimation combining CNN segmentation and morphological operations. *Med Phys.* 2019;46(8):3508–3519 [PMID: 31228267 PMID: PMC6692233. <https://doi.org/10.1002/mp.13675>].
41. Graffy PM, Sandfort V, Summers RM, Pickhardt PJ. Automated Liver Fat Quantification at Nonenhanced Abdominal CT for Population-based Steatosis Assessment. *Radiology* 2019;293(2):334–342 [PMID: 31526254 PMID: PMC6822771. <https://doi.org/10.1148/radiol.2019190512>].
42. Redman JS, Natarajan Y, Hou JK, et al. Accurate Identification of Fatty Liver Disease in Data Warehouse Utilizing Natural Language Processing. *Dig Dis Sci.* 2017;62(10):2713–2718 [PMID: 28861720. <https://doi.org/10.1007/s10620-017-4721-9>].
43. Forlano R, Mullish BH, Giannakeas N, Maurice JB, Angkathunyakul N, Lloyd J, Tzallas AT, Tsiouras M, Yee M, Thursz MR, Goldin RD, Manousou P. High-Throughput, Machine Learning-Based Quantification of Steatosis, Inflammation, Ballooning, and Fibrosis in Biopsies From Patients With Nonalcoholic Fatty Liver Disease. *Clin Gastroenterol Hepatol.* 2020;18(9):2081–2090.e9 [PMID: 31887451; PMID: PMC7397508. <https://doi.org/10.1016/j.cgh.2019.12.025>].
44. Neuschwander-Tetri BA, Loomba R, Sanyal AJ, et al. Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. *Lancet* 2015; 385(9972): 956–965 [PMID: 25468160 PMID: PMC4447192 DOI: [https://doi.org/10.1016/s0140-6736\(14\)61933-4](https://doi.org/10.1016/s0140-6736(14)61933-4)].
45. Qu H, Minacapelli CD, Tait C, Gupta K, Bhurwal A, Catalano C, Dafalla R, Metaxas D, Rustgi VK. Training of computational algorithms to predict NAFLD activity score and fibrosis stage from liver histopathology slides. *Comput Methods Programs Biomed.* 2021;207:106153 [PMID: 34020377. <https://doi.org/10.1016/j.cmpb.2021.106153>].
46. Feng G, Zheng KI, Li YY, Rios RS, Zhu PW, Pan XY, Li G, Ma HL, Tang LJ, Byrne CD, Targher G, He N, Mi M, Chen YP, Zheng MH. Machine learning algorithm outperforms fibrosis markers in predicting significant fibrosis in biopsy-confirmed NAFLD. *J Hepatobiliary Pancreat Sci.* 2021 [PMID: 33908180. <https://doi.org/10.1002/jhbp.972>].
47. Vanderbeck S, Bockhorst J, Kleiner D, Komorowski R, Chalasani N, Gawrieh S. Automatic quantification of lobular inflammation and hepatocyte ballooning in nonalcoholic fatty liver disease liver biopsies. *Hum Pathol.* 2015;46(5):767–775 [PMID: 25776030. <https://doi.org/10.1016/j.humpath.2015.01.019>].
48. Douali N, Abdenmour M, Sasso M, et al. Noninvasive diagnosis of non-alcoholic steatohepatitis disease based on clinical decision support system. *Stud Health Technol Inform.* 2013;192:1178 [PMID: 23920952].
49. Ma H, Xu CF, Shen Z, Yu CH, Li YM. Application of Machine Learning Techniques for Clinical Predictive Modeling: A Cross-Sectional Study on Nonalcoholic Fatty Liver Disease in China. *Biomed Res Int.* 2018;2018:4304376 [PMID: 30402478 PMID: PMC6192080. <https://doi.org/10.1155/2018/4304376>].
50. Helal KM, Taylor JN, Cahyadi H, et al. Raman spectroscopic histology using machine learning for nonalcoholic fatty liver disease. *FEBS Lett.* 2019;593(18):2535–2544 [PMID: 31254349. <https://doi.org/10.1002/1873-3468.13520>].
51. Katsiki N, Gastaldelli A, Mikhailidis DP. Predictive models with the use of omics and supervised machine learning to diagnose non-alcoholic fatty liver disease: A "non-invasive alternative" to liver biopsy? *Metabolism* 2019;101:154010 [PMID: 31711877 DOI: <https://doi.org/10.1016/j.metabol.2019.154010>].
52. Chiappini F, Desterke C, Bertrand-Michel J, Guettier C, Le Naour F. Hepatic and serum lipid signatures specific to nonalcoholic steatohepatitis in murine models. *Sci Rep.* 2016;6:31587 [PMID: 27510159 PMID: PMC4980672. <https://doi.org/10.1038/srep31587>].
53. Canbay A, Kälisch J, Neumann U, et al. Non-invasive assessment of NAFLD as systemic disease-A machine learning perspective. *PLoS One* 2019;14(3):e0214436 [PMID: 30913263 PMID: PMC6435145 declare. JPS and AC state that they received royalties for a scientific lecture, which was in part supported by TECOmedical group. This does not alter the authors' adherence to PLoS One policies on sharing data and materials. <https://doi.org/10.1371/journal.pone.0214436>].
54. Brattain LJ, Telfer BA, Dhyani M, Grajo JR, Samir AE. Objective Liver Fibrosis Estimation from Shear Wave Elastography. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:1–5 [PMID: 30440285. <https://doi.org/10.1109/embc.2018.8513011>].
55. Berná G, Romero-Gomez M. The role of nutrition in non-alcoholic fatty liver disease: Pathophysiology and management. *Liver Int.* 2020;40 Suppl 1:102–108 [PMID: 32077594. <https://doi.org/10.1111/liv.14360>].
56. Zeevi D, Korem T, Zmora N, et al. Personalized Nutrition by Prediction of Glycemic Responses. *Cell* 2015;163(5):1079–1094 [PMID: 26590418. <https://doi.org/10.1016/j.cell.2015.11.001>].
57. Segal E. Rich data sets could end costly drug discovery. *Nature* 2020;577(7792):S19 [PMID: 31996830. <https://doi.org/10.1038/d41586-020-00200-7>].
58. Cesaretti M, Brustia R, Goumar C, Cauchy F, Poté N, Dondoro F, Paugam-Burtz C, Durand F, Paradis V, Diaspro A, Mattos L, Scatton O, Soubrane O, Moccia S. Use of Artificial Intelligence as an Innovative Method for Liver Graft Macrosteatosis Assessment. *Liver Transpl.* 2020;26(10):1224–1232 [PMID: 32426934. <https://doi.org/10.1002/lt.25801>].
59. Nelson MR, Tipney H, Painter JL, et al. The support of human genetic evidence for approved drug indications. *Nat Genet.* 2015;47(8):856–860 [PMID: 26121088. <https://doi.org/10.1038/ng.3314>].
60. Doan S, Maehara CK, Chaparro JD, Lu S, Liu R, Graham A, Berry E, Hsu CN, Kanegaye JT, Lloyd DD, Ohno-Machado L, Burns JC, Tremoulet AH; Pediatric Emergency Medicine Kawasaki Disease Research Group. Building a Natural Language Processing Tool to Identify Patients With High Clinical Suspicion for Kawasaki Disease from Emergency Department Notes. *Acad Emerg Med.* 2016;23(5):628–36 [PMID: 26826020. <https://doi.org/10.1111/acem.12925>].
61. Xia J, Wang Z, Huan Y, et al. Pose Filter-Based Ensemble Learning Enables Discovery of Orally Active, Nonsteroidal Farnesoid X Receptor Agonists. *J Chem Inf Model.* 2020;60(3):1202–1214 [PMID: 32050066. <https://doi.org/10.1021/acs.jcim.9b01030>].
62. Yamada T, Iwasaki K, Maedera S, Ito K, Takeshima T, Noma H, Shojima N. Myocardial infarction in type 2 diabetes using sodium-glucose co-transporter-2 inhibitors, dipeptidyl peptidase-4 inhibitors or glucagon-like peptide-1 receptor agonists: proportional hazards analysis by deep neural network based machine learning. *Curr Med Res Opin.* 2020;36(3):403–409 [PMID: 31855074. <https://doi.org/10.1080/03007995.2019.1706043>].
63. Justice AC, Smith RV, Tate JP, McGinnis K, Xu K, Becker WC, Lee KY, Lynch K, Sun N, Concato J, Fiellin DA, Zhao H, Gelernter J, Kranzler HR. VA Million Veteran Program. AUDIT-C and ICD codes as phenotypes for harmful alcohol use: association with ADH1B polymorphisms in two US populations. *Addiction.* 2018;113(12):2214–2224 [PMID:29972609. <https://doi.org/10.1111/add.14374>].
64. Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. *Wiley Interdiscip Rev Data Min Knowl Discov.* 2019;9(4):e1312 [PMID: 32089788. <https://doi.org/10.1002/widm.1312>].

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.