

Artificial Intelligence in Nonalcoholic Fatty Liver Disease: A New Frontier in Diagnosis and Treatment

Pankaj Aggarwal, M.D., and Naim Alkhouri, M.D.

Artificial intelligence (AI) is an increasingly relevant field in medicine as clinicians incorporate technology into their daily practices. Nonalcoholic fatty liver disease (NAFLD) has risen to become the most common form of chronic liver disease globally. Furthermore, the extent of invasive and noninvasive objective metrics used in the diagnosis and management of NAFLD creates a seamless marriage between fatty liver disease and AI.

AI refers to the digital simulation and replication of human cognition. Machine learning (ML) is a popular subdiscipline of AI that refers to the process by which computers can be taught to independently collate data

without specific instruction. ML is further classified into supervised and unsupervised learning. Whereas the computer is trained with human feedback in supervised learning, unsupervised learning allows the computer to independently interpret data without specific human guidance or feedback. A subdiscipline of ML is deep learning (DL), which is based on deep neural networks (DNNs) inspired by the human brain (Fig. 1). DL refers to the computer's ability to conduct nonlinear analyses to identify predictive variables, assign them weights, and create prospective models for diseases. For example, a computer could be taught to use supervised ML

Abbreviations: AdaBoost, adaptive boosting; AI, artificial intelligence; ALT, alanine aminotransferase; AODE, aggregate one-dependence estimator; AST, aspartate aminotransferase; AUROC, area under the receiver operating characteristic curve; BMI, body mass index; DL, deep learning; DLRE, deep learning radiomic elastography; DNN, deep neural network; EMR, electronic medical record; F1, fibrosis stage 1; GGT, gamma-glutamyltransferase; HA, hyaluronic acid; HbA1c, hemoglobin A1c; ML, machine learning; MRI, magnetic resonance imaging; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; SWE, shear-wave elastography; TG, triglyceride; VCTE, vibration-controlled transient elastography.

From the Texas Liver Institute, University of Texas Health San Antonio, San Antonio, TX.

Potential conflict of interest: N.A. advises, is on the speakers' bureau for, and received grants from Gilead; advises and received grants from Allergan, Pfizer, and Zydus; advises and is on the speakers' bureau for Intercept; is on the speakers' bureau for AbbVie, Alexion, and Simply Speaking; and received grants from Akeru, Albireo, Axcella, BI, BMS, Celgene, Galmed, Galectin, Genfit, Enanta, Enyo, Hanmi, Inventiva, Madrigal, Merck, Novartis, Novo Nordisk, and Poxel.

Received April 3, 2020; accepted November 7, 2020.

View this article online at wileyonlinelibrary.com

© 2020 by the American Association for the Study of Liver Diseases

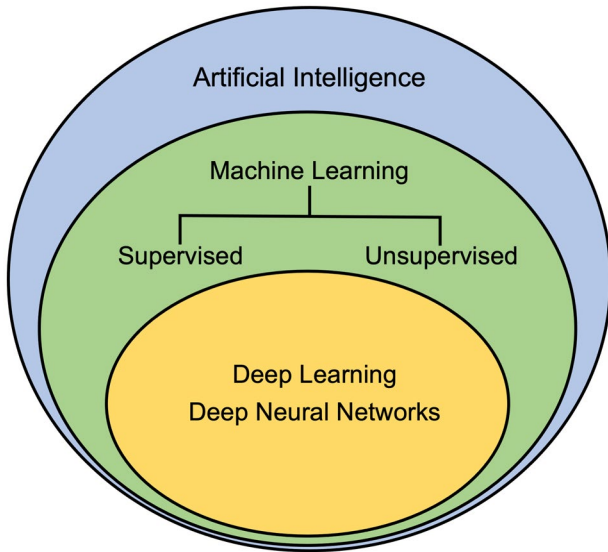


FIG 1 AI is an umbrella term that encapsulates ML. Within ML, DL models can be implemented to gain a more thorough understanding of the topic at hand.

algorithms to identify all patients with the diagnosis of diabetes mellitus in an electronic medical record (EMR). Using these patients as a test group, unsupervised ML algorithms can be used to identify the most common comorbidities or the utilization and response to specific classes of medications across a much larger population. This algorithm could be extended using DL to create a DNN that can predict long-term outcomes or socioeconomic determinants of care.

Thus, AI has the potential to elucidate and streamline several knowledge gaps within the field, specifically patient identification, determination of disease severity, and drug development. In this review, we aim to characterize how AI may close these gaps (Table 1).

IDENTIFYING PATIENTS AT RISK FOR NAFLD

Given the lack of efficient screening methods and the fact that most are asymptomatic, identifying patients with NAFLD is challenging. AI may solve this problem by identifying patients with explicit and implicit risk factors for NAFLD in EMRs. A recent study used ML to predict the presence of NAFLD in the general population who underwent screening with magnetic resonance imaging (MRI). Their model produced the NAFLD ridge score, which included six clinical variables (Table 2), with an area under the receiver operating characteristic curve (AUROC)

TABLE 1. CLOSING KNOWLEDGE GAPS IN NAFLD USING AI

Gap in Knowledge	Why It Matters	Utility of AI	Studies
Identifying patients at risk for clinically significant NAFLD	<ul style="list-style-type: none"> No established, efficient screening method Disease course is asymptomatic 	<ul style="list-style-type: none"> Screening EMR efficiently Using DL to recognize high-risk patients Identifying clinical predictors of disease with improved accuracy 	<ul style="list-style-type: none"> Yip et al.¹ developed the NAFLD ridge score based on weighted metabolic targets Ma et al.² identified predictors of NAFLD on a large scale in an EMR
Identifying patients with NAFLD who have advanced disease	<ul style="list-style-type: none"> If identified early, lifestyle modifications can alter disease progression Liver biopsy is invasive gold standard for diagnosis of fibrosis Patients may be indicated for pharmacotherapy 	<ul style="list-style-type: none"> Decreasing interobserver variability among pathologists Identifying patients with advanced fibrosis with significant accuracy to avoid invasive testing modalities 	<ul style="list-style-type: none"> Sowa et al.⁵ and Canbay et al.⁶ showed a combination of cellular markers to predict fibrosis Pokkalla et al.³ demonstrated uniformity in histopathological analyses via DNN Wang et al.⁹ showed DLRE could more consistently identify advanced fibrosis than SWE alone
Disease-specific drug targets for NAFLD	<ul style="list-style-type: none"> Limited options available for NAFLD pharmacotherapy Untapped potential for novel drug targets 	<ul style="list-style-type: none"> Identifying patient response to medications based on certain clinical parameters Repurposing previous drugs for novel purpose 	<ul style="list-style-type: none"> Hu et al.¹⁰ and Xie et al.¹¹ showed the potential to predict clinical response of medications Aliper et al.¹² demonstrated the potential to repurpose existing medications

TABLE 2. USING ML ALGORITHMS TO IDENTIFY PATIENTS AT HIGH RISK FOR NAFLD AND DETERMINE THE SEVERITY OF THE DISEASE

Study	No. of Patients	Purpose	Gold Standard	Algorithms Used	Key Variables	AUROC or Accuracy (%)	Sensitivity	Specificity
Yip et al. ¹	922	Patient identification using EMR	MRI	<ul style="list-style-type: none"> Logistic regression Ridge regression* Decision tree AdaBoost 	<ul style="list-style-type: none"> ALT High-density lipoprotein cholesterol TGs HbA1c White blood cells Hypertension 	AUROC: 0.88	0.923	0.904
Ma et al. ²	10,508	Patient identification using EMR	Ultrasound	<ul style="list-style-type: none"> Logistic regression* k-nearest neighbor Support vector machines* Naive Bayes Bayes network* Decision tree AdaBoost Bootstrap aggregating Random forest AODEs* 	<ul style="list-style-type: none"> Alkaline phosphatase GGT TGs Blood urea nitrogen Bilirubin Cholesterol Creatinine Fasting glucose Uric acid 	Accuracy: 83%	0.68	0.946
Sowa et al. ⁵	126	Determining severity of fibrosis	Liver biopsy	<ul style="list-style-type: none"> Logistic regression k-nearest neighbor Support vector machines Rule based Decision tree Random forest* Ensemble feature selection* Logistic regression* 	<ul style="list-style-type: none"> ALT AST M30 M65 HA 	AUROC: 0.67 Accuracy: 79%	0.60	0.77
Cambay et al. ⁶	164	Determine histological severity	Liver biopsy	<ul style="list-style-type: none"> Logistic regression* 	<ul style="list-style-type: none"> Age Sex BMI ALT AST M30 M65 HbA1c Fasting glucose Cholesterol Adiponectin 	AUROC: 0.73		

*Identifies most significant ML algorithms used in data analyses.

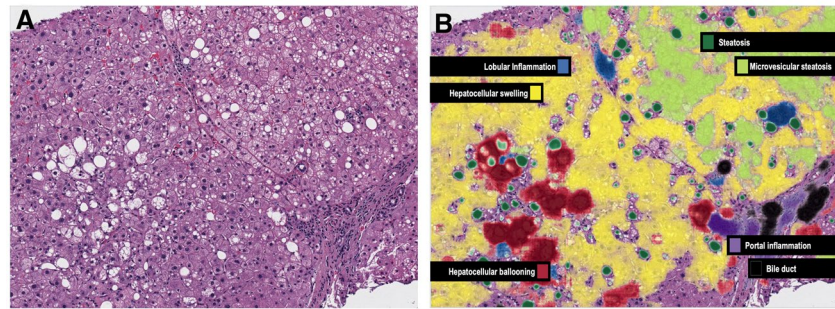


FIG 2 AI can improve accuracy and concordance in histological assessment of NAFLD severity. The PathAI System (B) identifies key histological features of NASH that are usually seen on traditional H&E staining (A), including steatosis, inflammation, hepatocyte ballooning, and fibrosis.

of 0.87 in the training cohort and 0.88 in the validation cohort.¹ Furthermore, in a large-scale study of more than 10,000 patients, researchers used AI to identify that triglycerides (TGs), alanine aminotransferase (ALT), gamma-glutamyltransferase (GGT), uric acid, and body mass index (BMI) were the most effective predictive markers of NAFLD.² These two studies highlight the potential utility of AI in patient identification. With the ability to screen and process exorbitant amounts of data, ML technology could help identify patients at risk for disease earlier and potentially effect change at the population level.

DIAGNOSING THE SEVERITY OF NAFLD

Similarly, AI is a potentially useful method of assessing the severity of NAFLD and identifying those with nonalcoholic steatohepatitis (NASH) and advanced fibrosis. Although liver biopsy is still considered the gold standard for the diagnosis of NASH, the histopathological evaluation of steatosis, inflammation, and ballooning is performed semiquantitatively, making the assessment subjective and prone to intra-observer and interobserver variability. In a recent presentation at The Liver Meeting, the PathAI research platform (PathAI; Boston, MA) was used to train a DNN with more than 20 layers and 8 million parameters from more than 600 liver biopsies. The ML models showed high concordance with the pathologist interpretations of liver histopathology, which may help to streamline the diagnosis and management of NASH³ (Fig. 2). Another similar example compared expert pathologist interpretation of NAFLD histopathology with a calculated collagen proportionate area in the staging of liver fibrosis. There was good-to-excellent agreement between pathologist interpretation and automated calculation, emphasizing the potential for programmed diagnosis of liver fibrosis.⁴ These examples lend credence to the

integration of AI in the interpretation of pathology and are favorable given increased uniformity, decreased human workload, and rapid processing of samples.

Alternatively, clinicians are more consistently turning to noninvasive tests based on clinical and serological biomarkers to determine the presence of NASH and the stage of liver fibrosis. For example, instead of comparing individual biomarkers, such as ALT or GGT, with fibrosis stage (F), ML algorithms can be used to compare dozens of biomarkers, identify the most predictive variables, and assign them relative weights to create novel scoring systems. In a study of morbidly obese patients, ML algorithms, such as logistic regression, random forest, and *K*-nearest neighbor, were used to compare serological tests and histopathology; the combination of liver enzymes (aspartate aminotransferase [AST] and ALT), cell death markers (M30 and M60), and hyaluronic acid (HA) was able to successfully identify stage F1 versus F2 (fibrosis) with 79% accuracy.⁵ Similarly, another study used ML algorithms to advocate for the addition of caspase cleaved serum CK-18 (M30) and adiponectin to noninvasive scoring systems⁶ (Table 2). The authors posit that the addition of the aforementioned markers has the added advantage of commenting on patients' overall metabolic status, and although this study is limited given the exclusion of patients with advanced fibrosis, these analyses are theoretically superior to conventional methodology. Furthermore, their ability to account for multiple features simultaneously in a robust multivariate approach eliminates the inherent biases of traditional linear analyses (Fig. 3).

Methods to stage liver fibrosis based on measuring liver stiffness, such as vibration-controlled transient elastography (VCTE) and shear-wave elastography (SWE), are increasingly being used. For example, VCTE

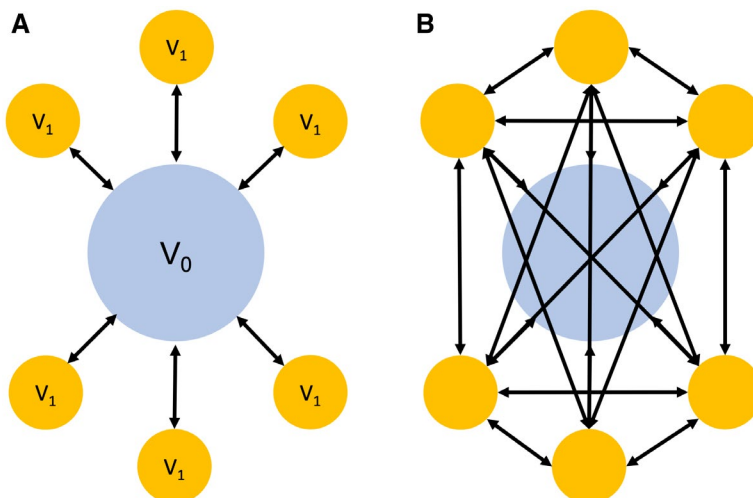


FIG 3 Comparison between hypothesis-based testing and ML. (A) The first graphic represents traditional hypothesis-based testing in which the control variable V_0 (blue circle) is compared with experimental variables V_1 (yellow) in sequential, linear fashion (six total tests). (B) The second graphic illustrates the potential for ML algorithms to conduct more robust analyses. As every experimental variable becomes its own control, the algorithm is able to conduct 3.5x more tests (21 total) to identify patterns beyond traditional cognition.

has demonstrated promise in the diagnosis of advanced fibrosis with AUROC of up to 0.89 for stage F4 fibrosis. However, VCTE is less sensitive in the detection of intermediate stages of fibrosis with an AUROC of 0.80 for F3 and 0.77 for F2.⁷ By applying DNN to radiomic data acquired through SWE, DL radiomic elastography (DLRE) was developed and improved the diagnostic accuracy of SWE with increased AUROC to 0.97 for F4, 0.98 for F3, and 0.85 for F2 when compared with liver biopsy.⁸ This suggests that ML algorithms such as DLRE could be used to diagnose cases of advanced fibrosis with excellent accuracy (Fig. 4).

DRUG DISCOVERY

The application of AI to drug discovery involves the ability of DNNs to incorporate and assemble large-scale data from genomic and transcriptomic analyses in ways that can inform drug design and predict both therapeutic and

toxic drug effects. Whereas genomics refers to the study of gene association with disease pathogenesis, transcriptomics focuses on gene expression and the biochemical processes that follow. More specifically, the “pharmacogenome” is defined as the sequence(s) of genes that will account for and reliably predict a patient’s response to a medication. AI may play different roles in the field of drug discovery, including, but not limited to, the following: (1) increasing the efficiency by which researchers identify drugs, (2) improving the precision of those drugs by recognizing patient-specific drug targets, and (3) repurposing previously discovered drugs for novel indications. A recent study demonstrated that DL computational models could more accurately identify metabolic pathways leading to the formation of toxic metabolites—a crucial first step in drug development.⁹ Preliminary studies have also shown that DL methods can more reliably elucidate and predict drug-target interactions compared with previous standards.^{10,11} In addition, ML has shown promise in the repurposing of existing agents, although this study was not specific to NAFLD.¹² Thus, the rise of AI presents a significant opportunity in the world of drug discovery, which is specifically relevant and applicable to the pharmacotherapy of NASH. Given the complex interplay between genetic background and environmental exposures, patients with NAFLD may have several unique pathophysiologies accounting for their disease state. Therefore, clinicians could rely on ML algorithms in the identification of drug targets and their efficacy in patient-specific disease processes.

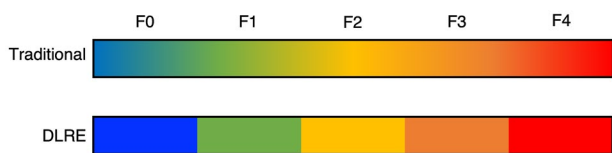


FIG 4 In traditional VCTE and SWE, minimal fibrosis (F0-F1) is easily distinguished from advanced fibrosis (F4). However, the diagnosis of intermediate stages of fibrosis is more difficult. DLRE using DNNs helps to distinguish progressive stages of fibrosis with remarkable accuracy.

CONCLUSION

The rise of AI in medicine has shown significant promise in the identification, diagnosis, and staging of patients with NAFLD. Furthermore, DL algorithms have the potential to provide target-specific medications, yielding efficacious pharmacotherapy in this biochemically complex disease. Notably, due to the lack of large-scale, randomized controlled trials, further research is necessary to demonstrate the utility of AI. However, given its potential to affect patient outcomes at a population-based level, AI may become increasingly relevant in the future diagnosis and management of NAFLD.

CORRESPONDENCE

Naim Alkhouri, M.D., Texas Liver Institute, University of Texas Health San Antonio, 607 Camden Street, San Antonio, TX 78215. E-mail: naim.alkhouri@gmail.com

REFERENCES

- 1) Yip TC-F, Ma AJ, Wong VW-S, et al. Laboratory parameter-based machine learning model for excluding non-alcoholic fatty liver disease (NAFLD) in the general population. *Aliment Pharmacol Ther* 2017;46:447-456.
- 2) Ma H, Xu CF, Shen Z, et al. 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.
- 3) Pokkalla H, Pethia K, Glass B, et al. Machine learning models accurately interpret liver histology in patients with nonalcoholic steatohepatitis (NASH). *Hepatology* 2019;70(Suppl. 1):121A-122A.
- 4) 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.
- 5) Sowa J, Heider D, Bechmann LP, et al. Novel algorithm for non-invasive assessment of fibrosis in NAFLD. *PLoS One* 2013;8:e62439.
- 6) Canbay A, Kälisch J, Neumann U, et al. Non-invasive assessment of NAFLD as systemic disease: a machine learning perspective. *PLoS One* 2019;14:e0214436.
- 7) Eddowes PJ, Sasso M, Allison M, et al. Accuracy of FibroScan controlled attenuation parameter and liver stiffness measurement in assessing steatosis and fibrosis in patients with nonalcoholic fatty liver disease. *Gastroenterology* 2019;156:1717-1730.
- 8) Wang K, Lu X, Zhou H, et al. 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;68:729-741.
- 9) Dang NL, Matlock MK, Hughes TB, et al. The metabolic rainbow: deep learning phase I metabolism in five colors. *J Chem Inf Model* 2020;60:1146-1164.
- 10) Hu S, Zhang C, Chen P, et al. Predicting drug-target interactions from drug structure and protein sequence using novel convolutional neural networks. *BMC Bioinformatics* 2019;20(Suppl. 25):689.
- 11) Xie L, He S, Song X, et al. Deep learning-based transcriptome data classification for drug-target interaction prediction. *BMC Genomics* 2018;19(Suppl. 7):667.
- 12) Aliper A, Plis S, Artemov A, et al. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. *Mol Pharm* 2016;13:2524-2530.