Kiani et al. BMC Medical Informatics and Decision Making (2025) 25:323
https://doi.org/10.1186/s12911-025-03168-4

BMC Medical Informatics and Decision Making

RESEARCH Open Access

Machine learning techniques in hepatic encephalopathy: a scoping review

Fatemeh Kiani[1], Farkhondeh Asadi[1*], Azamossadat Hosseini[1], Shahabedin Rahmatizadeh[1], Farhang Hosseini[2] and Behzad Kiani[3]

Abstract

Introduction Hepatic encephalopathy (HE) is defined as a specific type of cerebral dysfunction that encompasses a wide range of cognitive, psychomotor, and psychiatric disturbances. The burgeoning field of Artificial Intelligence (AI), particularly Machine Learning (ML), offers promising avenues for early detection and enhanced control of HE. This scoping review aims to provide a consolidated overview of AI's role in the diagnosis and management of HE, thereby informing and guiding future research endeavors in this domain.

Methods We followed Arksey and O'Malley's methodological framework to perform this scoping review, using PubMed, Web of Science, Scopus, ScienceDirect, and IEEE databases to find relevant articles. We also utilized the PRISMA standard to report our review in a standardized manner. Studies that focused on the applications of AI or ML techniques in relation to the prediction or diagnosis of HE disease were included.

Results Out of the 231 articles identified, 20 were ultimately included in this scoping review. The integration of artificial neural networks and expert systems represented an early and pioneering approach in applying AI to HE. Among supervised learning algorithms, Support Vector Machine emerged as the most frequently employed technique in HE research, based on our review of the selected studies. Notably, the primary application of AI in HE studies has been predictive modeling ($n=14$), followed by five studies focused on classifying HE stages and one study analyzing patient survival using AI methodologies.

Conclusions This scoping review highlights the growing use of AI and ML diagnostic models and predictive tools utilizing various data types. These advancements have the potential to positively impact patient outcomes. Future research should focus on validating and implementing these AI models in clinical settings to assess their real-world effectiveness in improving patient care.

Keywords Artificial intelligence, Machine learning, Hepatic encephalopathy, Prediction, Diagnostic tool

*Correspondence:
Farkhondeh Asadi
asadifar@sbmu.ac.ir
[1]Department of Health Information Technology and Management, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences (SBMU), Tehran, Iran
[2]Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[3]University of Queensland Centre for Clinical Research (UQCCR), Faculty of Health, Medicine, and Behavioural Sciences, The University of Queensland, Brisbane, Australia

Introduction

Hepatic encephalopathy (HE) is a brain disorder caused by liver failure and/or portal-systemic shunting. It presents as a distinct form of cerebral dysfunction, encompassing a wide range of cognitive, psychomotor, and psychiatric disturbances—ranging from subtle changes in mental status to coma. Figure 1 illustrates the relationship between liver cirrhosis and HE, as well as the associated types of symptoms [1–3]. This condition arises as a complication of both chronic and acute liver diseases [4, 5]. HE is classified into three major types based on its underlying causes: Type A, resulting from acute liver failure; Type B, associated with transjugular intrahepatic portosystemic shunt (TIPS); and Type C, linked to cirrhosis. In addition to these classifications, there exists Minimal HE (MHE), recognized as the earliest and mildest form of HE. MHE is prevalent in 80% of cirrhotic patients and significantly diminishes their quality of life [3, 6, 7]. Despite lacking clinical evidence for diagnosis, individuals with MHE exhibit gradual changes in psychomotor or neuropsychological functions [8]. Notably, MHE carries a poor prognosis, with predicted one and three-year survival rates of 42% and 23%, respectively, in the absence of liver transplantation [7]. The intensity of this condition underscores the need for comprehensive understanding and effective management strategies.

HE presents a significant diagnostic challenge due to its clinical overlap with various medical, neurological, and psychiatric conditions, complicating differential diagnosis and increasing vulnerability to additional brain injuries [9]. In response to this complexity, the International Society of Hepatic Encephalopathy and Nitrogen Metabolism introduced a revised classification system in 2011, dividing HE into two main categories: covert and overt [10]. MHE, as we discussed earlier, belongs to the covert category [11]. Covert HE is characterized by subtle neurocognitive impairments that often go undetected in routine clinical assessments, leading to underdiagnosis and inadequate treatment. Early identification of covert HE is critical for timely intervention, preventing disease progression and recurrence, improving patient quality of life, and potentially reducing mortality [12–14]. Given the high prevalence of MHE and the limitations of current screening methods, there is a growing need for the development and implementation of more sensitive and accessible diagnostic tools [15–17].

In contrast, overt HE is a more advanced and clinically evident manifestation of the condition, presenting marked neurological and psychiatric symptoms that are

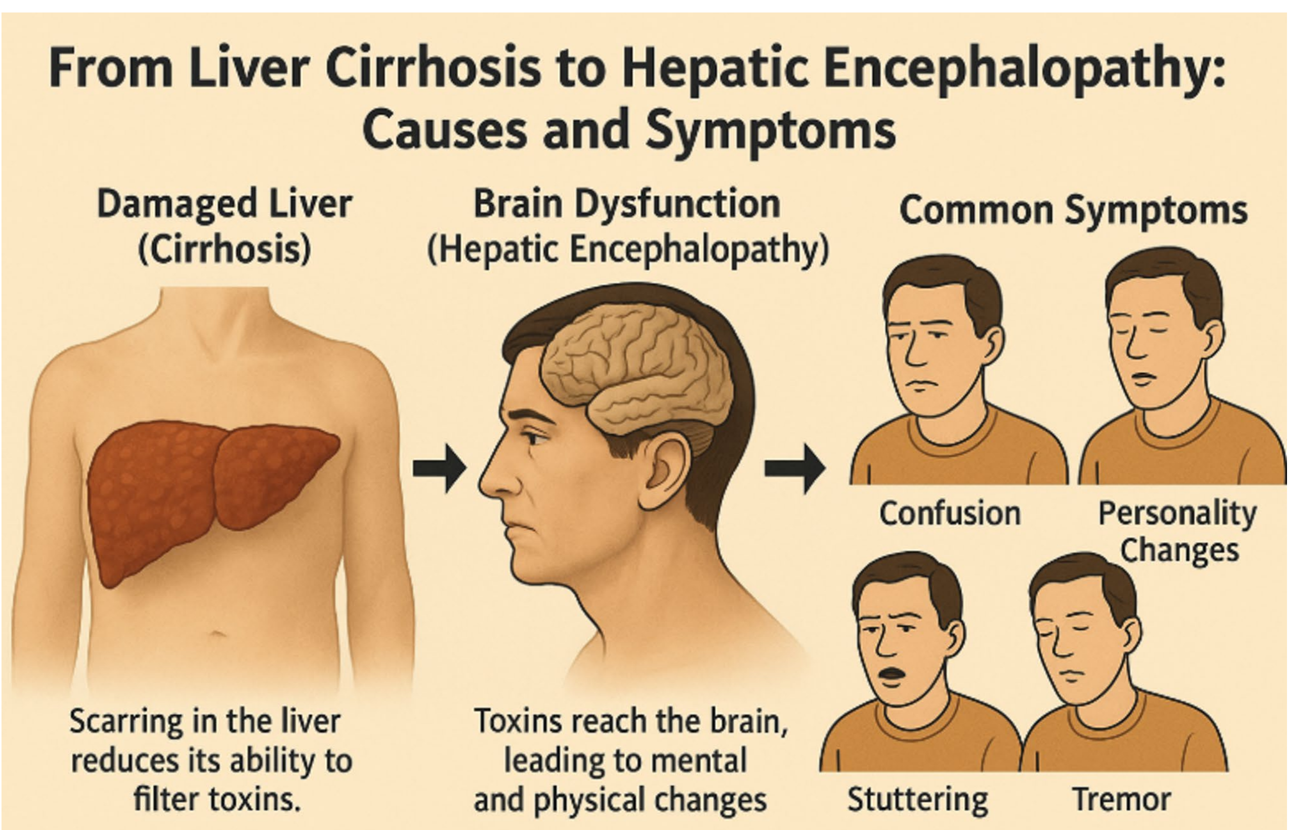

Fig. 1 Relationship between liver cirrhosis and HE, illustrating the associated symptom types

typically identifiable through standard clinical evaluation. It is a common and severe complication among individuals with cirrhosis, affecting approximately 30–50% of patients, with an estimated annual risk of 20% [18]. Recent advancements in artificial intelligence (AI) offer promising opportunities to enhance the diagnostic landscape of HE. AI-based methods can uncover latent patterns in laboratory results, clinical records, and electroencephalogram (EEG) data, surpassing traditional diagnostic approaches in both speed and accuracy [19, 20]. By overcoming current limitations in early detection, AI has the potential to significantly improve the prognosis and management of both covert and overt forms of HE, ultimately leading to better clinical outcomes.

AI constitutes a broad category encompassing diverse algorithms capable of discerning patterns within extensive datasets, providing valuable inferences and insights. Within this domain, machine learning (ML), a specialized scientific discipline and subset of AI, is focused on enabling computers to learn from data [21]. ML holds a significant role in the field of medical imaging, facilitating operations such as detection, segmentation, registration, integration, guided treatment, annotation, and recovery on images [22]. Originating in the 1950s and 1960s, the historical evolution of ML in medicine saw the development of algorithms for modeling and analyzing extensive datasets, with prominent contributions from Hunt et al. in symbolic learning [22], Nilsson in statistical methods [23], and Rosenblatt in neural networks [24]. One of the most transformative branches of ML is deep learning (DL), which employs multi-layered artificial neural networks (ANN) to model complex patterns and representations in data. By processing data through multiple layers of neurons, DL algorithms can automatically learn to extract features and make predictions with minimal human intervention, making them particularly effective for tasks such as image and speech recognition [25]. The ongoing trajectory of ML in medicine, with its promise to enhance accuracy and efficiency of diagnoses resulting in improved patient safety [26–29], has recently manifested in the application of AI, particularly ML, to HE research. Employing both clinical and laboratory data, as well as medical images, these approaches contribute to the in-time diagnosis, and enhanced control of HE.

There are a considerable number of studies using AI and ML in the field of HE. In one study, researchers employed demographic and clinical data as input for a prediction model, utilizing the ANN method, to predict HE [30]. Conversely, another investigation employed weighted support vector machine (SVM), weighted random forest (RF), and logistic regression algorithms for a similar purpose [19]. Notably, the majority of studies focusing on HE have extensively incorporated image processing methods, indicative of the widespread application of AI in this domain. For instance, researchers utilized diffusion White Matter Imaging and a ML technique based on Bayesian principles to classify and differentiate images of cirrhosis patients into two groups: those with and without MHE. This study identified two distinct areas in White Matter that effectively distinguished between these patient groups [31]. Additionally, the integration of EEG images with clinical data served as input for an ANN model and an expert system, aiding in the identification of EEG changes associated with HE patients [19]. Another study employed ML methods to predict 28-day mortality in patients [32]. These diverse approaches underscore the versatility of AI in leveraging various data sources for the diagnosis and prognostication of HE.

Although one previous review explored the use of AI for diagnosing MHE using handwriting and speech data [33], a comprehensive synthesis of AI and ML applications across the broader spectrum of HE is still lacking. Our scoping review aims to address that gap by mapping current research, summarizing methodologies, and identifying key limitations and opportunities. In doing so, it provides a foundation for future studies and supports the advancement of AI-driven approaches for improving HE diagnosis and management.

Methods

The method we have used is based on Arksey and O'Malley's methodological framework. In this framework, a five-step guideline including the following steps is provided [34]:

1. Identification of the research question
2. Identification of relevant studies
3. Selection of included studies
4. Charting of the key elements
5. Summarizing and reporting the results

We also used the PRISMAScR, a PRISMA extension intended to apply for reporting scoping reviews, which is included in Supplementary File 1. This standard, published in 2018, contains 20 essential and two optional items and helps to improve the reporting of scoping reviews [35].

Identifying the research question

What are the different applications of AI, especially ML and DL, in the field of HE? And what ML algorithms have been employed in the existing literature?

Identifying the relevant studies

Search strategy

A comprehensive search strategy was developed by combining relevant keywords to retrieve all studies on HE

Table 1 Search keywords used to create the search strategy

Concepts for AI/ML methods	Concepts for HE
"Machine Learning", "Artificial Intelligence", "Neural Network", "Deep Learning", "Computer-assisted", "Computer Vision", "Deep Network", "Computer-aided", "Convolutional Network", "Recurrent Network", "Graph Network", Backprop*, "Support vector", Ensemble*, "Random forest*", "Nearest neighbor*", "K-nearest neighbor*", "Gradient boost*", XGBoost*, Segmentation, "instance learning", "multi-instance learning", "Active Learning","Transfer Learning", "Reinforcement Learning", "Predictive Modeling", "Feature Engineering", "Hyperparameter Tuning", "Data Augmentation", "Expert system*", "Computational Intelligence", "Machine Intelligence", "Computer Reasoning", "Knowledge Representation", "Knowledge acquisition"," Computing Methodologies", "Long Short-Term Memor Network*", "Gated Recurrent Units", "Generative Adversarial Network*", "Deep Belief Networks", "Radial Basis Function Network*"	"Hepatic Encephalopath*", "Portal-Systemic Encephalopath*", "Hepatic Coma", "Hepatocerebral Encephalopathies", "Hepatic Stupor", "Fulminant Hepatic Failure", "Portosystemic Encephalopathy"

Table 2 PubMed database search strategy

PubMed search strategy
("hepatic encephalopath*"[Title/Abstract] OR "portal systemic encephalopath*"[Title/Abstract] OR "Hepatic Coma"[Title/Abstract] OR "Hepatic Stupor"[Title/Abstract] OR "Fulminant Hepatic Failure"[Title/Abstract] OR "Portosystemic Encephalopathy"[Title/Abstract] OR "Hepatic Encephalopathy"[MeSH Terms] OR "liver failure, acute"[MeSH Terms]) AND ("Machine Learning"[Title/Abstract] OR "Artificial Intelligence"[Title/Abstract] OR "Neural Network"[Title/Abstract] OR "Deep Learning"[Title/Abstract] OR "Computer-assisted"[Title/Abstract] OR "Computer Vision"[Title/Abstract] OR "Deep Network"[Title/Abstract] OR "Computer-aided"[Title/Abstract] OR "Convolutional Network"[Title/Abstract] OR "Recurrent Network"[Title/Abstract] OR "Graph Network"[Title/Abstract] OR "backprop*"[Title/Abstract] OR "Support vector"[Title/Abstract] OR "ensemble*"[Title/Abstract] OR "random forest*"[Title/Abstract] OR "nearest neighbor*"[Title/Abstract] OR "k nearest neighbor*"[Title/Abstract] OR "gradient boost*"[Title/Abstract] OR "xgboost*"[Title/Abstract] OR "Segmentation"[Title/Abstract] OR "instance learning"[Title/Abstract] OR "multi-instance learning"[Title/Abstract] OR "Active Learning"[Title/Abstract] OR "Transfer Learning"[Title/Abstract] OR "Reinforcement Learning"[Title/Abstract] OR "Predictive Modeling"[Title/Abstract] OR "Feature Engineering"[Title/Abstract] OR "Hyperparameter Tuning"[Title/Abstract] OR "Data Augmentation"[Title/Abstract] OR "expert system*"[Title/Abstract] OR "Computational Intelligence"[Title/Abstract] OR "Machine Intelligence"[Title/Abstract] OR "Computer Reasoning"[Title/Abstract] OR "Knowledge Representation"[Title/Abstract] OR "Knowledge acquisition"[Title/Abstract] OR "Computing Methodologies"[Title/Abstract] OR "Gated Recurrent Units"[Title/Abstract] OR "generative adversarial network*"[Title/Abstract] OR "Deep Belief Networks"[Title/Abstract] OR "radial basis function network*"[Title/Abstract] OR "Machine Learning"[MeSH Terms] OR "Artificial Intelligence"[MeSH Terms] OR "neural networks, computer"[MeSH Terms] OR "Deep Learning"[MeSH Terms] OR "diagnosis, computer assisted"[MeSH Terms])

involving AI or ML approaches. To ensure broad coverage in PubMed, a set of Medical Subject Headings (MeSH) terms was also incorporated. Table 1 presents the keywords used across all databases, while the complete search strategy for PubMed is provided in Table 2. Search strategies for the remaining databases are detailed in Supplementary File 2.

Information sources

We searched PubMed, Web of Science (WOS), Scopus, ScienceDirect, and IEEE Xplore to identify relevant articles published up to May 14, 2025. No geographic or temporal restrictions were applied, but searches were limited to studies published in English.

All search results were imported into the Zotero reference management software (version 6.0.30), and duplicate records were removed. The initial search yielded 231 citations across the five databases. After removing duplicates, 149 unique articles remained for screening. Title and abstract screening excluded 112 studies that did not meet the inclusion criteria. An additional 17 studies were excluded during the full-text review based on predefined exclusion criteria. Ultimately, 20 studies were included for data extraction (Fig. 2).

Study selection and eligibility criteria

Original studies published in peer-reviewed journals and written in English were considered for inclusion. Studies were eligible if they met both of the following criteria:

1. Applied at least one AI method, including ML or DL algorithms, in the context of HE.
2. Focused on the application of AI or ML techniques for the prediction or diagnosis of HE.

Studies were excluded if they met at least one of the following criteria:

1. Investigated AI applications in liver diseases other than HE.
2. Did not involve ML or DL methods specifically related to HE.
3. Were review articles or non-peer-reviewed publications, including editorials, conference abstracts, book chapters, or study protocols.

Screening process

The screening process was conducted by the first author, who reviewed the titles and abstracts of all retrieved

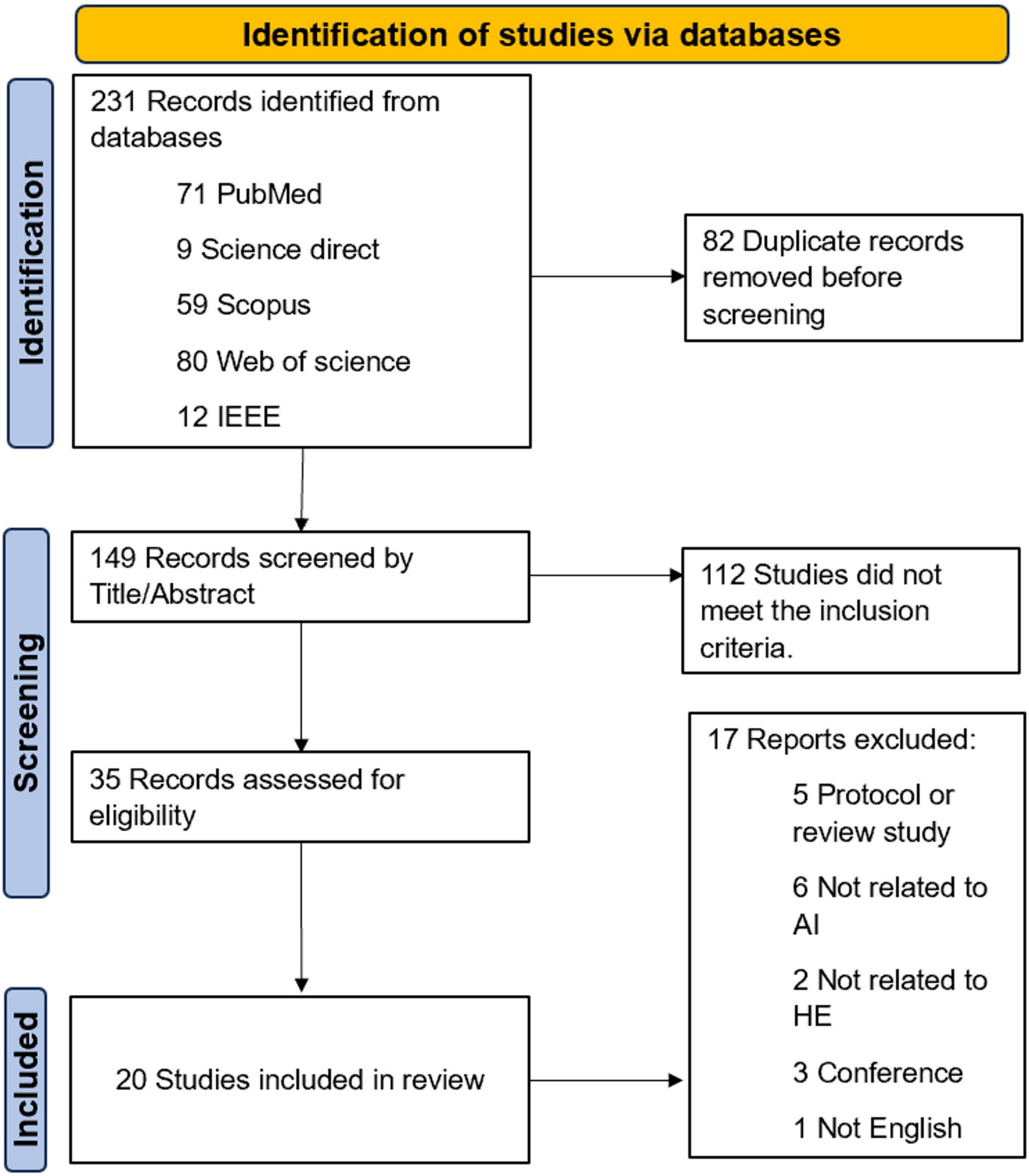

Fig. 2 Flowchart of study selection. Note: Some excluded studies may have met multiple exclusion criteria, but only the primary reason for exclusion is reported

articles based on the predefined inclusion and exclusion criteria. Irrelevant studies were excluded at this stage. For any articles where eligibility was uncertain, the first author consulted with the co-authors, and a consensus was reached through group discussion. Full-text articles of potentially relevant studies were then assessed by the first author to determine final eligibility for data extraction.

Charting the data

To guide the data extraction process, preliminary meetings were held to determine key variables, and a structured checklist was developed. The initial checklist was designed by the first author based on two comparable studies in the field [36, 37]. The checklist was reviewed and refined through group discussion with co-authors to ensure alignment with the study objectives and the specific context of HE. Given the limited number of AI-related studies in this field, HE-related articles were categorized into four groups based on the type of HE investigated: post-TIPS HE, overt HE, covert HE (including MHE), and studies encompassing all stages of HE.

Data extraction was conducted independently by the first author, who aggregated the results. A thorough review and verification of the extracted data were also performed by the same author to ensure accuracy and consistency. In terms of AI applications, studies were categorized into three major groups: classification, prediction, and survival analysis. The classification and prediction categories were further subdivided based on the primary type of input data used. For instance, when a study utilized multiple data sources, categorization was based on the data type that played the most critical role in the applied method. As a result, five types of input data were identified across the included studies:

1. Electroencephalography (EEG) data: EEG data are generated by recording the brain's electrical activity using electrodes placed on the scalp. These electrodes are connected to a device that amplifies and captures the brain's electrical signals as waveforms, allowing for analysis of neurological function [38, 39].
2. Magnetic Resonance Imaging (MRI) data: MRI data are obtained through a non-invasive imaging technique that uses powerful magnetic fields and radio waves to produce detailed anatomical and functional images of soft tissues, including the brain [40].
3. Clinical and Laboratory data: Clinical data encompasses a broad range of information pertaining to a patient's medical history, symptoms, physical examination findings, diagnostic results, treatments, and outcomes. Laboratory data, on the other hand, specifically refers to information obtained through the analysis of patient samples in a laboratory setting. This includes results from blood tests, urine tests, genetic tests, microbiology cultures, and other types of analyses performed on patient specimens.
4. Computed tomography (CT) images: Images that are created through a medical imaging procedure that utilizes X-rays to produce detailed cross-sectional images of the body [41].
5. Video-oculography: Video-oculography is a technique for capturing eye movement using digital video cameras. This represents a notable advancement from electronystagmography, which relies on the corneo-retinal potential, similar to a battery effect in the eye. As the eyes shift side to side and up and down, the positive and negative signals of the corneo-retinal potential are recorded [42].

Data items

For each eligible article, data items containing information regarding the characteristics of the studies and their data relevant to the purpose of our review were extracted and are shown in Table 3. Each study could have used more than one technique.

Collating, summarizing, and reporting results

We used different charts and tables to summarize and report the results.

Result

Of the 20 eligible articles, more than half of them ($n = 11$) were conducted in China [29–31, 43–50]. The first study was published in 2005 [19], and five studies were published by 2017 [19, 30, 43, 45, 51]. However, the number of studies increased, with 15 more published during 2020–2025 [16, 29, 31, 32, 44, 46–50, 52–56].

AI/ML techniques in research of HE

The combination of artificial neural network and expert system was the first technique used in AI studies related to HE [19, 51]. Studies from 2015 to 2017 focused on SVM and Bayesian algorithms [30, 43, 45]. However, the focus on various AI algorithms has increased since 2020, and other algorithms including RF, extreme gradient boosting (XGBoost), LR, K-nearest neighbors (KNN), Catboost, etc [16, 29, 32, 46, 49, 50, 52–56]. have also been used in recent years, but in these studies, the most widely used algorithm is the SVM, as one of the algorithms used in 6 studies was SVM [16, 50, 52, 54–56], and in 4 studies only the SVM algorithm was used [31, 44, 47, 48].

All these algorithms are listed by frequency and number of input data in Fig. 3. Figure 4 is a network diagram

Table 3 Performance data in studies of the application of AI in HE

Ref	Country	Application	HE type(s)	AI/ML algorithm(s)	Records (Training set, test set)	Validation method	Covariates (e.g.)	Performance: Accuracy (ACC), AUC
[46]	China	Prediction of HE	Post-TIPS HE	Random forest	168 (135, 33)	5-iterations 10-fold cross validation	CT and Clinical & laboratory data data: Chid-pugh score, PPG%, MELD, CRP, NH3, Bilirubin	ACC (90.9%) AUC (0.649)
[16]	Taiyuan	Prediction of HE	Overt HE	Random forest, support vector machine, logistic regression, weighted random forest and weighted support vector machine	1256 (829, 427)	Iterations100 times	Clinical & laboratory data: Constipation, Edema, Electrolyte disturbance, UGIB, Infection, Diuretic, WBC, RBC, HGB, NEP, AST, Na, Chloride, TB, ALB, DBIL, IBIL, PT, FIB, APTT	ACC (RF: 91.39%, SVM: 93.06%, logistic regression: 87.56%, weighted RF: 87.32%, weighted SVM: 83.49%) AUC (RF: 52.30%, SVM: 52.30%, logistic regression: 80.10%, weighted RF: 82.41% and weighted SVM: 80.34%)
[32]	Massachusetts	Survival analysis	Overt HE	Artificial neural networks, gradient boosting machine, random forest, and bagged trees	601 (422, 179)	Not mentioned	Clinical & laboratory data: Age, Male, Race, BMI, Coexisting disorders, Vital signs, Laboratory finding, Therapy strategy, Scoring system	AUC (ANN: 0.837, GBM: 0.769, RF: 0.789, and bagged trees: 0.741)
[30]	China	Prediction of HE	Covert HE	Bayesian ML technique called Graphical-model-based multivariate analysis	65	Not mentioned	MRI images	ACC (79.10%)
[43]	China	Prediction of HE	Covert HE	Support vector machine, multilayer perceptrons, and C4.5	74	Not mentioned	MRI images	Not mentioned
[52]	USA	Prediction of HE	Overt HE	Random forest, support vector machine, logistic regression and gradient boosting machine	269 (215, 54)	Cross-validation with 30 iterations	Clinical & laboratory data: Stool and saliva microbiota	AUC (RF: 73%, SVM: 72%, logistic regression: 71%, GBM: 62%)
[31]	China	Prediction of HE	Covert HE	Support vector machine	53	Not mentioned	MRI images	ACC (83.02%) AUC (94.0%)
[44]	China	Prediction of HE	Covert HE	Support vector machine	103 (102, 1)	A leave-one-out Cross-validation	MRI images	ACC (88.71%)
[45]	China	Prediction of HE	Covert HE	Support vector machine	35 (34, 1)	A leave-one-out cross-validation(LOOCV)	MRI images	ACC (82.90%)
[29]	China	Prediction of HE	After-TIPS	Artificial neural network	207 (137, 70)	Not mentioned	Clinical data & laboratory data: (Gender, Age, Etiology (Hepatitis B, Hepatitis C, Alcoholic, Others), ALBI grade, PALBI grade, Child–Pugh grade, MELD score, Diabetes, Presence of PVT, Sarcopenia, Pre-TIPS covert HE, TBil, ALB, ALT, AST, Platelet, Creatine, INR, Na)	Not mentioned
[51]	Ireland Ltd	Classification of HE	Covert HE	Artificial neural network and expert system	238	Not mentioned	EEG images	Not mentioned

Table 3 (continued)

Ref	Country	Application	HE type(s)	AI/ML algorithm(s)	Records (Training set, test set)	Validation method	Covariates (e.g.)	Performance: Accuracy (ACC), AUC
[53]	North American	Prediction of HE	Overt HE	Logistic regression	602	Not mentioned	Clinical & laboratory data: Age, Gender, MELD score, Serum Na, WBC count and Albumin levels and admission infection and prior HE, Metabolites serum (Aromatic amino acid, Branched chain amino acid, Serum metabolites (Benzoate, Short-chain fatty acid, Carbohydrates and Lipids)	AUC (0.75%)
[54]	Spain	Classification of HE	Covert HE	Support vector machine, k nearest neighbors, linear discriminant and ensembles method: subspace discriminant and bagged tree	47 (37, 10)	1000-iterations 5-fold cross validation	Video-oculography	ACC (SVM: 96.7%, KNN: 93.3%, Linear linear discriminant: 90%, subspace discriminant: 90% and bagged trees: 80%)
[55]	Italy	Prediction of HE	Overt HE	Decision tree, random forest, k-nearest neighbor, and support vector machine, multilayer perceptron	124 (100,24)	Not mentioned	MRI images	ACC (Decision tree: 58.82%, RF: 58.82%, KNN: 76.5%, SVM: 47% and MLP: 71%) AUC (Decision tree: 59%, RF: 60%, KNN: 77%, SVM: 50% and MLP: 71%)
[47]	China	Classification of HE	Covert HE	Support vector machine	95	500-iterations 10-fold cross validation	MRI images	ACC (96.02%) AUC (93.0%)
[48]	China	Classification of HE	Covert HE	Support vector machine	103	leave-one-out cross-validation	MRI images	ACC (72.5%)
[49]	China	Prediction of HE	After-TIPS	Random forest, extreme gradient boosting, and logistic regression	218 (152, 66)	10-fold cross-validation	Clinical & laboratory data: Age, Etiology, HE, Ascites, PVT, SMVT, SVT, BCS, CTPV, Pressure-related indicators, TBIL, ALB, CR, ALT, Prolonged PT, INR, Sodium, PLT, WBC, Child-Pugh score, Child-Pugh classification, MELD score, MELD-Na score, ALBI score, FIPS score, CLIF-C AD score, EBL, Partial splenectomy, EIS, GCAE, PSE	AUC (RF: 69%, XGBoost: 79%, and logistic regression: 82%)
[50]	China	Prediction of HE	After-TIPS	Logistic regression with regularization, support vector machine and random forest	106 (85, 21), External test: 24	10-fold cross-validation	CT Radiomic	ACC (LR:71%, SVM: 76% and RF: 76%) AUC (Logistic regression: 83%, SVM: 80% and RF: 69%)

Table 3 (continued)

Ref	Country	Application	HE type(s)	AI/ML algorithm(s)	Records (Training set, test set)	Validation method	Covariates (e.g.)	Performance: Accuracy (ACC), AUC
[56]	USA	Prediction of HE	After-TIPS	Support vector machine, logistic regression and CatBoost	327	5-fold cross-validation	Clinical & laboratory data: Age, Gender, BMI, Diuretics before TIPS, Cause of portal hypertension, Fluorotin, PSG, INR, AST, ALT, Albumin, Indirect bilirubin, Creatinine, Sodium, MELD, MELD-Na,–Pugh, Indication for TIPS, Portal vein site, Hepatic vein site, Complication, Prior HE, Length of stent, Diameter of stent, Presence of hydrothorax, Paracentesis at the time of TIPS, Variceal bleeding before TIPS	ACC (SVM: 74%, LR:75% and CatBoost: 73%) AUC (SVM: 82%,Logistic regression:83% and CatBoost: 83%)
[19]	Italy	Classification of HE	All stages	Artificial neural Network and expert system	100 (50, 50)	Not mentioned	EEG images	Not mentioned

BMI, body mass index, CRP, C-reactive protein, MELD, model for end-stage liver disease, PPG, portosystemic pressure gradient, UGIB, upper gastrointestinal bleeding, WBC, white blood cell count, RBC, red blood cell count, HGB, hemoglobin, NEP, neutrophilic granulocyte percentage, AST, glutamic-pyruvic transaminase, Na, sodium, TP, total protein, ALB, albumin, DBIL, direct bilirubin, IBIL, indirect bilirubin, PSG, portosystemic gradient gastric antral vascular ectasia, IVCP, inferior vena cava pressure, PVP, portal venous pressure, PT, prothrombin time, FIB, fibrinogen, APTT, activated partial thromboplastin time, ALBI, albumin-bilirubin, PALBI, platelet-albumin-bilirubin, PVT, portal vein thrombosis, TBil, total bilirubin, INR, international normalized ratio, ALT, alanine aminotransferase, BCS, Budd-Chiari syndrome, CLIF-C AD, chronic liver failure consortium acute decompensation,| CR, creatinine, CTPV, cavernous transformation of the portal vein, EBL, endoscopic band ligation, EIS, endoscopic injection sclerotherapy, FIPS, fibrosis-4 index for liver fibrosis, GCAE, gastric antral vascular ectasia, IVCP, inferior vena cava pressure, HE, hepatic encephalopathy, MELD-Na, model for end-stage liver disease sodium, OS, overall survival, PLT, platelet count, PSE, partial splenectomy, PVP, portal venous pressure, SMVT, superior mesenteric vein thrombosis, SVT, splenic vein thrombosis

of relationships between algorithm types, data types, and HE types.

AI applications in HE disease and research

The most common application of AI in HE research has been prediction ($n = 14$) [16, 29–31, 43–46, 49, 50, 52, 53, 55, 56], while five studies focused on classifying HE degrees [19, 47, 48, 51, 54], and only one study analyzed patient survival using AI techniques [32].

Classification of HE

EEG images Regarding EEG images, two pioneering studies in this field were conducted by A. Pellegrini and their colleagues. In the first study [19], demographic, clinical, and laboratory data of cirrhosis patients were input into an ANN and an expert system. These data were then classified into five classes: normal EEG, EEG with normal limits, mild signs of HE, distinctive features of HE, and signs of severe HE. The agreement between the system's classification and that of an expert electroencephalographer was subsequently measured. In the subsequent study [51], employing the same methodology as the previous one, classification was performed. Subsequently, with a one-year follow-up of the patients, the incidence of bouts of overt HE in each class was detected.

MRI images In a study, Yue Cheng et al. (2021) took resting-state MRI images of patients with liver cirrhosis with HE, patients with liver cirrhosis without HE, and healthy subjects, and then used the features of these images as data to classify these three groups using the SVM algorithm. The classification algorithm achieved an accuracy of 72.5% [48]. Two years later, in 2023, another study was conducted with a similar method and classification to the study by Yue Cheng et al. The SVM classification algorithm achieved an accuracy of 96% and an area under the curve (AUC) of 93% [47].

Video-oculography Only one study has used video-oculography technology to aid in the diagnosis of MHE by analyzing eye movements with ML tools. First, eye movements of subjects with different tests were recorded through an automated gaze tracking system and 150 features were extracted for each patient. Then, by performing statistical tests, 14 features were found to be significant and were selected for use in ML classification algorithms including SVM, KNN, linear discriminant and two algorithms from the ensemble algorithm category including subspace discriminant and bagged tree. The SVM algorithm had the highest performance with an accuracy of 96.7% [54].

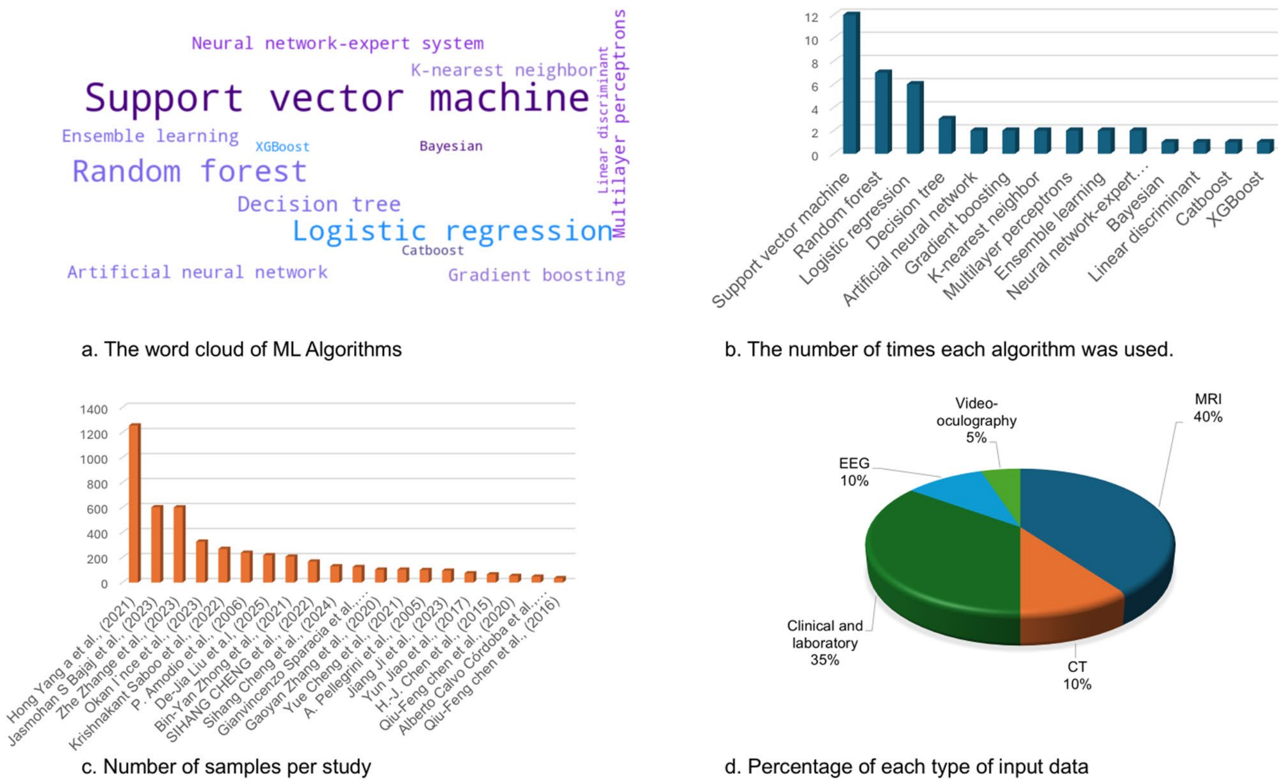

Fig. 3 AI/ML algorithms and their input data types

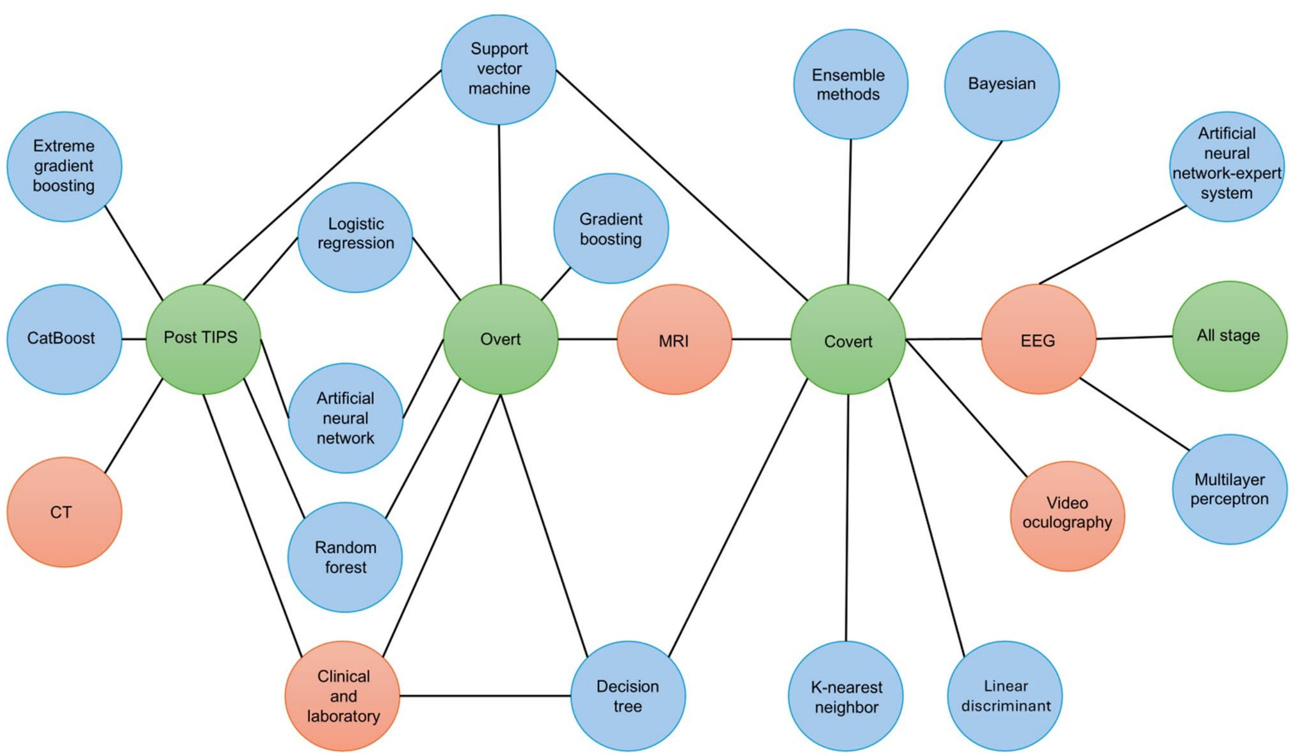

Fig. 4 Network diagram of relationships between algorithms, data types, and types of HE in included studies

Prediction of HE

MRI images In this subgroup of MRI images, five studies, conducted exclusively in China, utilized MRI techniques to generate input images for ML algorithms [30, 31, 43–45]. The objective across all studies was to distinguish patients with MHE from those without HE. Three studies solely employed the SVM algorithm [31, 44, 45], while one study utilized an ML algorithm based on Bayesian techniques known as Graphical model-based multivariate analysis (GAMMA) [30]. Another study employed several algorithms, including SVM, GAMMA, multilayer perceptron (MLP), and C4.5 [43]. Notably, in the study by Gaoyan Zhang et al., the SVM algorithm demonstrated the highest accuracy, reaching 88.71% [44].

Recently, in 2025, a study in Italy used MRI images to predict and grade cirrhotic patients with overt HE (grades 1 and 2), cirrhotic patients without HE, and controls (non-cirrhotic individuals). First, MRI images were preprocessed with the Principal Component Analysis (PCA) method and feature selection was performed. Then, various ML algorithms including decision tree, RF, KNN, SVM, and MLP were used for prediction. As the best classification algorithms between different classes, the MLP predicted patients versus controls with 100% accuracy, the KNN algorithm predicted patients with or without HE with 76.5% accuracy, and the MLP algorithm predicted the grade of HE (HE grade 1, HE grade ≥ 2) with 94.1% accuracy [55].

Clinical and laboratory data In recent years, there has been a growing emphasis on using AI techniques to predict HE using clinical and laboratory data, with two studies conducted in 2021 focusing on this approach [16, 29]. One study aimed to predict overt HE within the first three months after TIPS, utilizing an ANN algorithm and prognostic nomograms [29]. It achieved a concordance index (C-index) value of 0.816. The concordance index is a useful metric for assessing the predictive power of a model and comparing different models in terms of their ability to rank or order outcomes correctly [57].

Another study focused on predicting HE in cirrhosis patients [16]. Due to the imbalance in the number of patients with and without HE, and to enhance performance, the study employed weighted ML algorithm models, including weighted SVM and weighted RF. These models were compared with non-weighted models, such as SVM and RF. Although the SVM model demonstrated the highest accuracy in test data, with a value of 0.93, further examination revealed that the weighted RF model outperformed others, particularly in handling unbalanced HE data for external validation.

Additionally, a study employed AI techniques to compare microbial compositions in stool and saliva samples of cirrhosis patients and their association with the presence or absence of HE. Various algorithms were utilized, and ultimately, the RF algorithm, incorporating both microbial inputs, achieved the highest AUC value of 0.73 [52].

In 2023, a prospective, multicenter study used thyroid hormone levels collected and analyzed from patients' serum and followed during hospitalization to predict overt HE. The researchers achieved an AUC of 0.75% by training and finally testing a logistic regression algorithm [53].

In a retrospective study, researchers analyzed 327 patients who underwent TIPS surgery for liver cirrhosis. Feature selection was performed with a sequential feature selection model with five-way cross-validation, and 7 features were ultimately selected to build the models. Three machine learning models were developed using SVM, logistic regression, and CatBoost algorithms. The SVM algorithm performed best with an accuracy of 75% and an AUC of 83% [56].

A study was conducted on 218 patients with liver cirrhosis who were simultaneously diagnosed with acute variceal bleeding with the aim of developing a ML model to predict the occurrence of overt HE after TIPS. The dataset included basic demographic characteristics (gender, age, etiology) along with clinical, biochemical, and procedural data. Three algorithms including RF, XGBoost, and LR were validated through 10-fold cross-validation. Among the evaluated models, logistic regression showed the highest performance with an AUC of 82.5% [49].

CT images In a study, the RF algorithm was utilized to extract radiomics features from CT images and predict HE after portosystemic bypass or shunting, achieving an accuracy of 90% [46].

A multicenter retrospective study of 130 patients with cirrhosis who underwent TIPS surgery over a period of approximately five years was conducted. Pre-TIPS contrast-enhanced CT images were collected for VAT segmentation and radiomic feature extraction. Least absolute shrinkage and selection operator regression with ten-fold cross validation were performed for dimensionality reduction. Logistic regression with regularization, SVM, and RF were used to build the model. The mean AUC in the test sets was 84% [50].

Survival analysis

Regarding survival analysis, only one study conducted in Massachusetts aimed to predict the 28-day mortality of HE patients [32]. This study employed four different ML algorithms—ANN, gradient boosting machine (GBM), RF, and bagging trees—all utilizing clinical data as input. Notably, the ANN algorithm demonstrated the highest predictive performance, achieving an AUC value of 0.837.

Table 3 summarizes the model performance measures from the reviewed studies focusing on AI applications in HE disease. These studies underscore the potential of AI approaches in enhancing diagnostic accuracy and survival analysis, which can ultimately result in improved patient outcomes.

Discussion

HE poses a significant challenge, characterized by a dismal prognosis and low survival rates, underscoring the critical need for accurate prediction and identification of individuals at risk [6]. In recent years, the surge in AI and ML applications has prompted researchers to explore avenues for enhancing diagnostic precision and, ultimately, patient outcomes [58]. This scoping review represents the inaugural effort to elucidate the role of AI and ML techniques in HE care and research. Our analysis encompassed 20 studies spanning from 2005 to 2025, all leveraging AI/ML methodologies. Notably, the studies were frequently constrained by limited data availability, with 17 studies featuring between 35 and 327 entries [19, 29–31, 43–52, 54–56], while only three studies boasted larger datasets of 601, 602 and 1256 entries [16, 32, 53]. Critically, seven studies exhibited poorly defined training and test sets from 35 to 168 [19, 30, 31, 43–48, 50, 54, 55]. Algorithmic accuracy ranged between 47% and 96.7%, underlining the pivotal importance of ample dataset sizes for robust model training. Furthermore, 11 studies implemented validation methods [16, 44–50, 52, 54, 56]. Validation methods are employed to optimize model hyperparameters and estimate validation errors [59]. Noteworthy findings from our investigation revealed that the highest algorithmic accuracy was associated with the study with a small input volume but used a 5-fold validation method with 1000 iterations, demonstrating the importance of using validation methods [54]. This iterative approach facilitated the development of a model exhibiting superior performance [60].

We found that early studies primarily focused on employing a combination of ANN and expert systems to classify EEG images in HE staging [19, 51]. Despite working with relatively small datasets, these algorithms demonstrated noteworthy performance. ANN are commonly associated with contemporary ML techniques, while expert systems are considered part of traditional AI due to their rule-based and knowledge-oriented approach [61, 62]. The fusion of these methodologies into ANN and expert systems represents a hybrid approach, integrating aspects from both traditional and modern AI. The study conducted by P. Amodio et al. revealed that this combined algorithm provides classifications of HE stages that generally align with a significant level of accuracy compared to assessments by Electroencephalography specialists [19, 51]. However, there is a noticeable gap in information regarding the model's accuracy in the current literature, emphasizing the need for further investigation in future research endeavors.

Between 2015 and 2020, a notable shift was observed towards predicting diseases from MRI images, employing a variety of both traditional and emerging ML algorithms, such as MLP [43], C4.5 [43], notably SVM [31, 43–45] and Bayesian techniques like multivariate analysis based on graphical models [30, 43]. Among these, SVM emerged as the predominant algorithm in the studies, demonstrating its efficacy as a supervised learning technique [31, 44, 45]. Notably, in studies focusing on liver disease prediction, SVM has exhibited remarkable performance, as exemplified by Vijayarani et al.'s study, where SVM outperformed Naïve Bayes with an accuracy of 79.66%, which was about 19% more than Naïve Bayes algorithm [63]. Similarly, in another study comparing six supervised algorithms for liver disease classification, SVM ranked fourth in accuracy at 64% but displayed the highest sensitivity at 88% [64]. This consistent efficacy underscores SVM's status as a robust tool in disease prediction research, particularly in liver diseases [65], owing to its ability to handle high-dimensional data and high discrimination power with limited sample sizes [66]. This advantage becomes particularly crucial given our review findings, revealing a high number of input variables relative to observations in HE diagnosis studies, coupled with a limited number of HE patient samples.

In 2021 and 2022, a notable shift occurred towards leveraging AI techniques utilizing clinical and laboratory data for predicting HE [16, 52]. One study employed various weighted ML algorithms to address unbalanced data, resulting in enhanced prediction models [16]. Additionally, the RF algorithm demonstrated promising results in such studies [16, 52], with accuracy ranging from 62 to 93% in this group.

Between 2023 and 2025, the landscape of algorithms and data types utilized in predicting and classifying HE has evolved significantly. Recent studies have employed a diverse array of both traditional ML and modern techniques, such as logistic regression, SVM, RF, and more sophisticated models like XGBoost and CatBoost [47, 49, 50, 53–56]. Notably, a study introduced features extracted from video-oculography to classify covert HE, achieving impressive accuracy with the SVM algorithm at 96.7% [54]. This emphasizes the potential of integrating novel data types to enhance diagnostic precision. Furthermore, the incorporation of newer algorithms like CatBoost demonstrates a growing trend toward advanced ML techniques. CatBoost is particularly noteworthy as it is designed to handle categorical features efficiently while mitigating overfitting, making it well-suited for clinical datasets [67]. Zhe Zhange et al.'s study used ML algorithms to predict 28-day mortality in HE patients [32]. In this study, four traditional and advanced AI algorithms were used, and the

ANN algorithm obtained the highest prediction performance. It is worth noting that real-time prediction of mortality risk among HE patients admitted to the ICU could, in turn, optimize treatment to improve clinical prognosis [68].

Modern AI algorithms, including ML models like RF, show promise in the classification of medical MRI images, particularly in predicting HE disease. These models often achieve high accuracy rates, with studies reporting performance ranging from 58.82 to 96.02% for various algorithms [31, 43–45, 47, 48, 55]. For instance, the SVM model has been a standout, achieving accuracies of 96.7% in Spain and 96.02% in China for classifying covert HE [47, 54]. Additionally, logistic regression continues to demonstrate strong performance, with an accuracy of 87.56% reported for predicting HE using clinical and laboratory data [16]. In a comprehensive study from 2024, multiple algorithms—including RF and logistic regression—were evaluated for predicting HE after TIPS. The logistic regression model achieved an AUC of 83%, showcasing its effectiveness, while RF performed competitively as well [50]. Furthermore, the usage of ensemble methods, such as bagged trees and subspace discriminants, has shown effectiveness in increasing classification accuracy for covert HE [54]. Assessing the performance of both traditional and contemporary AI algorithms is crucial in the context of HE disease. Each approach offers distinct advantages yet determining the definitive superiority between traditional and contemporary methods remains challenging. The integration of various techniques, along with the emerging use of advanced models like CatBoost, indicates a progressive enhancement in HE diagnostics [56]. Overall, the growing trend toward integrating modern algorithms and varied data types illustrates a progressive enhancement in HE diagnosis methodologies while underscoring the necessity for more extensive datasets to bolster algorithmic reliability and application-specific efficacy.

Notably, DL algorithms were not explicitly implemented in the studies reviewed, indicating that these advanced techniques are still not fully leveraged in HE research. We can only consider the four studies that employed MLPs as focusing on DL [19, 43, 51, 55], because ANNs qualify as DL algorithms when they contain more than two hidden layers [69]. This is due to their ability to learn hierarchical representations from raw data, handle high-dimensional inputs, and perform well in image-based tasks [70]. However, the full potential of these algorithms is still largely unexploited, presenting a valuable opportunity for future studies to utilize their strengths for enhanced diagnostic precision.

The field of HE researches often struggles with limited sample sizes, which can significantly hinder the generalizability of findings and the robustness of predictive models. Small sample sizes not only restrict the diversity of patient populations represented in studies but also lead to increased variability in results, making it challenging to draw broad, reliable conclusions applicable to the wider clinical context [71]. This limitation is critical, as findings derived from narrow cohorts may not be representative of the larger population, thereby compromising the practical utility of predictive models in everyday clinical settings.

Moreover, inherent biases in the datasets used for training AI models could distort algorithmic predictions [72]. These biases may arise from factors such as demographic skewness, variations across medical centers, or inconsistent data quality. Such distortions could compromise the accuracy and reliability of the predictive capabilities of the proposed models, potentially leading to suboptimal clinical decisions based on these tools. Therefore, while some studies demonstrate promising results, these limitations highlight the urgent need for larger, more diverse datasets and a proactive approach to bias mitigation strategies in future research endeavors.

Another notable limitation identified from the synthesized studies is the pervasive lack of external validation. While the findings and algorithms developed within individual studies showcase potential utility and efficacy in managing HE, the absence of robust external validation is a significant barrier to their generalizability and real-world applicability. Without validation against independent datasets, the reliability and performance of these AI and ML applications remain vulnerable to overfitting or biases inherent in the training data [73]. This risks their seamless translation into clinical practice, where validated tools are essential for ensuring accurate predictions and interventions. Addressing this critical gap in experimental design and methodology will be imperative for fostering trust among practitioners and enhancing the clinical utility of AI and ML tools developed for HE.

Within the spectrum of AI and ML applications for HE, the predominant reliance on retrospective cohorts in the studies reviewed introduces another significant limitation: selection bias [16, 19, 29–32, 43–50, 54–56]. The retrospective nature of the data sources can lead to an overrepresentation of certain patient cohorts, treatment approaches, or outcomes. This skewing not only compromises the robustness and validity of the findings but also limits their applicability in real-world clinical scenarios [74]. It underscores the necessity for future research endeavors to encompass prospective cohort designs or randomized controlled trials, which can mitigate selection bias and enhance the reliability of results. This methodological evolution is instrumental in fortifying the credibility and impact of AI-driven solutions and ensuring their effective integration into clinical management strategies for this complex neurological condition.

Limitations

This scoping review has several limitations. First, the search was limited to studies published in English, which may have excluded relevant research in other languages. Second, the review did not assess the quality or risk of bias of included studies, as is standard in scoping reviews, which limits interpretation of the robustness of the evidence. Third, the field of machine learning in hepatic encephalopathy is still emerging, and many studies were exploratory or heterogeneous in methodology, making it difficult to draw consistent conclusions. Finally, due to rapid advances in machine learning, some recent developments may not have been captured if published after our final search date.

Future directions

Some studies either did not report or inadequately reported performance indicators of algorithms. It is imperative for future studies to comprehensively report their findings on AI algorithm performance in a robust manner, encompassing accuracy, F1 score, area under the receiver operating characteristic, and area under the precision-recall curve. Each of these metrics gauges distinct aspects of performance and may serve as better predictors than others in specific circumstances [75]. Future research should set clear guidelines for performance reporting, potentially adopting a standardized reporting framework to facilitate easier comparison and validation of results across studies.

A pivotal direction for future research involves addressing the intricate challenge of seamlessly integrating AI solutions into clinical practice. Future studies should dedicate efforts to investigating the standardization of datasets, with specific strategies such as collaborative data-sharing initiatives among institutions. Establishing a consortium of researchers and clinical centers may enhance data diversity and volume, thereby improving the generalizability of findings. Additionally, developing standardized data collection protocols could help mitigate biases and inconsistencies in datasets. Furthermore, refining regulatory processes and devising strategies for the harmonious assimilation of AI tools into real-world clinical workflows will be essential. Establishing robust frameworks for data interoperability, compliance with regulatory standards, and creating user-friendly integration interfaces will support the successful adoption and efficacy of AI-driven interventions in the domain of HE management.

Moreover, future research should incorporate Explainable AI (XAI) methodologies to enhance the interpretability and trustworthiness of AI-driven models [76]. Employing techniques such as feature importance analysis and local explanations will allow researchers to clarify the rationale behind algorithmic predictions, thereby identifying critical predictors of patient outcomes, such as lab values like ammonia levels. By integrating XAI, researchers can also tackle potential biases within datasets and bolster the external validation of models. Emphasizing XAI will not only mitigate selection bias inherent in retrospective cohorts but will also foster greater confidence among healthcare providers regarding the AI tools they employ. Ultimately, these initiatives will promote the reliable adoption of AI technologies in clinical practice for effective management of HE.

It is also crucial to transcend the exclusive focus on accuracy and sensitivity metrics and pivot towards enhancing the computational speed and responsiveness of AI algorithms. Future research should aim to develop benchmarks for computational efficiency alongside traditional performance metrics. By prioritizing efficiency in tandem with performance, researchers can foster the development of streamlined and agile computational models that can swiftly analyze complex datasets and deliver rapid insights. This focus on computational speed will optimize clinical decision-making processes and ensure that AI tools are not only accurate but also timely in their application.

Conclusion

This scoping review highlights the growing application of AI, particularly ML, in the diagnosis and prognosis of HE. Studies demonstrated the feasibility of developing accurate classification and prediction models using diverse data sources, including medical imaging, clinical variables, and laboratory findings. These models have shown encouraging performance in detecting disease and estimating survival outcomes. However, limitations such as small sample sizes and lack of external validation constrain generalizability. Future research should prioritize robust validation and real-world implementation to assess the clinical utility of AI-driven tools in HE management.

Abbreviations

HE	Hepatic encephalopathy
TIPS	Transjugular intrahepatic portosystemic shunt
MHE	Minimal hepatic encephalopathy
AI	Artificial intelligence
ML	Machine learning
DL	Deep learning
SVM	Support vector machine
EEG	Electroencephalogram or Electroencephalography
MRI	Magnetic resonance imaging
CT	Computed tomography
GAMMA	Graphical-Model-based Multivariate Analysis
RF	Random forest
MeSH	Medical Subject Headings
WOS	Web of Science
MLP	Multilayer perceptrons
ANN	Artificial neural network
XGBoost	Extreme gradient boosting
KNN	K-nearest neighbors
AUC	Area under the curve
GAMMA	Graphical model-based multivariate

PCA	Principal component analysis
C-index	Concordance index
GBM	Gradient boosting machine
ACC	Accuracy
XAI	Explainable AI

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-025-03168-4.

Supplementary Material 1

Supplementary Material 2

Acknowledgements
We would like to acknowledge all the authors we reviewed their papers in our study.

Author contributions
FK wrote the first draft, analyzed and interpreted the data. SR, AH and FH contributed to extracting the data and provided intellectual thinking on data charting. BK and FA supervised and conceptualized the study. All authors read and approved the final manuscript.

Funding
None.

Data availability
All data generated or analyzed during this study are included in this published article.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 10 December 2024 / Accepted: 25 August 2025
Published online: 01 September 2025

References

1. Hepatic encephalopathy 2018: a clinical practice guideline by the Italian association for the study of the liver (AISF). Dig Liver Dis. 2019;51(2):190–205.
2. Butterworth RF. Hepatic encephalopathy in cirrhosis: pathology and pathophysiology. Drugs. 2019;79(Suppl 1):17–21.
3. Vilstrup H, Amodio P, Bajaj J, Cordoba J, Ferenci P, Mullen KD, et al. Hepatic encephalopathy in chronic liver disease: 2014 practice guideline by the American association for the study of liver diseases and the European association for the study of the liver. Hepatol Baltim Md. 2014;60(2):715–35.
4. Prakash R, Mullen KD. Mechanisms, diagnosis and management of hepatic encephalopathy. Nat Rev Gastroenterol Hepatol. 2010;7(9):515–25.
5. Wijdicks EFM. Hepatic encephalopathy. N Engl J Med. 2016;375(17):1660–70.
6. Bustamante J, Rimola A, Ventura PJ, Navasa M, Cirera I, Reggiardo V, et al. Prognostic significance of hepatic encephalopathy in patients with cirrhosis. J Hepatol. 1999;30(5):890–5.
7. Minimal hepatic encephalopathy impairs quality of life. J Clin Exp Hepatol [Internet]. [cited 2023 Nov 20]. Available from: https://www.jcehepatology.com/article/S0973-6883(14)00428-9/fulltext
8. Nardelli S, Gioia S, Pasquale C, Pentassuglio I, Farcomeni A, Merli M, et al. Cognitive impairment predicts the occurrence of hepatic encephalopathy after transjugular intrahepatic portosystemic shunt. Off J Am Coll Gastroenterol ACG. 2016;111(4):523.
9. Bai M, Qi X, Yang Z, Yin Z, Nie Y, Yuan S, et al. Predictors of hepatic encephalopathy after transjugular intrahepatic portosystemic shunt in cirrhotic patients: a systematic review. J Gastroenterol Hepatol. 2011;26(6):943–51.
10. Meyer T, Eshelman A, Abouljoud M. Neuropsychological changes in a large sample of liver transplant candidates. Transpl Proc. 2006;38(10):3559–60.
11. Kappus MR, Bajaj JS. Covert hepatic encephalopathy: not as minimal as you might think. Clin Gastroenterol Hepatol. 2012;10(11):1208–19.
12. Subclinical hepatic encephalopathy predicts the development of overt hepatic encephalopathy - PubMed [Internet]. [cited 2023 Nov 20]. Available from: https://pubmed.ncbi.nlm.nih.gov/11569701/
13. Takikawa Y, Endo R, Suzuki K, Fujiwara K, Omata M, Fulminant Hepatitis Study Group of Japan. Prediction of hepatic encephalopathy development in patients with severe acute hepatitis. Dig Dis Sci. 2006;51(2):359–64.
14. Ferenci P, Lockwood A, Mullen K, Tarter R, Weissenborn K, Blei AT. Hepatic encephalopathy–definition, nomenclature, diagnosis, and quantification: final report of the working party at the 11th world congresses of gastroenterology, vienna, 1998. Hepatol Baltim Md. 2002;35(3):716–21.
15. Zhang G, Cheng Y, Liu B. Abnormalities of voxel-based whole-brain functional connectivity patterns predict the progression of hepatic encephalopathy. Brain Imaging Behav. 2017;11(3):784–96.
16. Yang H, Li X, Cao H, Cui Y, Luo Y, Liu J, et al. Using machine learning methods to predict hepatic encephalopathy in cirrhotic patients with unbalanced data. Comput Methods Programs Biomed. 2021;211:106420.
17. Zhang LJ, Wu S, Ren J, Lu GM. Resting-state functional magnetic resonance imaging in hepatic encephalopathy: current status and perspectives. Metab Brain Dis. 2014;29(3):569–82.
18. Bajaj JS. Review article: the modern management of hepatic encephalopathy. Aliment Pharmacol Ther. 2010;31(5):537–47.
19. Pellegrini A, Ubiali E, Orsato R, Schiff S, Gatta A, Castellaro A, et al. Electroencephalographic staging of hepatic encephalopathy by an artificial neural network and an expert system. Neurophysiol Clin Clin Neurophysiol. 2005;35(5–6):162–7.
20. Cancers. | Free Full-Text | Machine learning and radiomics applications in esophageal cancers using non-invasive imaging methods—A critical review of literature [Internet]. [cited 2023 Nov 20]. Available from: https://www.mdpi.com/2072-6694/13/10/2469.
21. Machine learning for medical applications [Internet]. [cited 2023 Nov 20]. Available from: https://www.hindawi.com/journals/tswj/2015/825267/
22. (PDF) Book Review. Experiments in Induction Earl B. Hunt, Janet Marin, and Phillip J. Stone, 1966; 247. (New York and London: Academic Press, 76s.) [Internet]. [cited 2023 Nov 21]. Available from: https://www.researchgate.net/publication/311742428_Book_Review_Experiments_in_Induction_Earl_B_Hunt_Janet_Marin_and_Phillip_J_Stone_1966_247_New_York_and_London_Academic_Press_76s
23. Learning Machines: Foundations of … by Nilsson, Nils J. [Internet]. [cited 2023 Nov 21]. Available from: https://www.amazon.com/Learning-Machines-Foundations-Trainable-Pattern-Classifying/dp/0070465703
24. Rosenblatt F. Principles of neurodynamics: perceptrons and the theory of brain mechanisms [Internet]. Vol. 55. Spartan books Washington, DC; 1962 [cited 2023 Nov 21]. Available from: https://apps.dtic.mil/sti/citations/AD0256582
25. What Are Machine and Deep Learning? | SpringerLink [Internet]. [cited 2025 May 31]. Available from: https://link.springer.com/chapter/10.1007/978-3-030-83047-2_1
26. Weisberg EM, Chu LC, Park S, Yuille AL, Kinzler KW, Vogelstein B, et al. Deep lessons learned: radiology, oncology, pathology, and computer science experts unite around artificial intelligence to strive for earlier pancreatic cancer diagnosis. Diagn Interv Imaging. 2020;101(2):111–5.
27. Burton RJ, Albur M, Eberl M, Cuff SM. Using artificial intelligence to reduce diagnostic workload without compromising detection of urinary tract infections. BMC Med Inf Decis Mak. 2019;19(1):171.
28. Nguyen A, Hassanzadeh H, Zhang Y, O'Dwyer J, Conlan D, Lawley M, et al. A decision support system for pathology test result reviews in an emergency department to support patient safety and increase efficiency. Stud Health Technol Inf. 2019;264:729–33.
29. Zhong BY, Wang WS, Shen J, Du H, Zhang S, Li WC, et al. Single-centre retrospective training cohort using artificial intelligence for prognostic prediction of encephalopathy, mortality, and liver dysfunction after early TIPS creation. Cardiovasc Intervent Radiol. 2021;44(10):1597–608.
30. Chen HJ, Chen R, Yang M, Teng GJ, Herskovits EH. Identification of minimal hepatic encephalopathy in patients with cirrhosis based on white matter imaging and bayesian data mining. AJNR Am J Neuroradiol. 2015;36(3):481–7.

31. Chen QF, Zou TX, Yang ZT, Chen HJ. Identification of patients with and without minimal hepatic encephalopathy based on Gray matter volumetry using a support vector machine learning algorithm. Sci Rep. 2020;10(1).
32. Zhang Z, Wang J, Han W, Zhao L. Using machine learning methods to predict 28-day mortality in patients with hepatic encephalopathy. BMC Gastroenterol. 2023;23(1):111.
33. Gazda J, Drotar P, Drazilova S, Gazda J, Gazda M, Janicko M, et al. Artificial intelligence and its application to minimal hepatic encephalopathy diagnosis. J Pers Med. 2021;11(11).
34. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 8(1) [Internet]. [cited 2023 Dec 5]. Available from: https://www.tandfonline.com/doi/abs/10.1080/1364557032000119616
35. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. 2018;169(7):467–73.
36. Machine learning applications for early detection of esophageal cancer: a systematic review | Research Square [Internet]. [cited 2023 Dec 5]. Available from: https://www.researchsquare.com/article/rs-2777227/v1
37. Sharma M, Savage C, Nair M, Larsson I, Svedberg P, Nygren JM. Artificial intelligence applications in health care practice: scoping review. J Med Internet Res. 2022;24(10):e40238.
38. Rayi A, Murr N. Electroencephalogram. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 [cited 2023 Dec 25]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK563295/
39. Niedermeyer E, da Silva FHL. Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams & Wilkins; 2005. 1342 p.
40. Hargreaves B, Worters PW, Pauly KB, Pauly JM, Koch KM, Gold GE. Metal induced artifacts in MRI. AJR Am J Roentgenol. 2011;197(3):547–55.
41. National Institute of Biomedical Imaging and Bioengineering [Internet]. [cited 2023 Dec 25]. Computed tomography (CT). Available from: https://www.nibib.nih.gov/science-education/science-topics/computed-tomography-ct
42. Gans RE. Video-oculography: a new diagnostic technology for vestibular patients. Hear J. 2001;54(5):40.
43. Jiao Y, Wang XH, Chen R, Tang TY, Zhu XQ, Teng GJ. Predictive models of minimal hepatic encephalopathy for cirrhotic patients based on large-scale brain intrinsic connectivity networks. Sci Rep. 2017;7(1):11512.
44. Zhang G, Li Y, Zhang X, Huang L, Cheng Y, Shen W. Identifying mild hepatic encephalopathy based on multi-layer modular algorithm and machine learning. Front Neurosci. 2020;14:627062.
45. Chen QF, Chen HJ, Liu J, Sun T, Shen QT. Machine learning classification of cirrhotic patients with and without minimal hepatic encephalopathy based on regional homogeneity of intrinsic brain activity. PLoS One. 2016;11(3):e0151263.
46. Cheng S, Yu X, Chen X, Jin Z, Xue H, Wang Z, et al. CT-based radiomics model for preoperative prediction of hepatic encephalopathy after transjugular intrahepatic portosystemic shunt. Br J Radiol. 2022;95(1132):20210792.
47. Ji J, Liu Y, Wu G, Hu Y, Liang C, Wang X. Changes in dynamic and static brain fluctuation distinguish minimal hepatic encephalopathy and cirrhosis patients and predict the severity of liver damage. Front Neurosci. 2023;17.
48. Cheng Y, Zhang G, Zhang X, Li Y, Li J, Zhou J, et al. Identification of minimal hepatic encephalopathy based on dynamic functional connectivity. Brain Imaging Behav. 2021;15(5):2637–45.
49. Liu DJ, Jia LX, Zeng FX, Zeng WX, Qin GG, Peng QF, et al. Machine learning prediction of hepatic encephalopathy for long-term survival after transjugular intrahepatic portosystemic shunt in acute variceal bleeding. World J Gastroenterol [Internet]. 2025;31(4). Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85213695733&doi=10.3748%2fwjg.v31.i4.100401&partnerID=40&md5=a839b1923d9e12a2315a953a47772ec3
50. Cheng S, Hu G, Jin Z, Wang Z, Xue H. Prediction of hepatic encephalopathy after transjugular intrahepatic portosystemic shunt based on CT radiomic features of visceral adipose tissue. Acad Radiol. 2024;31(5):1849–61.
51. Amodio P, Pellegrini A, Ubiali E, Mathy I, Piccolo FD, Orsato R, et al. The EEG assessment of low-grade hepatic encephalopathy: comparison of an artificial neural network-expert system (ANNES) based evaluation with visual EEG readings and EEG spectral analysis. Clin Neurophysiol Off J Int Fed Clin Neurophysiol. 2006;117(10):2243–51.
52. Saboo K, Petrakov NV, Shamsaddini A, Fagan A, Gavis EA, Sikaroodi M, et al. Stool microbiota are superior to saliva in distinguishing cirrhosis and hepatic encephalopathy using machine learning. J Hepatol. 2022;76(3):600–7.
53. Bajaj JS, Tandon P, O'Leary JG, Reddy KR, Garcia-Tsao G, Thuluvath P, et al. Admission serum metabolites and thyroxine predict advanced hepatic encephalopathy in a multicenter inpatient cirrhosis cohort. Clin Gastroenterol Hepatol Off Clin Pract J Am Gastroenterol Assoc. 2023;21(4):1031–e10403.
54. Calvo Córdoba A, García Cena CE, Montoliu C. Automatic video-oculography system for detection of minimal hepatic encephalopathy using machine learning tools. Sensors [Internet]. 2023;23(19). Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85174053536&doi=10.3390%2fs23198073&partnerID=40&md5=2b283ca417ce9ffed739e515ef5db139
55. Sparacia G, Colelli G, Parla G, Mamone G, Maruzzelli L, Lo Re V, et al. Brain magnetic resonance imaging radiomic signature and machine learning model prediction of hepatic encephalopathy in adult cirrhotic patients. Life-Basel. 2025;15(3).
56. İnce O, Önder H, Gençtürk M, Golzarian J, Young S. Machine learning insights: predicting hepatic encephalopathy after TIPS placement. Cardiovasc Intervent Radiol. 2023;46(12):1715–25.
57. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval Estimation. Stat Med. 2004;23(13):2109–23.
58. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Heal J. 2019;6(2):94–8.
59. Maleki F, Muthukrishnan N, Ovens K, Reinhold C, Forghani R. Machine learning algorithm validation. Neuroimaging Clin N Am. 2020;30(4):433–45.
60. Tougui I, Jilbab A, El Mhamdi J. Impact of the choice of cross-validation techniques on the results of machine learning-based diagnostic applications. Healthc Inf Res. 2021;27(3):189–99.
61. Zou J, Han Y, So SS. Overview of artificial neural networks. In: Livingstone DJ, editor. Artificial neural networks: methods and applications [Internet]. Totowa, NJ: Humana Press; 2009 [cited 2024 Apr 16]. p. 14–22. Available from: https://doi.org/10.1007/978-1-60327-101-1_2
62. Gill TG. Early expert systems: where are they now? MIS Q. 1995;19(1):51–81.
63. Mohan V. Liver disease prediction using SVM and naïve bayes algorithms. 2015.
64. Rahman AS, Shamrat FJM, Tasnim Z, Roy J, Hossain SA. A comparative study on liver disease prediction using supervised machine learning algorithms. Int J Sci Technol Res. 2019;8(11):419–22.
65. Modhugu VR, Ponnusamy S. Comparative analysis of machine learning algorithms for liver disease prediction: SVM, logistic regression, and decision tree. Asian J Res Comput Sci. 2024;17(6):188–201.
66. Peng C, Cheng Q. Discriminative ridge machine: a classifier for high-dimensional data or imbalanced data. IEEE Trans Neural Netw Learn Syst. 2021;32(6):2595–609.
67. Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. J Big Data. 2020;7(1):94.
68. Ding Y, Wang Y, Zhou D. Mortality prediction for ICU patients combining just-in-time learning and extreme learning machine. Neurocomputing. 2018;281:12–9.
69. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to machine learning, neural networks, and deep learning. Transl Vis Sci Technol. 2020;9(2):14.
70. Deep learning in medical image analysis. Annu Rev [Internet]. [cited 2025 Jun 1]. Available from: https://www.annualreviews.org/content/journals/10.1146/annurev-bioeng-071516-044442
71. Kokol P, Kokol M, Zagoranski S. Machine learning on small size samples: a synthetic knowledge synthesis. Sci Prog. 2022;105(1):00368504211029777.
72. Ahsen ME, Ayvaci MUS, Raghunathan S. When algorithmic predictions use human-generated data: a bias-aware classification algorithm for breast cancer diagnosis. Inf Syst Res. 2019;30(1):97–116.
73. Beam AL, Manrai AK, Ghassemi M. Challenges to the reproducibility of machine learning models in health care. JAMA. 2020;323(4):305–6.
74. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. JNCI J Natl Cancer Inst. 1959;22(4):719–48.
75. F1 score vs ROC AUC vs accuracy vs PR AUC: which evaluation metric should you choose? [Internet]. [cited 2024 Apr 2]. Available from: https://neptune.ai/blog/f1-score-accuracy-roc-auc-pr-auc
76. Ali S, Abuhmed T, El-Sappagh S, Muhammad K, Alonso-Moral JM, Confalonieri R, et al. Explainable Artificial Intelligence (XAI): what we know and what is left to attain trustworthy artificial intelligence. Inf Fusion. 2023;99:101805.

Publisher's note