The OSCAR PUBLISHING Services
American Journal of Applied Science and Technology

VOLUME Vol.05 Issue 06 2025
PAGE NO. 8-12
DOI 10.37547/ajast/Volume05Issue06-02

Opportunities of Artificial Intelligence in The Detection and Prognosis of Viral Hepatitis in Children

Qodirova Dilafruz Abdusamat qizi
2nd-year PhD student at Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

Received: 11 April 2025; **Accepted:** 07 May 2025; **Published:** 09 June 2025

Abstract: This article provides a comprehensive overview of the clinical and laboratory diagnostics of viral hepatitis types (A, B, C, D, E) in children, methods for evaluating viral load, and the potential of artificial intelligence (AI) models for prognosis. It discusses the use of machine learning algorithms like LSTM, GRU, and random forest for analyzing, forecasting, and classifying hepatitis dynamics based on viral load data obtained through serological and molecular testing. Key aspects such as data preparation, platforms, and clinical integration of AI models are also considered.

Keywords: Viral hepatitis, children, artificial intelligence, machine learning, diagnostics, prognosis, medical AI.

Introduction:

Viral hepatitis in children is a serious health issue, as these infections can lead to chronic liver diseases, cirrhosis, and hepatocellular carcinoma. Hepatitis types A, B, C, D, and E exhibit different clinical courses in children and require molecular and immunological tests for diagnosis. In this regard, qPCR, dPCR, ELISA, CLIA, and newer methods such as NGS and LAMP play a central role. In addition, the use of artificial intelligence and machine learning algorithms is increasing in analyzing viral load over time, automating result interpretation, and enabling forecasting. Below is a brief description of the types of viral hepatitis to provide an overview.

Hepatitis A (HAV): HAV is usually transmitted through contaminated water or food. It is mostly asymptomatic in children. It does not progress to a chronic form and lifelong immunity develops after recovery. According to WHO, 90% of children under 10 in countries with poor sanitation are infected with HAV.

Hepatitis B (HBV): Mainly transmitted perinatally or during childhood through blood and body fluids. It becomes chronic in 90% of infants and 30–50% of children aged 1–5. Chronic HBV can lead to cirrhosis and hepatocellular carcinoma. Its spread is decreasing thanks to vaccination.

Hepatitis C (HCV): Transmitted through blood. Often asymptomatic in children, but it progresses to a chronic stage. Approximately 3.5–5 million children live with chronic HCV. There are medications for treatment, but no vaccine is available.

Hepatitis D (HDV): Only infects alongside HBV. It aggravates chronic HBV and increases the risk of liver cancer. HDV can be prevented by HBV vaccination.

Hepatitis E (HEV): Transmitted via the fecal-oral route. It is usually mild in children. Fulminant cases are rare, mainly dangerous for pregnant women. HEV vaccines are available in some countries.

Related Research

In recent years, a number of studies have been conducted on the use of modern technologies, including artificial intelligence (AI) models, in the diagnosis and prognosis of viral hepatitis. According to data from the World Health Organization (WHO), millions of children worldwide are infected with hepatitis B and C viruses, with most cases being asymptomatic and progressing to chronic forms. Therefore, early detection and monitoring are of great importance.

Tang et al. (2018) described methods for accurately determining HBV DNA levels using digital PCR (dPCR) and real-time qPCR in their research. Similarly,

Alavian et al. (2016) presented results on the early detection and monitoring of hepatitis C infections in children using serological tests.

In terms of AI-based approaches, LSTM (Long Short-Term Memory) networks proposed by Hochreiter and Schmidhuber (1997) have been successfully used in medicine for forecasting time-varying data. Cho et al. introduced the GRU (Gated Recurrent Unit) model, which is considered lighter and more efficient compared to LSTM.

Ahmad et al. (2021) developed a random forest model based on HBV data that predicted the likelihood of disease progression with high accuracy (AUC ~0.85). Additionally, it was found that SVM and LSTM models produced significantly lower error rates in predicting hepatitis incidence compared to the ARIMA model.

All of this demonstrates that AI technologies can enhance medical diagnostics and provide effective management of hepatitis infection dynamics. However, for these models to be widely applied in clinical practice, further research is needed concerning data quality, ethical standards, and interpretability challenges.

METHODS

Determining viral hepatitis A, B, C, D, and E in children and assessing their viral load requires a complex and multi-step diagnostic process. Traditional serological and molecular tests provide important clinical indicators of the patient's condition, but they are not sufficient to fully analyze or predict changes in viral load over time. There is a need to use artificial intelligence models to detect viral load early and predict the stages of disease development.

Molecular diagnostics (PCR methods): Quantitative PCR (real-time PCR, qPCR) is used to accurately detect viral load (such as Hepadnavirus DNA or Flavivirus RNA) in the blood. qPCR measures the increase in DNA during the reaction by tracking fluorescence. Digital PCR (dPCR) divides the sample load into thousands of individual mini-reactions, each of which determines whether the virus is present and calculates the exact copy number. These methods are used to detect not only HBV DNA, but also HCV and HDV RNA.

Serological markers: Immunological tests are used to confirm hepatitis at an early stage. For example, in hepatitis A and E, the detection of anti-HAV IgM and anti-HEV IgM antibodies (rapid ELISA/ CLIA) indicates a recent infection. In hepatitis B, the main markers are HBsAg (viral surface antigen) and HBeAg (a marker of high viral activity), the presence of HBsAg indicates the presence of infection. IgM anti-HBc indicates a recent B infection. In the diagnosis of hepatitis C, an anti-HCV antibody test is first used; if it is positive, an HCV-RNA (PCR) test is performed. That is, if anti-HCV is positive, the presence of an active infection (blood virus) is determined using RNA-PCR. The clinic also monitors the level of liver enzymes (ALT, AST), as they indicate liver damage. Together, these biomarkers provide the basis for determining viral infection and assessing viral load.

Test type	Biomarkers	Main purpose	Advantages	Disadvantages
Serologik testlar (ELISA, CLIA)	HBsAg, HBeAg, Anti-HCV, Anti-HAV IgM, Anti-HEV IgM	To detect the presence of the virus	Fast, relatively affordable, widely used	Does not measure viral load precisely, potential for false positive/negative
qPCR (Quantitative PCR)	HBV DNK, HCV RNK, HDV RNK	Quantitative measurement of viral load	Highly accurate, real-time measurements, allows monitoring of viral dynamics	Expensive, requires laboratory infrastructure
dPCR (Digital PCR)	HBV DNK, HCV RNK	Detection of very low viral loads	High precision, detects low loads, highly sensitive	Very expensive, time-consuming

NGS (Next-Generation Sequencing)	HBV, HCV, HDV genomes	Genomic profiling of the virus, identification of mutations	Provides complete genomic data, enables detection of resistance mutations	Very expensive, requires analysis of large datasets
LAMP (Loop-mediated Isothermal Amplification)	HCV RNK, HBV DNK	Rapid detection of viral load	Affordable, quick, does not require complex instruments	Less sensitive, prone to false positives
ALT/AST o'lchovlari	ALT, AST (liver enzymes)	Evaluation of liver damage	Affordable, easy to use	Does not detect viral load, only indicates liver damage
FibroScan / FibroTest	Liver fibrosis stage	Assessment of liver tissue damage	Non-invasive, easy to apply	Does not measure viral load, only assesses liver condition

Table 1. Serological and molecular diagnostics of viral hepatitis types

In addition, there are several artificial intelligence models for the detection of viral hepatitis. LSTM (Long Short-Term Memory) network: This is a type of recurrent neural network specifically designed for working with sequential data (time series). An LSTM network learns long-term relationships between data over time. For example, LSTM is useful for predicting future viral levels based on previously obtained viral loads and clinical indicators in children. An LSTM consists of a single cell, in which the "input", "forget" and "output" capabilities control the flow of information. With the help of these capabilities, the network remembers the important signal and discards unnecessary information, thus making predictions based on the time series.

GRU (Gated Recurrent Unit): GRU is also a type of RNN designed for sequential data like LSTM, but its internal structure is somewhat simpler. It consists only of "update" and "forget" capabilities, so it has fewer parameters. GRU also learns long-term relationships between data, but is computationally lighter than LSTM.

Regression models: Simple regression (linear regression) is used to study linear relationships in data. For example, it is possible to predict the amount of virus (count) based on the ALT level of children, age, and other clinical parameters. Logistic regression predicts the outcome in terms of binary (infected/not infected). Regression models are intuitive and fast, but may sometimes not be sufficient to model complex time series well.

Other ML methods: Classical algorithms such as decision trees and random forests, SVM, and k-NN have also been applied to hepatitis data. For example, in one study, a random forest model based on laboratory parameters predicted the development of HBV disease with high accuracy (AUC ~0.85). However, nonlinear models such as LSTM and SVM performed better in predicting hepatitis incidence than classical ARIMA (LSTM showed the lowest error). Since the problem of viral load prediction is usually related to the analysis of time-varying data, RNN models such as LSTM/GRU are preferable.

Model type	Task type	Advantages	Disadvantages	Best application area
LSTM / GRU	Time series analysis, forecasting	Learns long-term dependencies, understands dynamic	Learns long-term dependencies, understands dynamic	Forecasting viral load over time

Model type	Task type	Advantages	Disadvantages	Best application area
		changes	changes	
CNN (Convolutional Neural Network)	Image analysis, segmentation	Automatic feature extraction, high accuracy	Requires large datasets and GPUs	Detecting liver diseases from biopsy images
Transformer (BERT, GPT)	Text analysis, NLP	Parallel processing, deep contextual understanding	Requires massive data and computational resources	Analyzing clinical notes related to viral load
Random Forest	Classification, regression	Low risk of overfitting, easy to interpret	Slows down with many trees	Classifying hepatitis types based on clinical features
XGBoost / LightGBM	Classification, regression, forecast	Very fast, high accuracy, handles noise well	Hard to learn, many parameters	Exploring relationships between parameters for viral load prediction
AutoML (H2O.ai, AutoKeras)	Model creation and optimization	Easy to use, no-code ML, rapid prototyping	Limited for complex datasets	Building simple models to predict hepatitis viral load
ARIMA / SARIMA	Time series analysis	Simple, intuitive, effective for small datasets	Not suitable for non-linear data	Month-by-month forecasting of hepatitis viral load
SVM (Support Vector Machine)	Classification, regression	Good for non-linear data, high accuracy	Slow with large datasets	Classifying viral load into low/high levels
DNN (Deep Neural Network)	Complex classification and prediction	Learns complex relationships, high accuracy	Requires very large datasets	Predicting hepatitis based on multiple biomarkers
FastAPI / Flask	Model deployment	Easy API development, convenient integration	Not suitable for complex scaling	Creating APIs for viral load prediction
Docker / Kubernetes	Model containerization and scaling	Easy integration, independent execution	Complex to learn	Deploying AI models for large-scale data environments

Table 2. Types of Artificial Intelligence Models and their Applications

RESULTS

Data Types: The data used in the prediction model consist of clinical and laboratory parameters. Clinical parameters include children's age, gender, medical history, symptoms, and physical examination results; laboratory parameters include liver enzymes (ALT, AST), hematological parameters (hemoglobin, plt), viral biomarkers (HBV DNA, HCV RNA), and other tests; and time-series variables include the above

parameters obtained from multiple measurements. Longitudinal data (e.g., viral levels at hospital admission) are processed using a time-series model.

Data Preparation: First, beginners need to clean the data: fill in missing or incorrect values (e.g., with previous or subsequent data), remove unnecessary columns, and check for anomalous values. Numerical data (e.g., ALT scores) are normalized (standardized, normalized) – this is important for machine learning. Categorical data (e.g., gender, symptoms) should be converted to a numeric format (one-hot encode or numeric codes). If there is a time series, the data is sorted in time order and, if necessary, processed together with indicators obtained over a period of time. If the database is large, it is also useful to first select the "most important" columns (feature selection).

Tools for building and training AI models

AutoML tools: For beginners in the software industry, there are AutoML platforms (e.g., H2O.ai, TPOT, AutoKeras, AutoGluon, Ludwig, and others). These tools automatically analyze the data, select the most suitable model, and make it easy to adjust its parameters (hyperparameter tuning). For example, TPOT offers tree-based optimization, and AutoKeras offers auto-building of graph block-type neural networks.

Python libraries: The most commonly used libraries for AI and ML model building are scikitlearn (simple regression, classification, preprocessing), pandas and NumPy (data analysis and preparation), TensorFlow and Keras or PyTorch (building differentiated neural network architectures). For new learners, scikit-learn has a lighter interface and is convenient for quickly testing statistical machine learning and regression models. It is easy to experiment with these tools in the Jupyter Notebook environment. It is also recommended to use matplotlib or seaborn for data visualization.

Brief programming environments: If you do not want to write code, there are also graphical interface platforms such as Google AutoML, Azure AutoML for beginners. They allow you to visually control the machine learning process. However, for advanced research, it is usually preferable to work with the above Python libraries.

CONCLUSION

Molecular and serological tests play a crucial role in the analysis of viral load, but their predictive value is limited. Artificial intelligence models, especially LSTM and GRU, show high efficiency in forecasting taking into account time variability. Classical ML algorithms such as Random Forest can be a tool for automating diagnostics and assisting doctors. However, the application of AI models in healthcare poses challenges in terms of data security, clinical validation, and interpretation.

In the future, AI-based decision-making systems are expected to be widely implemented in clinical practice, but this process should be carried out with caution.

REFERENCES

World Health Organization (WHO). Hepatitis A. https://www.who.int/news-room/fact-sheets/detail/hepatitis-a

World Health Organization (WHO). Hepatitis B. https://www.who.int/news-room/fact-sheets/detail/hepatitis-b

World Health Organization (WHO). Hepatitis C. https://www.who.int/news-room/fact-sheets/detail/hepatitis-c

WHO. Hepatitis D and E. https://www.who.int/news-room/fact-sheets/detail/hepatitis-d

Tang et al., "Digital PCR and real-time PCR in detecting HBV DNA," Journal of Clinical Virology, 2018.

Alavian SM et al. "Epidemiology of Hepatitis C in children." Pediatrics Infect Dis J, 2016.

Hochreiter & Schmidhuber. "Long Short-Term Memory." Neural Computation, 1997.

Cho et al., "Learning Phrase Representations using RNN Encoder–Decoder," arXiv, 2014.

Breiman, L. "Random Forests." Machine Learning, 2001.

Ahmad et al., "Machine learning methods for predicting hepatitis." Health Informatics Journal, 2021.

AutoML platforms. Google, H2O.ai, AutoKeras documentation.

Chollet F. Deep Learning with Python. Manning, 2018.

Pedregosa et al., "Scikit-learn: Machine Learning in Python." JMLR, 2011.