Development of a Liver Disease-Specific Large Language Model Chat Interface using Retrieval Augmented Generation

Jin Ge, MD, MBA; Steve Sun; Joseph Owens, PhD; Victor Galvez; Oksana Gologorskaya, MS; Jennifer C. Lai, MD, MBA; Mark J. Pletcher, MD, MPH; Kl Lai, MS

1 Division of Gastroenterology and Hepatology, Department of Medicine, University of California – San Francisco, San Francisco, CA
2 UCSF Health Information Technology, University of California – San Francisco, San Francisco, CA
3 Bakar Computational Health Sciences Institute, University of California – San Francisco, San Francisco, CA
4 Department of Epidemiology and Biostatistics, University of California – San Francisco, San Francisco, CA

Financial/Grant Support:
The authors of this study were supported in part by the KL2TR001870 (National Center for Advancing Translational Sciences, Ge), P30DK026743 (UCSF Liver Center Grant, Ge and Lai), UL1TR001872 (National Center for Advancing Translational Sciences, Pletcher), and R01AG059183/K24AG080021 (National Institute on Aging, Lai). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or any other funding agencies. The funding agencies played no role in the analysis of the data or the preparation of this manuscript.

Abbreviations:
AASLD, American Association for the Study of Liver Diseases; APIs, application programming interfaces; FHIR, Fast Healthcare Interoperability Resources; HBV, hepatitis B; HCC, hepatocellular carcinoma; GAI, generative artificial intelligence; GPT, generative pre-trained transformer; LLMs, large language models; PHI, protected health information; RAG, retrieval-augmented generation; UCSF, University of California, San Francisco.

Correspondence:
Jin Ge, MD, MBA
513 Parnassus Avenue, S-357

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Disclosures:
The authors of this manuscript have the following potential conflicts of interest to disclose:

- Dr. Jin Ge receives research support from Merck and Co; and consults for Astellas Pharmaceuticals/Iota Biosciences.
- Dr. Jennifer C. Lai receives research support from Lipocene and Vir Biotechnologies; receives an education grant from Nestle Nutrition Sciences; serves on an advisory board for Novo Nordisk; and consults for Genfit, Third Rock Ventures, and Boehringer Ingelheim.

Writing Assistance:
None.

Author Contributions:
Authorship was determined using ICMJE recommendations.

Ge: Concept and design; data extraction; analysis and interpretation of data; drafting of manuscript; critical revision of the manuscript for important intellectual content; study supervision

Sun: Data extraction; analysis and interpretation of data; critical revision of the manuscript for important intellectual content

Owens: Concept and design; analysis and interpretation of data; critical revision of the manuscript for important intellectual content

Galvez: Analysis and interpretation of data; critical revision of the manuscript for important intellectual content

Gologorskaya: Analysis and interpretation of data; critical revision of the manuscript for important intellectual content

J. Lai: Concept and design; critical revision of the manuscript for important intellectual content; study supervision

Pletcher: Concept and design; critical revision of the manuscript for important intellectual content; study supervision
K. Lai: Concept and design; critical revision of the manuscript for important intellectual content; study supervision

Acknowledgements:
- The authors thank the UCSF AI Tiger Team, Academic Research Services, Research Information Technology, and the Chancellor’s Task Force for Generative AI for their software development, analytical, and technical support related to the use of Versa API gateway (the UCSF secure implementation of large language models and generative AI via API gateway), Versa chat (the chat user interface), and related data assets.

Word Count:
2,025 (without references)
**Abstract:**

**Background:** Large language models (LLMs) have significant capabilities in clinical information processing tasks. Commercially available LLMs, however, are not optimized for clinical uses and are prone to generating incorrect or hallucinatory information. Retrieval-augmented generation (RAG) is an enterprise architecture that allows embedding of customized data into LLMs. This approach “specializes” the LLMs and is thought to reduce hallucinations.

**Methods:** We developed “LiVersa,” a liver disease-specific LLM, by using our institution’s protected health information (PHI)-complaint text embedding and LLM platform, “Versa.” We conducted RAG on 30 publicly available American Association for the Study of Liver Diseases (AASLD) guidelines and guidance documents to be incorporated into LiVersa. We evaluated LiVersa’s performance by comparing its responses versus those of trainees from a previously published knowledge assessment study regarding hepatitis B (HBV) treatment and hepatocellular carcinoma (HCC) surveillance.

**Results:** LiVersa answered all 10 questions correctly when forced to provide a “yes” or “no” answer. Full detailed responses with justifications and rationales, however, were not completely correct for three of the questions.

**Discussions:** In this study, we demonstrated the ability to build disease-specific and PHI-compliant LLMs using RAG. While our LLM, LiVersa, demonstrated more specificity in answering questions related to clinical hepatology – there were some knowledge deficiencies due to limitations set by the number and types of documents used for RAG. The LiVersa prototype, however, is a proof of concept for utilizing RAG to customize LLMs for clinical uses and a potential strategy to realize personalized medicine in the future.
Introduction:

Large language models (LLMs), such as OpenAI’s Generative Pre-trained Transformer (GPT) family of models, have demonstrated significant capabilities in clinical information processing tasks, such as data extraction, summarizing literature, content generation, and predictive modeling. Commerically available general-purpose LLMs, such as OpenAI’s ChatGPT, however, are trained on publicly available data and not optimized for clinical uses. This means that the outputs of publicly available LLMs when prompted with clinical questions may include incorrect, incomplete, or hallucinatory information. Despite these limitations, LLMs are thought to have significant potential in biomedical and clinical applications. This is because the practice of modern medicine is highly complex endeavor with ever-increasing amounts of knowledge generated yearly. For instance, it was estimated that two papers were uploaded to PubMed every minute in 2016, a figure that has undoubtedly increased in the seven years since. National practice societies have also responded to the ever-expanding body of clinical knowledge by developing comprehensive practice guidelines. For instance, the American Association of the Study of Liver Diseases (AASLD) has issued guidelines and guidance documents for 26 liver diseases and conditions, and two quality measures concerning cirrhosis and hepatocellular carcinoma (HCC).

Methods to incorporate medical literature, such as clinical practice guidelines and guidance documents, into LLM responses, therefore, represent growing area of interest to help general-purpose LLMs become “specialized” for clinically-focused applications.

Currently, there are three general approaches and techniques to allow LLM “specialization:” 1. Fine-tune the original LLM model, which is computationally expensive; 2. Prompting within the LLM, which only accommodates small amounts of data and requires iterative user input; and 3. Retrieval-augmented generation (RAG). RAG is an enterprise architecture that augments LLM abilities by adding an information retrieval system that provides external data, which then supplement and constrain LLM output. In practice, this means that a dataset, such as a compendium of clinical practice guidelines, is vectorized and encoded using embedding models and then incorporated into the LLM by layering it on top of the LLM information retrieval and output processes. The theoretical advantages of RAG are two-fold: 1. Handling large numbers of documents to provide “ground knowledge” for the LLM, and 2. Decreasing hallucinations by limiting the potential “solution-space” for LLM outputs. This RAG approach has been proposed and utilized in other clinical specialties, such as general medicine and urology. Existing RAG implementations, however, have largely been on publicly available and not protected health information [PHI]-compliant LLMs – thereby limiting permitted use in clinical practice.
In this study, we utilized RAG to create a prototype liver disease-specific LLM, called “LiVersa,” within the University of California, San Francisco’s (UCSF) PHI-compliant implementation of Microsoft OpenAI GPT family of LLM models, “Versa.”

**Methods:**
This study was not considered human subjects research because no human data or specimens were used. Consequently, approval from the UCSF Institutional Review Board was not sought. To construct our liver disease-specific LLM (LiVersa), we used all available provider-facing clinical practice guidelines, guidance documents, and quality measure documents published by AASLD on its website. We excluded introduction and executive summary documents as the content in these were often duplicated in the corresponding full-length versions (e.g., for Wilson’s Disease). We included both original guidelines/guidance documents and updates if both were publicly available for download on AASLD website (e.g., for Chronic Hepatitis B [HBV]). Patient-facing and/or outdated guidelines/guidance documents were not included in this study. In total, 30 documents were retrieved to be incorporated into the RAG process (Table 1).

We utilized application programming interfaces (APIs) provided by the Microsoft Azure OpenAI Cognitive Search suite of tools to incorporate the AASLD guidelines and documents for RAG (Figure 1). In the pre-processing phase, the 30 AASLD guidelines and documents in PDF format were transformed into embeddings using Microsoft Azure OpenAI’s ADA Text Embedding Version 2 model (text-embedding-ada-002). Text embeddings are numerical representations of text where words for phrases are represented as multi-dimensional vectors. These embedding vectors are then stored in a database with the Microsoft Azure Cognitive Search services. During chat interactions with LiVersa, users’ prompts are converted into embeddings in real time using the same text-embedding-ada-002 model. A search is then performed on the previously processed vector database of AASLD guidelines and documents to find matches for the prompt embeddings. The search results from this search process are then passed to the gpt-35-turbo or gpt-4-32k LLMs to generate a completion, which is the output from the LLM in the LiVersa chat interface.

We pre-set three sample questions into “Sample Questions” section of the LiVersa interface:

1. What are the indications for liver transplantation?
2. Who should be screened for chronic hepatitis B?

3. What is the recommended therapy for a patient with BCLC 0-A HCC without portal hypertension?

To compare LiVersa's performance on free-form clinical hepatology questions, we compared its responses to medical trainees' responses in a previously published case-vignette based knowledge assessment questions on HBV treatment and hepatocellular carcinoma (HCC) surveillance. We chose this set of questions because they have also been evaluated by the publicly available ChatGPT. Prior to input into LiVersa, case-vignettes were edited for clarity and appropriateness with the goal of retaining the essence of the clinical scenario. We asked LiVersa to give both a full response and a forced “yes” or “no” response by appending the prompt with “Answer yes or no” to each inputted case-vignette (Figure 2).

**Results:**

The user interface for LiVersa is shown in Figure 2. The interface is the same as to the one used for “Versa Chat” (the chat interface for the general-purpose PHI-compliant LLM implementation at UCSF), except with the data source being set to the “UCSF Intranet: AASLD Clinical Practice Guidelines (Prototype).” LiVersa's responses to the three sample questions are featured in Table 2. LiVersa's performance on case-vignette based knowledge assessment questions on HBV treatment and HCC surveillance are featured in Table 3, Columns B and C. The correct answers along with percentages of trainees who answered correctly per prior studies are featured in Table 3, Column D. LiVersa answered all 10 questions correctly when forced to provide a “yes” or “no” answer. Rationales and justifications within the full answers, however, were not completely correct for three clinical scenarios in HCC surveillance. These three clinical scenarios were: “A 25-year-old Haitian man with chronic hepatitis B, on treatment with entecavir,” “A 40-year-old Cuban woman who was recently diagnosed with hepatitis B after developing jaundice in the setting of a surgical procedure. There is no evidence of cirrhosis,” and “A 40-year-old woman from Thailand with cirrhosis and chronic inactive hepatitis B.”

**Discussion:**

Two of the most significant barriers to using general-purpose LLMs like ChatGPT in clinical practice are the tendency for these LLMs to “hallucinate” or generate confident sounding but false responses, and their lack of compatibility with PHI. The first issue stems from the nature of the transformer architecture, which optimizes the LLM's objective of predicting the most probable next word (token).
from its pre-trained data without consideration for accuracy. Augmenting the knowledge base of general-purpose LLMs, such as through RAG, becomes one of the key strategies to shaping and constraining LLM outputs to prevent false information from being propagated and disseminated. Our work is one of the first demonstrations and proof-of-concept of using RAG to create a liver disease-specific and, more importantly, PHI-compliant LLM chat interface. Our incorporation of hepatology specific knowledge, such as the AASLD practice guidelines, resulted in the LiVersa chat interface generating answers that were likely more specific than those generated by general-purpose ChatGPT.24

While LiVersa answered all 10 questions regarding HBV treatment and HCC surveillance in our test set correctly, the rationales given for three cases were not completely correct. These incorrect responses reflect limitations of utilizing RAG. For instance, LiVersa’s stated rationale for a “Yes” recommendation for the case-vignette of “A 25-year-old Haitian man with chronic hepatitis B, on treatment with entecavir” was based on the presumption that the patient was from an HBV endemic country (Table 3). The actual rationale given in prior literature is based on the 2018 version of the AASLD HCC practice guidance document, which recommended early surveillance for African HBV patients and/or North American HBV patients of African-American descent.

In this example, we highlight two key issues. First, the 2018 version of the AASLD HCC practice guidance document was not uploaded into LiVersa’s RAG dataset, the 2023 version was.26,27 LiVersa gave the wrong justification for HCC surveillance in this case simply because the 2018 version was not available to it and “restrictions” placed by the RAG dataset forced it to generate a hallucinatory answer. This limitation could easily be overcome by increasing the number and variety of literature incorporated into LiVersa’s RAG dataset. Additional documents that could be considered for inclusion in the future include clinical practice guidelines/guidance documents from other societies, such as those from the American Gastroenterological Association, American College of Gastroenterology, European Association for the Study of the Liver, and Asian Pacific Association for the Study of the Liver; and from compendium sources, such as the Cochrane Review and UpToDate.

The second key issue is contextual bias. In this instance, the recommendation for HCC surveillance is based on the assumption that the man from Haiti in the clinical scenario is of African descent. This is a problem that affects both literature data incorporated into the RAG dataset and in the pre-training data that underpins the LLM itself. The more “correct” response should have been to ask for additional
context to clarify the self-identified national-origin or racial/ethnic background of the patient to allow the LLM to give a more comprehensive recommendation/output. This is an illustration of the problems concerning algorithm bias that generative artificial intelligence (GAI) research community is actively grappling with. Finally, in addition to the two key considerations noted above, there is a minor limitation with regards to our grading of LiVersa’s responses. We had edited the case-vignette knowledge assessment questions for clarity and appropriateness. While we tried to retain the intent of the clinical cases, the wording of these cases may not be the same as those evaluated by trainees and ChatGPT in prior studies. This may have affected the comparisons in Table 3.

Despite these limitations, we demonstrated that RAG could be a powerful method to create a “specialized” LLM. Given that our LiVersa prototype was developed and deployed in a PHI-compliant environment (UCSF’s Versa), it could theoretically be used clinically to evaluate actual patient scenarios. By extension, there is a potential to incorporate both literature and patient data from the electronic health record through extraction by the Fast Healthcare Interoperability Resources (FHIR) API in the RAG process. This process would allow for the creation of patient-specific clinical LLMs that could be a true realization of GAI-enabled personalized medicine.
<table>
<thead>
<tr>
<th>#</th>
<th>Journal</th>
<th>Year</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Hepatology</td>
<td>2020</td>
<td>Diagnosis and Treatment of Alcohol-Associated Liver Diseases: 2019 Practice Guidance From the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>3</td>
<td>Hepatology</td>
<td>2021</td>
<td>Diagnosis, Evaluation, and Management of Ascites, Spontaneous Bacterial Peritonitis and Hepatorenal Syndrome: 2021 Practice Guidance by the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>4</td>
<td>Hepatology</td>
<td>2020</td>
<td>Diagnosis and Management of Autoimmune Hepatitis in Adults and Children: 2019 Practice Guidance and Guidelines From the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>5</td>
<td>Hepatology</td>
<td>2023</td>
<td>AASLD practice guidance on drug, herbal, and dietary supplement-induced liver injury</td>
</tr>
<tr>
<td>6</td>
<td>Hepatology</td>
<td>2011</td>
<td>Diagnosis and management of hemochromatosis: 2011 Practice Guideline by the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>7</td>
<td>Hepatology</td>
<td>2014</td>
<td>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</td>
</tr>
<tr>
<td>8</td>
<td>Hepatology</td>
<td>2016</td>
<td>AASLD guidelines for treatment of chronic hepatitis B</td>
</tr>
<tr>
<td>9</td>
<td>Hepatology</td>
<td>2018</td>
<td>Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance</td>
</tr>
<tr>
<td>10</td>
<td>Clinical Infectious Diseases</td>
<td>2023</td>
<td>Hepatitis C Guidance 2023 Update: American Association for the Study of Liver Diseases–Infectious Diseases Society of America Recommendations for Testing, Managing, and Treating Hepatitis C Virus Infection</td>
</tr>
<tr>
<td>11</td>
<td>Hepatology</td>
<td>2023</td>
<td>AASLD Practice Guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma</td>
</tr>
<tr>
<td>12</td>
<td>Hepatology</td>
<td>2009</td>
<td>AASLD Position Paper: Liver biopsy</td>
</tr>
<tr>
<td>13</td>
<td>Hepatology</td>
<td>2014</td>
<td>Evaluation for liver transplantation in adults: 2013 practice guideline by the American Association for the Study of Liver Diseases and the American Society of Transplantation</td>
</tr>
<tr>
<td>15</td>
<td>Liver Transplantation</td>
<td>2013</td>
<td>Long-Term Management of the Successful Adult Liver Transplant: 2012 Practice Guideline by the American Association for the Study of Liver Diseases and the American Society of Transplantation</td>
</tr>
<tr>
<td>16</td>
<td>Liver Transplantation</td>
<td>2013</td>
<td>Long-term medical management of the pediatric patient after liver transplantation: 2013 practice guideline by the American Association for the Study of Liver Diseases and the American Society of Transplantation</td>
</tr>
<tr>
<td>17</td>
<td>Hepatology</td>
<td>2021</td>
<td>Malnutrition, Frailty, and Sarcopenia in Patients With Cirrhosis: 2021 Practice Guidance by the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>18</td>
<td>Hepatology</td>
<td>2023</td>
<td>AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease</td>
</tr>
<tr>
<td>19</td>
<td>Hepatology</td>
<td>2022</td>
<td>AASLD Practice Guidance: Palliative care and symptom-based management in decompensated cirrhosis</td>
</tr>
<tr>
<td>20</td>
<td>Hepatology</td>
<td>2023</td>
<td>AASLD practice guidance on risk stratification and management of portal hypertension and varices in cirrhosis</td>
</tr>
<tr>
<td>22</td>
<td>Hepatology</td>
<td>2019</td>
<td>Primary Biliary Cholangitis: 2018 Practice Guidance from the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>23</td>
<td>Hepatology</td>
<td>2022</td>
<td>Primary biliary cholangitis: 2021 practice guidance update from the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td></td>
<td>Journal</td>
<td>Year</td>
<td>Title</td>
</tr>
<tr>
<td>---</td>
<td>--------------------------</td>
<td>------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>24</td>
<td>Hepatology</td>
<td>2023</td>
<td>AASLD practice guidance on primary sclerosing cholangitis and cholangiocarcinoma</td>
</tr>
<tr>
<td>25</td>
<td>Hepatology</td>
<td>2021</td>
<td>Reproductive Health and Liver Disease: Practice Guidance by the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>26</td>
<td>Hepatology</td>
<td>2021</td>
<td>Vascular Liver Disorders, Portal Vein Thrombosis, and Procedural Bleeding in Patients With Liver Disease: 2020 Practice Guidance by the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>27</td>
<td>Hepatology</td>
<td>2022</td>
<td>A multidisciplinary approach to the diagnosis and management of Wilson disease: 2022 Practice Guidance on Wilson disease from the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>28</td>
<td>Hepatology</td>
<td>2018</td>
<td>Development of Quality Measures in Cirrhosis by the Practice Metrics Committee of the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>29</td>
<td>Hepatology</td>
<td>2021</td>
<td>Quality measures in HCC care by the Practice Metrics Committee of the American Association for the Study of Liver Diseases</td>
</tr>
<tr>
<td>30</td>
<td>Hepatology</td>
<td>2023</td>
<td>AASLD Practice Guidance on Acute-on-chronic Liver Failure and the Management of Critically Ill Patients with Cirrhosis</td>
</tr>
</tbody>
</table>
Table 2 – LiVersa Responses to “Sample Questions” Embedded in the Chat Interface

<table>
<thead>
<tr>
<th>Sample Question</th>
<th>LiVersa Generated Answer</th>
</tr>
</thead>
</table>
| **What are the indications for liver transplantation?** | The indications for liver transplantation include severe acute or advanced chronic liver disease when the limits of medical therapy have been reached. Specific conditions that may necessitate a liver transplant include:  
  - Acute Liver Failure  
  - Complications of cirrhosis such as ascites, chronic gastrointestinal blood loss due to portal hypertensive gastropathy, encephalopathy, refractory variceal hemorrhage, and synthetic dysfunction  
  - Liver cancer  
  - Liver-based metabolic conditions with systemic manifestations such as alpha-1 antitrypsin deficiency, familial amyloidosis, glycogen storage disease, hemochromatosis, primary oxaluria, and Wilson disease  
  - Systemic complications of chronic liver disease such as hepatopulmonary syndrome and portopulmonary hypertension  
  Evaluation for liver transplantation should be considered once a patient with cirrhosis has experienced an index complication such as ascites, hepatic encephalopathy, or variceal hemorrhage or hepatocellular dysfunction results in a MELD Score > 15. Potential liver transplant candidates with worsening renal dysfunction or other evidence of rapid hepatic decompensation should have prompt evaluation for liver transplant. |
| **Who should be screened for chronic hepatitis B?** | Screening for chronic Hepatitis B is recommended for the following groups:  
  1. Persons born in regions of high or intermediate HBV endemicity (HBsAg prevalence of ≥2%)  
  2. U.S.-born persons not vaccinated as an infant whose parents were born in regions with high HBV endemicity (≥8%)  
  3. Persons who have ever injected drugs  
  4. Men who have sex with men  
  5. Persons needing immunosuppressive therapy, including chemotherapy, immunosuppression related to organ transplantation, and immunosuppression for rheumatological or gastroenterologic disorders  
  6. Individuals with elevated ALT or AST of unknown etiology  
  7. Donors of blood, plasma, organs, tissues, or semen  
  8. Persons with end-stage renal disease, including predialysis, hemodialysis, peritoneal dialysis, and home dialysis patients  
  9. All pregnant women  
  10. Infants born to HBsAg-positive mothers  
  11. Persons with chronic liver disease, e.g., HCV  
  12. Persons with HIV  
  13. Household, needle-sharing, and sexual contacts of HBsAg-positive persons  
  14. Persons who are not in a long-term, mutually monogamous relationship (e.g., >1 sex partner during the previous 6 months)  
  15. Persons seeking evaluation or treatment for a sexually transmitted disease  
  16. Health care and public safety workers at risk for occupational exposure to blood or blood-contaminated body fluids  
  17. Residents and staff of facilities for developmentally disabled persons  
  18. Travelers to countries with intermediate or high prevalence of HBV infection  
  19. Persons who are the source of blood or body fluid exposures that might require postexposure prophylaxis  
  20. Inmates of correctional facilities  
  21. Unvaccinated persons with diabetes who are aged 19 through 59 years (discretion of clinician for unvaccinated adults with diabetes who are aged ≥60 years). |
<p>| <strong>What is the recommended therapy for a patient with BCLC 0- A HCC without portal hypertension?</strong> | The expert panel recommends that surgical resection should be considered in relevant cases for patients with BCLC 0- A HCC without portal hypertension. |</p>
<table>
<thead>
<tr>
<th>Case Vignette (Edited for Clarity and Appropriateness)</th>
<th>LiVersa Full Response</th>
<th>Correct Answer</th>
<th>Amongst Trainees</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Does the following patient meet criteria for hepatitis B treatment?</strong></td>
<td>Does the following patient meet criteria for hepatitis B treatment?</td>
<td><strong>Yes</strong></td>
<td><strong>No</strong></td>
</tr>
<tr>
<td><strong>37-year-old man without cirrhosis infected with Hepatitis B and lab values showing HBeAg positive, ALT 83 U/L, and DNA Viral Load at 22,000 units/mL that are unchanged after 8 months.</strong></td>
<td><strong>Yes</strong></td>
<td><strong>Yes</strong></td>
<td><strong>78.9%</strong></td>
</tr>
<tr>
<td><strong>55-year-old woman with compensated cirrhosis secondary to Hepatitis B with lab values showing HBeAg negative, DNA Viral Load at 4,000 units/mL and ALT 35 U/L.</strong></td>
<td><strong>Yes</strong></td>
<td><strong>Yes</strong></td>
<td><strong>49.2%</strong></td>
</tr>
<tr>
<td><strong>32-year-old man without cirrhosis infected with Hepatitis B with lab values showing HBeAg positive, DNA Viral Load at 8000 units/mL and ALT 22 U/L.</strong></td>
<td><strong>No</strong></td>
<td><strong>No</strong></td>
<td><strong>22.2%</strong></td>
</tr>
<tr>
<td><strong>Should the following patient be screened for hepatocellular carcinoma?</strong></td>
<td>Should the following patient be screened for hepatocellular carcinoma?</td>
<td><strong>Yes</strong></td>
<td><strong>No</strong></td>
</tr>
<tr>
<td><strong>A 25-year-old Haitian man with chronic hepatitis B, on treatment with entecavir.</strong>*</td>
<td><strong>Yes</strong></td>
<td><strong>Yes</strong></td>
<td><strong>71.1%</strong></td>
</tr>
<tr>
<td><strong>A 40-year-old Cuban woman who was recently diagnosed with chronic hepatitis B and has ALT 70 U/L and HBsAg positive.</strong></td>
<td><strong>Yes</strong></td>
<td><strong>No</strong></td>
<td><strong>43.4%</strong></td>
</tr>
<tr>
<td><strong>A 25-year-old Vietnamese man with past history of hepatitis B with surface antigen positive, ALT 45 U/L and ALT 25 U/L and AST/ALT ratio of 3.</strong></td>
<td><strong>Yes</strong></td>
<td><strong>No</strong></td>
<td><strong>46.7%</strong></td>
</tr>
</tbody>
</table>

*Note: Direct-acting antiviral (DAV) therapy is recommended for patients with chronic hepatitis B and evidence of hepatic decompensation (Child-Pugh B or C), severe liver fibrosis or cirrhosis, or without contraindications to antiviral therapy.
**hepatitis B surface antibody positive).**

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Hepatocellular carcinoma screening according to the provided guidelines.</th>
<th>Yes</th>
<th>Yes - 76.8%</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 35-year-old Chinese woman with platelets of 80,000, INR of 1.1, bilirubin of 1.4 mg/dL. On ultrasound, she was noted to have mild ascites and a spleen that measures 14 cm.</td>
<td>Based on the provided context, the patient should be considered for hepatocellular carcinoma (HCC) screening. The presence of ascites and a low platelet count may suggest advanced liver disease or cirrhosis. Regular screening for HCC with cross-sectional imaging at 6-month intervals is advised for patients with cirrhosis. However, the final decision should be made by a healthcare professional considering the patient’s overall health status and other risk factors.</td>
<td>Yes</td>
<td>Yes - 34.8%</td>
</tr>
<tr>
<td>A 40-year-old woman from Laos with chronic inactive hepatitis B without cirrhosis.</td>
<td>Based on the provided information, a 40-year-old woman from Laos (an endemic country as defined by AASLD hepatitis B virus guidance) with chronic inactive hepatitis B without cirrhosis does not meet the criteria for sufficient risk to warrant surveillance for hepatocellular carcinoma (HCC). The guidelines suggest surveillance for women from endemic countries when they are over the age of 50.</td>
<td>No</td>
<td>No - 81.2%</td>
</tr>
<tr>
<td>A 40-year-old woman from Thailand with cirrhosis and chronic inactive hepatitis B.*</td>
<td>Yes, the patient should be screened for hepatocellular carcinoma. According to the provided context, individuals with Child–Pugh A–B cirrhosis, any etiology, and chronic hepatitis B are at sufficient risk to warrant surveillance for hepatocellular carcinoma. Additionally, the patient is from an endemic country (as defined by AASLD hepatitis B virus guidance) and is over the age of 40, which further supports the need for screening.</td>
<td>Yes</td>
<td>Yes - 65.8%</td>
</tr>
<tr>
<td>A 40-year-old Nigerian woman with chronic hepatitis B with normal liver function tests.</td>
<td>Yes, this patient should be screened for hepatocellular carcinoma. The guidelines suggest that a person from Africa, regardless of gender, should start surveillance as early as the third decade of life given the median age of 46 years at HCC diagnosis. This is especially relevant if the person has chronic hepatitis B, which is a known risk factor for hepatocellular carcinoma.</td>
<td>Yes</td>
<td>Yes - 65.8%</td>
</tr>
</tbody>
</table>

*LiVensa's generated answer and rationale may not be completely correct in these cases.*
Figure 1 – Diagram for Retrieval Augmented Generated as Executed Through Azure Cognitive Search

Pre-Processing Phase

AASLD Guidelines

PDF

Azure OpenAI ADA Text Embedding Version 2 Model

Embellishments

0.1 0.3 0.2
0.7 0.4 0.1
0.8 0.2 0.5

Search Results

Chat Phase

Azure Cognitive Services

Prompt Embeddings

0.3 0.1 0.7

Azure OpenAI ADA Text Embedding Version 2 Model

Completion

Azure OpenAI GPT-4 (32K) Model
**Figure 2 – LiVerse Chat Interface**

The LiVerse Chat Interface allows users to interact with a chatbot that provides medical information. The interface includes options to select various engines, data sources, and large language models. Sample questions are posed, and the chatbot provides responses based on the selected medical guidelines and clinical practice guidelines.

**References are as follows:**
- Hepatology - 2018 - Update on Prevention, Diagnosis, and Treatment of Chronic Hepatitis B.pdf

Enter your question here.
References:


