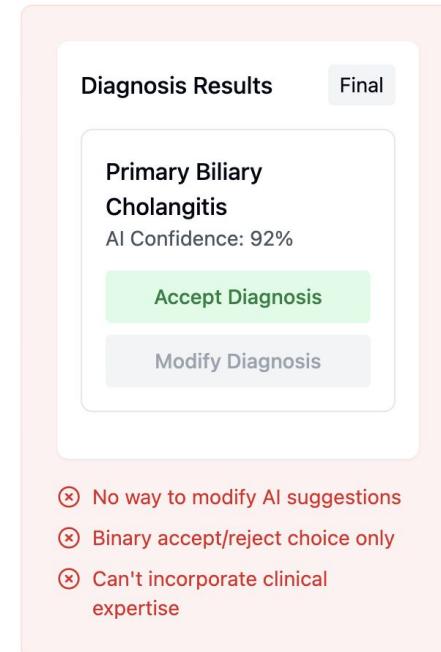
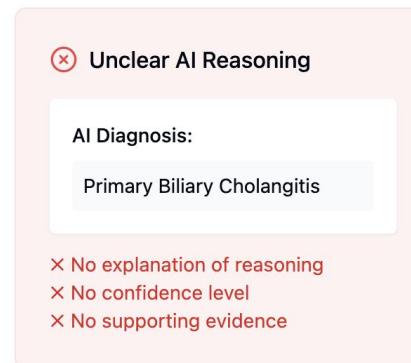


Hepnovate Experimental Findings

AI-Assisted Liver Disease Diagnosis

Recap: HFAI Problems and Avenues for Improvement

- **Lack of time efficiency**
 - Manually typing diagnosis features into ChatGPT or OpenEvidence
- **Black box problem**
 - Lack of explainability
 - Trust in AI



... so we have Hepnovate

Hepnovate

Hepnovate

Liver Disease Diagnostics, Powered by Human + AI.

Current Symptoms

- Pale skin
- Abdominal pain
- Fever

Current Vital Signs

Temperature	102°F
Blood Pressure	90/120 mmHg
Pulse	Value

Live Transcription

You have a temperature of 102°F and blood pressure of 90 over 120. How are you feeling? I'm feeling unwell. What are your symptoms? I have pale skin and abdominal pain. Thanks. We'll run some tests.

Start Playback **Reset**

Abnormal Lab Results

Test Date: 2023-02-15

	Value	Unit	Status
ALT	78 U/L	U/L	High
AST	85 U/L	U/L	High
ALP	137 U/L	U/L	High
Albumin	3.28 g/dL	g/dL	Low
Bilirubin	1.33 mg/dL	mg/dL	High
GGT	78 U/L	U/L	High

15 normal lab results not shown (Debug: 6 abnormal results showing)

Medical History

Active Conditions

- Chronic Hepatitis B: Diagnosed 2023-09-23
- Fatty Liver Disease: Diagnosed 2023-04-21

Current Medication

Furosemide	40mg daily
Entecavir	0.5mg daily

CT Scan

CT Report info...

Select Scan

- Color
- Draw
- Circle Region
- Measure
- Add Text
- Erase

Save edits

The image shows a grayscale axial CT scan of the abdomen. Several regions of interest are highlighted with colored circles and lines: a large yellow circle highlights a portion of the liver; a red circle highlights a region near the spine; a red line points to a structure in the upper right quadrant; and another red line points to a structure in the lower right quadrant. These likely correspond to the findings described in the medical history and differential diagnosis.

Hepnovate

Liver Disease Diagnostics, Powered by Human + AI.

AI suggestions are meant to support, not replace, clinical judgment. Please review all findings carefully.

Diagnostic Suggestions

Acute Hepatitis Exacerbation with Possible Liver Cirrhosis

AI Confidence: 75%

Clinical Findings

- Enlarged liver with a heterogeneous echotexture, suggestive of chronic liver disease
- Mild splenomegaly (enlarged spleen)
- Possible ascites (fluid accumulation) in the abdominal cavity
- Mild gallbladder wall thickening
- No significant abnormalities in the pancreas or kidneys

Physician Assessment

Add your clinical assessment, concerns, or plan...

Request AI Review **Confirm Diagnosis**

Hepnovate

Liver Disease Diagnostics, Powered by Human + AI.

[← Back to Diagnosis](#) [Edit](#) [Download](#)

Medical Write-Up

1. Chief Concern (CC):
The patient presents with a chief concern of acute exacerbation of chronic hepatitis B, accompanied by abdominal pain, fever, and pale skin.

2. History of Present Illness (HPI):
The patient reports a sudden onset of abdominal pain, fever, and pallor over the past week. The abdominal pain is persistent and localized in the upper right quadrant, with a dull, aching quality. The fever has been intermittent, reaching temperatures up to 102°F. The patient also notes a yellowish discoloration of the skin and sclera, as well as fatigue and loss of appetite.

3. Past Medical History (PMH):
The patient has a history of chronic hepatitis B, diagnosed on 2023-09-23, and fatty liver disease, diagnosed on 2023-04-21. The patient is a former smoker, having quit in 2019, and has a history of social alcohol use. The patient is an IT professional.

4. Past Surgical History (PSH):
The patient has undergone a cholecystectomy on 2018-03-15 and an appendectomy on 2005-II-22.

5. Medications:
The patient is currently taking Furosemide 40mg daily and Entecavir 0.5mg daily.

6. Allergies:
The patient is allergic to Penicillin, causing a rash, and Sulfa, causing anaphylaxis.

7. Social History:
The patient is a former smoker, having quit in 2019, and has a history of social alcohol use. The patient's occupation is an IT professional.

8. Family History:
The patient's father has a history of hepatitis B. The mother has Type 2 Diabetes.

TL;DR

Through online demonstrations with eight hepatologists, we will assess how our physician-controlled AI workflow impacts **diagnostic efficiency** and **trust** compared to traditional methods. We will do this through a Likert scale and 1 targeted open-ended **pre and post survey** sections.



Exemplar paper

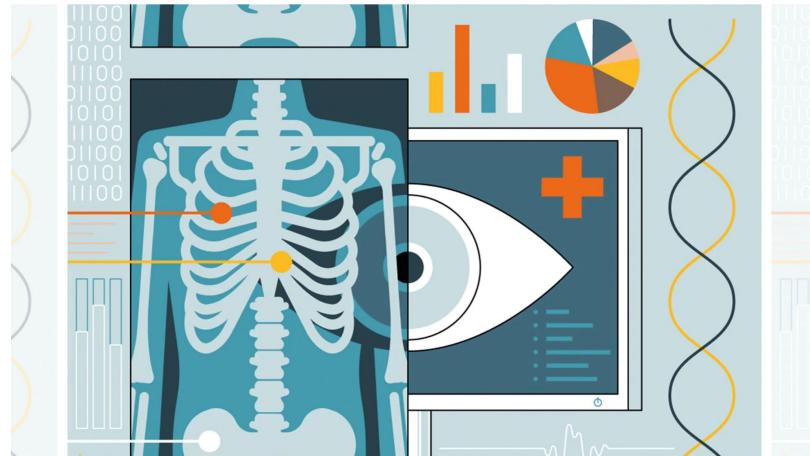
- Enhancing Clinical Documentation Workflow with Ambient Artificial Intelligence: Clinician Perspectives on Work Burden, Burnout, and Job Satisfaction

<https://www.medrxiv.org/content/10.1101/2024.08.12.24311883v1.full>

- Clinical UI+AI tool helping clinicians with documentation
- **Initial survey** (8 items) before the first use of the AI tool
- **Follow-up survey** (11 items) after at least 5 clinical encounters using the tool.
- Assessing:
 - Ease of use
 - Patient care
 - Time taken for old and new workflows
 - Increased satisfaction of workflow

Research Questions

- How does Hepnovate affect diagnosis time and efficiency for diagnosing liver diseases?
- How does physician interaction with Hepnovate affect diagnostic confidence and trust in AI?
- Does allowing physicians to guide and modify AI-generated diagnoses increase trust in clinical settings?
- Which UI elements most improve physician trust in AI and time efficiency?



Hypotheses

- Physicians using Hepnovate can **diagnose liver disease faster** than with traditional methods.
- **Automated symptom extraction** and **documentation** will be the most valuable feature for time efficiency.
- An organized AI diagnosis and review page with confidence scores will **increase trust** in AI.
- Physicians will trust AI more when using interfaces that allow them to **modify AI outputs** compared to viewing static AI suggestions.
- **Interactive image annotation** and **AI guidance** will be the most valuable feature for diagnostic confidence.

Hepnovate
Liver Disease Diagnostics, Powered by Human + AI.

Patient ID: P1000

Current Symptoms: Pale skin, Abdominal pain, Fever

Live Transcription: You have a temperature of 102F and blood pressure of 90 over 120. How are you feeling? I'm feeling unwell. What are your symptoms? I have pale skin and abdominal pain. Thanks. We'll run some tests.

Start Playback | Reset

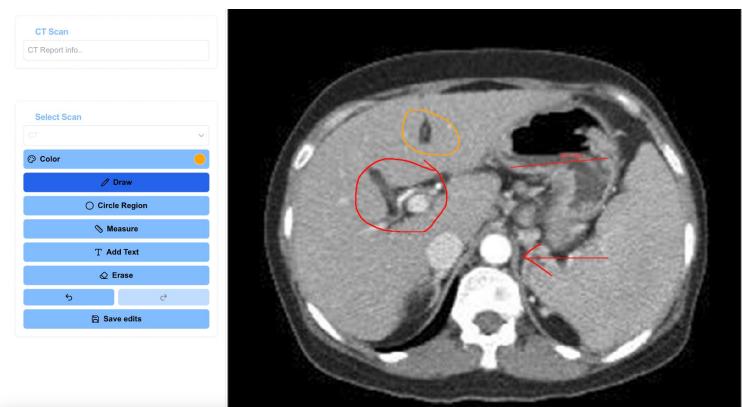
Current Vitals: Temperature 102°F, Blood Pressure 90/120 mmHg, Pulse Value

Abnormal Lab Results: ALT 78 U/L (High), AST 85 U/L (High), ALP 137 U/L (High), Albumin 3.28 g/dL (Low), Bilirubin 1.33 mg/dL (High), GGT 76 U/L (High). 15 normal lab results not shown. (Display 6 abnormal results showing)

Medical History: Chronic Hepatitis B, Deposited 2023-09-23; Fatty Liver Disease, Deposited 2023-04-21

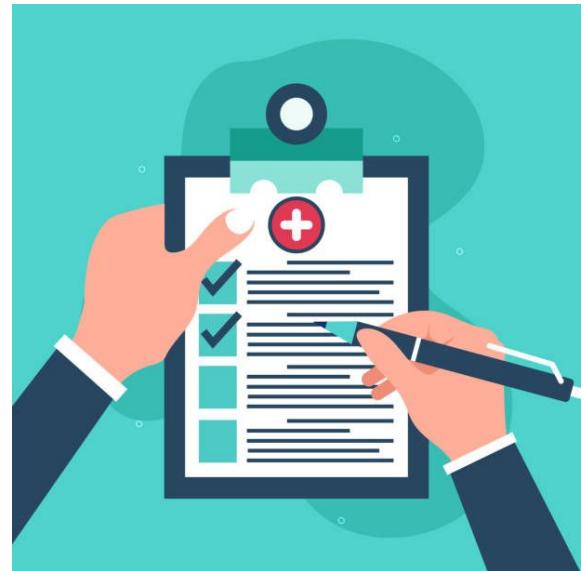
Active Conditions: Furosemide 40mg daily, Entecavir 0.5mg daily

Current Medication: Furosemide 40mg daily, Entecavir 0.5mg daily



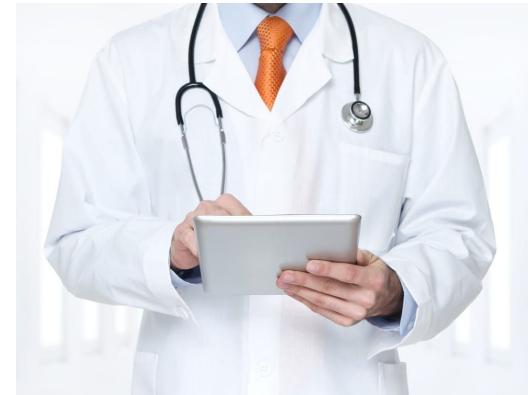
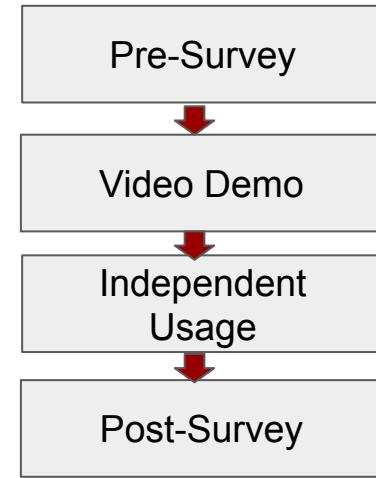
Design

- Study Components:
 - **Pre-demo survey section**
 - **Video demonstration**
 - **Independent usage of Hepnovate**
 - **Post-demo survey section**
- Survey Design:
 - 5-point Likert scales and open-ended questions
 - Assessment of time savings in liver disease diagnosis
 - Evaluation of trust in AI (guided annotations, confidence, explanations)
 - Ranking of Hepnovate's features by clinical value and usability



Tasks & Procedure – Experiment Workflow

- **Email invitation/Zoom:**
 - Doctors receive a link to the form including all study components.
- **Pre-Demo Survey Section:**
 - Collects data on current workflows and pain points.
- **Video Demonstration:**
 - Showcases Hepnovate's interface for liver disease diagnosis.
- **Independent Usage of Hepnovate:**
 - Doctors access the UI and interact with simulated patient cases.
- **Post-Demo Survey Section:**
 - Evaluates efficiency, usability, and trust.
 - Captures feedback on Hepnovate vs. current workflows.



Tasks & Procedure – Survey Information

Comprehensive online survey comparing Hepnovate to daily process

22 questions split between pre and post survey

<https://forms.gle/MSuiT5jRvwDr1bWi9>

How intuitive was Hepnovate's user interface? *

1 2 3 4 5

Very unintuitive Very intuitive

The following features would help me save time *

1- Strongly Disagree
2- Slightly Disagree
3- Neither Disagree/Agree
4- Slightly Agree
5- Strongly Agree

	1	2	3	4	5
Speech-Symp...	<input type="radio"/>				
Scan Annotation	<input type="radio"/>				
AI Suggestions	<input type="radio"/>				
AI Review	<input type="radio"/>				
Write-Up	<input type="radio"/>				

Hepnovate's ability to let the doctor guide the analysis and provide feedback allows me to trust AI in my workflow more. *

1 2 3 4 5

Strongly disagree Strongly agree

I would be comfortable using Hepnovate as a standard part of my diagnostic and documentation process. *

1 2 3 4 5

Strongly disagree Strongly agree

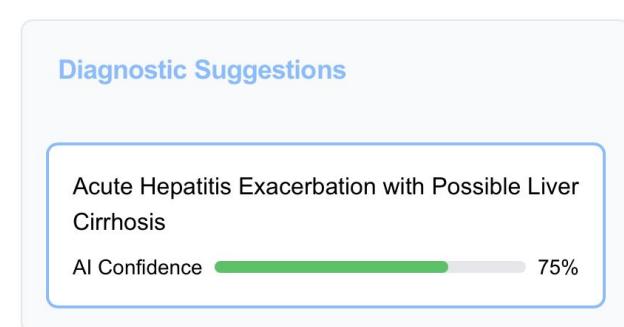
Apparatus

- **Survey Tool:**
 - Google Forms (Likert scale and open-ended questions).
- **Demonstration Materials:**
 - Youtube Video walkthrough of Hepnovate.
 - Link to the Hepnovate UI for simulated patient case.
- **Communication Method:**
 - Email with study form link.
 - Zoom meeting with doctors.
- **Hepnovate System:**
 - Vercel deployed web-based interface.
 - React <-> Local Database <-> HuggingFace API
 - Key features:
 - **Speech-to-text symptom extraction**
 - **AI-assisted image annotation**
 - **Differential diagnosis generation**
 - **Report formatting**



Metrics & Measurement

- **Comparative Survey Metrics**
- **Time Efficiency:**
 - Estimated time savings (5-point scale)
 - Ranking of features by time-saving potential
 - Comparison to current workflow efficiency
- **Trust:**
 - Confidence in AI-generated outputs (5-point Likert scale)
 - Comfort with AI decision-making
 - Transparency of AI reasoning
 - Value of modification for trust
- **Feature Evaluation:**
 - Usefulness rating for each key feature (5-point scale)
 - Most valuable and least valuable components
 - Likelihood to use each feature in daily practice
- **Implementation Feasibility:**
 - Ease of integration (5-point scale)
 - Adoption barriers
 - Training requirements



Analysis

- Survey Data Analysis:
 - Descriptive statistics for Likert-scale responses
 - Mean ratings for each feature and workflow component
 - Identification of highest-value features based on ratings
- Comparative Analysis:
 - Gap analysis between current workflow and Hepnovate perception
 - Analysis of projected time savings across diagnostic tasks
 - Analysis of AI CDSS trust
- Qualitative Analysis:
 - Thematic analysis of open-ended survey responses
 - Identification of common concerns and enhancement requests
- Prioritization Framework:
 - Mapping physician feedback to feature enhancement priorities
 - Cross-referencing time savings with trust impacts
 - Development of optimization recommendations

Participants

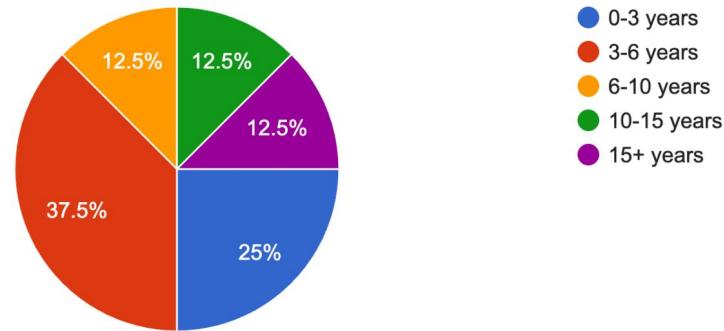
- 8 physicians specializing in hepatology
- Inclusion Criteria:
 - Board-certified hepatologists with active clinical practice
 - Experience in liver disease diagnosis
 - Familiarity with digital health records and liver imaging
- Exclusion Criteria:
 - Currently involved in competing AI diagnostic tool development.



Respondent Demographics

How many years have you been practicing medicine?

8 responses



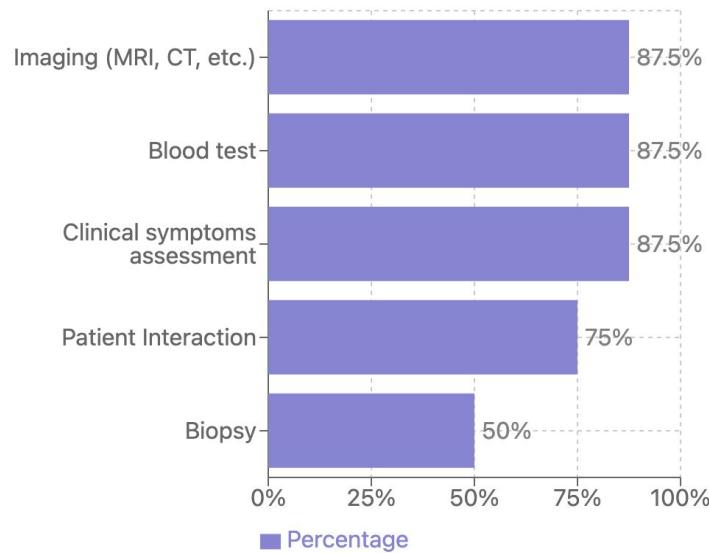
Analysis and Findings

Pre-Survey Findings

Findings- Medical Diagnosis

Current Diagnostic Methods

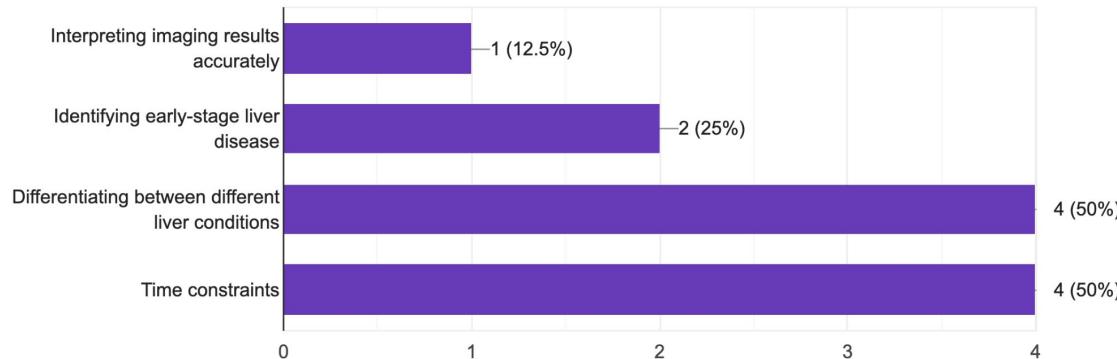
The visualization highlights the relative importance of different diagnostic approaches in current clinical practice.



Findings- Medical Diagnosis

Primary Diagnostic Challenges

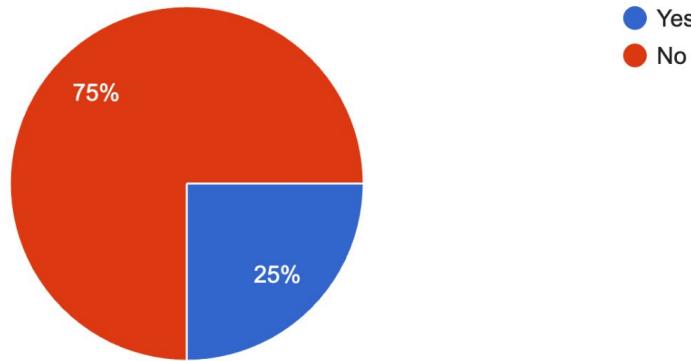
This visualization highlights the critical bottlenecks in the diagnostic process that may need targeted interventions.



Findings- AI

This visualization highlights the usage of AI in medical practices by the doctors

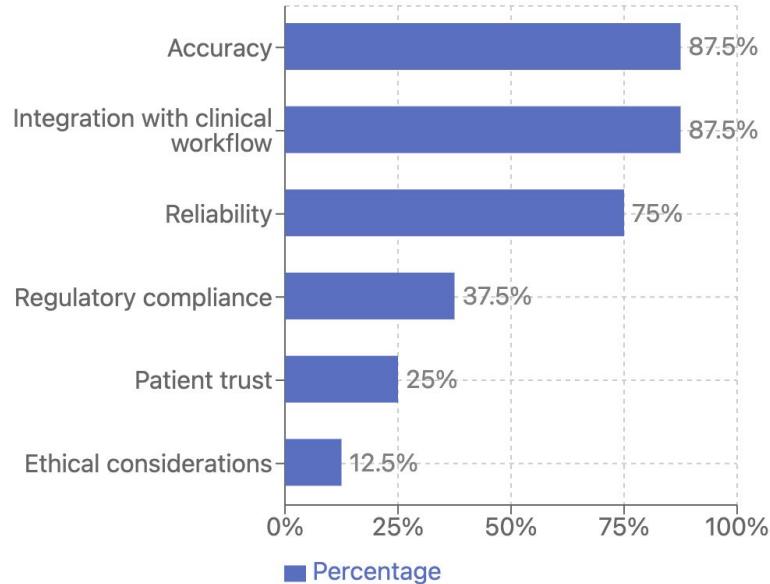
Have you previously used any AI-based diagnostic tools in your practice?
8 responses



For those who chose yes, they mentioned using ChatGPT and OpenEvidence

Findings- AI Trust

This visualization highlights the Key Concerns in Clinical Diagnostics Ranked by Frequency.



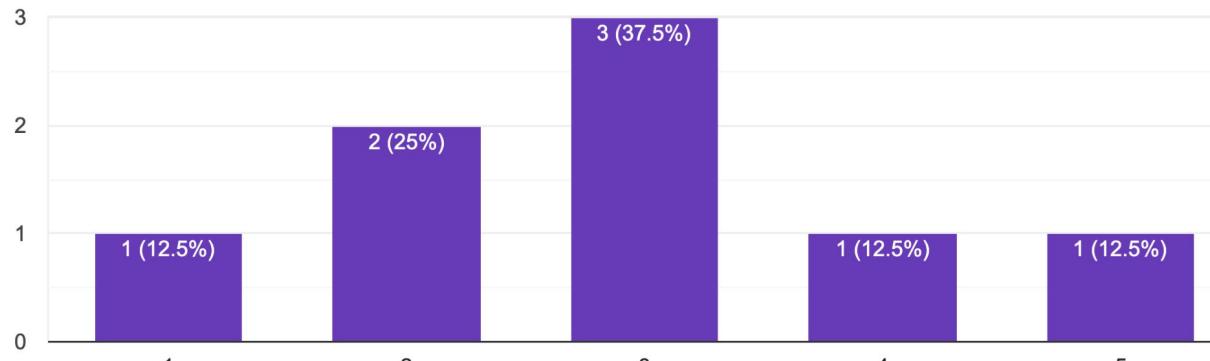
Accuracy and integration with clinical workflow are the most prevalent concerns, each identified by 87.5% of respondents, followed closely by reliability at 75.0%. There is a significant drop to the next tier of concerns, with regulatory compliance at 37.5%, patient trust at 25.0%, and ethical considerations being the least frequently mentioned at 12.5%.

Findings- AI Trust

This chart illustrates varying degrees of openness to AI integration among 8 healthcare professionals.

How open are you to using AI-based tools in your practice?

8 responses



Strongly Disagree

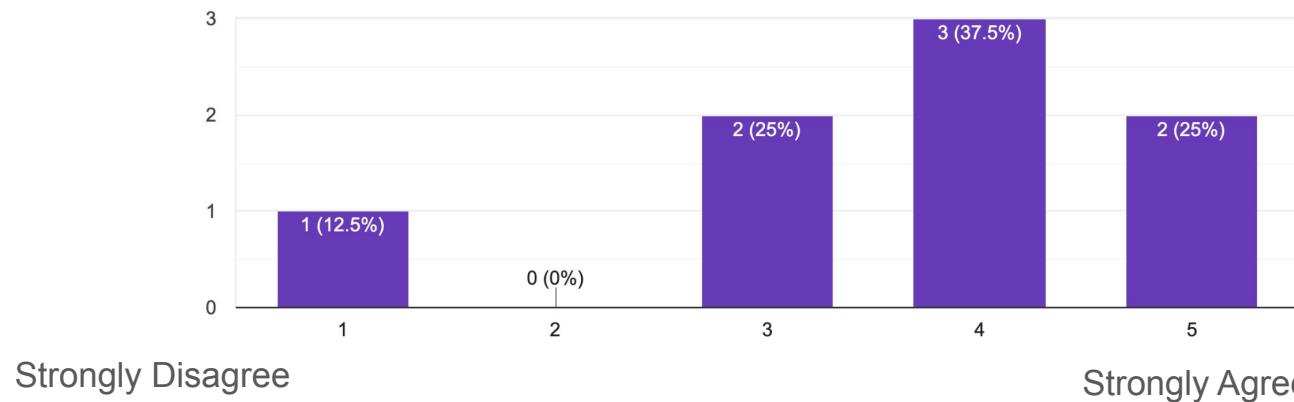
Strongly Agree

Time Efficiency

This visualisation highlights the time burden of documentation and data management in clinical workflows

I spend large amounts of time inputting/gathering data and creating documentation.

8 responses



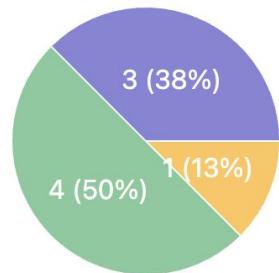
This distribution suggests that documentation and data entry remain significant time commitments for the majority of healthcare professionals surveyed, highlighting a potential area for workflow optimization.

Post-Survey Findings

Hepnovate - The UI

User Interface Intuitiveness

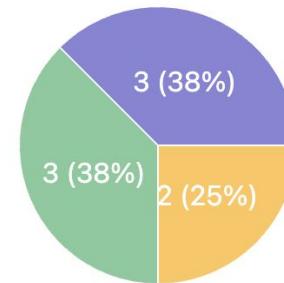
Average rating: **3.75/5**



■ Rating: 3 ■ Rating: 4 ■ Rating: 5

Ease of Use

Average rating: **3.88/5**



■ Rating: 3 ■ Rating: 4 ■ Rating: 5

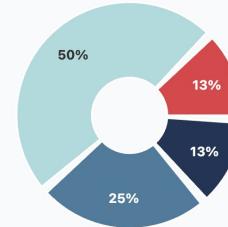
Note: All respondents rated both aspects 3 or higher.

Overall Speed and Trust Metrics

How physicians perceive Hepnovate's impact on diagnostic speed

Diagnostic Speed Perception	Count	Percentage
● Yes, slightly faster	4	50%
■ No difference	2	25%
● No, slightly slower	2	25%

Hepnovate's ability to let the doctor guide the analysis and provide feedback allows me to trust AI in my workflow more.



● 1 - Strongly Disagree ● 3 - Neutral ● 4 - Slightly Agree
● 5 - Strongly Agree

Key Statistics

Average Rating
3.25/5

Positive Responses
37.5%
(Rating 4-5)

Neutral Responses
50%
(Rating 3)

Negative Responses
12.5%
(Rating 1-2)

Feature-Specific Metrics

Feature	Time-Saving	Trust Building
Write-Up	4.4	N/A
Speech-Symptom	4.4	N/A
AI Review	3.4	3.6
AI Suggestions	3.5	2.9
Scan Annotation	3.5	3.4

Feature performance comparison showing Time-Saving and Trust Building metrics across five healthcare tools. Write-Up and speech-symptom excels in time efficiency (4.4), while AI Review demonstrates the strongest trust-building capabilities (3.6). >3 indicates a positive impact, <3 indicates a negative impact.

Survey Correlational Analysis

Diagnostic speed perception & adoption willingness **+0.75**

Strong positive correlation

Perceived speed improvements drive adoption decisions

Trust & adoption willingness **+0.62**

Strong positive correlation

Higher trust leads to greater adoption willingness

Years of practice & AI adoption **-0.18**

Weak negative correlation

More experienced physicians slightly less likely to use AI

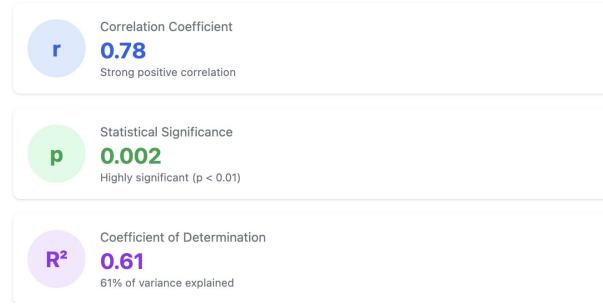
AI experience & concern level **-0.13**

Slight negative correlation

Experienced AI users have fewer concerns

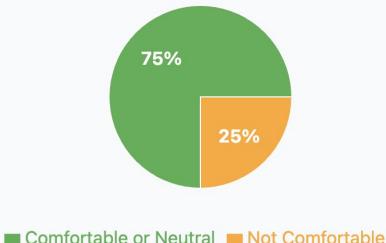
Adoption and Recommendation

AI Openness & Hepnovate Integration Correlation



The strong correlation ($r = 0.78$) indicates that physicians who reported higher openness to AI tools in general also tended to give higher ratings for Hepnovate's potential to integrate into their workflow. This relationship was statistically significant ($p = 0.002$), suggesting it's unlikely to be due to random chance.

Majority Would Use or Consider



3 in 4 physicians would use or consider using Hepnovate

75%

Would recommend or were neutral

37.5%
Strongly Recommend

12.5%
Somewhat Recommend

25%
Neutral

Open-ended feedback

Most Positive Comments

"The documentation feature would save me significant time."
"I like the ability to guide the AI analysis process."
"Voice recognition would greatly improve my workflow."

Constructive Feedback

"Need better editing capabilities for AI-generated content."
"Integration with existing EMR systems is essential."
"Would prefer more customizable templates for my specialty."

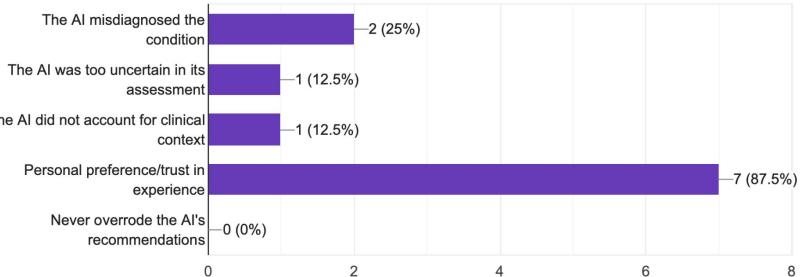
Key Insights from Textual Feedback

Top Positive Themes

- Documentation tools seen as biggest time-saver
- Time-saving potential recognized across multiple features
- Voice recognition highlighted as valuable for efficiency
- Clinical decision support appreciated for complex cases

If you overrode the AI's recommendations, what was the main reason? Select all that apply. If you selected "Other," please specify.

8 responses



Key Findings

Key Findings - Negative

- **Autonomy:**
 - Better integration from **in-person Zoom trials** vs. online video
 - Confusion from online users, **poor integration ratings**
 - Complexity vs. capability gap highlights **need for training**
 - One participant said more time with tool → learned trust
 - Need for **more fine-grained editing**
- **AI Model:**
 - Still produces **errors** despite being top 5 (according to MediConfusion UCLA talk) for clinical diagnosis
 - **Doctor skepticism** towards AI-driven diagnoses.
 - Strong model ≠ guaranteed trust
- **Scan Annotation:**
 - **Underutilized**
 - Feedback shows **limited usefulness**
 - Slows workflow for some users (annotations are rarely used)
- **Potential Outlier**
 - Ranked negatively in almost all questions
 - Suggested it may help non-hepatologists more than specialists



Key Findings - Positive

- **Promising Clinical Impact**
 - **50%** of physicians reported **faster disease detection** with Hepnovate
 - **87.5%** of physicians see workflow integration as the biggest obstacle
 - **75%** expressed neutral to positive **willingness to adopt into practice**
- **Feature Performance & Time Savings**
 - **Write-Up & Speech Transcription**: highest time-saving rating (4.4/5)
 - **87.5%** rated at least one feature as time-saving (4+/5)
 - **75%** of time-constrained physicians identified time savings
- **Trust-Building Elements**
 - AI Review: highest trust rating (3.8/5)
 - **62.5%** positive feedback for **physician-guided analysis**
 - **26.7% higher trust ratings** among physicians managing complex diagnoses
 - 2/3 hesitant users said both **AI review** and **scan annotations** increased trust



THANK YOU