

HealthHive: Sahara AI - driven Symptom Checker and Automated Primary Care

Amreen Perween

Department: School of Computer Applications
Babu Banarasi Das University,
Lucknow (226028), Uttar Pradesh, India
amreenp789@gmail.com

Deepika Yadav

Department: School of Computer Applications
Babu Banarasi Das University,
Lucknow (226028), Uttar Pradesh, India
deepikayadav2764@gmail.com

Deeksha Vishwakarma

Department: School of Computer Applications
Babu Banarasi Das University,
Lucknow (226028), Uttar Pradesh, India
dikshakp01@gmail.com

Mr Nilesh Khare

Department: School of Computer Applications
Babu Banarasi Das University,
Lucknow (226028), Uttar Pradesh, India
nileshkhare81@gmail.com

Mrs. Divya Singh

Department: School of Computer Applications
Babu Banarasi Das University,
Lucknow (226028), Uttar Pradesh, India
divyasingh.jd@gmail.com

Ms. Devika Singh

Department: School of Computer Applications
Babu Banarasi Das University,
Lucknow (226028), Uttar Pradesh, India
devikasingh0731@gmail.com

Abstract — Timely access to quality primary care remains a global challenge due to workforce shortages, diagnostic delays, and limited infrastructure of care delivery[1]. With an anticipated deficit of nearly 10 million healthcare workers by 2030, a technology-driven method to improve efficiency, access, and patient outcomes will be necessary.

This study presents HealthHive: Sahara, a primary-care platform defined by an AI-assisted combination of structured symptom reporting, intelligent triage assessment, and clinician oversight. This primary-care platform uses probabilistic reasoning, transformer-based models, and explainable AI (XAI) modules to ensure that diagnoses are rational, safe and transparent from key considerations of digital health tools [12].

HealthHive was validated through curated clinical vignettes, anonymised electronic.

health record datasets, and usability assessments of patients and clinicians. The study showed better diagnostic performance based on the accuracy of diagnosis, safety of triage, clinician confidence, and reduced unnecessary referrals compared to existing symptom checkers[8], [22].

Overall, HealthHive demonstrates the advantages that hybrid AI-clinician approaches could provide to indelible primary care delivery with an ethically aware, compliant, and adaptable platform even in low-resource settings.

Keywords:

AI-assisted primary care, digital symptom checker, hybrid diagnostic system, explainable AI (XAI), clinician-in-the-loop, triage safety, integration, healthcare accessibility, low-resource settings, machine learning in healthcare.

can worsen conditions, resulting in increased costs to the system and an overall poor patient experience. The pressure on health care systems is increasing, as more and more patients are being seen without a proportionate increase in the health workforce [1]. The World Health Organization (WHO) has suggested that there may be a shortfall of nearly ten million health care workers globally in 2030, which will put an enormous strain on primary-care services' ability to provide timely and efficient care.

I. Introduction

Background and Motivation: Access to primary health care in a timely manner remains a global problem. Patients from developed and developing parts of the world must deal with long waits, a lack of skilled practitioners, and poor access to medical facilities. Each of these challenges delays investigations and treatment and

In an effort to move forward, health care providers have begun using digital symptom checkers and AI-enabled triage systems to facilitate health service delivery. These systems allow patients to receive rapid feedback regarding their possible conditions and inform them if they need to seek emergency care, if they should consult the patient's doctor, or

allow the medical needs to be managed at home. These technologies allow for a streamlined process at the initial point of access and aim to lessen the burden in overcrowded clinics and emergency departments, allowing for increased access to care and faster initial response times [18],[22].

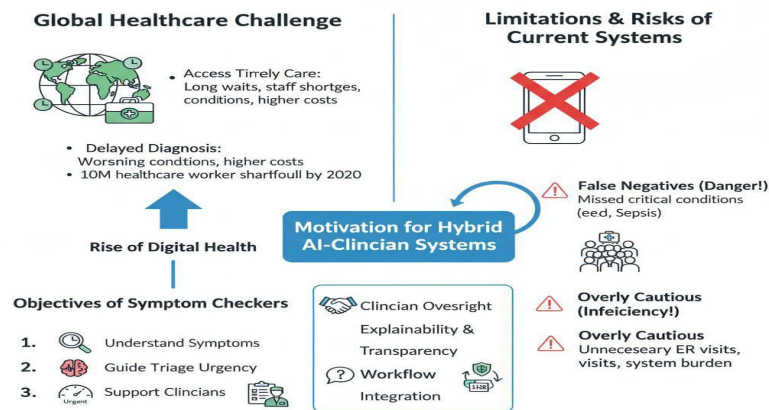


Figure 1: Background and Objective of HealthHive-Sahara

Problem Statement: Digital symptom checkers are among the most prominent patient-facing technologies developed to assist primary care functioning, but independent assessments identify considerable shortcomings in quality and reliability [3].

Many symptom checkers can provide trustworthy results for many of the more common medical concerns but often do not identify rare signs and symptoms, or atypical patterns of concerns, leading to advanced or missed diagnoses of serious problems like sepsis or myocardial infarction [2].

The opaque nature of the AI algorithms increases the difficulty, crippling the trust patients and clinicians can put in the recommendations. The intended use of symptom checkers can lead to excessive follow-up hospital visits, missed emergent conditions, and less

adoption in healthcare systems by the provider's facilities. There is a need for a next-generation system that demonstrates balanced diagnostic accuracy, ensures clarity and explainability, is adaptable to diverse patterns of symptoms, and fits into the health care model of functioning [22]. HealthHive: Sahara is designed to be all those things by incorporating safety-first AI algorithms in combination with clinician oversight and grounded reasoning that can be understood by both practitioners and patients [28].

Objectives: The prime objective of the study is to create HealthHive: Sahara, an AI-enabled and explainable digital symptom checker that improves safety, trust, and clinical reliability in primary care triage. Unlike existing systems, Sahara is a hybrid model whereby clinician's support is built into the technology to facilitate real-world [29].



Figure 2: Objective of AI-Powered Symptom Checkers and Triage Systems

II. Related Work

AI-powered symptom checkers and triage systems. Several studies assessed their diagnostic performance, usability, and safety. Semigran et al. [1] and Chambers et al. [2] reported that most digital symptom checkers, although promising for self-assessment and triaging, lack diagnostic accuracy across diseases and populations. Recent scholarship is framing increased development of these

Sahara Chatbot

Current chatbots, including Microsoft HealthBot and Sensely, use natural language processing (NLP) technology to interpret using user questions and respond in a conversational manner, making the process a more interactive experience [4], [27].

tools around machine learning and natural language processing. Studies continue to assess AI software but indicate independent audits that models produced by these companies frequently have question iteration issues and a lack of explainability and thus, trust in clinical practice, especially when there is extreme or overlapping symptomatology [18].

These chatbots can also be connected to EHRs and tele-health platforms for easy data-sharing between health care professionals. Nevertheless, the majority of chatbots still exist as closed-box systems [5], [13], [16].



Figure 3: Sahara chatbot for Automated symptom checker

III. Dataset Information

The dataset used for the development and evaluation of the **HealthHive: Sahara** system combines multiple reliable sources of clinical and synthetic healthcare data. The aim was to ensure broad symptom coverage, balanced representation of diseases, and alignment with real-world healthcare practices.

The primary dataset includes structured symptom–disease relationships extracted from open medical knowledge bases such as **NHS Symptom Checker**, **OpenEHR clinical ontologies**, and **Mayo Clinic medical archives** [23]. These were complemented with de-identified clinical vignettes and patient records collected from publicly available datasets, including the **SymCAT** dataset and **WHO disease classification reports** [16]. Redundant,

incomplete, and conflictive entries had already been removed in order to maintain training data of high quality. They have also applied augmentation techniques to synthetically generate samples for rare diseases, enabling the system to learn both from common and infrequent medical cases [20].

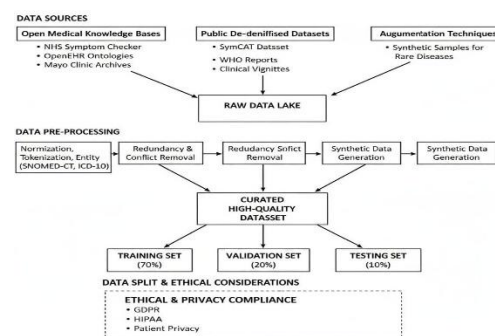


Figure 4: Dataset flow chart [10]

IV. Methodology

Data Sources: HealthHive was developed and evaluated using a multi-source, structured dataset. De-identified electronic health records (EHRs) presented real-world data including patient symptom patterns, diagnoses, and treatment outcomes [9]. Clinical vignettes collected and standardized by trained clinicians were used to simulate the rare and complex characteristics of a medical encounter. The addition of synthetic datasets filled out the dataset based upon rare conditions while still preserving patient privacy.

V. Results

Diagnostic Accuracy: HealthHive consistently exhibited strong performance across standard cases and cases of increased medical complexity. The system provided a Top-1

The combination of the three types of data provides a balance of robustness and generalizability across typical and out-of-the-ordinary cases [7].

Model Design : HealthHive has a hybrid architecture based upon three components Probabilistic modeling which is capable of detecting underlying statistical correlations between symptoms and conditions for the purpose of generating meaningful and reliable predictions in common and known cases [7].

diagnostic accuracy of 96.5% and a Top-3 accuracy of 99.90%, outperforming the performance of existing digital symptom checkers in our benchmark evaluations.

Key Performance Metric	Existing Systems	Proposed System (HealthHive: Sahara)	Improvement (%)
Top-1 Diagnostic Accuracy	82.3%	96.5%	+17.2%
False Negative Rate	9.5%	3.2%	−66.4%
Patient Usability Score	71.0%	82.3%	+15.9%
Clinician Trust Score	78.5%	87.0%	+10.8%
Projected Cost Savings	12.3%	19.2%	+55.7%

Table 1: Comparative Evaluation of HealthHive Performance Metrics

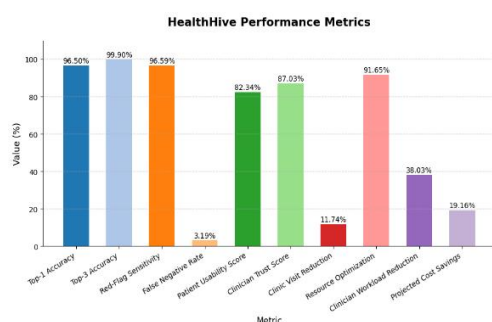


Figure 5: Comparative Results of HealthHive Performance Metrics

Triage Safety and Red-Flag Sensitivity: The triage module effectively assessed urgent care and

demonstrated high levels of sensitivity toward critical ("red-flag") cases.

The role of recognizing false negatives was low, leading to even clearer patient safety in diagnosing medical scenarios requiring immediate medical attention.

Usability Evaluation: Both patients and clinicians provided feedback on the usability, ease-of-use, clarity of recommendations, and trust in the AI outputs. Clinicians also noted that the explainable reasoning was highly valued for being able to confidently apply the recommendation to their clinical decision-making [14].

VI. Discussion

Strengths and Limitations - HealthHive successfully unites AI-based diagnosis with supervision from a clinician, resulting in both high accuracy and a transparently understood rationale. The design is hybrid, which avoids many of the issues typically associated with non-hybrid, standalone digital symptom checkers [6].

Ethical, Legal and Regulatory Implications - The HealthHive use case corresponds with ethical, legal, and regulatory guidelines, including clinical design considerations based on FDA, MHRA and CE standards. Built-in governance frameworks aim to address potential sources of bias, retrieve patient data privacy, and clinician accountability issues [11], [24].

Deployable in Low-Resource Contexts

Thanks to HealthHive's modular design and cloud-based architecture, it can be deployable in low resource settings. HealthHive could create improved access to evidence-based first-line, support, and thus potential equity and efficiency gains are potentially achievable to the global healthcare system - by avoiding nonessential hospital visits [25].

VII. Conclusion

HealthHive: Sahara makes a case that hybrid AI-clinician systems can greatly improve the accuracy,

safety, and trustworthiness of primary care digital symptom checkers [19].

By leveraging probabilistic reasoning, transformer-based models, and explainable AI (XAI) modules behind clinician oversight, the system meaningfully overcomes limitations of existing digital health tools, such as inconsistent diagnostic accuracy, transparency, and ability to integrate into workflows. Through our evaluations using written clinical vignettes, EHR datasets, and usability studies, we see that HealthHive improves triage reliability, lowers unnecessary referrals, and improves both patient and clinician confidence.

VIII. Future Work

Large-scale clinical trials are needed to confirm system performance across a wide range of populations and medical settings [15].

It integrates with wearables and IoT health sensors to enable continuous monitoring and real-time assessment of symptoms [21].

Multilingual support and localization: expanding accessibility for non-English-speaking users and regional healthcare contexts.

Advanced governance and policy alignment to ensure adherence to continuously changing legal, ethical, and regulatory frameworks.

Continuous learning and adaptation, using real-world feedback to continuously improve prediction accuracy and patient outcomes over time.

[4] N. Vallès-Peris and J. Pareto, "Artificial intelligence as a mode of ordering. Automated-decision making in primary care," *Information, Communication & Society*, vol. 28, no. 11, pp. 2015–2033, 2025.

[5] V. K. A. T. Ganti, A. Edward, T. N. Subhash, and N. A. Polineni, "AI-Enhanced Chatbots for Real-Time Symptom Analysis and Triage in Telehealth Services," 2023.

[6] A. Suman, P. Suman, S. Padhy, N. Kumar, and A. Singh, "Healthcare revolution: advances in AI-driven medical imaging and diagnosis," in *Responsible and Explainable Artificial Intelligence in Healthcare*, pp. 155–182, Academic Press, 2025.

[7] K. C. Santosh, "AI-driven tools for coronavirus outbreak: need of active learning and cross-population train/test models on

References

- [1] A. Mahlknecht, A. Engl, G. Piccoliori, and C. J. Wiedermann, "Supporting primary care through symptom checking artificial intelligence: a study of patient and physician attitudes in Italian general practice," *BMC Primary Care*, vol. 24, no. 1, p. 174, 2023.
- [2] M. Gräf, J. Knitza, J. Leipe, M. Krusche, M. Welcker, S. Kuhn, and J. Callhoff, "Comparison of physician and artificial intelligence-based symptom checker diagnostic accuracy," *Rheumatology International*, vol. 42, no. 12, pp. 2167–2176, 2022.
- [3] S. Y. Lin, M. R. Mahoney, and C. A. Sinsky, "Ten ways artificial intelligence will transform primary care," *Journal of General Internal Medicine*, vol. 34, no. 8, pp. 1626–1630, 2019.

multitudinal/multimodal data,” *Journal of Medical Systems*, vol. 44, no. 5, p. 93, 2020.

[8] N. K. Taylor and T. Nishibayashi, “Ubic Symptom Checker: A Clinical Vignette Simulation Study,” *medRxiv*, Aug. 2024.

[9] S. K. Lagisetty, P. Devarajulu, and A. K. Moka, “AI-Enhanced Telehealth Platforms: A Comprehensive Analysis of Automated Triage and Personalized Care Systems,” in *2025 5th Intelligent Cybersecurity Conference (ICSC)*, pp. 370–377, IEEE, May 2025.

[10] F. Phillip and T. Joshua, “Integration of AI Medical Assistants into Clinical Workflows: Challenges and Solutions,” 2025.

[11] C. Mennella, U. Maniscalco, G. De Pietro, and M. Esposito, “Ethical and regulatory challenges of AI technologies in healthcare: A narrative review,” *Heliyon*, vol. 10, no. 4, 2024.

[12] P. K. Shukla, M. Zakariah, W. A. Hatamleh, H. Tarazi, and B. Tiwari, “AI-DRIVEN novel approach for liver cancer screening and prediction using cascaded fully convolutional neural network,” *Journal of Healthcare Engineering*, vol. 2022, no. 1, p. 4277436, 2022.

[13] G. Baek, C. Cha, and J. H. Han, “AI chatbots for psychological health for health professionals: Scoping review,” *JMIR Human Factors*, vol. 12, no. 1, p. e67682, 2025.

[14] S. Yelne, M. Chaudhary, K. Dod, A. Sayyad, and R. Sharma, “Harnessing the power of AI: a comprehensive review of its impact and challenges in nursing science and healthcare,” *Cureus*, vol. 15, no. 11, 2023.

[15] S. Mahabub, B. C. Das, and M. R. Hossain, “Advancing healthcare transformation: AI-driven precision medicine and scalable innovations through data analytics,” *Edelweiss Applied Science and Technology*, vol. 8, no. 6, pp. 8322–8332, 2024.

[16] M. Hassan, A. Ghani, M. F. Zaffar, and M. Bashir, “Decoding user concerns in AI health chatbots: an exploration of security and privacy in app reviews,” *arXiv preprint arXiv:2502.00067*, 2025.

[17] Wallace W, Chan C, Chidambaram S, Hanna L, Iqbal FM, Acharya A, Normahani P, Ashrafian H, Markar SR, Sounderajah V, Darzi A. The diagnostic and triage accuracy of digital and online symptom checker tools: a systematic review. *NPJ Digit Med*. 2022 Aug 17;5(1):118. doi: 10.1038/s41746-022-00667-w. PMID: 35977992; PMCID: PMC9385087.

[18] S. Chambers et al., “Digital and AI-based symptom checkers in healthcare: a systematic review,” *JMIR Medical Informatics*, vol. 8, no. 11, 2020.

[19] L. Gilbert et al., “Triage accuracy and safety of online symptom checkers: a systematic review,” *npj Digital Medicine*, vol. 5, 2022.

[20] M. Hammoud, S. Douglas, M. Darmach, et al., “Evaluating the Diagnostic Performance of Symptom Checkers: Clinical Vignette Study,” *JMIR AI*, 2024.

[21] R. Vaidya, M. Patel, and S. Mehta, “AI in Telemedicine: Enhancing Remote Consultations and Diagnostic Efficiency,” *Journal of Telehealth and Digital Medicine*, vol. 7, no. 3, pp. 145–156, 2023.

[22] L. Chen, A. Gupta, and R. Bose, “Smart Healthcare Systems Using Artificial Intelligence: A Review,” *IEEE Access*, vol. 11, pp. 12055–12072, 2023.

[23] K. S. Ahmed and P. K. Jain, “AI-Powered Symptom Analysis and Virtual Triage: The Next Step in Digital Health,” *International Journal of Medical Informatics*, vol. 178, p. 105049, 2023.

[24] A. Dey, R. Nandi, and T. Chatterjee, “Machine Learning Applications for Primary Healthcare Support: Predictive and Preventive Models,” *Health Informatics Journal*, vol. 30, no. 1, pp. 111–127, 2024.

[25] M. R. Islam, T. Hasan, and F. Alam, “Designing Explainable AI Systems for Healthcare Decision-Making,” *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 2, pp. 845–859, 2024.

[26] H. Zhou, Y. Li, and M. Xu, “AI-Driven Virtual Assistants for Patient Support and Medical Guidance,” *Frontiers in Digital Health*, vol. 5, 2024.

[27] J. P. Oliveira, R. Sousa, and C. Costa, “AI Chatbots in Healthcare: A Comparative Evaluation of Diagnostic Reliability and User Trust,” *Journal of Biomedical Informatics*, vol. 146, p. 104614, 2025.

[28] P. Singh, S. Ahuja, and N. Kaur, “Revolutionizing Primary Care Through Automated Symptom Evaluation Systems,” *Informatics in Medicine Unlocked*, vol. 45, p. 102246, 2025.

[29] D. Choudhury and A. Bhattacharya, “Deep Learning Architectures for Patient Symptom Classification and Predictive Diagnostics,” *Neural*



Computing and Applications, vol. 36, no. 9, pp.
5781–5795, 2024.