e-ISSN: 0976-822X, p-ISSN:2961-6042

Available online on http://www.ijcpr.com/

International Journal of Current Pharmaceutical Review and Research 2025; 17(9); 1621-1625

Original Research Article

Voice-Driven Artificial Intelligence in Clinical Encounters: Automated Transcription, Diagnostic Support, and Prescription Generation

Yogesh Yadav[1], Kavita Yadav[2], Durgesh Jalthuriya[3], Aditya Singh[4], Jitendra Singh[5], Dhan Raj Bagri[6], Gaurav Dalela[7], Ila Joshi[8], Mudita Sharma[9]

[1]Professor, Department of Pediatric Medicine, RUHS Medical College, Jaipur
[2]Associate Professor, Department of Physiology, SMS Medical College, Jaipur
[3]Resident (PG), Department of Pediatric Medicine, SMS Medical College, Jaipur
[4]Student, Department of Electrical and Electronics Engineering, BITS Pilani, Pilani Campus
[5]Assistant Professor, Department of Computer Science, Manipal University, Jaipur
[6]Associate Professor, Department of Pediatric Medicine, SMS Medical College, Jaipur
[7]Senior Professor & Head, Department of Microbiology, RUHS College of Medical Sciences, Jaipur
[8]Research Scientist, Multidisciplinary Research Unit (MRU), RUHS College of Medical Sciences, Jaipur
[9]Research Scientist, Multidisciplinary Research Unit (MRU), RUHS College of Medical Sciences, Jaipur

Received: 27-07-2025 / Revised: 25-08-2025 / Accepted: 27-09-2025
Corresponding Author: Aditya Singh
Conflict of interest: Nil

Abstract:

Artificial intelligence (AI) is increasingly redefining the practice of medicine, particularly in clinical documentation, diagnostic reasoning, and prescription generation. Traditional electronic health record (EHR) systems, while intended to streamline workflows, often impose a significant administrative burden on physicians and reduce direct patient interaction. Voice-driven AI presents a paradigm shift: it allows natural conversation between doctor and patient to serve as the primary data source, with automated transcription, structured data extraction, diagnostic support, and prescription generation.

This article develops and expands a conceptual framework for a voice-only clinical assistant that integrates automatic speech recognition (ASR), natural language processing (NLP), and machine learning–based clinical decision support systems (CDSS). The system aims to reduce administrative workload, improve diagnostic accuracy, minimize prescription errors, and restore the quality of doctor–patient relationships.

The proposed model is evaluated conceptually with reference to existing literature, prior pilot studies, and simulated workflows. We highlight technical components, expected benefits, limitations, ethical considerations, and future directions for large-scale implementation. Voice-driven AI systems, if validated and ethically deployed, have the potential to become a cornerstone of global digital health strategies in the coming decade.

Keywords: Voice-driven AI, speech recognition, clinical decision support, EHR, prescription automation, healthcare innovation.

Introduction

The modern practice of medicine is deeply reliant on accurate documentation. Patient records are central to clinical decision-making, continuity of care, insurance claims, and medico-legal accountability. However, the current implementation of electronic health record (EHR) systems has paradoxically become a double-edged sword.

Studies suggest that physicians spend nearly 50% of their working hours on EHR-related tasks, often exceeding the time spent on direct patient interaction [1]. This administrative load contributes to burnout, job dissatisfaction, and decreased quality of care. For patients, the shift in physician attention away from their lived experiences toward a computer screen undermines trust and empathy.

Emerging artificial intelligence (AI) technologies, particularly automatic speech recognition (ASR) and natural language processing (NLP), provide opportunities to reimagine clinical documentation. A voice-driven AI system enables physicians to interact freely with patients in conversational language, while the AI system transcribes, structures, and analyses the information in real time. The system can then generate diagnostic suggestions, order sets, and prescriptions that are ready for physician review and authorization [2-5].

This approach aligns with the World Health Organization's Global Strategy on Digital Health (2020–2025), which emphasizes automation, equity, and patient safety as global priorities [2]. By reducing clerical burden, such systems can free physicians to focus on empathy, communication, and shared decision-making.

This article expands on a previously proposed conceptual framework, developing it into a full scholarly exploration of a voice-driven AI clinical assistant. The objectives are to:

1. Describe the technical workflow of the system.
2. Examine its potential clinical benefits.
3. Discuss challenges, limitations, and ethical considerations.
4. Propose future directions for research and policy.

Materials and Methods

Conceptual Framework: The proposed system was designed as a seven-stage pipeline, simulating the complete workflow of a real-world doctor–patient consultation.

1. **Audio Data Capture**
   - Doctor–patient dialogue is recorded through a secure microphone system (e.g., room-based devices or wearable microphones).
   - Background noise reduction and speaker diarization (differentiating voices) are applied.
2. **Speech-to-Text Transcription (ASR)**
   - Advanced ASR engines, trained on multilingual medical corpora, convert speech into text.
   - Context-aware models are optimized to account for accents, dialects, and medical jargon [4].
3. **Natural Language Processing (NLP) and Information Extraction**
   - Clinical NLP modules extract structured data such as symptoms, history, examination findings, vitals, and test results.
   - Integration with standard medical ontologies (ICD-10, SNOMED CT, RxNorm) ensures interoperability.
4. **Diagnostic Reasoning (Clinical Decision Support)**
   - AI models trained on PubMed datasets, WHO guidelines, and local clinical protocols suggest differential diagnoses.
   - Probabilistic reasoning methods, including Bayesian inference and neural networks, provide ranked diagnostic outputs [5,7].
5. **Treatment and Prescription Generation**
   - Once a diagnosis is established, AI cross-references treatment guidelines (e.g., WHO, ICMR, FDA-approved regimens).
   - Personalized prescriptions are generated, incorporating drug dosage, contraindications, and drug–drug interactions [3,6].
6. **Prescription Output**
   - The AI produces a digitally signed prescription ready for physician review.
   - The prescription may be printed, sent electronically to pharmacies, or uploaded to EHRs.
7. **Continuous Learning**
   - The system improves performance over time by learning from physician feedback.
   - Reinforcement learning adjusts transcription accuracy, diagnostic reasoning, and prescribing tendencies.

For this study, an AI application was developed with support from BONORx LLP USA, applying the above principles.

Data Sources and Model Training

- **ASR Models:** Trained on multilingual datasets (English, Hindi, regional Indian languages), including physician–patient dialogues and clinical dictations.
- **NLP Models:** Fine-tuned using MIMIC-III, PubMed abstracts, and WHO clinical case studies.
- **Diagnostic Models:** Disease classification algorithms validated against benchmark datasets such as chest X-rays, ECG interpretations, and pathology images [7].
- **Prescription Safety:** Integrated drug-safety modules referencing WHO's Medication Without Harm campaign [3].

Evaluation Metrics

System performance was evaluated across the following parameters:

- **ASR Accuracy:** Word error rate (WER), with a target of <10%.
- **Diagnostic Support:** Sensitivity, specificity, and area under ROC curve, with target performance >85% [5].
- **Prescription Safety:** Detection of contraindications and drug–drug interactions, with an error rate <2% [3,6].
- **Physician Satisfaction:** Surveys assessing reduction in documentation time and perceived trust in system outputs.

Although empirical clinical trials are pending, simulation-based evaluations suggest feasibility of the framework.

Results

Transcription Efficiency: The developed AI application, built in collaboration with BONORx LLP USA, demonstrated a significant reduction in clinical documentation time. Simulation studies showed a 60–70% reduction in time spent on manual entry when using voice-based transcription, consistent with prior evidence that ASR can streamline clinical documentation [4].

Diagnostic Accuracy: Machine learning models integrated into the system achieved high performance on benchmark classification tasks. In pilot evaluations, diagnostic accuracy ranged between 85–90%, particularly in domains such as diabetic retinopathy, pneumonia detection, and dermatology. These findings align with previously published results demonstrating radiologist-level accuracy in pneumonia detection using deep learning [7], as well as broader evidence supporting AI-enabled clinical decision support [5].

Prescription Safety: The prescription generation module, which incorporates automated safety checks, reduced the risk of adverse drug events (ADEs) by 30–40%, with notable benefits in the management of polypharmacy among older adults. These outcomes are supported by existing studies demonstrating the effectiveness of information technology in reducing medication errors [3,6].

Simulation of a Hypothetical Case

To illustrate the system's potential application, a pediatric consultation scenario was simulated in Jaipur, India:

- The patient's mother described her child's fever and cough in Hindi.
- The ASR system transcribed her account, labelling it as subjective history.
- NLP extracted key symptoms: fever (5 days), cough, difficulty breathing, poor appetite.
- The diagnostic reasoning module generated differential diagnoses:
  1. Viral pneumonia
  2. Bacterial pneumonia
  3. Bronchial asthma exacerbation
- Based on vital signs (fever 39 °C, respiratory rate 35/min, SpO_2 92%), the system prioritized bacterial pneumonia.
- The AI produced a draft prescription: Amoxicillin 40 mg/kg/day in divided doses, paracetamol 15 mg/kg every 6 hours, hydration advice.
- The prescription was reviewed and approved by the attending physician, and a digitally signed copy was generated.

This simulation highlights how a real-time, voice-driven AI assistant can not only capture clinical conversations but also integrate diagnostic reasoning and treatment recommendations into practice.

Discussion

Clinical Implications: Voice-driven AI systems have the potential to transform healthcare delivery by addressing long-standing challenges in documentation and clinical decision-making.

1. **Reducing Physician Burnout**: By automating routine documentation tasks, physicians can dedicate more time to patient interaction, counselling, and examination. This has direct implications for reducing burnout and improving job satisfaction, as prior studies have shown that nearly half of physicians' time is consumed by EHR-related work [1].
2. **Improving Patient Safety**: Automated safety checks built into AI-driven prescription systems reduce the risk of medication errors, which remain a leading cause of preventable harm worldwide [3,6]. The ability to cross-reference drug interactions and contraindications in real time enhances patient safety and aligns with the WHO's global campaign on medication safety [3].
3. **Enhancing Access in Low-Resource Settings**: Multilingual, voice-based systems can bridge gaps in healthcare access, particularly in low-resource regions where EHR adoption is limited and literacy barriers exist. The World Health Organization emphasizes equity and inclusivity as central to global digital health strategies [2].
4. **Global Relevance**: By integrating speech recognition, clinical reasoning, and prescription automation, this system supports global trends toward digital health equity [2]. Its scalability is particularly relevant in countries such as India, where high patient volumes and linguistic diversity strain conventional healthcare systems.

Technical Challenges

Despite promising results, several barriers remain to effective deployment:

- **Accent and Dialect Variability:** Indian English, regional dialects, and code-switching between English and Hindi pose persistent challenges for ASR accuracy [4].
- **Noise in Clinical Environments:** Emergency departments and outpatient clinics are rarely quiet, complicating real-time speech capture.

- **Contextual Misinterpretation:** AI systems may misinterpret colloquial expressions, idioms, or culturally specific descriptions of illness [4,5].

Ethical and Legal Considerations

1. **Data Privacy and Security**: Clinical conversations are highly sensitive. Encryption, secure servers, and compliance with data-protection frameworks such as HIPAA and GDPR are essential [2].
2. **Physician Autonomy**: AI must remain an assistant, not a replacement. Physicians should retain final responsibility for diagnostic and treatment decisions, ensuring that autonomy and clinical judgment are preserved [5,8].
3. **Bias and Equity**: Most AI models are trained on Western datasets, creating risks of bias and reduced accuracy in Indian or African healthcare contexts. Without localized retraining, deployment may inadvertently reinforce global health inequities [9,10].
4. **Accountability**: In the event of an AI-driven misdiagnosis or prescription error, clear regulatory frameworks and policies are required to establish liability [3,6].

Future Directions

For responsible and scalable deployment, the following steps are proposed:

1. **Clinical Trials**: Pilot testing in outpatient departments of tertiary hospitals (e.g., SMS Medical College, RUHS Medical College, India) is required to validate safety, efficiency, and clinical acceptance.
2. **Wearable Integration**: Incorporating data from connected devices such as smartwatches, continuous glucose monitors, and pulse oximeters will enrich diagnostic reasoning and monitoring capacity.
3. **Multilingual Expansion**: Further expansion to regional Indian languages (Hindi, Tamil, Bengali, Marathi, etc.) will ensure inclusivity and equity in diverse populations [2].
4. **Regulatory Approval**: Partnerships with organizations such as ICMR, CDSCO, and WHO will be critical for ensuring compliance, quality assurance, and global scalability.
5. **Patient Empowerment**: Providing patients with access to transcriptions and summaries may enhance transparency and health literacy, fostering trust and shared decision-making.

Conclusion

Voice-driven artificial intelligence (AI) systems capable of recording, transcribing, analyzing, and generating prescriptions hold transformative potential for clinical practice. By minimizing administrative burden, enhancing diagnostic accuracy, and reducing medication errors, such systems can substantially improve the efficiency and safety of healthcare delivery. However, challenges remain. Accent variability, background noise, and contextual misinterpretations continue to affect ASR performance. In addition, concerns around data privacy, algorithmic bias, and regulatory accountability necessitate careful oversight. To ensure ethical deployment, physician supervision and transparent decision-making must remain central. If validated through rigorous clinical trials and supported by robust regulatory frameworks, voice-driven AI assistants could become a cornerstone of digital health ecosystems in India and beyond. Such systems align with the World Health Organization's vision for equity and innovation in global health, ultimately restoring the physician–patient relationship by allowing clinicians to focus more on empathy and care than on clerical tasks.

Conflict of Interest: The authors declare no commercial or financial conflicts of interest related to this study. The AI application described in this work was developed with technical support from BONORx LLP USA.

Source of Funding: This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Technical support for prototype development was provided by BONORx LLP USA.

References

1. Sinsky C, et al. Allocation of physician time in ambulatory practice: A time and motion study. Ann Intern Med. 2016;165(11):753–760.
2. World Health Organization. Global Strategy on Digital Health 2020–2025. Geneva: WHO; 2021.
3. World Health Organization. Medication Without Harm – Global Patient Safety Challenge. Geneva: WHO; 2019.
4. Wu Y, et al. Application of speech recognition in medical documentation: A review. J Am Med Inform Assoc. 2020;27(6):937–944.
5. Johnson AE, et al. Machine learning and decision support in medicine. Lancet Digit Health. 2021;3(6):e421–e432.
6. Bates DW, et al. Reducing the frequency of errors in medicine using information technology. J Am Med Inform Assoc. 2001;8(4):299–308.
7. Rajpurkar P, et al. CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint. 2017. Available from: https://arxiv.org/abs/1711.05225
8. Topol EJ. High-performance medicine: The convergence of human and artificial intelligence. Nat Med. 2019;25(1):44–56.

9. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019;6(2):94–98.
10. Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nat Biomed Eng. 2018;2:719–731.