

Background

Healthcare professionals face increasing pressure to navigate complex, emotionally difficult conversations with patients, yet current training methods such as role-play or faculty feedback are subjective, costly, and unable to scale to the demands of modern clinical education. More than **53% of U.S. physicians report burnout [1]**, with communication-related stress cited as a major contributor, and **poor communication accounts for 30% of malpractice cases [2]**. Simultaneously, higher perceived clinician empathy significantly improves patient satisfaction and treatment adherence.

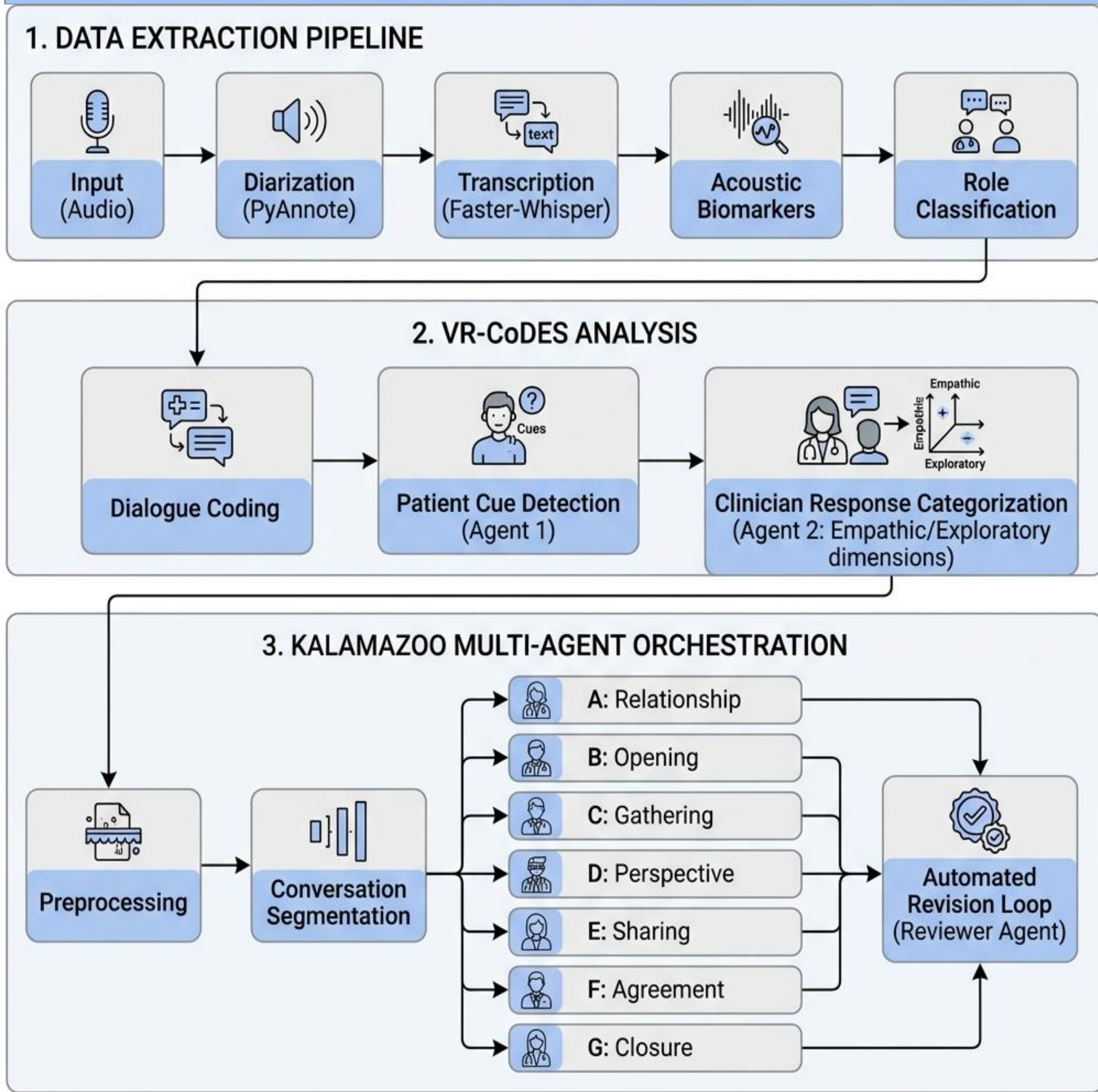
References



Mission Statement

Our mission is to create an AI-powered, privacy first, communication coach that delivers fast, objective, and personalized feedback to clinicians, empowering them to strengthen empathy, improve patient trust, and reduce burnout.

Final Technical Model



Prototype

The prototype interface includes several key components:

- Conversation Analysis:** A radar chart showing interaction metrics like 'Provides Closure' (1.0), 'Builds a Relationship' (3.0), 'Opens the Discussion' (2.0), 'Reaches Agreement' (2.0), 'Gathers Information' (2.0), 'Shares Information' (2.0), and 'Understands Perspective' (4.0).
- Patient Frustration Scenario:** A detailed view of a conversation transcript with analysis overlays, such as 'Emotional Cues' and 'Patient Expressed: Vague Words'.
- Session Activity:** A dashboard showing session statistics (12 sessions, 1h 5m total time) and performance metrics like '13 WPM Increase in Pacing' and '10% Shift in Balance'.
- Key Metrics:** Graphs for 'Emotional Dynamics (VAD)' showing Avg. Balance (48%), Avg. Response Time (1.44s), Avg. Pacing (165 WPM), and Avg. Interruptions (0.8).

Final Design Specifications

CORE CAPABILITY	SPECIFICATION	CLINICAL GOAL
ASR	Faster-Whisper (Large) WER < 9%	Verbatim Records
Diarization	PyAnnote 3.1	Interaction Tracking
Affective	Wav2Vec2-MSP-Dim	Emotional Awareness
Acoustic	OpenSMILE eGeMAPS	Tone Analysis
Role-ID	Neural Role AI	Attribution
Privacy	De-ID & Local Execution	HIPAA Compliance
Analysis	Qwen 3:14b	Kalamazoo Feedback

Design Status

This iteration represents our fully functional Version 1, with early validation already proving MedComAI is both intelligent and effective. While the core architecture and UI are deployable, we are awaiting IRB approval to begin the final phase. Once secured, we will conduct rigorous clinical validation and fine-tuning to ensure the platform perfectly aligns with professional workflow standards and privacy requirements.

Future Steps

- Secure Formal Clearance:** Obtain the necessary formal IRB approval from the Mayo Clinic to proceed to the next phase.
- Execute Rigorous Validation:** Conduct the final stages of comprehensive testing and validation for the platform.
- Fine-Tune Models:** Complete the necessary technical model fine-tuning processes.
- Align with Clinical Workflows:** Ensure the system fully integrates with and supports established clinical workflow standards.
- Verify Privacy Compliance:** Confirm that the platform strictly adheres to all required privacy regulations and requirements.

Acknowledgements

We would like to thank Aman Verma, Molly Svendsen, Asif Salekin, Elia Lima-Walton for their guidance, expertise, and invaluable feedback throughout the development of this project. We also gratefully acknowledge Mayo Clinic and ASU SBHSE for providing us with the opportunity to pursue this project.

Verification Results

Scenario	WER (%)	DER (%)
Mental Health (11:00)	0.00%	3.6%
Frustrated Patient (4:30)	0.00%	1.5%
Discussing Costs (5:13)	0.13%	19.0%
Heart Symptoms (13:24)	0.39%	19.8%
Breast Cancer (4:24)	0.60%	3.2%
AVERAGE	0.22%	11.7%

Figure 1: Benchmark of transcript word error rate (WER) and diarization error rate (DER)

Diarization Performance Analysis:

- Exceptional transcription performance with a near-perfect aggregate **Word Error Rate (WER) of 0.22%**.
- Impact of Crosstalk:** Error rates are significantly higher in "Crosstalk" scenarios, reaching 19.8% in multi-speaker history taking

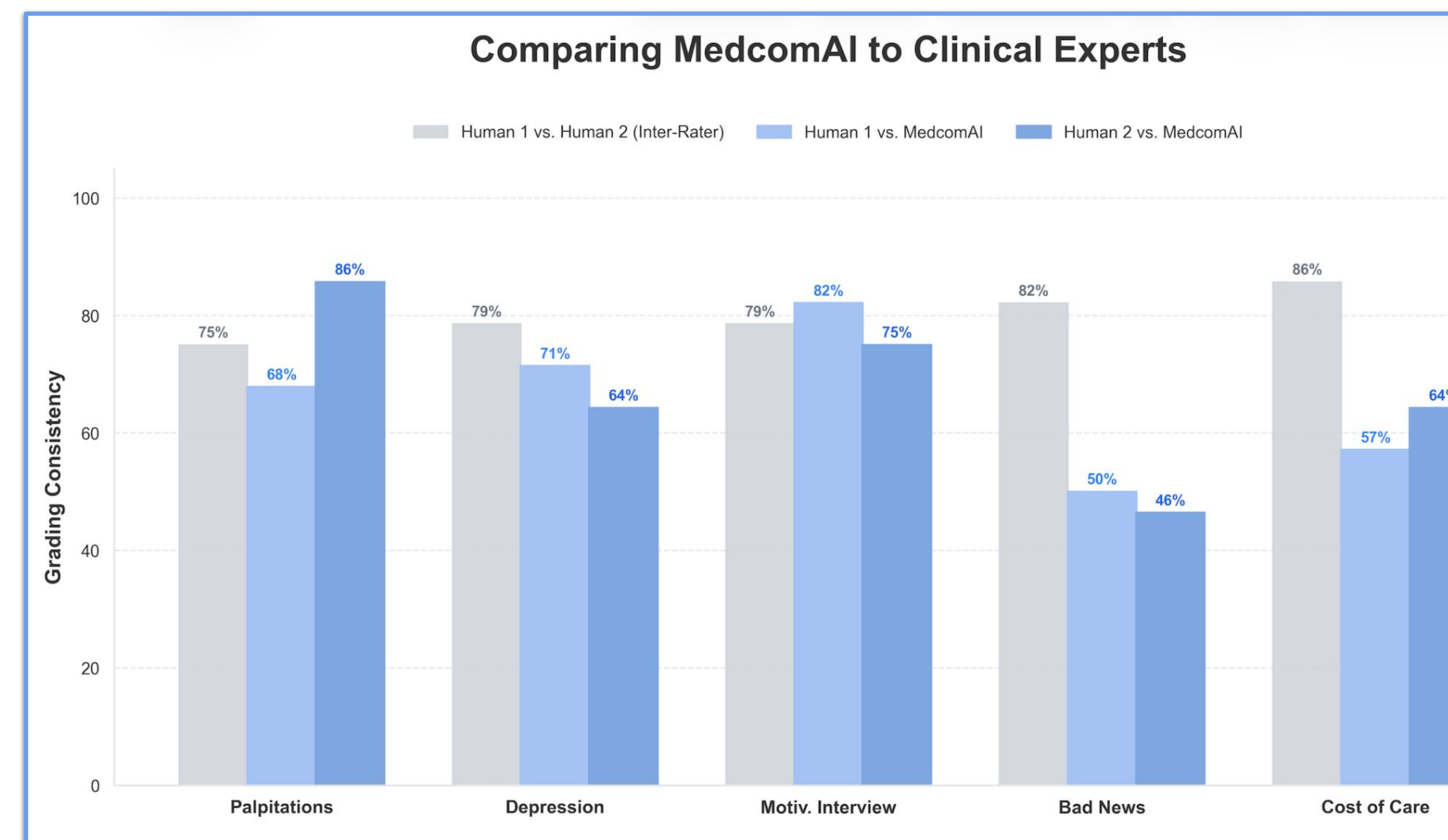


Figure 2: Performance of MedcomAI AI-generated Kalamazoo scores against Human Experts.

- Human-to-Human Benchmarking:** In structured clinical scenarios, human experts showed a baseline agreement of 79%. MedcomAI achieved a **higher consistency index (82%)** with expert raters than the experts did with each other.
- Validated Scalability:** Validated across diverse scenarios (Motivational Interviewing, Depression, Palpitations)

Contribute to Validation >

