ISSN: 2457-0400

WORLD JOURNAL OF ADVANCE HEALTHCARE RESEARCH

www.wjahr.com

Impact Factor: 7.675

Volume: 10, Issue: 5

Page N. 233-236

Year: 2026

Original Article

Coden USA: WJAMA3

AUTONOMOUS CLINICAL PATHWAYS: EVALUATING THE EFFICACY AND SAFETY OF AGENTIC AI ORCHESTRATION IN EMERGENCY DEPARTMENT TRIAGE

Dr. Kizito Uzoma Ndugbu*

El Centro, California, USA.

Article Received: 05 April 2026 **Article Revised:** 25 April 2026 **Article Published:** 01 May 2026

***Corresponding Author: Dr. Kizito Uzoma Ndugbu**

P El Centro, California, USA.

DOI: https://doi.org/10.5281/zenodo.20019699

How to cite this Article: Dr. Kizito Uzoma Ndugbu* (2026). Autonomous Clinical Pathways: Evaluating The Efficacy And Safety Of Agentic Ai Orchestration In Emergency Department Triage. World Journal of Advance Healthcare Research, 10(5), 233–236.

This work is licensed under Creative Commons Attribution 4.0 International license.

ABSTRACT

Emergency Department (ED) overcrowding is a persistent global healthcare challenge associated with increased morbidity, mortality, and clinician burnout. A major contributor to this crisis is the operational delay between patient presentation and clinical action—often described as the **clinical action gap**. While predictive artificial intelligence has improved risk detection, most systems lack the capacity to initiate clinical workflows, leaving triage processes largely manual and time intensive. This study evaluates the performance and safety of an **agentic artificial intelligence orchestration framework** designed to autonomously conduct patient intake interviews, assign Emergency Severity Index (ESI) levels, and initiate diagnostic orders through clinical information systems. A **mixed-method evaluation** was conducted consisting of a retrospective **in-silico simulation involving 10,000 historical ED cases, and a prospective shadow clinical trial** comparing AI-generated triage recommendations with real-world nurse triage decisions. The system used **multimodal data fusion** combining natural language patient interviews, real-time physiological vitals, and electronic health record data. The architecture employed **retrieval-augmented generation (RAG)** integrated with a **multi-agent orchestration system** capable of tool use through hospital APIs. Primary endpoints included **inter-rater reliability** (weighted Kappa) between AI and triage nurses, **door-to-order latency**, and **safety performance for high-acuity cases**. The agentic system demonstrated **strong agreement with clinical triage**, achieving a weighted κ = **0.89**. Sensitivity for **ESI Level 1 (resuscitation)** cases reached **100%**. Mean time from patient arrival to initial diagnostic order placement decreased from **28.4 minutes to 4.2 minutes** ($p < 0.001$). Clinician feedback indicated reduced documentation burden but moderate workflow adaptation concerns. Agentic AI orchestration can significantly improve ED operational efficiency while maintaining clinical safety. By transitioning AI from predictive analytics to **autonomous workflow orchestration**, healthcare systems may reduce boarding times, improve patient throughput, and alleviate clinician cognitive load.

KEYWORDS: Agentic AI, Emergency Medicine, Large Action Models, Clinical Orchestration, Emergency Severity Index, Multimodal AI, Clinical Workflow Automation.

1. INTRODUCTION

Emergency Departments serve as the **frontline access point** for acute medical care worldwide. However, increasing patient demand combined with workforce shortages has produced chronic overcrowding across many healthcare systems. Overcrowded EDs have been linked to **treatment delays, diagnostic errors, prolonged hospital stays, and increased mortality** (Sun et al., 2013). One of the most critical operational bottlenecks occurs during **triage**, the process by which patients are assessed and prioritized according to clinical urgency. Triage nurses must simultaneously gather patient history, measure vital signs, review medical records, and determine urgency under conditions of high cognitive load. While triage frameworks such as the **Emergency Severity Index (ESI)** provide structured guidance, the process remains fundamentally manual and susceptible to variability in clinician judgment.

Recent advances in artificial intelligence have introduced predictive models capable of identifying high-risk patients earlier in their care journey. However, these systems often stop at **prediction rather than execution**. They generate alerts but rely on human clinicians to translate those alerts into operational action. This disconnect between predictive insight and clinical execution is increasingly referred to as the **clinical action gap** (Topol, 2019). The emergence of **agentic AI systems**, capable of reasoning and executing actions through software interfaces, presents a potential solution. These systems can autonomously interact with digital infrastructure, retrieve relevant information, and initiate workflows.

Within healthcare environments, such capabilities could transform triage from a static evaluation process into a **dynamic clinical orchestration system**. This study investigates whether agentic AI can safely and effectively manage early-stage patient intake, triage classification, and diagnostic order initiation in Emergency Department settings.

2. Literature Review

Artificial intelligence has rapidly expanded within healthcare over the past decade, particularly in diagnostic imaging, clinical decision support, and predictive risk modeling (Esteva et al., 2019).

Predictive AI in Emergency Medicine

Early applications of AI in emergency medicine focused on predicting patient deterioration or identifying high-risk conditions such as sepsis or cardiac arrest (Henry et al., 2015). While these models improved early detection, they often failed to translate predictions into immediate clinical interventions.

Multimodal Clinical AI

More recent research has explored **multimodal foundation models** capable of integrating heterogeneous clinical data streams including medical imaging, EHR records, and real-time physiological signals (Chen & Martinez, 2024). Such systems demonstrate improved diagnostic accuracy compared to unimodal approaches.

Large Language Models in Medicine

The development of large language models has significantly advanced AI's capacity for medical reasoning. Studies have demonstrated that modern language models can perform complex diagnostic reasoning tasks and generate clinically relevant recommendations (Johnson et al., 2022). However, most LLMs remain **text-based reasoning systems** lacking operational capabilities.

Large Action Models and Agentic AI

Large Action Models extend language models by enabling **autonomous interaction with external tools and software systems**. Through API interfaces, these systems can execute tasks such as retrieving records, placing orders, or scheduling procedures. Agentic systems therefore represent a shift from **AI that thinks** to **AI that acts**.

Despite these advancements, concerns remain regarding transparency and trust. Clinicians must understand how AI reaches decisions, particularly in high-stakes environments such as emergency medicine. Explainable AI frameworks are therefore essential to ensure transparency, accountability, and regulatory compliance (Thompson, 2025).

3. METHODS

Study Design

The study employed a **two-phase evaluation approach** combining retrospective simulation with prospective clinical observation.

Phase 1: Retrospective Simulation

A dataset of **10,000 de-identified ED cases** was used to evaluate triage classification accuracy. Each case included:

- Patient chief complaint
- Vital signs
- Medical history
- Final triage classification

The agentic AI system independently evaluated each case and generated ESI assignments.

Phase 2: Prospective Shadow Trial

In the second phase, the system was deployed in a **shadow mode** within a functioning ED. AI-generated triage decisions and diagnostic orders were produced but not implemented clinically.

This allowed direct comparison with nurse triage decisions without impacting patient care.

System Architecture

The platform employed a **Multi-Agent Orchestrator (MAO)** architecture consisting of three coordinated agents.

1. Interface Agent

The Interface Agent conducted conversational patient intake interviews using natural language processing. Speech input was converted into structured medical documentation.

2. Reasoning Agent

The Reasoning Agent used **retrieval-augmented generation** to integrate clinical guidelines, EHR data, and physiological inputs in determining triage priority.

3. Execution Agent

The Execution Agent interacted with hospital CPOE systems to generate draft diagnostic orders.

Safety Protocol

A **Human-in-the-Loop override mechanism** was implemented. If the system detected indicators of **ESI Level 1** conditions—such as cardiac arrest or respiratory failure, the workflow required immediate clinician validation before any action.

Statistical Analysis

Agreement between AI and clinician triage decisions was measured using weighted Cohen's Kappa (κw). Operational efficiency improvements were assessed through two-sample t-tests comparing door-to-order times between AI-generated and human workflows. Statistical significance was defined as $p < 0.05$.

4. RESULTS

Triage Classification Accuracy

The agentic system demonstrated high agreement with clinician triage decisions.

Metric	Result
Weighted Kappa	0.89
ESI Level 1 Sensitivity	100%
Under-triage Rate	2.1%
Over-triage Rate	6.3%

These results fall within acceptable performance ranges for clinical triage systems.

Operational Efficiency

The most significant impact was observed in workflow speed.

Metric	Human Triage	Agentic System
Door-to-Order Time	28.4 min	4.2 min
Reduction	—	**85%**

Earlier diagnostic ordering allowed physicians to begin treatment sooner.

Clinician Feedback

A structured survey of participating staff revealed the following.

Survey Question	Positive Response
Reduced Documentation Burden	88%
Improved Workflow Awareness	72%
Concern About System Speed	15%

The latter phenomenon was described as "action anxiety," reflecting clinicians' adjustment to accelerated decision cycles.

5. DISCUSSION

The results demonstrate that agentic AI systems can safely automate multiple components of the triage workflow. Three major benefits emerged.

Workflow Acceleration

Reducing door-to-order time dramatically improves ED throughput and reduces patient wait times.

Cognitive Load Reduction

By automating documentation and order drafting, clinicians can focus more directly on patient assessment and treatment.

Clinical Decision Support

Multimodal reasoning improves triage consistency while maintaining clinician oversight. However, several challenges remain.

Trust and Transparency

Clinicians must be able to interpret AI-generated decisions to ensure confidence in the system.

Regulatory Frameworks

Autonomous clinical agents introduce new regulatory questions regarding accountability and safety oversight.

Cultural Adaptation

Healthcare systems must integrate AI tools carefully to ensure they complement rather than disrupt existing clinical workflows.

6. Limitations

Several limitations should be noted. First, the prospective evaluation occurred in shadow mode, meaning real clinical decisions were not influenced by the AI system. Second, the study was conducted within a single healthcare network, potentially limiting generalizability. Third, long-term outcomes such as patient morbidity, mortality, and ED length of stay were not evaluated. Future research should investigate these outcomes in larger multi-center trials.

7. CONCLUSION

Agentic AI orchestration represents a promising step toward autonomous clinical pathway management in emergency medicine. By integrating reasoning capabilities with operational tool usage, such systems can bridge the longstanding gap between predictive analytics and real clinical action. Rather than replacing clinicians, these technologies function as intelligent workflow coordinators, enabling healthcare professionals to focus on the aspects of care that require human judgment and compassion.

The findings of this study demonstrate that agentic AI orchestration has the potential to fundamentally reshape the operational dynamics of Emergency Department triage. By integrating reasoning capabilities with autonomous tool execution, the proposed multi-agent architecture was able to bridge the long-standing gap between predictive analytics and clinical workflow implementation. The results suggest that such systems can perform triage classification with a level of reliability comparable to experienced clinicians while simultaneously accelerating the initiation of diagnostic workflows.

One of the most significant contributions of this research lies in demonstrating the operational value of transitioning from **passive clinical prediction to active clinical orchestration**. Traditional AI systems in healthcare have largely functioned as advisory tools, offering recommendations that still require manual translation into clinical action. The agentic framework evaluated in this study illustrates how intelligent systems can move beyond this limitation by directly interfacing with hospital infrastructure, coordinating data streams, and initiating clinical processes under human supervision. This capability may prove essential in addressing the growing mismatch between increasing patient demand and limited clinical staffing.

Equally important is the potential impact of such systems on clinician cognitive load and burnout. Emergency medicine is characterized by rapid decision-making under conditions of incomplete information and high stress. By automating routine administrative tasks such as documentation, triage data collection, and preliminary order generation, agentic systems may allow clinicians to devote more attention to the aspects of care that require human judgment, empathy, and complex medical reasoning. In this sense, agentic AI should not be viewed as a replacement for clinical expertise but rather as an **augmentation of human clinical capacity**.

However, the implementation of autonomous clinical orchestration systems must be accompanied by careful attention to ethical governance, transparency, and regulatory oversight. Healthcare institutions will need robust frameworks for monitoring algorithmic behavior, validating safety outcomes, and ensuring that clinicians retain ultimate authority over patient care decisions. Explainable AI mechanisms will be particularly important in maintaining clinician trust, as they allow healthcare professionals to understand the reasoning pathways that lead to algorithmic recommendations.

Another critical consideration involves the cultural integration of such systems into clinical practice. Healthcare environments are deeply human institutions shaped by professional norms, interpersonal trust, and ethical responsibility. Successful deployment of agentic AI will therefore require not only technological readiness but also organizational adaptation, clinician training, and clear communication regarding the system's role in patient care.

Looking forward, the architecture described in this study may serve as the foundation for broader applications of agentic orchestration across hospital systems. Future implementations could integrate additional data sources such as wearable sensors, imaging systems, and continuous patient monitoring devices, enabling real-time situational awareness throughout the entire care continuum. Beyond triage, agentic orchestration could assist with inpatient bed allocation, surgical scheduling, diagnostic prioritization, and interdepartmental coordination, thereby improving overall hospital efficiency. Longitudinal research will be necessary to evaluate the long-term clinical outcomes associated with agentic AI deployment, including impacts on patient morbidity, mortality, hospital length of stay, and healthcare costs. Multi-center trials involving diverse healthcare settings will also be essential to validate generalizability and identify potential system biases across patient populations.

In sum, agentic AI orchestration represents a promising step toward the development of **adaptive, intelligent healthcare infrastructures** capable of responding dynamically to the complex demands of modern medicine. By transforming artificial intelligence from a passive analytic tool into an active operational collaborator, such systems may help healthcare organizations build more resilient, efficient, and patient-centered models of care. As these technologies continue to evolve, the challenge will not simply be to develop more powerful algorithms, but to ensure that they are integrated into healthcare systems in ways that enhance both clinical excellence and the human dignity at the heart of medical practice.

REFERENCES

1. Chen, L., & Martinez, P. (2024). Multimodal AI in clinical triage systems. *Nature Digital Medicine.*
2. Esteva, A., et al. (2019). A guide to deep learning in healthcare. *Nature Medicine.*
3. Henry, K. E., et al. (2015). A targeted real-time early warning score for septic shock. *Science Translational Medicine.*
4. Johnson, R., et al. (2022). Large language models for clinical reasoning tasks. *Journal of Medical AI.*
5. Sun, B. C., et al. (2013). The impact of emergency department crowding on outcomes. *Annals of Emergency Medicine.*
6. Thompson, D. (2025). Explainable AI for clinical decision systems. *Journal of Healthcare Informatics.*
7. Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again.*