



THE EVOLVING ROLE OF AI IN GENERAL SURGERY AND EMERGENCY GENERAL SURGERY: A SYSTEMATIC REVIEW



Arshpreet Kaur, MBBS, MS; Samreen Shaikh, MD; Sebastian Valdivieso, MD; Faran Bokhari, MD, FACS,MBA*; Vishnu R Mani, MD, FACS*

OSF Saint Francis Medical Center; University of Illinois College of Medicine -Peoria

Introduction

Artificial intelligence (AI) has demonstrated significant applications in medicine. Yet, its integration into surgery remains limited due to procedural variability, challenges in data standardization, and a lack of targeted research compared with imaging-driven fields. In general, emergency general surgery (EGS), AI models show promise for perioperative risk prediction and surgical planning, but concerns persist about accountability, overreliance, and potential deskilling.

Methodology

A systematic review was conducted according to PRISMA guidelines (2020–2025) across PubMed, Embase, Medline, and Cochrane Library. Six eligible studies were synthesized, including a large-scale validation of the POTTER (Predictive Optimal Trees in Emergency Surgery Risk) calculator in EGS, a prospective comparison of POTTER versus surgeon gestalt, an international cross-sectional survey of 650 trauma surgeons on AI adoption, and an ethical analysis addressing barriers and concerns in surgical AI implementation.

Results

In 59,955 EGS patients (2017 NSQIP dataset), POTTER demonstrated strong predictive accuracy for mortality (c-statistic 0.93) and morbidity (0.83), particularly for septic shock, respiratory failure requiring prolonged ventilation, and acute renal failure. In 150 prospective EGS cases, surgeon interaction with POTTER improved the prediction of mortality, bleeding, and pneumonia, though performance for septic shock and ventilator dependence was not statistically significant. Despite these findings, survey results revealed that most of the 650 surgeons across 71 countries preferred traditional decision aids, with AI rated as only modestly relevant to decision-making, reflecting knowledge gaps; thereby reinstating the black–box effect leading to heterogeneous levels of trust.

Conclusion

AI demonstrates high potential to enhance risk prediction, precision, and efficiency in general surgery, particularly in EGS, as supported by large-scale validation and prospective comparisons with POTTER models. However, real-world adoption remains limited, hinging on multicenter validation, workflow integration, surgeon education, and ethical-legal clarity. Long-term reliance on AI raises concerns regarding accountability and potential cognitive decline in surgical decision-making, underscoring the need for AI to function as an adjunctive tool rather than a substitute for clinical judgment.