Predictive Power in GI Surgery: Machine Learning Models and the AI-Driven Healthcare Revolution

Olivia Rennie[1], Nour Helwa[1]

1. FluidAI Medical (Formerly NERv Technology Inc.), Kitchener, Ontario, Canada

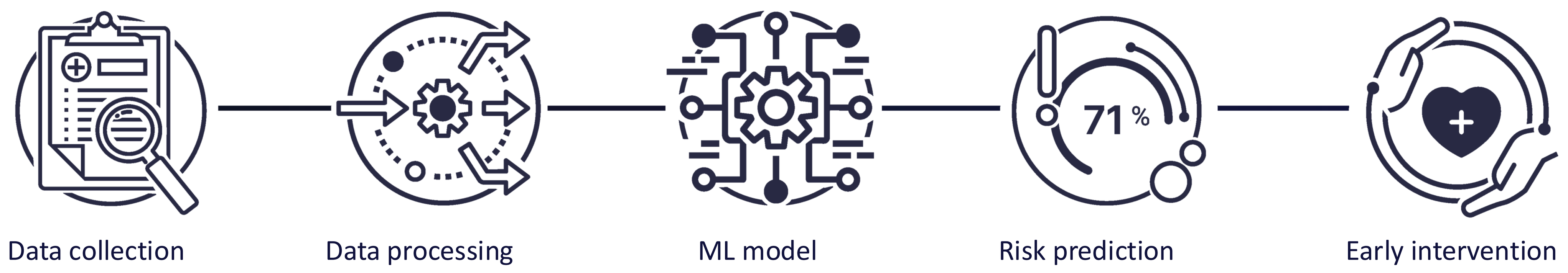
Background

The integration of Artificial Intelligence (AI) and Machine Learning (ML) into gastrointestinal (GI) surgery has garnered considerable attention, demonstrating **notable promise in predicting postoperative complications.**

Anastomotic leaks, infections, and thromboembolic events remain key concerns in GI surgery, **contributing to increased morbidity, mortality, and healthcare costs.**

This review evaluates current AI and ML technologies and their predictive capabilities, particularly focusing **on their role in predicting adverse outcomes and improving patient care.**

The review also addresses current gaps in the application of AI, **highlighting areas ripe for future research that could further revolutionize surgical care through earlier and more precise prediction of postoperative complications.**



Methods

A comprehensive literature review was conducted by examining scientific databases (PubMed, Web of Science, OVID Embase, Google Scholar, and Cochrane library), using the following search terms: **artificial intelligence/AI, machine learning/ML, gastrointestinal/GI surgery, postoperative complications, and outcomes.**

Studies involving the use of ML models for predicting postoperative outcomes in GI surgery were selected, with **specific attention to performance metrics such as accuracy, sensitivity, and area under the receiver operating characteristic curve (AUC-ROC).**

The review also examined **current limitations and proposed future directions to enhance the clinical applicability of these models.**

Conclusion

This study provides a comprehensive literature review of the current landscape of AI/ML in the postoperative gastrointestinal surgery space, revealing the compelling work that has been completed thus far, as well as future directions of the field, to further push the boundaries of exceptional care.

These advancements signify a new era in surgery, where data-driven insights empower clinicians to deliver more efficient, targeted, and cost-effective healthcare.

Key Next Steps in AI Predictive Capabilities

Improving dataset quality
Expanding AI model validation

Scan our QR code for references

RESULTS

Current predictive models range in performance, with some of the best ranging between **83%-89% accuracy, and AUC-ROC scores around 0.85.**

In the gastrointestinal surgery space, research remains **heavily weighted towards colorectal patient populations, with far less research in the hepatobiliary (HPB) and upper GI (UGI) space.**

Use of ML in the postoperative space was concluded to **significantly outperform traditional clinical methods.**

By enabling early intervention, these technologies have the potential to **reduce complication rates and shorten hospital stays, contributing to improved patient outcomes and optimized healthcare resource utilization.**

Beyond improving clinical results, AI and ML bring **notable economic benefits by lowering costs associated with extended hospital stays and readmissions.** New technologies continue to emerge, designed to push the boundaries of postoperative care, such as providing earlier prediction for a wide range of significant surgical complications, thus offering healthcare providers a proactive approach to managing patient care.

Colorectal

Shen et al., (2024)
LASSO-logistic model, AUC: 0.790

Liang et al. (2024)
Multimodal approach: integration of clinical data with imaging results - Accuracy: 0.84, Recall rate: 0.82, F1 score: 0.81, AUC: 0.85

Taha-Mehlitz et al. (2024)
Random forest - AUC: 0.78; Accuracy: 0.82; F1 Score: 0.58
Logistic regression - AUC: 0.69; Accuracy: 0.81; F1 Score: 0.53

Ingwersen et al. (2022)
Artificial neural network - AUC: 0.85, Sensitivity: 0.93, Specificity: 0.57

Soguero-Ruiz et al. (2016)
Support Vector Machine - Sensitivity: 100%, Specificity: 72%

Colorectal, HPB

Chen et al. (2022)

- Neural networks
  - Colorectal AL - AUC: 0.676
  - Bile Leaks - AUC: 0.750
  - POPF - AUC: 0.746
- Logistic regression
  - Colorectal AL - AUC: 0.633
  - Bile Leaks - AUC: 0.722
  - POPF - AUC: 0.713

Gastric

Shao et al. (2021)

- Logistic regression
- Random forest
- Support vector machine
- XGBoost

AUC: 0.89, Sensitivity: 81.8%
Specificity: 82.2%

Bariatric

Nudel et al. (2021)

- Artificial neural networks - AUC: 0.75
- Gradient Boosting Machine (XGB) - AUC: 0.70
- Logistic regression - AUC: 0.63

Esophageal, gastric

Van Kooten et al. (2022)

- LASSO regression
- Logistic regression
- k-Nearest neighbors
- Neural networks
- Support vector machine
- Random Forest
- Adaboost
- Super learner

Linear regression had the highest predictive value, with AUC values varying between 0.619 - 0.68, but the difference between ML models did not reach statistical significance

Esophageal

Klontzas et al. (2024)
XGBoost - AUC: 0.792, Specificity: 77.46%, Sensitivity: 65.22%, PPV: 48.39%, NPV: 87.3%, F1-score: 56%

Zhao et al. (2021)

- Decision tree
- Random forest
- Naive Bayes
- Logistic regression with least absolute shrinkage and selection operator

AUC: 0.72