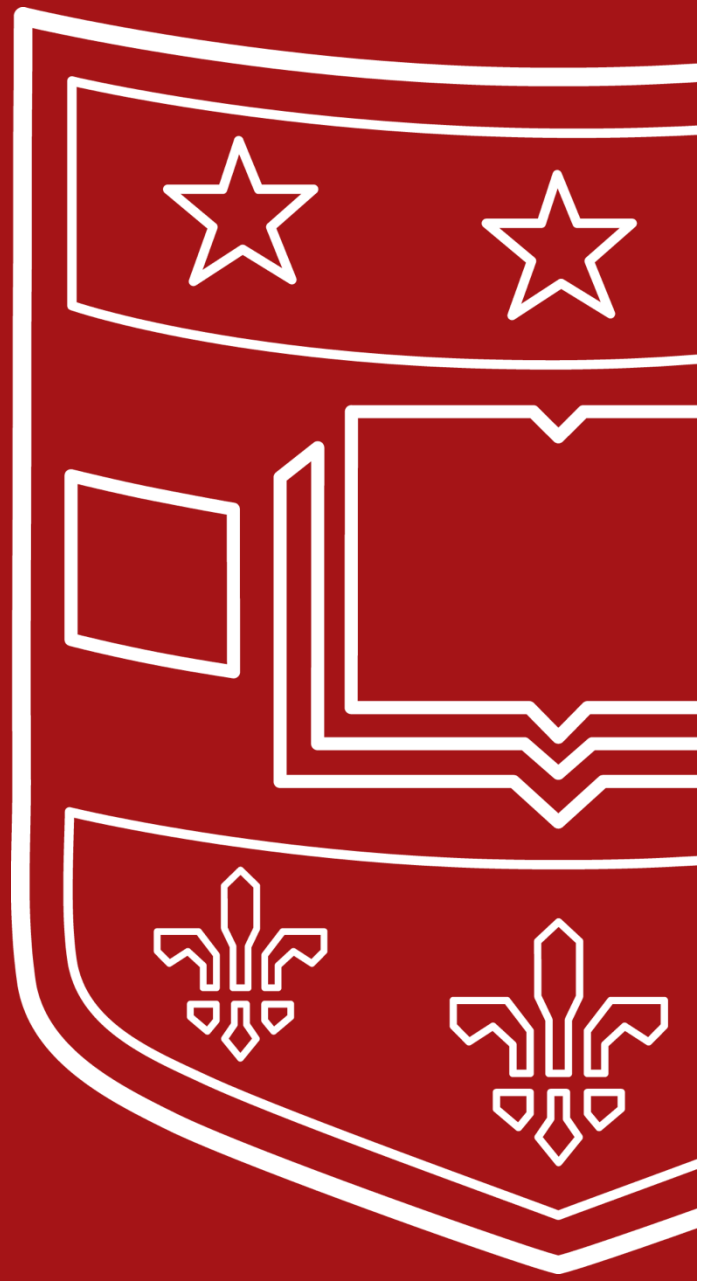
Predicting Postoperative Complications from Clinical Notes

Chenyang Lu
Director, AI for Health Institute (AIHealth)
Fullgraf Professor
Computer Science · Anesthesiology · Medicine · Neurosurgery · Psychiatry · Public Health



Washington University in St. Louis

AI for Surgery

PHENOTYPE
complex diseases

DETECT EARLIER
identify risk sooner

PREDICT OUTCOMES
improve prognostication

GUIDE DECISIONS
support personalized care



CLINICAL DATA

EHR | Clinical notes | Imaging | Genomics

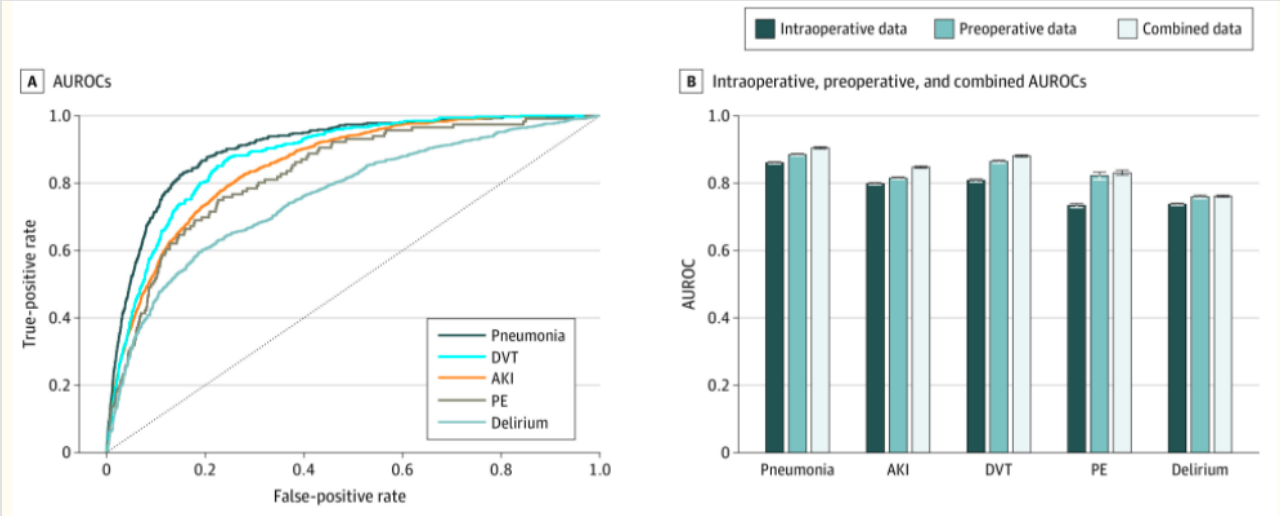
PATIENT-GENERATED DATA

Wearables | Smartphones

AI transforms data into actionable insights to improve surgical care and patient outcomes.

Predict Postoperative Complications

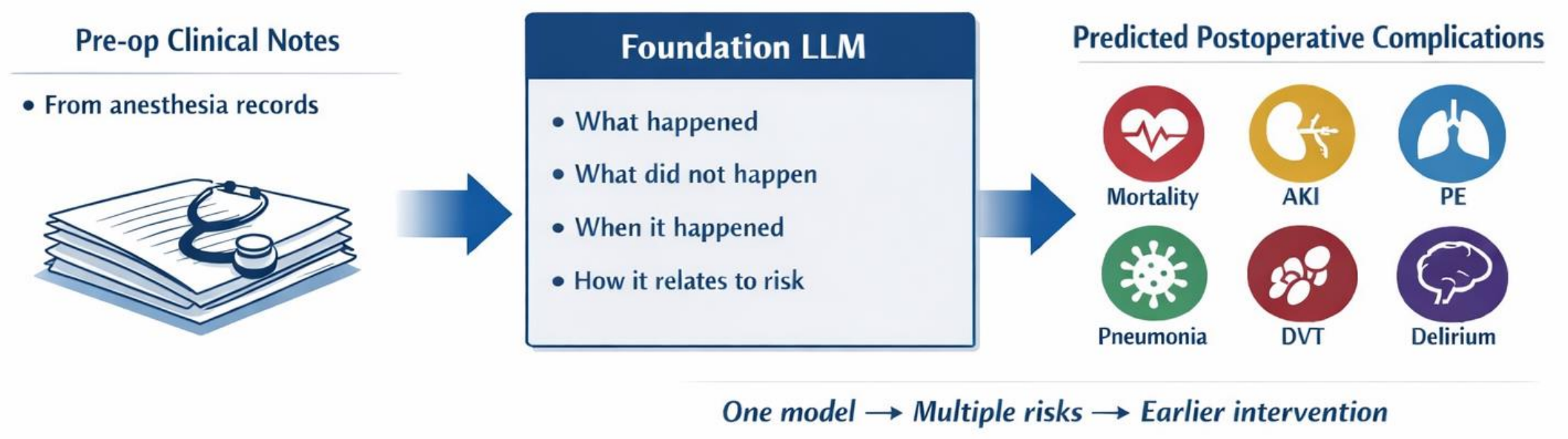
- Benefits of **predictive** models
  - Early identification of risk factors
  - Anticipatory management for perioperative contingency planning
  - Shortened hospital stays and reduced complication risks
- Most existing models use **structured** EHR data



B. Xue, D. Li, C. Lu, C.R. King, T. Wildes, M.S. Avidan, T. Kannampallil, and J. Abraham, Use of Machine Learning to Develop and Evaluate Models Using Preoperative and Intraoperative Data to Identify Risks of Postoperative Complications, *JAMA Network Open*, 4(3):e212240, 2021.

Unlocking Postoperative Risk from Clinical Notes

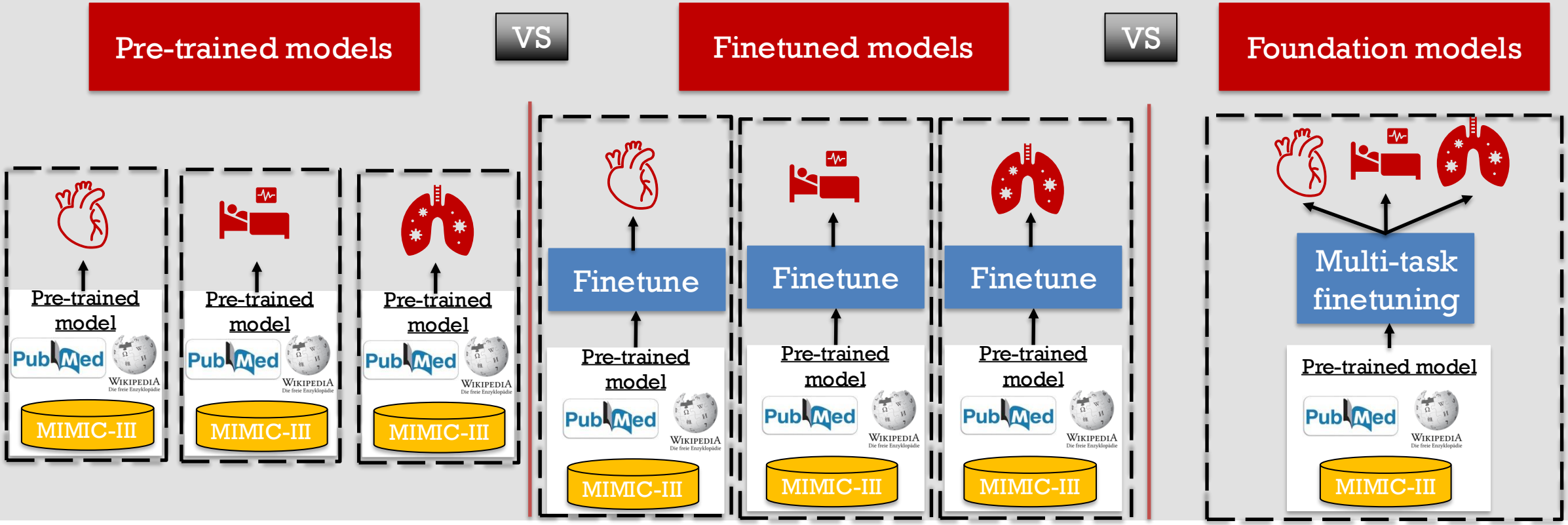
- Clinical notes capture nuanced clinical judgment that structured data cannot.
- **Large Language Models (LLMs)** translate narrative documentation into risk signals.



- Dataset: 84,875 pre-operative anesthesia notes from Barnes-Jewish Hospital (BJH)

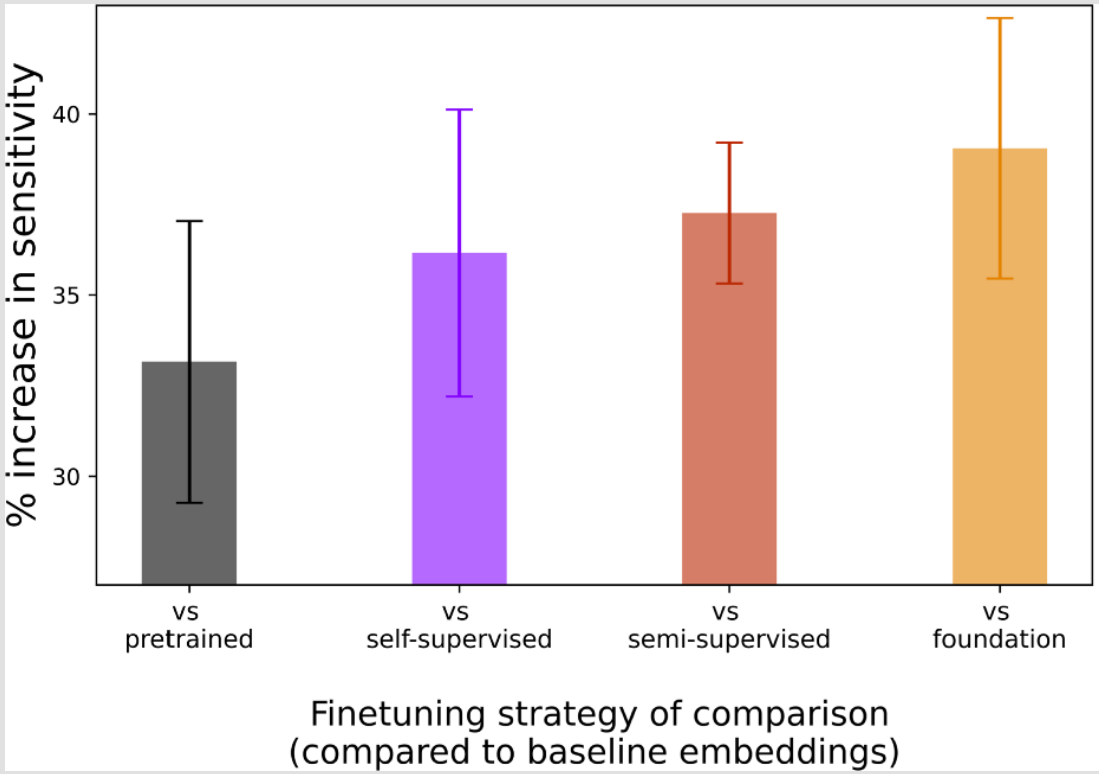
LLMs for Predicting Postoperative Risks

1. Do **pre-trained** LLMs outperform traditional NLP?
2. Does **fine-tuning** to surgical notes improve performance?
3. Can a single **foundation** model predict multiple risks?



Foundation LLMs Improve Surgical Risk Prediction

- Pretrained: **+33** percentage point increase over NLP in sensitivity (at 95% specificity)
- **One** foundation model predicts **multiple** complications
  - **Outperforms** individually fine-tuned models
  - **Simplifies** clinical deployment



The number of high-risk patients correctly identified per 100 cases that were missed by baseline models.

Alba C, Xue B, Abraham J, Kannampallil T, Lu C. The foundational capabilities of large language models in predicting postoperative risks using clinical notes. *npj Digital Medicine*. 8, 95 (2025).

From Model to Clinical Impact



Embedding risk prediction into pre-op assessment



Transparent risk estimates that support clinical judgment



Efficient fine-tuning for resource-constrained systems

Thank You

Learn more and stay connected

AI for Health Institute

aihealth.wustl.edu

Chenyang Lu

cse.wustl.edu/~lu