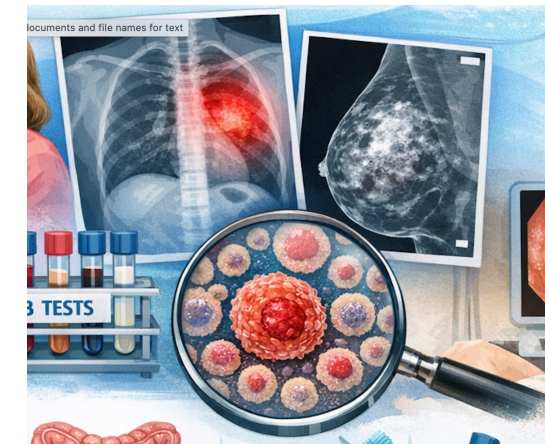




Background



➤ The average cost of cancer screening, such as mammogram or colonoscopy, can range from hundreds to over a thousand dollars. This study explores the potential of building a cancer foundation model based on Generative Pre-trained Transformers (GPT) to predict future outcomes for patients. We assess the prediction accuracy and feasibility of leveraging the foundation model to inform when screening can be prioritized, thereby reducing the associated burden of unnecessary procedures.

Methods

➤ Patient population

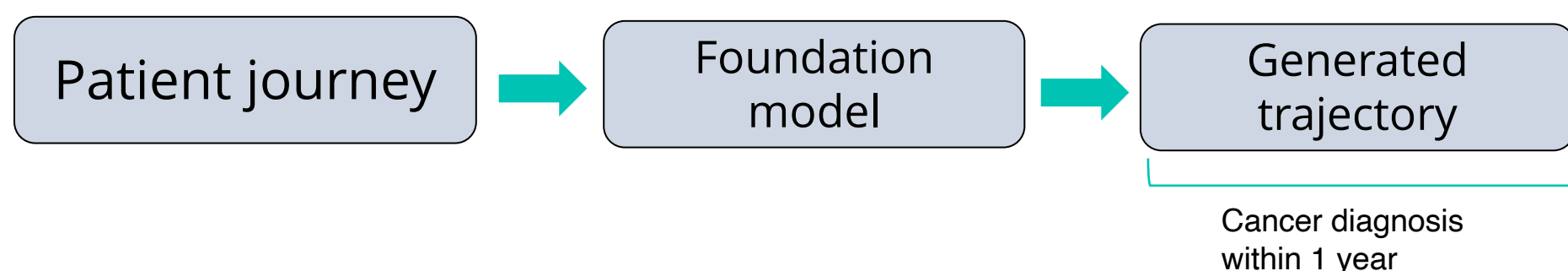
- 2.6 millions diagnosed with 4 types of cancers (lung, breast, colorectal, prostate)
- Train dataset: 1.4 millions
- Test dataset: 500 (study 1), 10,000 (study 2)

	Breast	Prostate	Lung	Colorectal
Count	976340 (37.77%)	686250 (26.55%)	443077 (17.14%)	479486 (18.55%)
Average Age	68.82 ±18.26	74.80 ±12.80	72.80 ±14.73	69.28 ±15.55

➤ Training approach

EHR data → Patient journeys → tokenization → model pretraining

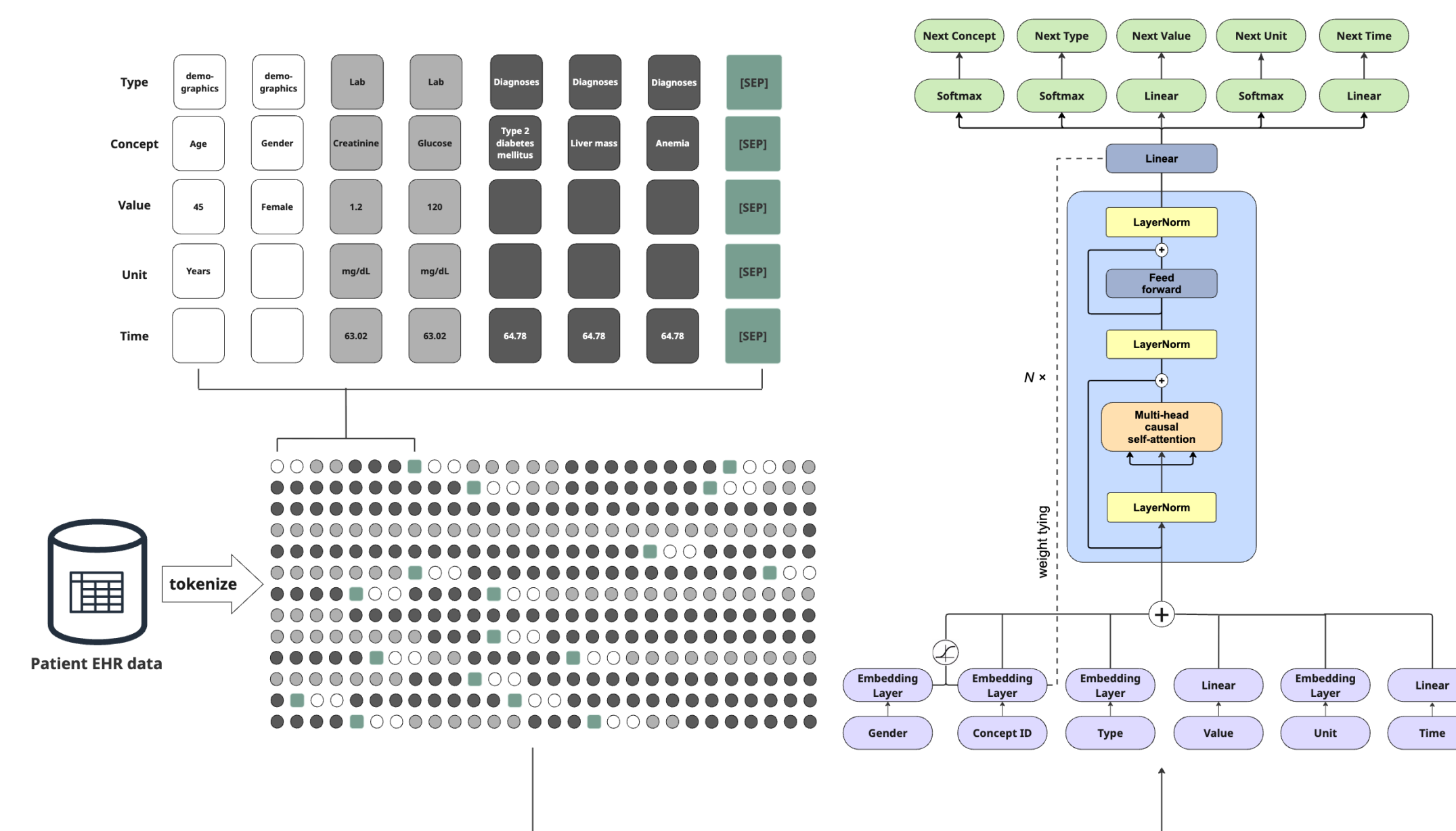
➤ Cancer outcome prediction



Conclusions

- **High specificity and NPV:** Suggest the feasibility of applying the patient journey foundation model to predict negative cancer outcomes with high accuracy (Study 1)
- **Comparable performance:** supervised ML models and PJFM showed generally similar performance, while no labelling data are needed for training PJFM (Study 2)
- **Gender specific diseases:** Tree based ML models can be biased by confounding features, such as gender (Study 2)

Patient journey data processing and model architecture



Results

Study 1 (include conditions and diagnoses in patient journeys)

	Breast	Prostate	Lung	Colorectal
Sensitivity	0.71	0.79	0.57	0.38
Specificity	0.94	0.89	0.97	0.98
PPV	0.83	0.67	0.86	0.85
NPV	0.89	0.94	0.85	0.86

Note: Test Set: 500 patients

Study 2 (include conditions, diagnoses and lab results in patient journeys)

Model	Metric	Breast	Prostate	Lung	Colorectal
MLR (Age+Gender)	AUROC	0.847 (±0.01)	0.883 (±0.01)	0.580 (±0.01)	0.603 (±0.01)
	AUPRC	0.617 (±0.02)	0.915 (±0.01)	0.366 (±0.02)	0.376 (±0.02)
MLR	AUROC	0.883 (±0.01)	0.915 (±0.01)	0.716 (±0.01)	0.660 (±0.01)
	AUPRC	0.706 (±0.02)	0.733 (±0.02)	0.366 (±0.02)	0.376 (±0.02)
RF	AUROC	0.878 (±0.01)	0.916 (±0.06)	0.721 (±0.01)	0.650 (±0.02)
	AUPRC	0.687 (±0.02)	0.741 (±0.02)	0.392 (±0.02)	0.371 (±0.02)
XGBoost	AUROC	0.916 (±0.01)	0.943 (±0.01)	0.807 (±0.01)	0.763 (±0.01)
	AUPRC	0.793 (±0.02)	0.829 (±0.02)	0.542 (±0.03)	0.525 (±0.02)
Out PJFM	AUROC	0.866 (±0.01)	0.866 (±0.03)	0.846 (±0.02)	0.760 (±0.03)
	AUPRC	0.733 (±0.01)	0.702 (±0.06)	0.557 (±0.06)	0.469 (±0.05)

Note: Test Set: 10,000 patients

Future potential for research

- Pretrain with a larger general population
- Include other clinical events, e.g. observations, procedures, medications
- Evaluate different disease outcome predictions