

Can machine learning improve survival prediction after colorectal cancer surgery?

Mohammad Asghari-Jafarabadi, Simon Wilkins, Katherine Goodall, Veena Suresh, Rathin Gosavi, Paul J. McMurrick

Cabrini Research | Monash University | Cabrini Monash University Department of Surgery



Background

Postoperative prognosis after colorectal cancer surgery is important for counselling, surveillance and follow-up. Machine-learning methods may capture complex tumour, patient and treatment information more flexibly than traditional survival models.



Methods

Patients from the Cabrini Monash Colorectal Neoplasia Database were analysed. Conventional Cox-based survival modelling was compared with machine-learning approaches.

Outcomes: overall survival (OS), cancer-specific survival (CSS), and progression-free survival (PFS).

Primary emphasis: 5-year prediction, discrimination and clinical reliability of risk estimates.

 Cabrini Monash Colorectal Neoplasia Database

 Cox models vs machine learning

 5-year prediction, discrimination and reliability

Cohort summary (colorectal cancer surgery)

Overall survival (OS)

 **n = 3,340**
 **724 deaths**
5-year survival 76.0%

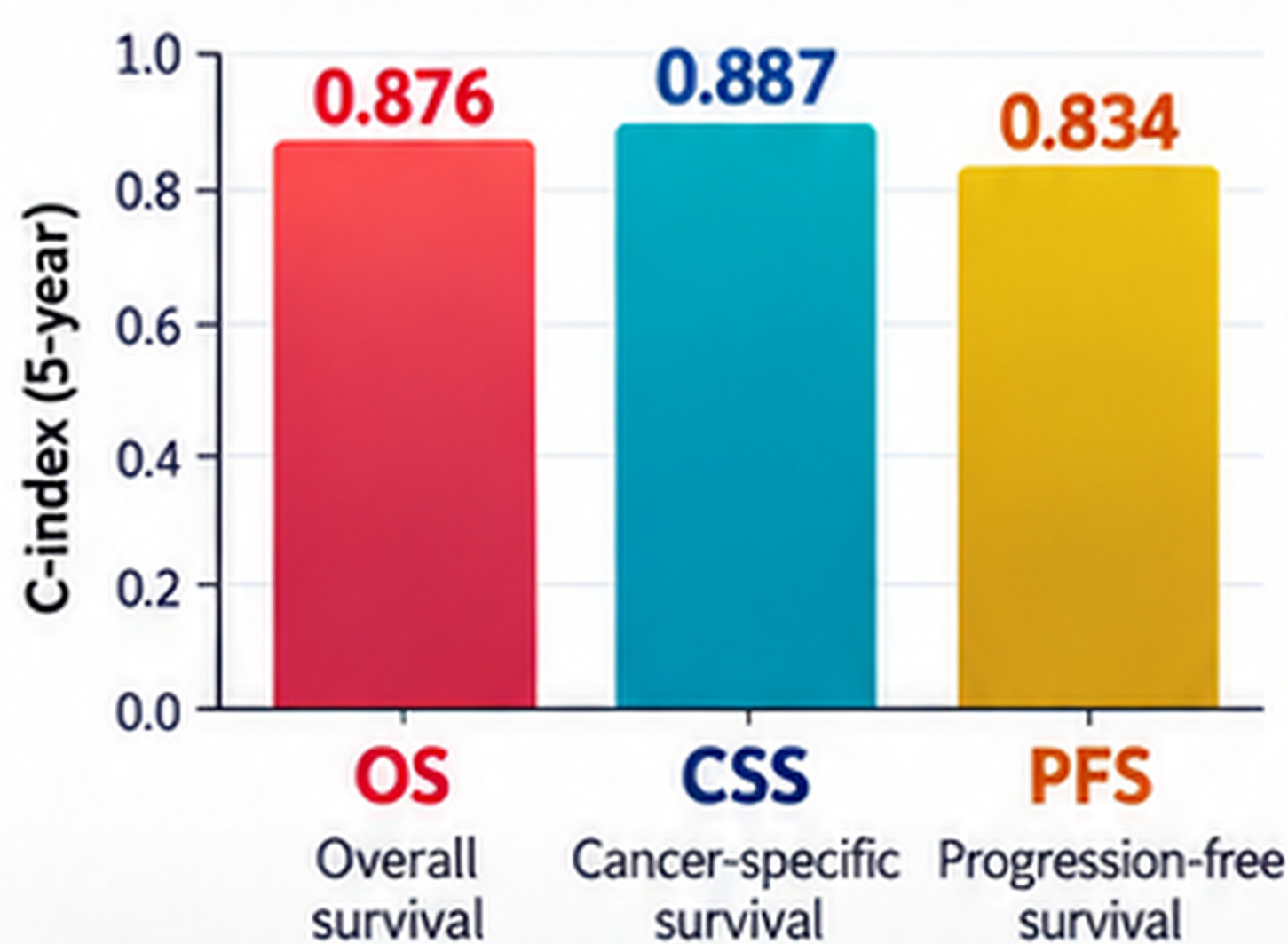
Cancer-specific survival (CSS)

 **n = 3,331**
 **515 cancer-related deaths**
5-year survival 82.4%

Progression-free survival (PFS)

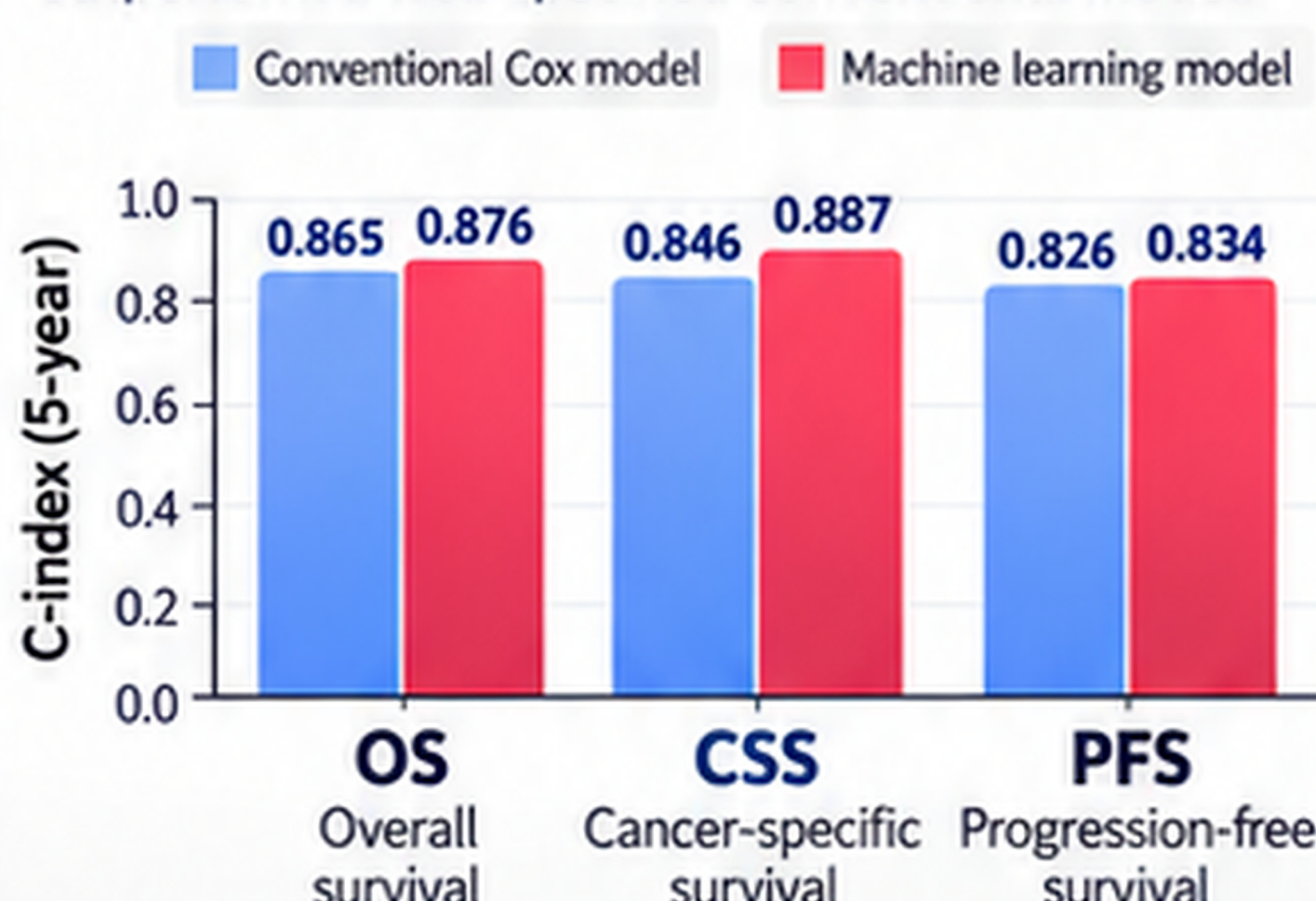
 **n = 3,339**
 **776 progression/death events**
5-year survival 74.4%

Best 5-year discrimination

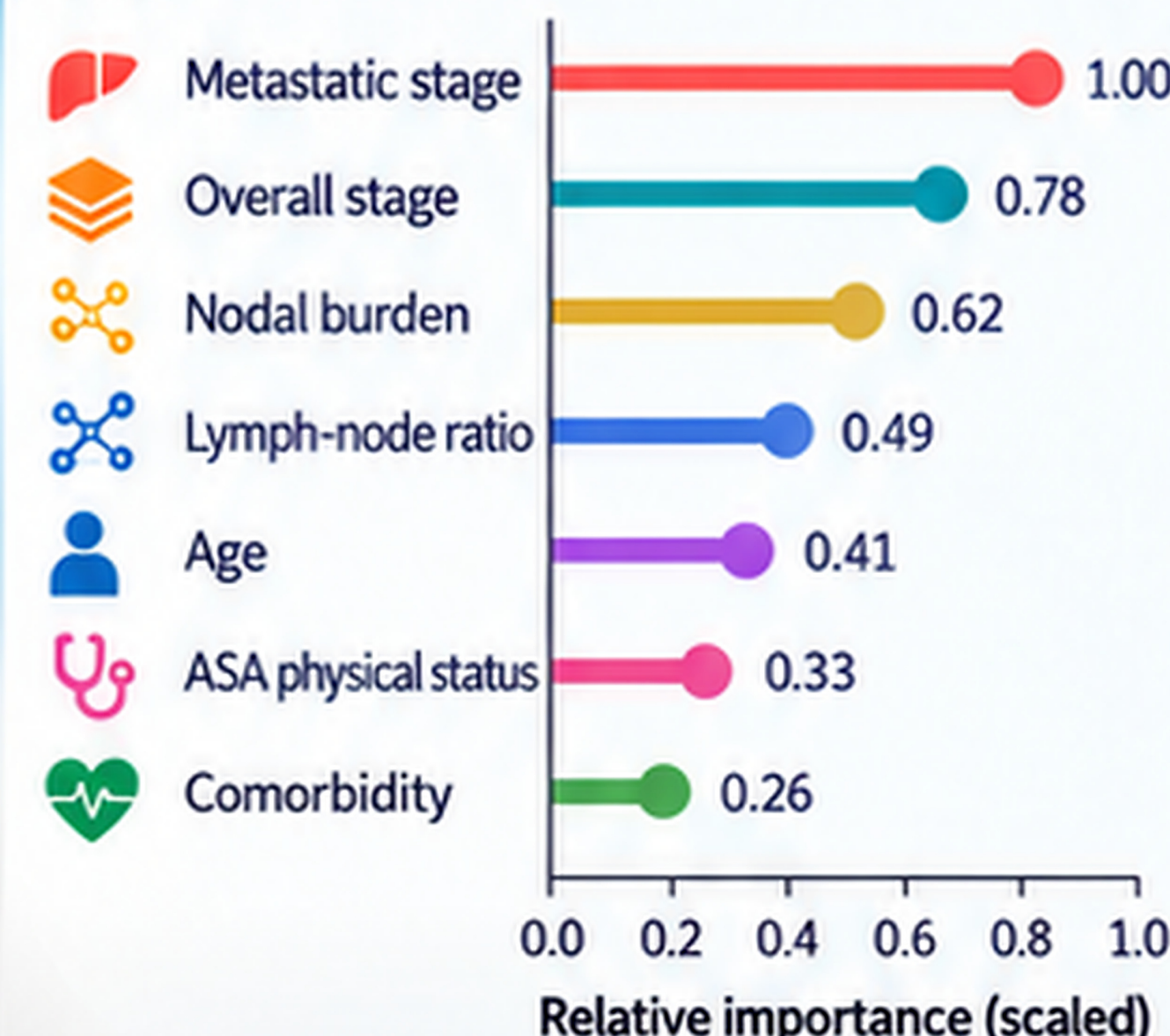


Machine learning vs conventional model

Machine learning improved prediction for some outcomes, especially CSS, but did not consistently outperform a well-specified conventional model.



Key predictive domains



Key results

- ✓ Best-performing models separated higher- and lower-risk patients well.
- ✓ Five-year discrimination was approximately 0.83-0.89 across outcomes.
- ✓ Machine learning showed outcome-specific gains, especially for CSS.
- ✓ Prognosis reflected both cancer burden and underlying patient health.



Conclusion / Take-home message

Machine-learning methods can strengthen postoperative risk assessment, but more complex models are not automatically better. Clinically useful tools must combine accuracy with reliable absolute-risk estimates. External validation is required before routine use.

References

- 1 Sung H, et al. Global Cancer Statistics 2020. CA Cancer J Clin. 2021;71;209-249.
- 2 Steyerberg EW. Clinical Prediction Models. 2nd ed. Springer; 2019.

Abbreviations

OS = overall survival
CSS = cancer-specific survival

PFS = progression-free survival
ASA = American Society of Anesthesiologists physical status