

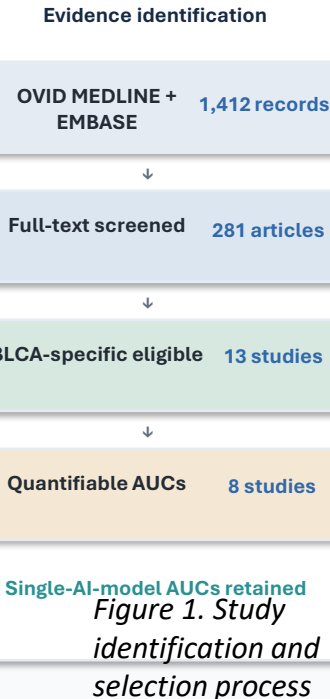
R. Das¹, B. Gonsalves¹, V. Ravindran², Y. Agarwala³, D. Khanna⁴, A. Tiwari⁵, A. Ghose⁶, G. Luigi Banna⁷, R. Dylan Frazer⁸, S. Adeleke⁹, A. Maniam⁷

(1) Faculty of Life Sciences and Medicine, King's College London, UK • (2) UCL Cancer Institute, UK • (3) Faculty of Medicine, Imperial College London, UK • (4) University of Buckingham Medical School, UK • (5) Dept. of Oncology, Princess Alexandra Hospital NHS Trust, UK
(6) Cancer Division, University College London, UK • (7) Dept. of Medical Oncology, Portsmouth Hospitals University NHS Trust, UK • (8) Velindre Cancer Centre, Cardiff, UK • (9) Curenetics, London, UK

BACKGROUND

- BLCA is common, recurrent and biologically heterogeneous.
- Immune checkpoint inhibitors (ICIs) have transformed advanced BLCA treatment.
- Predicting ICI response and overall survival (OS) remains challenging.
- AI-based models are increasingly proposed for response and prognosis.
- Aim: benchmark model performance and reporting quality across both outcomes.

METHODS



7

AUCs from 4 studies

0.752
best internal AUC

11–384
patient cohort sizes

6/7

AUCs from IMvigor210

Figure 2. Cohort size and results

AUC performance ranges

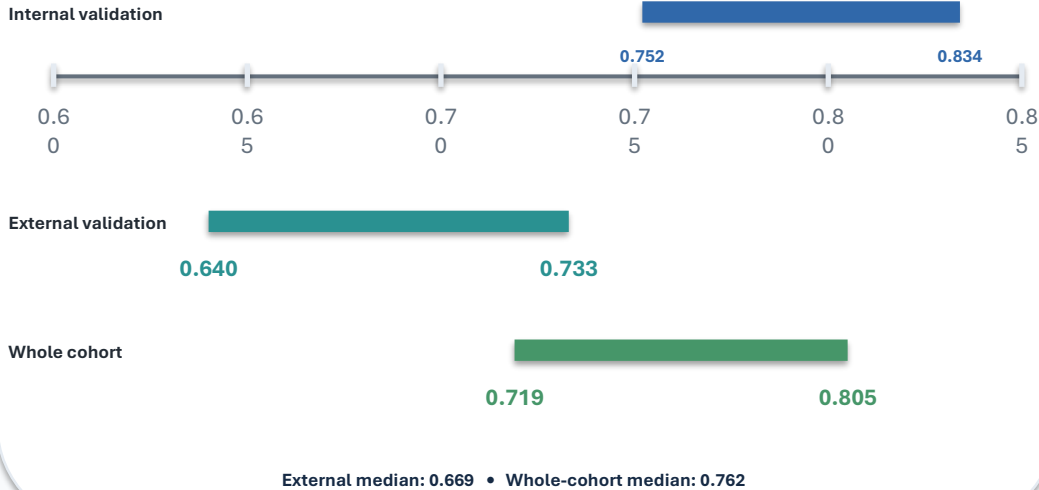


Figure 3. ICI-response prediction performance (category A)

Training median AUC

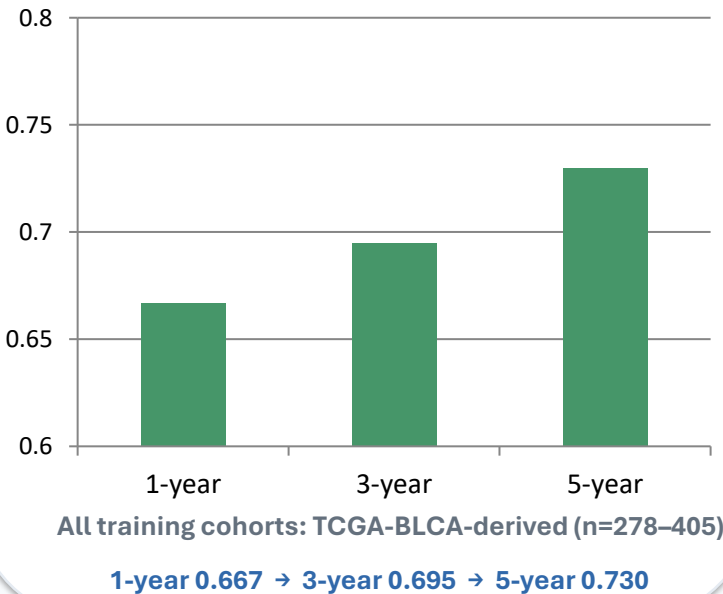


Figure 4. Overall-survival prediction performance (category B)

Validation evidence

External validation	1-year 0.670	3-year 0.710	5-year 0.726
Internal validation	AUC 0.751–0.763		
Whole cohort	AUC 0.751–0.763		

Critical limitation: external validation comes from one study (Zhang X et al., 2025), while internal/whole-cohort estimates are likewise from one study

Figure 5. Validation performance of AI models for overall-survival prognosis

RESULTS

7

ICI-response AUCs

24

OS AUC estimates

0.752

Best internal AUC

0.669

External ICI median

Key evidence patterns

- 6/7 Category A AUCs used IMvigor210.
- Category B training cohorts were all TCGA-BLCA-derived.
- OS validation was corroborated by only one study per setting.

CONCLUSIONS

AI models show moderate-to-good discrimination for ICI response and OS prognosis.

BUT

- Single-cohort dependence limits generalisability.
- Uncertainty reporting was inconsistent.
- Prospective, larger, externally validated studies are needed.