

Deep Learning for Prostate Cancer BRCA Mutation Detection and Phenotypic Characterization from Histopathology

Carlos Macarro¹, Pablo Cresta-Morgado¹, Sara Simonetti¹, Anupama Pasam², Daniel Navarro¹, Gisela Mir¹, Haitham Alatoom¹, Christina Zatse¹, Natalie Brajanovski², Francesca Zacchi³, Andrea Zivi³, Paolo Nuciforo¹, Ana Vivancos¹, David Olmos⁴, Elena Castro⁴, Shahneen Sandhu², Joaquin Mateo^{1,5}, Raquel Perez-Lopez¹

¹Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain

²Peter MacCallum Cancer Centre, The University of Melbourne, Melbourne, Australia

³Section of Innovation Biomedicine - Oncology Area, Department of Engineering for Innovation Medicine, University of Verona and University and Hospital Trust (AOUI) of Verona, Verona, Italy.

⁴12 de Octubre University Hospital, Madrid, Spain

⁵Vall d'Hebron University Hospital, Barcelona, Spain

Background

BRCA1/2 alterations are associated with poor prognosis in prostate cancer (PC) but show increased response to PARP inhibitors and platinum-based chemotherapy. Access to genomic testing remains limited in routine clinical practice, and technical challenges arise with next-generation sequencing of small tumour biopsies. Phenotypic biomarkers of *BRCA1/2* status can support the deployment of precision medicine strategies in the clinic. Recent advances in AI for digital pathology render great promise in oncology. Here, we investigate whether deep learning can assess risk of a tumor harboring *BRCA1/2* alterations in PC from routine haematoxylin and eosin whole slide images (WSIs).

Methods

Data and tumor specimens from PC patients were collected across six academic studies in Spain, Italy and Australia and TCGA-PRAD to develop and validate a DL model. *BRCA1/2* genomic status was determined by whole exome and/or targeted panel sequencing. We performed class-wise data augmentation to balance the training cohort. WSIs were tessellated into 224x224 tiles and slide level features embeddings were extracted using Prov-GigaPath. Slide-level embeddings fed a 5-fold cross-validated linear-probe model. Kaplan Meier analysis was performed to study the biomarker's prognostic value (overall survival (OS) and time to CRPC (tCRPC)). Nuclear morphology features and pathologist's annotations were extracted for *BRCA* phenotypic characterization and model's explainability.

Results

A multi-center, multi-study, cohort of 884 patients with PC was assembled for model development. The prevalence of *BRCA1/2* alterations (BRCAm) was 12.48%, increasing to 27.31% after data augmentation. The PROREPAIR-A study (n=161) was used as independent cohort for external validation.

The AI model estimated a BRCArisk score for each individual case. The model achieved an AUC of 0.70 in the external validation cohort; using the cut-offs defined in the development cohort, the model had a 95% sensitivity for identifying BRCAm tumors. Also, our AI-based biomarker was prognostic for OS and tCRPC; cases identified as BRCArisk-high by AI had lower median OS (HR: 2.14, p: 0.017) and tCRPC than BRCArisk-low cases (HR: 2.83, p: 0.022) (Fig. 1).

Compared with BRCA wild-type (BRCAwt), BRCAm tumors displayed nuclei with smaller area and higher hematoxylin intensity variability. These features may indicate a more aggressive cellular

phenotype, where smaller nuclear size could reflect faster replication and greater heterogeneity is captured by hematoxylin variability. Such nuclear phenotypes are also identified in AI-based BRCARisk groups. Importantly, BRCAwt cases with high AI BRCARisk scores featured as mBRCA tumors and BRCAm cases with low AI BRCARisk scores featured as BRCAwt tumors (Fig. 2).

Conclusions

Digital pathology can identify phenotypic biomarkers associated with *BRCA1/2* alterations in prostate cancer, supporting precision medicine approaches. These findings may guide patient management and optimize the use of NGS testing in high-risk populations. The AI model demonstrates prognostic value, and explainability analyses indicate that it captures biologically relevant nuclear alterations characteristic of *BRCA* deficiency.

Funding Acknowledgements

RPL and JM are supported by a CRIS Foundation Talent Award (TALENT19-05 and TALENT20-10), Asociación Española Contra el Cáncer (PRYCO211023SERR and LABAE20019MATE) and the FERO Foundation. EC is supported by Asociación Española Contra el Cáncer (CLSEN223433CAS). FZ is supported by an AIRC fellowship Italy Post-Doc ID 29854- 2023. Vall d'Hebron Institut d'Oncologia (VHIO) authors would like to acknowledge the Spanish State Agency for Research (Agencia Estatal de Investigación) for financial support as a Centre of Excellence Severo Ochoa (CEX2020-001024-S/AEI/ 10.13039/501100011033), the Cellex Foundation for providing research facilities and equipment, and the Generalitat de Catalunya CERCA Programme for their support of this research. We would like to acknowledge all the patients who participate in this and other clinical research projects for their contribution to science.

Conflicts of Interest

E. Castro: consultant or advisory role for AstraZeneca, Bayer, Full Life, Johnson & Johnson, Lilly, Medscape, Merck, MSD Oncology, Novartis and Pfizer; travel/accommodation support from AstraZeneca, Bayer, Janssen and Pfizer; honoraria from Astellas Pharma, AstraZeneca, Bayer, Johnson & Johnson, Medscape, Peerview, Pfizer and Telix; and research funding from Bayer, Johnson & Johnson and Pfizer. J. Mateo has served as advisor for AstraZeneca, Amunix/Sanofi, Daichii-Sankyo, Janssen, MSD; Pfizer and Roche; he is member of the scientific board for Nuage Therapeutics and is involved as investigator in several pharma-sponsored clinical trials. He is also the PI of grants funded by AstraZeneca, Amgen and Pfizer to VHIO (institution).

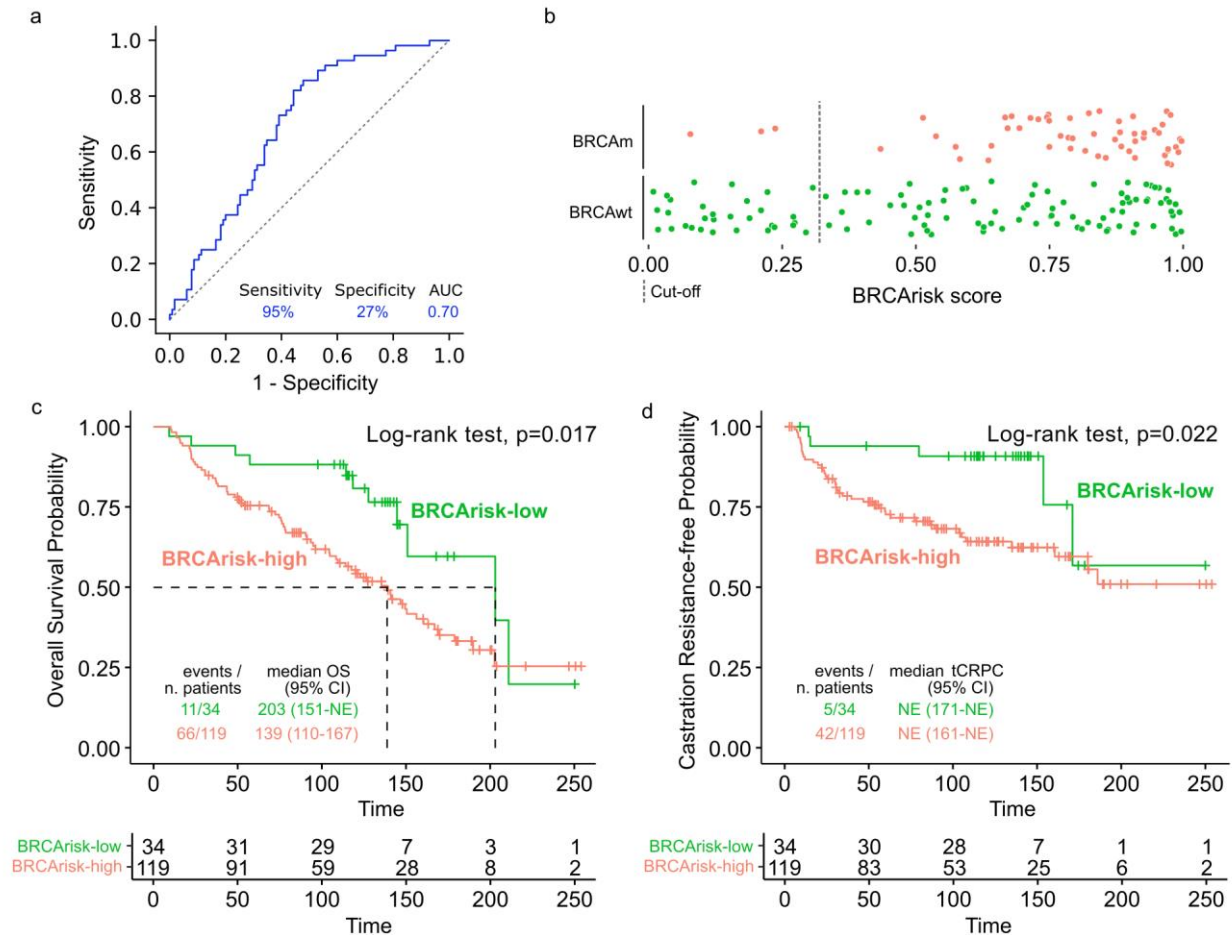


Figure 1. Deep Learning predicts PC BRCA1/2 alterations from histopathology and shows prognostic value. (a) Model's performance on the external validation cohort. (b) Distribution of predicted BRCArisk score on the external validation WSIs (c-d) Clinical validation of the AI-based BRCArisk status as prognostic biomarker for overall survival (c) and time to CRPC (d). AUC: Area Under the Receiver Operating Curve

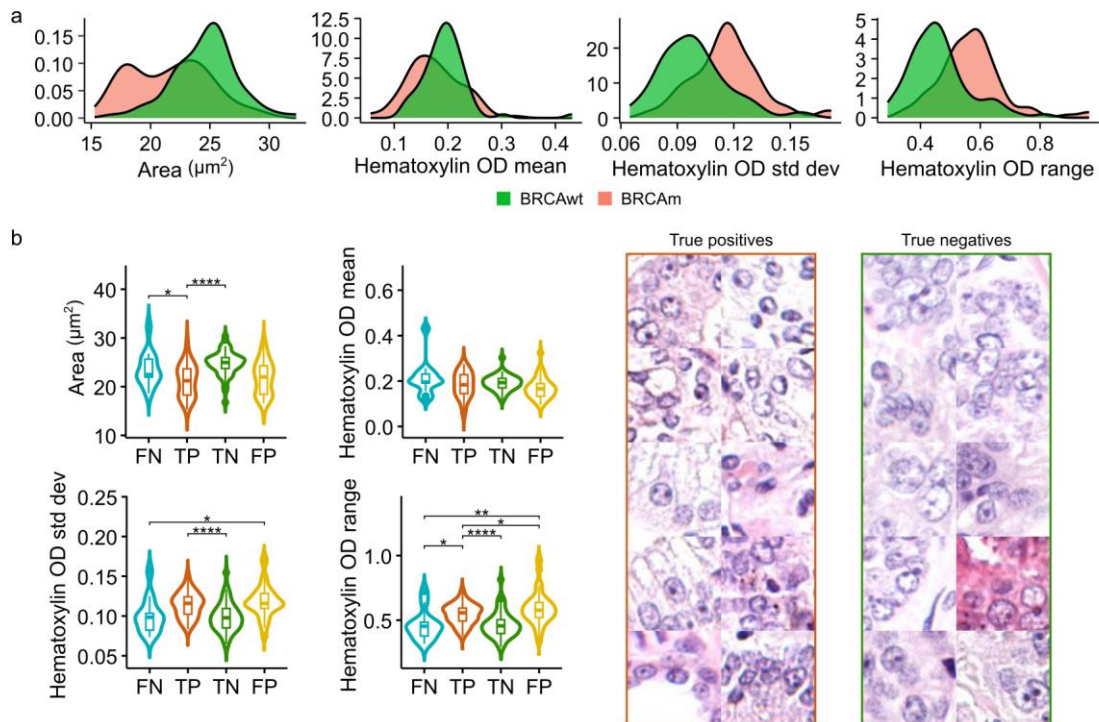


Figure 2. Nuclear morphology associates with BRCA status. (a) Distribution of relevant nuclear features in BRCAm and BRCAwT tumors, labelled by NGS. (b) Nuclear tumor features of BRCAm cases predicted BRCArisk-high (TP) or BRCArisk-low (FN) and BRCAwT cases predicted BRCArisk-low (TN) or BRCArisk-high (FP). TP: true positive; TN: true negative; FP: false positive; FN: false negative.