CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Machine learning methodologies for modelling of time-to-event endpoints in prostate cancer
Authors	O'Donnell, Autumn;Moghaddam, Shirin;Cronin, Michael;Wolsztynski, Eric
Publication date	2022-08-01
Original Citation	O'Donnell, A., Moghaddam, S., Cronin, M. and Wolsztynski, E. (2022) 'Machine learning methodologies for modelling of time-to-event endpoints in prostate cancer', ISCB 2022 - 43rd Annual Conference of the International Society for Clinical Biostatistics Newcastle, UK, 21-25 Aug.
Type of publication	Conference item
Rights	© 2022 - https://creativecommons.org/licenses/by/4.0/
Download date	2026-11-06 07:06:22
Item downloaded from	https://hdl.handle.net/10468/14071

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Machine learning methodologies for prognostic modelling of time-to-event endpoints in Prostate Cancer

Prostate cancer is among the most common cancers for men globally, accounting for 13% of cancer diagnoses in the male population each year. Radical prostatectomy (RP) fails in 20-40% of patients, who subsequently develop biochemical recurrence (BCR). This study focuses on machine learning methodologies for prediction of major prognostic endpoints, including in continuous time such as time-to-BCR prediction. An overarching aim of this work is the inclusion of genetic attributes alongside routine clinical information.

Analyses were carried out on a clinical cohort of 198 patients with BCR-free time information, 140 of which also had mRNA data. Predictive models used in the benchmarks included tree-based methods and other learning strategies such as boosting and regularization applied to Cox proportional hazards modelling. Other steps of the machine learning pipelines, including the choice of prefiltering technique and the relative contribution of candidate feature for prediction of BCR-free survival from both pre-RP and post-RP perspectives, were also assessed via comparative analyses. Potential interpretation of these features was examined in the context of such multivariate models.

Preliminary cross-validated results indicated that inclusion of mRNA information yielded increased prognostic performance with a 5% increase in C-index (95% Confidence Interval: (1%, 11%)), different performances across prognostic modelling strategies, and at various clinical time points (e.g., post-operative prediction outperformed pre-operative performance (C-index: 0.80 vs 0.73; 95% CI: (-0.14, -0.01)).

These levels of performance were also considered against routine predictive nomograms to assess their clinical potential for integration in routine practice.