Substantiating Predictive Models for Metastatic Relapse in Newly Diagnosed Upper Tract Urothelial Carcinoma Patients Utilizing a Multi-Cancer Machine Learning Approach

Brandon Ye[1], Sankalp Mrutyunjaya[1], George Courcoubetis[1], Hooman Djaladat[2,3], Alireza Ghoreifi[2,3], Jeremy Mason[1,2,3]

[1]*Convergent Science Institute in Cancer, University of Southern California, Los Angeles, CA*

[2]*Catherine and Joseph Aresty Department of Urology, Institute of Urology, Keck School of Medicine, University of Southern California, Los Angeles, CA*

[3]*Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA*

USC Michelson Center
Convergent Science Institute in Cancer

Introduction

Upper tract urothelial carcinoma (UTUC) is a rare disease of the urinary tract with few prognostic determinants. While low-grade tumors are noninvasive and unlikely to spread, high-grade UTUCs are aggressive and highly invasive, with great potential to recur elsewhere in the body. Recently, machine learning (ML) has emerged as an approach to studying disease progression. However, it is shown that the rarity of UTUC provides insufficient data for the accurate modeling of patient outcomes via ML with high precision and sensitivity, indicating a need to improve upon the clinical paradigm for UTUC treatment. Bladder cancer (BCa) and UTUC share distinct biological similarities marked by the homologous cell lining within the bladder, ureters, and urethra, and thus, could prove BCa data to be a valuable supplement for UTUC predictive models.

Motivation

To incorporate available BCa data into existing UTUC classifiers to construct a robust unified ML model that accurately predicts the recurrence outcomes of UTUC patients at timepoints of 6, 12, 24, 36, and 60 months.

Methods

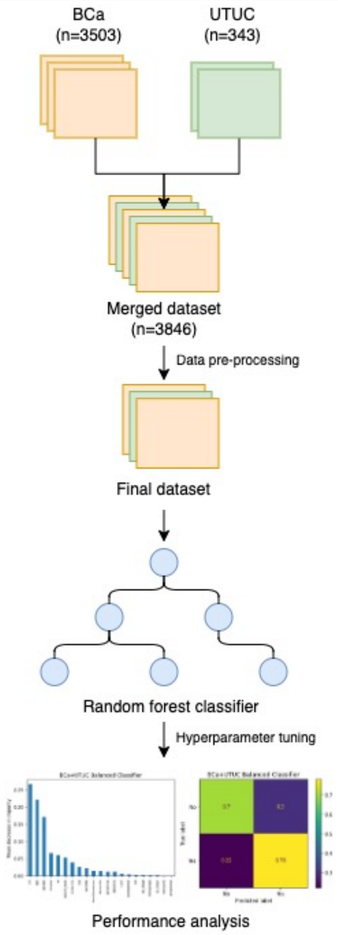

Figure 1. Methods Overview.

- Combine existing clinical and demographic datasets for BCa and UTUC based on clinically relevant shared variables.
- Clean and preprocess merged dataset to standardize variable measurement techniques.
- Implement timepoint thresholds to measure recurrence outcomes.
- Split dataset into training (80%) and withheld test (20%) sets.
- Address class imbalance between BCa and UTUC patients in dataset.
  - Balance the number of BCa and UTUC patients in training data.
  - Balance the recurrence outcome frequency within BCa and UTUC patient training data.
- Fit random forest classifier (RFC) on training data. Cross validate.
- Optimize hyperparameters using RandomSearch and GridSearch cross validation techniques.
- Fit RFC on withheld test set data.
- Compile classifier performance and feature importance metrics.
- Evaluate model utility by comparing performance of baseline UTUC classifier with performance of balanced BCa+UTUC classifier trained on the combined dataset.

Results

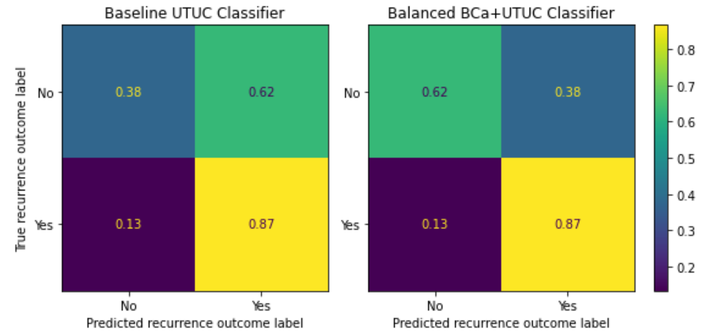

Figure 2. Normalized confusion matrices for baseline UTUC and balanced BCa+UTUC 24-month classifiers. Comparison of overall classifier accuracy on the withheld test set; proportions denote percentage of total patients per outcome category.

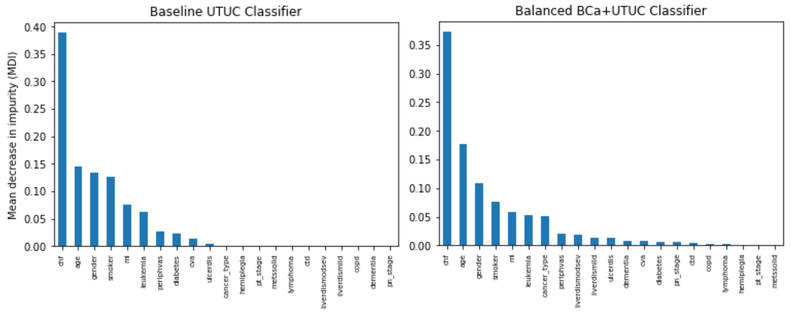

Figure 3. Feature importance charts for baseline UTUC and balanced BCa+UTUC 24-month classifiers. Dataset features are ranked based on importance, where mean decrease in impurity (MDI), or Gini importance, effectively approximates the proportion of samples reaching a specific node in the RFC decision tree.

Time Threshold	Baseline UTUC	BCa+UTUC Balanced
6 months	Accuracy: 0.794 Precision: 0.50 Sensitivity: 0.429 Specificity: 0.889	Accuracy: 0.735 (-0.059) Precision: 0.25 (-0.25) Sensitivity: 0.143 (-0.286) Specificity: 0.889 (0)
12 months	Accuracy: 0.69 Precision: 0.50 Sensitivity: 0.778 Specificity: 0.65	Accuracy: 0.724 (+0.034) Precision: 0.538 (+0.038) Sensitivity: 0.778 (0) Specificity: 0.70 (+0.05)
24 months	Accuracy: 0.652 Precision: 0.684 Sensitivity: 0.867 Specificity: 0.25	Accuracy: 0.826 (+0.174) Precision: 0.824 (+0.14) Sensitivity: 0.933 (+0.066) Specificity: 0.625 (+0.375)
36 months	Accuracy: 0.70 Precision: 0.733 Sensitivity: 0.846 Specificity: 0.429	Accuracy: 0.80 (+0.10) Precision: 0.765 (+0.032) Sensitivity: 1.0 (+0.154) Specificity: 0.429 (0)
60 months	Accuracy: 0.647 Precision: 0.688 Sensitivity: 0.917 Specificity: 0.0	Accuracy: 0.706 (+0.059) Precision: 0.733 (+0.045) Sensitivity: 0.917 (0) Specificity: 0.20 (+0.20)

Figure 4. Performance of baseline UTUC and unified BCa+UTUC classifiers across 6, 12, 24, 36, and 60-month time thresholds. Comparisons between metrics of classifier accuracy, precision, sensitivity, and specificity. Parenthetical values denote Δ between baseline and unified balanced classifier performance metrics.

Discussion

- In the 24-month recurrence classifier, we observe an overall increase in accuracy driven by improved performance in patients that did not experience recurrence of the disease within the threshold (Figure 2).
- Relative feature importance is largely conserved between baseline and balanced classifiers; patient gender and smoking status are marginally more important for the former and age for the latter (Figure 3).
- Changes in feature importance between baseline and balanced classifiers can be indicative of underlying similarities between BCa and UTUC, which can improve understanding of disease progression.
- Supplementing existing UTUC classifiers with available BCa dataset in a balanced approach confers improved performance in a majority of metrics across 12, 24, 36, and 60-month time thresholds (Figure 4).
- For 12, 36, and 60-month threshold classifiers, overall improvement in accuracy when implementing BCa data is primarily driven by increased specificity, sensitivity, and specificity, respectively (Figure 4).
- Positive results suggest that such a data supplementation approach as taken here may be applicable to other rare cancer indications. Supplementing poorly performing predictive models with surplus data from biologically similar diseases can serve to improve performance.

Future Directions

- Generate pair-wise correlations between combined dataset variables to identify those highly related between cancer types.
- Identify the most significant recurrence predictors for BCa and UTUC across timepoints to evaluate their clinical uniqueness.
- Begin to understand the biological significance of highly important contributors as identified by feature importance analyses.
- Given the utility of the liquid biopsy in potentially guiding clinical decision-making for UTUC (Shishido et al. 2022), integrate high-definition single cell assay (HDSCA) liquid biopsy analytes with the existing classifier to construct a multi-cancer liquid biopsy model.
- With noted improvement in UTUC classifier accuracy when integrating BCa data, investigate transfer learning approaches to apply pre-trained BCa models directly to UTUC data.

Acknowledgements

This project is generously supported by funding from the USC Provost Undergraduate Research Fellowship and USC Dornsife Summer Undergraduate Research Fund. Additional thanks to Dr. George Courcoubetis, Dr. Jeremy Mason, and Sankalp Mrutyunjaya for their insightful feedback, scientific guidance, and unending support.

References

Shishido, S. N., Ghoreifi, A., Sayeed, S., Courcoubetis, G., Huang, A., Ye, B., Mrutyunjaya, S., Gill, I. S., Kuhn, P., Mason, J., & Djaladat, H. (2022). Liquid Biopsy Landscape in Patients with Primary Upper Tract Urothelial Carcinoma. *Cancers*, *14*(12), 3007. https://doi.org/10.3390/cancers14123007

Contact

Brandon Ye: branye@usc.edu
USC Convergent Science Institute in Cancer: https://kuhn.usc.edu/