Development of a Machine Learning Algorithm for Hybrid Interstitial Needle Prediction in High-Dose-Rate Cervical Brachytherapy

K. Stenhouse[1,2], M. Roumeliotis[1,2,3], P. McGeachy[1,2,3]
[1] University of Calgary, Department of Physics and Astronomy
[2] Tom Baker Cancer Centre, Department of Medical Physics
[3] University of Calgary, Department of Oncology

Purpose

Develop a **machine learning model for predicting the optimal hybrid interstitial (HIS) needle configuration** for high-dose-rate (HDR) cervical brachytherapy based on high-risk clinical target volume (HR-CTV) geometry metrics and expand upon previous work developing a machine learning applicator selection model for HDR cervical brachytherapy.

Introduction

- HIS applicators used in HDR brachytherapy for cervical cancer **(Figure 1)**
- Use of interstitial needles results in more invasive procedure
  - Avoid use of excessive needles
- No explicit guidelines for optimal needle selection
  - Selection based on extent of high-risk clinical target volume (HR-CTV)
  - Largely dependent on clinician's experience

Figure 1. Vienna hybrid interstitial applicator (Elekta AB, Stockholm, Sweden)

Methods – Data Preparation

- **Dataset:** 86 historical patient treatment fractions using HIS applicators
  - Suffered from label imbalance, some needle positions used infrequently
  - To account for imbalance, model performance evaluated using micro-averaged metrics that calculate the metric globally
- **Extracted Features:** Mean and maximum HR-CTV lateral and vertical extent, volume, and offset of the HR-CTV center of mass (**Figure 2**)
  - Features divided into 33° sections centered over applicator needle channels to express directional HR-CTV geometry

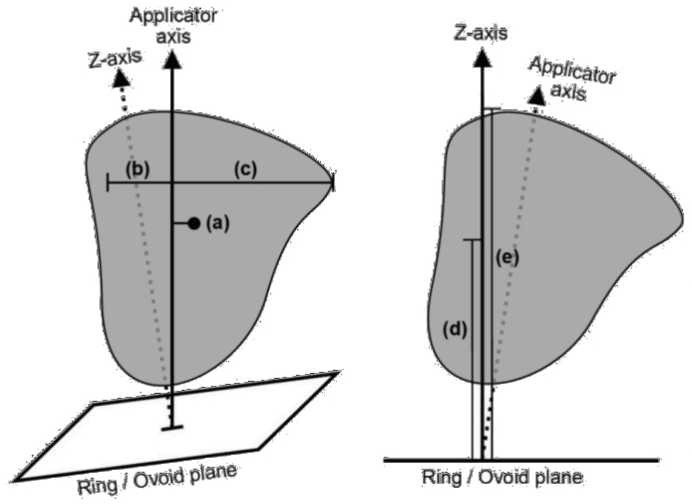

***Figure 2**. HR-CTV geometry features. Centre of mass offset (a), mean (b) and maximum (c) lateral extent are defined perpendicular to tandem axis. HR-CTV mean (d) and maximum (e) vertical extent are defined perpendicular to ring/ovoid plane.*

Methods – Machine Learning Model

- Multi-label K-nearest neighbors algorithm selected to predict the use of hybrid interstitial needles based on directional geometry
  - Multi-label classification can select the use of multiple needles simultaneously
- Model training/testing was repeated for 1,000 random selections of training/testing data (75%/25%) to evaluate performance (**Figure 3**)

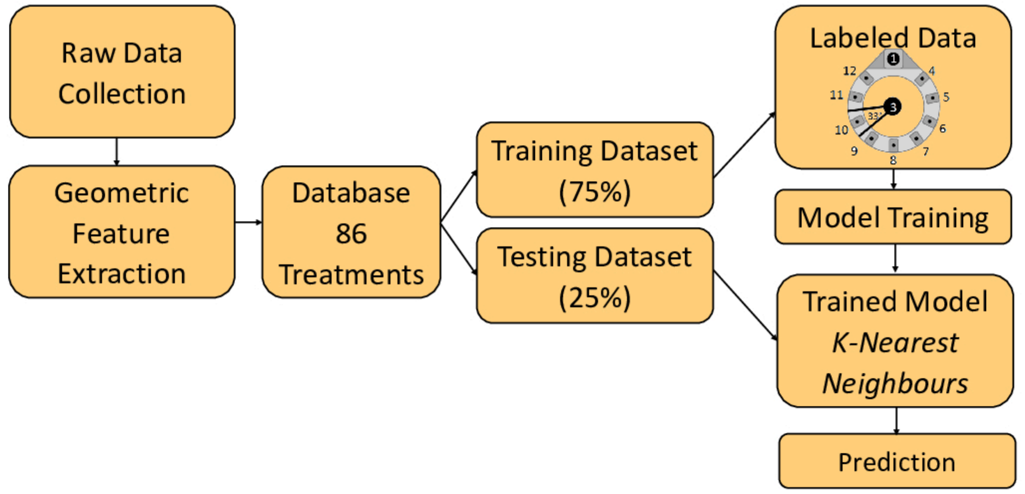

***Figure 3**. Machine learning model development workflow. The process of splitting the database into training and testing datasets to train model and evaluate performance was repeated for 1,000 iterations.*

Results

- Classification metrics evaluated over 1,000 iterations (**Table 1**), micro-averaged for data imbalance
  - F1 Score: harmonic mean of precision and recall
  - Hamming Loss: measure of disagreement between the needles predicted by machine learning and the clinical selection
- F1 Scores for each needle calculated to determine individual classification performance (**Figure 4**)
- Lower F1 scores for certain needles can be attributed to the infrequent use of these needles clinically

Metric	Definition	Micro-Average
Precision	$\frac{TP}{TP+FP} \cdot 100$	84.2% ± 4.2%
Recall	$\frac{TP}{TP+FN} \cdot 100$	83.9% ± 4.7%
F1 Score	$2 \cdot \frac{Precision \cdot Recall}{Precision+Recall} \cdot 100$	83.9% ± 3.5%
Hamming Loss	$\frac{FN+FP}{TN+TP+FN+FP} \cdot 100$	12.8% ± 2.8%

***Table 1**. Average model performance metrics. Micro-averaged metrics are computed by aggregating all samples and computing the average, as opposed to calculating the metric for each label independently, then computing the average. TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.*

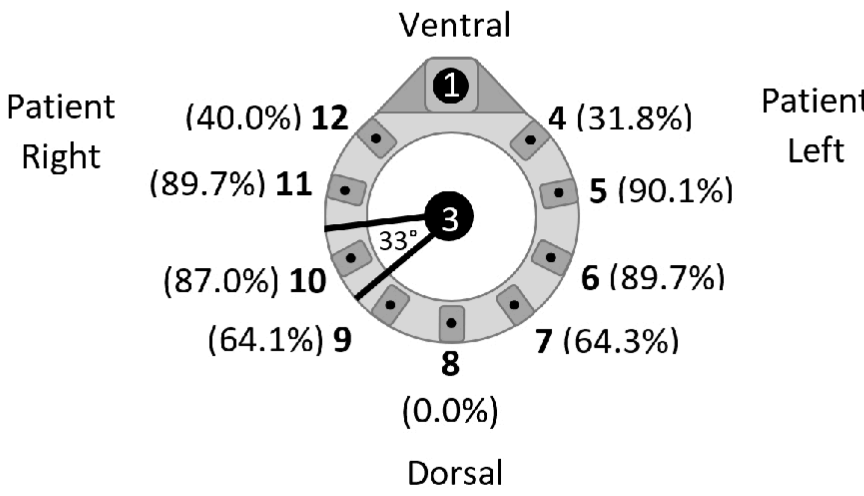

***Figure 4**. Individual needle F1 Scores (indicated in % in parentheses) for all needle positions (labelled 4-12).*

Conclusion

Model demonstrated **high predictive accuracy (83.9% F1 Score),** with most incorrect predictions coming from infrequently used needles (12.8% for Hamming Loss). This illustrates the **potential for machine learning to be a powerful predictive tool for guiding needle selection** but highlights the need for more data.

In combination with an applicator selection model, the addition of needle selection capabilities will **aid in developing a comprehensive applicator selection framework** that aims to increase uniformity in the decision-making processes involved in HDR cervical brachytherapy.

Contact Information: kjstenho@ucalgary.ca