Development of a Precision Medicine Model for Sinonasal Cancer using Artificial Intelligence

Edmund Zhi, BS[1]; Bishoy M Galoaa MS[2], Andrew G Girgis BS[1], Andrew Tsao BS[1]; Anand Devaiah, MD[3]
[1]Boston University Chobanian and Avedisian School of Medicine, [2]Northeastern University College of Engineering [3]Department of Otolaryngology, Boston Medical Center

Introduction

- Sinonasal cancers are a broad, heterogenous group of malignancies which frequently present significant challenges for clinicians to diagnose and manage due to their rarity and diversity, variability around treatment regimens, and challenges in conducting large-scale studies. Thus, we aimed to develop an AI model and explore its potentials in assisting clinicians in personalized decision-making in sinonasal cancer management.

Methods and Materials

Study Design and Data Collection

- Data was extracted from 2,854 patients diagnosed with sinonasal cancers from the Surveillance, Epidemiology and End Results (SEER) dataset, including information on:
  - Demographics
  - Clinical characteristics
  - Treatment modalities
  - Pathological features

Machine Learning Architecture

- Our predictive model utilizes a hybrid deep learning approach combining:
  1. Feature Engineering:
     - Clinical and pathological features were standardized using StandardScaler
     - Textual data from medical records were processed using BERT-based embeddings
  2. Neural Network Architecture:
     - Multi-layer perceptron with dropout layers (p=0.2) for regularization
     - Batch normalization between fully connected layers
     - ReLU activation functions in hidden layers
     - Sigmoid activation in the output layer for probability estimation
  3. Model Training:
     - Implemented in PyTorch
     - Binary cross-entropy loss function
     - Adam optimizer
     - Early stopping based on validation loss

Statistical Analysis

- Performance Metrics: AUC-ROC, sensitivity, specificity, accuracy
- 5-fold cross-validation
- DeLong's test for comparing ROC curves
- Kaplan-Meier analysis for analyzing survival outcomes

Figure 1. The steps of developing the AI model

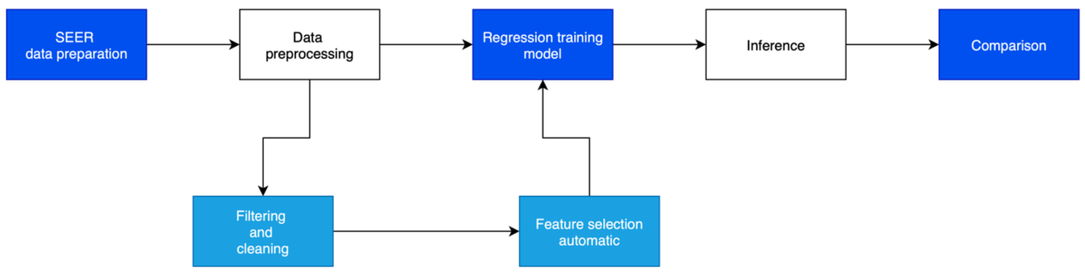


Figure 2. inference phase of inputting new patient data to provide a predictive optimum treatment

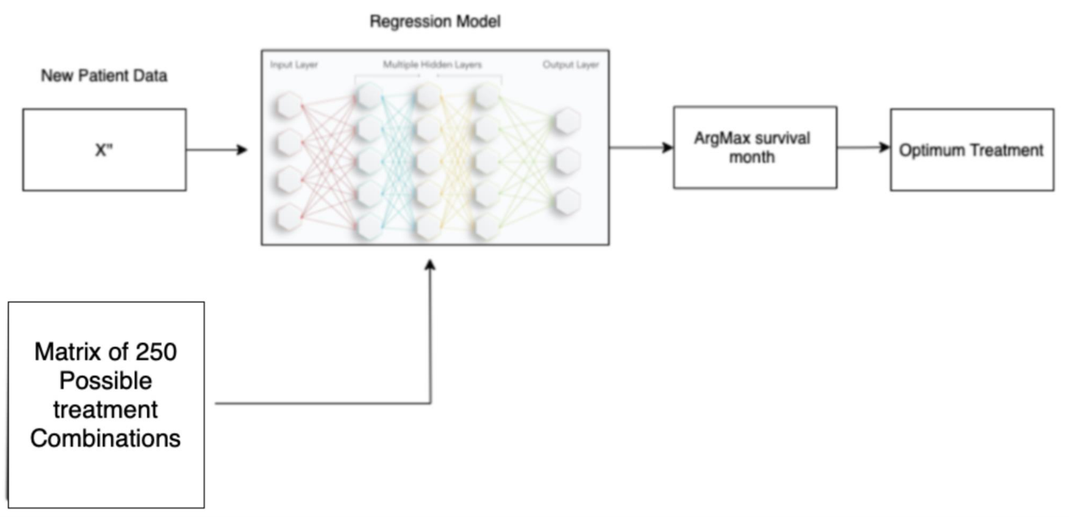
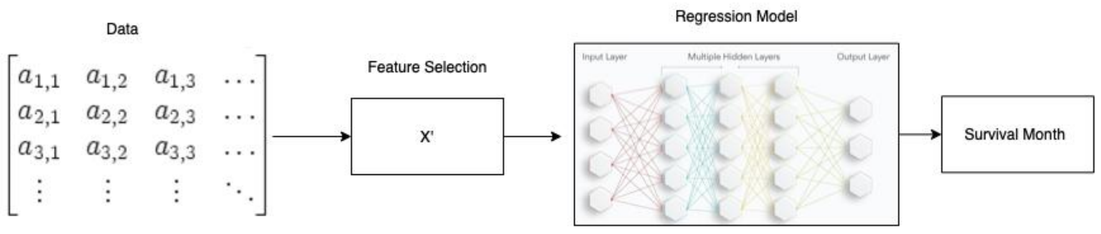


Figure 3. Regression Model Training



Results

Table 1. Cohort Characteristics

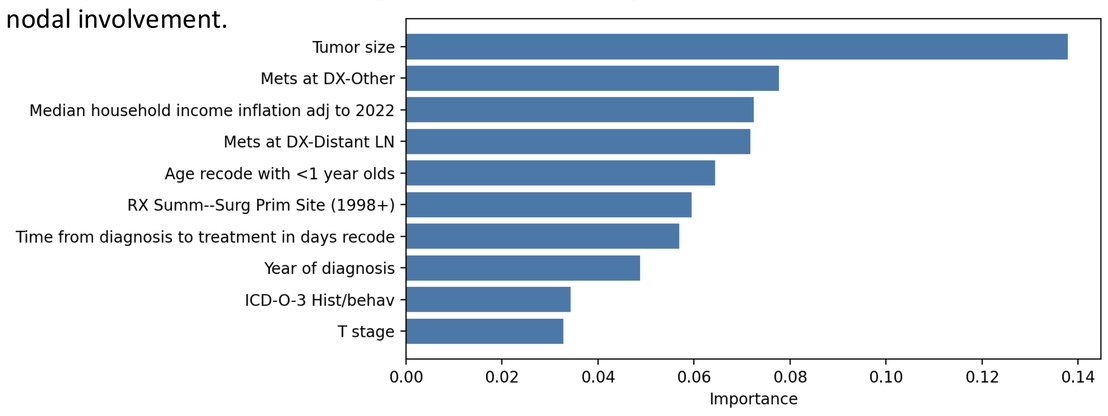
Characteristic	N = 2,854[1]
Age	
15–34	15 (0.5%)
30–49	288 (10%)
50–69	1,404 (49%)
70–84	943 (33%)
85+	204 (7.1%)
Sex	
Male	1,915 (67%)
Female	939 (33%)
Race	
White	2,053 (72%)
Hispanic/Latino	299 (10%)
Black	254 (8.9%)
Asian or Pacific Islander	218 (7.6%)
American Indian/Alaska Native	17 (0.6%)
Unknown	13 (0.5%)
Median Household Income Adjusted for Inflation to 2022	
<$40,000	67 (2.3%)
$40,000–$79,999	1,762 (62%)
$80,000–$119,999	925 (32%)
$120,000+	100 (3.5%)
Tumor Primary Site	
C30.0-Nasal cavity	1,450 (51%)
C31.0-Maxillary sinus	1,112 (39%)
C31.1-Ethmoid sinus	157 (5.5%)
C31.2-Frontal sinus	44 (1.5%)
C31.3-Sphenoid sinus	51 (1.8%)
C31.9-Accessory sinus, NOS	40 (1.4%)
Tumor Histology	
Squamous Cell Neoplasms	2,754 (96%)
Adenomas and Adenocarcinomas	34 (1.2%)
Complex Epithelial Neoplasms	30 (1.1%)
Transitional Cell Papillomas and Carcinomas	24 (0.8%)
Complex Mixed and Stromal Neoplasms	12 (0.4%)
Tumor Grade	
Grade I; Well differentiated	294 (10%)
Grade II; Moderately differentiated	805 (28%)
Grade III; Poorly differentiated	697 (24%)
Grade IV; Undifferentiated; anaplastic	45 (1.6%)
Unknown	1,013 (35%)

[1]n (%); Mean ± SD

Table 2. Model Performance with Feature Sets. The models showed good discriminative ability (AUC > 0.80), particularly at the 60-month time point (AUC > 0.92). However, the negative R^2 values and high mean absolute error (MAE) indicate that the model has room for improvement in survival month accuracy.

				Time-Dependent AUC (Months)			
	Mean Absolute Error (Months)	Root Mean Square Error (Months)	R^2 Score	12	24	60	120
Demographic Features	50.10	66.11	-0.35	0.808	0.832	0.921	0.846
No Demographic Features	50.90	66.35	-0.36	0.799	0.820	0.922	0.829

Figure 4. Feature Importance Results Including Demographic Features. The model's most important features (displayed in descending order) is generally consistent with prior literature on prognostic factors. Interestingly, tumor size outweighed importance of metastases and nodal involvement.



Discussion

- Although the model is successful in predicting a treatment regimen that optimizes the survival months, the low accuracy in predicting exact survival length suggests **room for improvement in model specification, feature engineering, or inclusion of more variables.**
- This further highlights the heterogeneity of sinonasal cancers and the current challenges in treatment; **other clinically-relevant variables may have yet to been explored and integrated into large databases**
- The minimal difference in performance between models with and without demographic features suggests that **clinical characteristics may be the primary drivers of survival outcomes in this population.**
- External validation on an independent cohort would be valuable to assess model generalizability.

Conclusions

- Our model has significant potential in assisting clinicians by identifying the most important prognostic features of sinonasal cancers and then use these to identify a precision medicine personalized regimen that provides the **best individual patient survival outcomes**.
- The model is not intended to replace clinical judgment, but to **augment clinical decision-making and patient stratification.**