Using Machine Learning to Predict Delays in Adjuvant Radiation Following Surgery for Head & Neck Cancer

KU MEDICAL CENTER
The University of Kansas

Matthew Shew, Jacob New, Andrés Bur
Department of Otolaryngology – Head and Neck Surgery
-the authors have no conflicts of interest-

INTRODUCTION

- Adjuvant radiation for head and neck cancer is well established for patients with advanced stage disease[1]
- Long intervals between surgery and adjuvant radiation therapy (>50 days) are associated with significantly worse clinical outcomes[2].
- Machine learning is built to make real-time predictions, compared to traditional statistics that provide various odds ratios that infer causation.
- We generated a predictive algorithm using the NCDB head and neck cancer data set to identify those patients that will experience delays in starting adjuvant radiation following head and neck oncologic surgery.

METHODS

- Data Source
  - The NCDB is a jointly sponsored initiative led by the American College of Surgeons and the American Cancer Society sourced from hospital registry data[3]. It captures over 70% of newly diagnosed cancers nationwide. Patients were identified from 2004 to 2013 who had head and neck squamous cell carcinoma treated with surgical resection followed by adjuvant radiation.
- Study Population
  - Inclusion required patients that underwent definitive resection by site specific procedures within the head and neck.
  - Patients who received 45 to 76 Gy and time from surgical resection to radiation was between 21 to 90 days were included.
  - Patients were then categorized as "delay in radiation" by a binary system of before or after 50 days.
- Machine Learning
  - We constructed a supervised ML classification model using opensource Azure Machine Learning Studio (Microsoft Corporation).
  - Data were randomly split in an 80/20 distribution, with 80% of data as a training set, and the remaining 20% as a test set.
  - The permutation feature importance scores were determined from the model to analyze which clinical factors were used in the model's prediction.

RESULTS

- A total of 76,573 patients were included for analysis (Table I & II).
- Patients that were treated with adjuvant therapy before 50 days had a significantly higher 5-year overall survival rate (66.3% vs 61.7%; $p < 0.001$).
- The algorithm predicted patients that started adjuvant radiation therapy 50 days or more after surgery with overall accuracy of 64.41% and precision of 58.5%.
- Using permutation feature importance scoring, we gained insight into how our ML algorithm weighs different factors in creating its algorithm (Table III).

DISCUSSION

ML enables computers to analyze and learn from large amounts of data, understand pattern recognition, and make predictions without specifically being programmed to do so. What makes ML an exciting prospect is that our ML algorithm can have real time applications. Our predictive model can identify patients that are "high risk" for delayed adjuvant therapy. Using the same variables that are already collected for the NCDB registry, the algorithm will predict who is 'high risk' for delayed follow up, with the understanding that 'high risk' is defined as 64.4% probability of showing up for adjuvant radiation after 50 days. We are currently exploring options within our clinic on how to properly intervene on these identified patients.

CONCLUSION

Using the National Cancer Database, we used over 75,000 patients to build a predictive ML algorithm to identify patients that are at risk of showing up 50 days or later for adjuvant therapy. We anticipate ML will prove to be a useful clinical tool in years to come, but its immediate role out as a reliable tool has several limitations to overcome.

Table 1. Patient Demographics

	Radiation < 50 days	Radiation ≥ 50 days
Patients	47,301	29,271
Age	60.86 (±12.53)	60.40 (±12.35)
Male	34,115 (72.1%)	20,066 (68.6%)
Urban / Rural		
≥ 1 Million	22,315 (48.5%)	14,806 (49.8%)
250K - 1 Million	10,397 (22.6%)	16,507 (22.2%)
20K - 250K	8,206 (17.9%)	12,715 (17.3%)
< 20K	5,779 (11.0%)	8,013 (10.7%)

Table 2. Patient Characteristics

	Radiation < 50 days	Radiation ≥ 50 days
Site		
Lip	343 (0.7%)	298 (1.0%)
Gum of Mouth	3,972 (8.3%)	4,059 (13.9%)
Mouth of Floor	1,624 (3.4%)	2,065 (7.1%)
Tongue	11,252 (23.8%)	5,430 (18.6%)
Tonsil	11,252 (23.8%)	4,686 (16.0%)
Oropharynx	804 (1.7%)	565 (1.9%)
Pharynx	183 (0.4%)	138 (0.5%)
Hypopharynx	834 (1.8%)	714 (2.4%)
Larynx	9,087 (19.2%)	4141 (14.1%)
Nasal	2,962 (6.3%)	2,471 (8.4%)
Nasopharynx	503 (1.1%)	263 (0.9%)
Salivary Gland	7,897 (16.7%)	4,441 (15.2%)
Stage		
I	4,392 (9.3%)	2,214 (7.6%)
II	4,108 (8.7%)	2,544 (8.7%)
III	6,445 (13.6%)	4,221 (14.4%)
IV	17,502 (68.4%)	13,219 (79.6%)
Surgical Margins		
Negative	26,083 (55.1%)	18,508 (63.2%)
Positive	13,591 (28.7%)	7,705 (26.3%)
Unknown	7,627 (16.1%)	3,058 (10.4%)

Table 3. Top NCDB Variables Used to Build Algorithm

Weighted Score	Feature Definition	Description
0.0037	PUF_FACILITY_ID	The treating facility participating in NCDB registry
0.0024	UR_CD_03	Urban vs rural patient demographic
0.0022	RX_SUMM_SCOPE_REG_LN_SUR	Number of lymph nodes removed at time of surgery or separate surgical procedure
0.0022	RX_SUMM_SURG_PRIM_SITE	Surgical procedure performed at the primary site
0.0021	DX_RX_STARTED_DAYS	Days from diagnosis to treatment
0.0020	DX_DEFSURG_STARTED_DAYS	Days from diagnosis to definitive surgery
0.0019	CS_SITESPECIFIC_FACTOR_9	Extracapsular spread of lymph node
0.0018	CS_METS_DX_LIVER	Distant metastasis to the liver at time of diagnosis
0.0016	REGIONAL_NODES_POSITIVE	Exact number of regional lymph nodes positive for SCCa examined by pathologist
0.0015	CS_SITESPECIFIC_FACTOR_11	Measured Thickness (Depth)
0.0013	DX_STAGING_PROC_DAYS	Days from diagnosis to staging procedure performed
0.0011	CS_SITESPECIFIC_FACTOR_7	Upper and Lower Cervical Node Levels
0.0011	RX_SUMM_SURGICAL_MARGINS	Final status of surgical margins
0.0010	TNM_PATH_STAGE_GROUP	Pathologic final stage based on AJCC TNM staging

REFERENCES

1. Bernier J, et al. Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. New England Journal of Medicine. 2004;350(19):1945-1952.
2. Harris JP, et al. Association of Survival With Shorter Time to Radiation Therapy After Surgery for US Patients With Head and Neck Cancer. JAMA Otolaryngology Head Neck Surg. 2018;144(4):349-359
3. Bilimoria KY, et al. The National Cancer Data Base: a powerful initiative to improve cancer care in the United States. Ann Surg Oncol. 2008;15(3):683-690.