A Data-Mining Algorithm for Large Scale Analysis of Dose-Outcome Relationships in a Database of Irradiated Head and Neck Cancer Patients

Scott Robertson, PhD
Presented at the 2014 John R. Cameron Young Investigator Symposium
Disclosure

- This work supported by:
  - Philips
  - Commonwealth Foundation
Introduction

• Sophisticated treatment planning software enables highly customized dose distributions
• But what dose do we want?
Oncospace: An analytic oncology database

- Performance Status
- Clinical assessments
- Toxicities
- Disease Status
- Lab Values
- Quality of Life
- Palliation
- Follow-up
- Consult
- Treatment
- Surgery
- Radiotherapy
- Chemotherapy
- Pathology
- Diagnosis
- Performance Status
- Comorbidities
- Patient History
- Survival
Purpose

• **Analysis**: Develop framework for dose-toxicity analysis and review

• **Data-Mining**: Search Oncospace for notable dose-toxicity relationships
Data Mining – Overview

1. General analysis parameters
2. Specific analysis parameters
3. Analysis
4. Review
Data Mining
1. General Parameters

Oncospace

SQL: Structured Query Language

Structures
... Parotids Mandible Oral mucosa Larynx ...

Outcomes
... Aspiration Dysphagia Trismus Xerostomia ...

Time Interval
Acute Intermediate Long-term All times

Severity
≥ Grade 1 ≥ Grade 2 ≥ Grade 3 ...
Data Mining
2. Analysis-Specific Parameters

- Query and process dose data

Parotids
Xerostomia
Acute
≥ Grade 2

3D Dose
Data Mining

2. Analysis-Specific Parameters

- Query and process dose data

- Parotids
- Xerostomia
- Acute
- $\geq$ Grade 2

DVH Curves

3D Structures

3D Dose
Data Mining

2. Analysis-Specific Parameters

• Query and process dose data
• Process toxicity data

Parotids
Xerostomia
Acute
≥ Grade 2
Data Mining

2. Analysis-Specific Parameters

Combined Parotids
All DVH Curves (N=361)
Data Mining

2. Analysis-Specific Parameters

Combined Parotids - Xerostomia
All DVH Curves (N=361)

Normalized Volume

Dose (cGy)
Data Mining

2. Analysis-Specific Parameters

Combined Parotids - Xerostomia
3-6 Months Post-Treatment (N=243)
Data Mining

2. Analysis-Specific Parameters

Combined Parotids - Xerostomia
3-6 Months Post-Treatment (N=243)
Data Mining

2. Analysis-Specific Parameters

Combined Parotids - Xerostomia
3-6 Months Post-Treatment (N=243)

Normalized Volume vs. Dose (cGy)
Data Mining

2. Analysis-Specific Parameters

Combined Parotids - Xerostomia
3-6 Months Post-Treatment (N=243)

<table>
<thead>
<tr>
<th>Normalized Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
</tr>
<tr>
<td>0.2</td>
</tr>
<tr>
<td>0.3</td>
</tr>
<tr>
<td>0.4</td>
</tr>
<tr>
<td>0.5</td>
</tr>
<tr>
<td>0.6</td>
</tr>
<tr>
<td>0.7</td>
</tr>
<tr>
<td>0.8</td>
</tr>
<tr>
<td>0.9</td>
</tr>
</tbody>
</table>

Dose (cGy)
Data Mining

3. Analysis

Combined Parotids - Xerostomia
3-6 Months Post-Treatment (N=243)
Data Mining

3. Analysis

Combined Parotids - Xerostomia
3-6 Months Post-Treatment (N=243)

Probability

Normalized Volume

Dose (cGy)
Data Mining

3. Analysis

Combined Parotids - Xerostomia
3-6 Months Post-Treatment (N=243)
Data Mining

3. Analysis

Combined Parotids - Xerostomia
3-6 Months Post-Treatment (N=243)
Data Mining

3. Analysis

Combined Parotids - Xerostomia
3-6 Months Post-Treatment (N=243)
Data Mining
3. Analysis

- Univariate logistic regression:

\[
\frac{\Delta(\text{log odds})}{\Delta(\text{Odds})} \text{ per unit dose}
\]

- Reflects curve steepness
- \(\Delta(\text{log odds})\) per unit dose

- Higher “( )” reflects stronger influence of dose on toxicity
At 99% Volume, $1.224 = 22.4\%$ change in odds per Gy
At 5% Volume, $1.121 = 12.1\%$ change in odds per Gy
<table>
<thead>
<tr>
<th>RISK STRUCTURES</th>
<th>TOXICITIES / OUTCOMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>brachial-plexus-l</td>
<td>brain</td>
</tr>
<tr>
<td>brachial-plexus-r</td>
<td>brainstem</td>
</tr>
<tr>
<td>combo-brachial-plexus</td>
<td>chiasm</td>
</tr>
<tr>
<td>combo-const-rhucle</td>
<td>combo-ear</td>
</tr>
<tr>
<td>combo-ear-inner</td>
<td>combo-ear-inner</td>
</tr>
<tr>
<td>combo-masticater</td>
<td>combo-masticater</td>
</tr>
<tr>
<td>combo-c</td>
<td>combo-c</td>
</tr>
<tr>
<td>combo-sternocleidomastoid</td>
<td>combo-subtract</td>
</tr>
<tr>
<td>constr-muscle</td>
<td>constr-muscle</td>
</tr>
<tr>
<td>constr-muscle</td>
<td>cord-cricopharynge</td>
</tr>
<tr>
<td>larynx</td>
<td>ear</td>
</tr>
<tr>
<td>masticator</td>
<td>ear</td>
</tr>
<tr>
<td>masticator</td>
<td>fop</td>
</tr>
<tr>
<td>masticator</td>
<td>op</td>
</tr>
<tr>
<td>masticator</td>
<td>oral</td>
</tr>
<tr>
<td>masticator</td>
<td>parotid</td>
</tr>
<tr>
<td>masticator</td>
<td>pituitary</td>
</tr>
<tr>
<td>masticator</td>
<td>skin</td>
</tr>
<tr>
<td>masticator</td>
<td>sternocleidomastoid-muscle-l</td>
</tr>
<tr>
<td>masticator</td>
<td>sternocleidomastoid-muscle-r</td>
</tr>
<tr>
<td>masticator</td>
<td>submandibular-l</td>
</tr>
<tr>
<td>masticator</td>
<td>submandibular-r</td>
</tr>
<tr>
<td>masticator</td>
<td>thyroid</td>
</tr>
<tr>
<td>masticator</td>
<td>tmjoint-l</td>
</tr>
</tbody>
</table>

Color scale from 1.01 to 1.1 for toxicity levels.
4. Review

Data Mining

- Default DVH plot
- Update DVH Plot
- Update DVH List
- Update Outcome List

1. Logistic Regression of Dose Profiles

**combo parotid - xerostomia (N=287)**

Normalization of Volume vs. Dose:
- < G2 (N=141)
- ≥ G2 (N=146)

Probability of ≥ Grade 2 xerostomia

- Normalization of Volume vs. Dose
- Color scale ranging from 0 to 1
## Data Mining
### 4. Review

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Grade</th>
<th>Organ at Risk</th>
<th>Time Interval</th>
<th>Total Patients</th>
<th>Patients with Outcome, N (%)</th>
<th>Normalized Volume Threshold</th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xerostomia</td>
<td>≥2</td>
<td><strong>Parotid glands (L+R)</strong></td>
<td>On treatment</td>
<td>406</td>
<td>270 (67%)</td>
<td>1.00</td>
<td>1.173</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Acute</td>
<td>328</td>
<td>168 (51%)</td>
<td>1.00</td>
<td>1.166</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chronic</td>
<td>300</td>
<td>128 (43%)</td>
<td>1.00</td>
<td>1.153</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mucositis</td>
<td>≥3</td>
<td><strong>Oral mucosa</strong></td>
<td>On treatment</td>
<td>295</td>
<td>159 (54%)</td>
<td>0.99</td>
<td>1.051</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Acute</td>
<td>229</td>
<td>16 (7%)</td>
<td>0.99</td>
<td>1.039</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chronic</td>
<td>215</td>
<td>4 (2%)</td>
<td>0.99</td>
<td>1.022</td>
<td>0.657</td>
</tr>
</tbody>
</table>
## Data Mining

### 4. Review

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Grade</th>
<th>Organ at Risk</th>
<th>Time Interval</th>
<th>Total Patients</th>
<th>Patients with Outcome, (N) (%)</th>
<th>Normalized Volume Threshold</th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dysphagia</td>
<td>≥2</td>
<td>Larynx</td>
<td>On treatment</td>
<td>94</td>
<td>26 (28%)</td>
<td>0.02</td>
<td>1.091</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>≥2</td>
<td></td>
<td>Acute</td>
<td>134</td>
<td>31 (23%)</td>
<td>0.03</td>
<td>1.160</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>≥1</td>
<td></td>
<td>Chronic</td>
<td>133</td>
<td>68 (51%)</td>
<td>0.04</td>
<td>1.061</td>
<td>0.014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Grade</th>
<th>Organ at Risk</th>
<th>Time Interval</th>
<th>Total Patients</th>
<th>Patients with Outcome, (N) (%)</th>
<th>Normalized Volume Threshold</th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dysphagia</td>
<td>≥2</td>
<td>Pharyngeal Constrictors</td>
<td>On treatment</td>
<td>165</td>
<td>50 (30%)</td>
<td>0.94</td>
<td>1.031</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>≥2</td>
<td></td>
<td>Acute</td>
<td>186</td>
<td>35 (19%)</td>
<td>0.03</td>
<td>1.105</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>≥1</td>
<td></td>
<td>Chronic</td>
<td>176</td>
<td>77 (44%)</td>
<td>0.00</td>
<td>1.105</td>
<td>0.003</td>
</tr>
</tbody>
</table>
# Data Mining

## 4. Review

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Grade</th>
<th>Organ at Risk</th>
<th>Time Interval</th>
<th>Total Patients</th>
<th>Patients with Outcome, N (%)</th>
<th>Normalized Volume Threshold</th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dehydration</td>
<td>≥2</td>
<td></td>
<td>On treatment</td>
<td>409</td>
<td>71 (17%)</td>
<td>1.00</td>
<td>1.075</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Dysgeusia</td>
<td>≥2</td>
<td></td>
<td>On treatment</td>
<td>428</td>
<td>338 (79%)</td>
<td>1.00</td>
<td>1.137</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mucositis</td>
<td>≥3</td>
<td>Mandible</td>
<td>On treatment</td>
<td>431</td>
<td>242 (56%)</td>
<td>1.00</td>
<td>1.074</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Nausea</td>
<td>≥2</td>
<td></td>
<td>On treatment</td>
<td>432</td>
<td>167 (39%)</td>
<td>1.00</td>
<td>1.095</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Xerostomia</td>
<td>≥2</td>
<td></td>
<td>On treatment</td>
<td>431</td>
<td>273 (63%)</td>
<td>1.00</td>
<td>1.120</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Bad DVH!

- DVH assumes that every sub-region of an OAR has the same radiosensitivity and functional importance to the related toxicity.
- DVH assumes that each OAR is uniquely responsible for the overall human function related to the toxicity.
3D Dose-Toxicity Modeling

The Parotid Gland and Facial Nerve
- Deep Lobe
- Superficial Lobe

Parotid Duct
Parotid Gland
Sublingual Gland
Submandibular Gland
3D Dose-Toxicity Modeling

Oncospace

3D Structures + 3D Dose

Machine Learning
Artificial Neural Networks
Random Forests
Clustering Algorithms
Summary

• Oncospace
  – An informatics platform for next-generation analytics and decision support
  – A goldmine for analysis of normal tissue complications

• Data-Mining
  – Framework supports 2D and 3D dose-toxicity analysis
  – Exploratory studies validate existing knowledge and generate new hypotheses
Thank You!

- Dr. Todd McNutt
- Dr. Harry Quon
- Dr. Ana Kiess
- Dr. Joseph A. Moore
- Dr. Wuyang Yang
- Dr. Zhi Cheng
- Dr. Andrew Sharabi