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
AAPM Young Investigator Symposium 2014
Disclosure

• This work supported by:
  – Philips
  – Commonwealth Foundation
Introduction

- High accuracy in modern radiation treatment planning and delivery
- Large variability in normal tissue response
Oncospace:
An analytic oncology database
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

4. Review
Data Mining
1. General Parameters

**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

SQL: Structured Query Language
Data Mining

2. Analysis-Specific Parameters

- Query and process dose data
- Parotids
- Xerostomia
- ≥ Grade 2 Acute

Combined Parotids
All DVH Curves (N=361)

Normalized Volume

Dose (cGy)
Data Mining

2. Analysis-Specific Parameters

- Query and process dose data

Parotids
Xerostomia
Acute
≥ 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

Xerostomia Score

Days from Start of Treatment

459 Patients
4805 Assessments
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)
Data Mining

2. Analysis-Specific Parameters

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

2. Analysis-Specific Parameters
Data Mining

2. Analysis-Specific Parameters

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

- < G2 (N=116)
- ≥ G2 (N=127)
Data Mining

2. Analysis-Specific Parameters

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

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

3. Analysis

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

- Univariate logistic regression:
  \[ f(x) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 x)]} \]

- \( \beta_1 \)
  - Reflects curve steepness
  - \( \Delta \) (log odds) per unit dose

- \( \exp(\beta_1) \)
  - \( \Delta \) (Odds) per unit dose
  - Higher “\( \exp(\beta_1) \)” reflects stronger influence of dose on toxicity
At 99% Volume, $1.224 = 22.4\%$ change in odds per Gy
Data Mining

4. Review

At 5% Volume, 1.121 = 12.1% change in odds per Gy
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
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