LEARNING HEALTH SYSTEMS FOR RADIATION ONCOLOGY: NEEDS AND CHALLENGES FOR FUTURE SUCCESS

Todd McNutt PhD
Associate Professor
Radiation Oncology
Johns Hopkins University

Disclosures

This work has been partially funded with collaborations from:

- Philips Radiation Oncology Systems
- Elekta Oncology Systems
- Toshiba Medical Systems
- as well as
- Commonwealth Foundation
- Maritz Foundation

Which patient will do better?

69-year-old man with Stage Squamous cell carcinoma, NOS of the Right Malignant neoplasm of larynx

65-year-old man with T3 N2b M0 Stage IVa Squamous cell carcinoma, NOS of the Malignant neoplasm of tonsil
What we need to get there?

• A means of quantifying the patient experience
• A system to capture that knowledge in routine clinical care
• Validated data science models to predict outcomes for the individual patients
• Incorporate models into treatment plan generation and clinical decisions

Types of Clinical Data

• Clinician Assessments
• Patient Reported
  – Quality of life
  – Toxicity and complications
• Biospecimen
  – Labs
  – Pathology
• Image derived features (Radiomics)
• Treatment
  – Radiation Dosimetry
  – Surgery
  – Chemotherapy
• Symptom management
  – Nutritional support
  – Pain medications

Learning health system

Facts

Knowledge Database

Predictive Modeling

Presentation of Predictions

Data Feedback (Facts, Outcomes)

Decisions

Predictive Modeling

Presentation of Predictions

Facts

Controls

Outcomes

Predicted Outcomes

Decision Point

Time

Facts

Controls

Outcomes

Predicted Outcomes

Decision Point

Time
Oncospace tables and schema

- Patient
- Private Health Info (access restricted)
- Family History
- Social History
- Medical History
- Medications (scheme)
- Surgical Procedures
- Test Results (labs)
- Assessments (Radiology)
- Clinical Events
- Radiation Summary
- Oncospace Consortium Repository

Oncospace Consortium Repository
(It’s all about the data)

- Johns Hopkins
- U. Washington
- U. Toronto
- Sunnybrook
- U. Virginia
- Johns Hopkins SOM

Consortium Status
Michael Bowers MS

University of Washington
University of Virginia
University of Toronto
University of Virginia
University of Toronto
Johns Hopkins SOM

Combined Analysis

Michael Bowers MS
Viability and Value

- Predictive factors must be accessible for new patients
- Prediction must be clinically valuable and extend the knowledge of the clinician
- Predictive models must be consistent with existing knowledge

Precision Radiotherapy Treatment

OVH: serial vs parallel

For parallel organs, OAR2 is more easily spared.
For serial organs, OAR1 is more easily spared.
Mandible vs PTV_7000
pt: 300

8/3/2016

Mandible vs PTV_7000
pt: 822

8/3/2016

Mandible vs PTV_7000
pt: 295

8/3/2016
8/3/2016

Mandible vs PTV_7000
pt: 258

Mandible vs PTV_7000
pt: 234

Shape-dose relationship for radiation plan quality

For a selected Organ at Risk and %V, find the lowest dose achieved from all patients whose %V is closer to the selected target volume?
Currently, shape (knowledge) based auto-planning…

- has demonstrated improved quality
- removed human variability for standard cases
- can learn as we improve our techniques and change our practices.
- is now advancing commercially

Promote Culture of Data Collection

Data collected over entire treatment

At what time point do we have enough data to make decision based on future prediction?

Input Variables => Prediction?

MOSAIQ for Clinical Assessment
Data Collection in Clinic

Clinical Assessment
- FACT HN
- SSQ
- SHIM
- IPSS
- PAN26

Quality of life

Disease Status

Extract, Transform, Load
- SQL Query
- Lab, Toxicity, Assessments
- Scripts, Python, DICOM
- DVH, OVH, Shapes

Head and Neck Inventory
- ~1000pts up to 6 yr follow up
Head and Neck Inventory

Organs at risk with full 3D dosimetry

Prostate Inventory

~1800 pts - ~700 with dose
Toxicity Prevalence

(P. Lakshminarayanan)

8/3/2016

Dysphagia < 1
Xerostomia < 2
Weight Loss < 1
Taste (Dysgeusia) < 1

DVH, Toxicities and Grade distributions

Number of patients by grade at D 50%

Toxicity Grade 0, 1, 2, 3, 4, 5

Mean and stddev of D X% at grade

Trismus
Mandible 20% Volume
Dysphagia
Superior Constrictor 50% Volume

Voice Change
Larynx 30% Volume
Dysphagia
Larynx Edema 30% Volume

DVH, Toxicities and Grade distributions

Number of patients by grade at D 20%

Toxicity Grade 0, 1, 2, 3, 4, 5

Mean and stddev of D X% at grade
Toxicity and Dose Volume Histogram
(Scott Robertson et al.)

Spatially dependent features of dose in the structures (F. Marungo et al.)

Method | Voice dysfunction \( \text{AUC}_{0.915} = 0.743 \) | Xerostomia \( \text{AUC}_{0.900} = 0.734 \)
--- | --- | ---
Bagged Naïve Bayes (1000 iterations) | 0.915 | 0.743
Bagged Linear Regression (1000 iterations) | 0.905 | 0.737
Naïve Bayes | 0.900 | 0.734
Linear Regression | 0.896 | 0.731
Random Forest (1000 trees) | 0.724 | 0.683

NTCP

• Predictors:
  - (1: Diagnosis) ICD-9 code
  - (2: Dosimetry) dose to swallowing muscles, larynx, parotid
  - (3: Patient) age

• Prediction result: High negative predictive value
  - The model can screen out patients without weight loss
  - Physicians can focus on patients with high probability of weight loss

Results: Weight loss prediction at planning
(Scott Robertson et al.)

Endpoint: > 5kg loss at 3 months post RT

Sierra Zhi Cheng MD MS
Minoru Nakatsagawa PhD

Prediction result

<table>
<thead>
<tr>
<th></th>
<th>AUC</th>
<th>Sensitivity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.773</td>
<td>0.765</td>
<td>0.426</td>
<td>0.805</td>
</tr>
</tbody>
</table>

End of document.
**Results: Weight loss prediction during RT**

- **Predictors:**
  - (1) QOL: patient reported oral intake
  - (2) Diagnosis and staging: ICD-9, N stage
  - (3) Dosimetry: dose to larynx, parotid
  - (4) Toxicity: skin toxicity, nausea, pain
  - (5) Geometry: minimum distance between PTV, larynx

**Prediction result**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.821</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.977</td>
</tr>
<tr>
<td>PPV</td>
<td>0.462</td>
</tr>
<tr>
<td>NPV</td>
<td>0.986</td>
</tr>
</tbody>
</table>

**Endpoint:** > 5kg loss at 3 months post RT

---

**Pancreas Resectability**

[Diagram and data table as shown in the image]

---

**Xerostomia Prediction**

(3-6 Months post RT)

[Diagram and data table as shown in the image]
Xerostomia prevalence separated by age = 51

Improving Care: Predicting radiation toxicities (Robertson et al.)

Grades 0-1 xerostomia
Grades 2-3 xerostomia
QUANTEC
Salivary (Deasy et al.)

“To best define xerostomia, we recommend that an observer-based system (e.g., the Common Terminology Criteria for Adverse Events) be supplemented by a validated QOL measurement device (e.g., the XQ (xerostomia questionnaire) [7]) and/or salivary measurements (e.g., whole mouth-stimulated measurements).”

We concur! And will add that CTCAE may not have the necessary resolution at all.

Can’t measure – Can’t predict

- Can we find viable methods to refine our clinician assessed outcomes in the clinical setting?
- What is resolution of the data?
- Patient reported outcomes can validate clinician assessments at somewhat low cost (SSQ etc...)
- Direct measurements tend to be more costly.
- Can natural language processing of our current documentation achieve the depth and granularity necessary?
- Must our culture change to more quantitative documentation of the patient condition?
Needs…

• For the vision of a learning health system, significantly improved user interfaces are required
• In order to present a prediction, we must first present the “quantitative” patient state
• More continuous assessment of patient condition is needed through mobile devices
• Stronger linkages between genomic, pathology and clinical databases

Summary

• We can quantify the patient experience and are improving our capabilities rapidly
• It is possible to collect and house RT data/knowledge in a clinical setting
• Current shape-based auto-planning utilizes a learning health system
• Data science models are maturing that can convert the knowledge to clinical predictions
• Sharing data across institutions allows for experience and expertise sharing
...we have work to do which requires real partnerships between clinicians and informaticists

Acknowledgments

JHU-RO
- Sierra Cheng MD
- Michael Bowers MD
- Joseph Moore PhD
- Scott Robertson PhD
- Praneet Ladehirmayaram
- Jian Hu MD
- John Wing PhD
- Theodore DeWeese MD
- GI Team
- Joseph Herman MD
- Amy Hocker-Previte PA
- HN Team
- Harry Quinn MD
- Ana Koets MD
- Tommie Geer/Mark
- William Song PhD
- Patrick Keck

JHU-CC
- Rana Taylor PhD
- Mohit Khadakar PhD
- Fumiyasu Murongo BS

Philips PROS
- Karl Bozok BSS

Toshiba
- Minoru Nakatsugawa PhD
- Bobby Davey PhD
- Rachel Louise Kohless
- John Haller

Elekta
- Bob Hubbell

University of Washington
- Kim Evans MD
- Mark Philips PhD
- Krist Hendrickson PhD

JHU-RO
- Sierra Cheng MD
- Michael Bowers MD
- Joseph Moore PhD
- Scott Robertson PhD
- Praneet Ladehirmayaram
- Jian Hu MD
- John Wing PhD
- Theodore DeWeese MD
- GI Team
- Joseph Herman MD
- Amy Hocker-Previte PA
- HN Team
- Harry Quinn MD
- Ana Koets MD
- Tommie Geer/Mark
- William Song PhD
- Patrick Keck

JHU-CC
- Rana Taylor PhD
- Mohit Khadakar PhD
- Fumiyasu Murongo BS

Philips PROS
- Karl Bozok BSS

Toshiba
- Minoru Nakatsugawa PhD
- Bobby Davey PhD
- Rachel Louise Kohless
- John Haller

Elekta
- Bob Hubbell

University of Washington
- Kim Evans MD
- Mark Philips PhD
- Krist Hendrickson PhD
Consent/Ethics

- It is our duty to learn from every patient we treat (experience-wise or electronically)
- Quantifying patient experience provides easier recall and enhances and enables sharing of that experience
- If we are capturing the data on every patient the same way, then isn’t it the standard of care for that service?
- Are we doing research or quality management?
- When does it become research?
  - Intent to publish?
  - When a group of patients is separated from standard of care?

Acknowledgements

- Harry Quon MD
- Joseph Herman MD
- Scott Robertson PhD
- Joseph Moore PhD
- Sierra Cheng MD
- Ana Keiss MD
- Minoru Nakatsagawa PhD
- John Wong PhD
- Theodore DeWeese MD

Are current radiobiology models good enough?

Current NTCP models are too simplistic, and based on a small amount of trial data.

…we treat patients every day with radiation, we just fail to capture the impact on all of them…

~60K HN cancer per year in US