Learning Objective

• Discuss technical and process issues related to pooling data among institutions in the context of collaborative studies among the presenting institutions.

• In particular, I will be discussing an on-going data pooling project between my institution and the Mayo Clinic.
  – Focusing on rationale for data pooling, challenges in data collection, understanding results, and long-term goals.

Interest in RT-related cardiac effects grew almost exponentially after a publication in NEJM by Darby et al. in 2013.
Linear Dose Response Model

- Rate of major coronary events according to mean radiation dose to the heart compared with the rate with no radiation exposure to the heart.

But how were mean heart doses determined?

Based on patients’ individual RT records, but were retrospectively reconstructed treatment plans using CT scan of a “representative” patient.

Many More Unanswered Questions

- Only mean heart dose was considered.
- Considered historic type BRT and reported a mean heart dose for the population ≈ 6 Gy
- Risk associated with partial volume effects remains unknown…..
- Are those data relevant in the context of modern BRT?

Motivation for Data Pooling in BRT

- Currently no large database exists that reports standard dose metrics of modern BRT for the heart and lungs.
- Normal tissue complications after BRT have been widely published, including cardiac and pulmonary toxicities.
  - However, these are based on early treatment trials or population studies and with no normal tissue doses or dose estimations only.
Objectives

- Primary objectives of our study were
  1. To establish a bi-institutional database in order to ascertain baseline values for typical heart and lung doses for modern BRT.
  2. To evaluate the effect of various treatment techniques on those doses.

Basic Requirements for Study

- Consensus on what data to collect
- How many patients to include from each institution
- DVH data extraction from TPS

How many patients?

- How many could we reasonably accrue in about 6 months?
  - Institutions 1 and 2 treat approximately 1000 and 350 breast patients with EBRT each year, respectively.
- Agreed to collect data for 100 right and 100 left BRT patients from each institution.
- Data were collected using consecutive sampling and included patients who received definitive EBRT for a primary breast cancer and who completed RT.
How did we decide what data to collect?

- We wanted to of course collect DVH metrics that we thought may be related to normal tissue toxicities, but also….

Needed to collect sufficient additional details about the patients, the prescriptions, and treatment beams so that we could interpret any differences that we might observe in data from collected at different institutions.

Data Collected for Each Patient

<table>
<thead>
<tr>
<th>Treatment Technique</th>
<th>Demographics</th>
<th>DVH Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right, Left, Bilateral</td>
<td>- DOB</td>
<td>Heart</td>
</tr>
<tr>
<td>For Left: FB or DIBH</td>
<td>- Gender</td>
<td>Max[Gy]</td>
</tr>
<tr>
<td>Breast or CW</td>
<td>- Age at RT</td>
<td>Max[Gy]</td>
</tr>
<tr>
<td>Prescription Dose(s)</td>
<td>- Race/Ethnicity</td>
<td>Min[Gy]</td>
</tr>
<tr>
<td>Field type: T, SC, IMC (upper/lower/support)</td>
<td>- BMI</td>
<td>Mean[Gy]</td>
</tr>
<tr>
<td>Field Energies</td>
<td></td>
<td>Mean[Gy]</td>
</tr>
</tbody>
</table>

DVH Data Extraction

- DVH data extraction generally includes the tasks of manually retrieving the treatment plan from storage, navigating its contents, and transcribing plan information into the analysis software.

- This process can be time-consuming and error-prone, especially when information needs to be extracted from a large number of treatment plans for clinical studies.
Methods

• How were DVH data extracted in our study?
  – In-house developed system, PlanDB, that can store, organize, and present radiation treatment plan data.
  – In use at our institution since 2008.


PlanDB Components and Implementation

• Components
  – Data source: Pinnacle³ (Philips Medical Systems, Milpitas, CA)

• Implementation
  – The system was implemented through a combination of the internal scripting language in the TPS and 3 externally executed codes.

Kantor, Starkschall, and Balter, J Radiat Oncol Inform 2013;5:1:1-10

PlanDB Code Details

• The internal scripting language in the TPS is used to access and output the DVH data.

• 3 additional externally executed codes are used to parse ASCII data and interact with a networked database server:
  1. File parser capable of converting structured TPS files into hierarchical objects (Python).
  2. Abstraction function to take these raw objects as input and convert them to processed objects more suited for database input (MATLAB).
  3. General function that accepts processed objects as input and queries the database (Python).

Kantor, Starkschall, and Balter, J Radiat Oncol Inform 2013;5:1:1-10
DVH Data Exported to Database

- Pinnacle hot script “Darkroom” create pdf, also includes option to send DVH data to PlanDB.
- Script is run for every approved plan that is sent to Mosaic.

PLANDB Web-based Interface

- Easily accessible from our physics intranet.

Customizable Database Queries

- Create study
- Define authorized study users
- Indicate desired dose and volume levels + default min, max, mean
### Customizable Database Queries

- **Select plan(s) of interest**

<table>
<thead>
<tr>
<th>MRN</th>
<th>Name</th>
<th>Plan</th>
<th>Trial</th>
<th>ROI</th>
<th>Volume</th>
<th>( V_{0.5} )</th>
<th>( V_{13} )</th>
<th>( V_{20} )</th>
<th>( D_{1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>123456</td>
<td>Doe, Jane</td>
<td>DIBH</td>
<td>Lt Breast</td>
<td>Total Lung</td>
<td>3704.70</td>
<td>0.09</td>
<td>4.01</td>
<td>39.35</td>
<td>7.42</td>
</tr>
<tr>
<td>123456</td>
<td>Doe, Jane</td>
<td>DIBH</td>
<td>Lt Breast</td>
<td>Lt Lung</td>
<td>2062.48</td>
<td>0.02</td>
<td>0.12</td>
<td>1.01</td>
<td>0.00</td>
</tr>
<tr>
<td>123456</td>
<td>Doe, Jane</td>
<td>DIBH</td>
<td>Lt Breast</td>
<td>Heart</td>
<td>577.70</td>
<td>0.35</td>
<td>0.53</td>
<td>3.36</td>
<td>33.12</td>
</tr>
</tbody>
</table>

Data is sorted by ROI name

Standardized roi nomenclature is beneficial

### Now the non-automated part

- **Limitations of current method**
  - Format of PlanDB data not ideal for our data pooling study.
    - For pilot study, we manually sorted data to best format
    - Next steps, automate this process.
  - In our pilot study, we only used PlanDB to extract DVH data.
    - Patients’ demographics and prescription information were abstracted from electronic medical records and Mosaic.
Results

- Between our the 2 participating institutions, we pooled data for 350 BRT patients:
  - 200 with left-sided cancers (74% treated with DIBH) and
  - 150 with right-sided cancers

### Results - Pooled Data

<table>
<thead>
<tr>
<th>Organ</th>
<th>Dose Metric</th>
<th>Tangent (T)</th>
<th>T + SCN</th>
<th>T + SCN + IMN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart</td>
<td>Mean (Gy)</td>
<td>N = 93</td>
<td>N = 10</td>
<td>N = 91</td>
</tr>
<tr>
<td></td>
<td>V40(Gy%)</td>
<td>1.1±0.2</td>
<td>0.0±0.0</td>
<td>2.9±1.6</td>
</tr>
<tr>
<td></td>
<td>V50(Gy%)</td>
<td>0.5±0.0</td>
<td>0.1±0.1</td>
<td>1.9±1.6</td>
</tr>
<tr>
<td>Lungs</td>
<td>Mean (Gy)</td>
<td>N = 77</td>
<td>N = 80</td>
<td>N = 75</td>
</tr>
<tr>
<td></td>
<td>V40(Gy%)</td>
<td>10.7±5.7</td>
<td>10.1±5.1</td>
<td>23.9±10.6</td>
</tr>
<tr>
<td></td>
<td>V50(Gy%)</td>
<td>5.8±3.4</td>
<td>6.2±2.9</td>
<td>14.2±5.3</td>
</tr>
<tr>
<td>Ipsilateral Lungs</td>
<td>Mean (Gy)</td>
<td>N = 64</td>
<td>N = 60</td>
<td>N = 63</td>
</tr>
<tr>
<td></td>
<td>V40(Gy%)</td>
<td>21.8±9.2</td>
<td>21.7±9.5</td>
<td>43.5±19.0</td>
</tr>
<tr>
<td></td>
<td>V50(Gy%)</td>
<td>11.9±5.8</td>
<td>11.3±4.9</td>
<td>23.6±5.5</td>
</tr>
</tbody>
</table>

### Results – by Institution

- [Graph showing dose distribution for heart and ipsilateral lung for left and right cases]
Understanding the Differences in Data between Institutions

• It is important not to jump to conclusions about differences in the data between the institutions, but ……

rather to try to understand differences in treatment techniques and data collection processes that may be the underlying source(s) of the differences.

Differences

• Potential sources of differences in dose metrics between our two institutions include the following:
  – Initial plan versus composite plan DVH data.
  – Different percent of left BRT treated with DIBH and FB.
  – May be using different target definition
  – Different patient populations and breast cancer staging.

Future Implications of Cooperative Prospective Data Pooling for BRT

• Define standard dose metrics.
• Correlation of actual lung and heart doses with long-term outcomes.
• Establishment of more accurate NTC dose-response models based on:
  – actual patient doses, and
  – dose volume effects not simply mean dose.
• Evaluation of treatment techniques across multiple institutions for quality comparisons.
Acknowledgements

**M.D. Anderson Cancer Center**
- Simona Shaitelman, MD
- Wendy Woodward MD, PhD
- Michael Kantor
- Mohammad Salehpour, PhD
- Mary Martel, PhD

**Mayo Clinic Rochester**
- Charles S. Mayo, PhD
- Ivy A. Petersen, MD
- Elizabeth S. Yan, MD
- Sean S. Park, MD
- Robert W. Mutter, MD

Thank you

Extra Slides
Evidence of Radiation Related Cardiac Effects

- Until recently, there was a general belief that radiation related cardiac effects were only associated with high doses, i.e., >30 Gy.
- More recently, evidence is emerging that cardiac toxicity can occur at much lower doses.
  - A bomb survivors (Preston et al. 2003)
  - Patients treated for peptic ulcers (Carr et al. 2005)
  - Childhood cancer survivors (Mulroony et al. 2009, Tukenova et al 2010)

Absolute Risk of a RT-Related Major Coronary Event

- For a 50 year old woman with no preexisting heart disease, the absolute risk of death from ischemic heard disease would increase from 1.9% to 2.4 % and 3.4 %, respectively for mean heart doses of 3 Gy and 10 Gy, respectively:

  **Absolute increase in risk is relatively small, between 0.5 and 1.5%**.

Darby et al. NEJM 2013

Why was the Darby et al. study such an important study?

- Estimated Number of Cancer Survivors in the United States (Millions)
  - Male: 14.5 Million Survivors
  - Female: 18.9 Million Survivors
  - 2014: 18.9 Million Survivors
  - 2024: 18.9 Million Survivors

- Chart showing the increase in estimated number of cancer survivors.

11
Why was the Darby et al. study such an important study?

• By 2024, it's estimated that there will be $> 3.9 \times 10^6$ breast cancer survivors.
• Assuming 50% receive RT, radiation may be related to 10,000 to 30,000 deaths from major coronary events for mean heart doses between 3 and 10 Gy.