Functional Imaging Quantitation for Radiation Therapy Planning in Head and Neck Cancer

John Buatti, MD
Department of Radiation Oncology
University of Iowa
Specific Aims

• Specific Aim 2:
• Develop novel semi-automated tools for reproducible tumor definitions applicable to quantitative image-based response assessment that will be compared with manual methods.
1. Volume of interest (VOI) masks for structures with uptake:
   - Reference regions (similar to PERCIST\(^+\))
   - Primary cancer
   - Lymph nodes
   - Etc.

2. Quantitative indices derived from segmented VOIs:
   - SUVmax, SUVpeak, SUVaverage, Volume, Metabolic Tumor Volume (MTV), …
   - Indices may also be based on CT data

Advantages of Approach

• Flexibility:
  – switch from
    • MTV $\rightarrow$ SUVpeak
    • reference region: liver $\rightarrow$ cerebellum
    • …

• New quantitative indices can be calculated retrospectively
• If outcome data or surrogate end points are available:
  – Select indices utilizing machine learning techniques
• Transparent process
• Digital volumetric models for RT targeting
• …

Q: Why not frequently utilized?
**Segmentation Approach**

→ key for success (usability)

- **Automated**
  - High variability of results
  - Zero user effort

- **Semi-Automated**
  - Zero variability of results
  - High user effort

- **Manual**
  - Zero variability of results
  - High user effort

IDEAL

Reference VOI Generation

Lesion VOI Generation

“Just Enough Interaction” Approach

Zero*
Automated Reference VOI Generation

• Methods for
  • Cerebellum (full)
  • Liver (tri-axial ellipsoid)
  • Aortic arch (tube, CT image)
• “Search regions” based on a brain segmentation
  • Gray-value threshold, morphology, size analysis
Cerebellum VOI Algorithm

• Based on a **Robust Active Shape Model**
• Learn shapes of cerebella → model
• Match model to new image data → VOI

1st mode of shape variation

Mean Shape

Training Examples
Example of a Resulting Cerebellum VOI
Validation – Image Data

- 134 PET/CT scans from 49 subjects with H&N cancer
- F-18 FDG (370 MBq +/- 10%)
- Uptake time 90 min +/- 10%
- Subjects fasted >4h
- Blood glucose <200 mg/dl
- Arms down
- 128x128 pixel matrix (3.5 x 3.5 x 3.4 mm) or
  168 × 168 pixel matrix (3.4mm × 3.4mm × 2.0mm)
Validation - Uptake in Cerebellum

• Independent reference standard:
  • Experts manually traced the cerebellum in 4 cross-sections (1 axial, 1 coronal, 2 sagittal [left & right hemisphere])
  • 2 experts: 134, 1 expert: 44, and 1 expert: 20
  • Average SUV from all 4 cross-sections
  • Consensus-true SUV model based on expert results + statistical analysis approach
Validation - Uptake in Cerebellum

• Independent reference standard:
  • Experts manually traced the cerebellum in 4 cross-sections (1 axial, 1 coronal, 2 sagittal [left & right hemisphere])
  • 2 experts: 134, 1 expert: 44, and 1 expert: 20
  • Average SUV from all 4 cross-sections
  • Consensus-true SUV model based on expert results + statistical analysis approach
Liver VOI Generation

PET volume

Distance transform

> 1 SUV

Hole closing

VOI
Validation - Uptake in Liver

- Reference: 2 experts
- Tracing in one axial, sagittal and coronal slice
- Same 134 scans
- Same analysis steps

<table>
<thead>
<tr>
<th>Site</th>
<th>Auto</th>
<th>Intercept $\beta_0$ (SUV)</th>
<th>Slope $\beta_1$ (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver</td>
<td>Volume</td>
<td>0.02 (-0.23, 0.15)</td>
<td>0.97 (0.94, 1.00)</td>
</tr>
</tbody>
</table>
Examples of Automatically Defined Reference VOI (1)
Decrease in total variability if the automated method was used instead of the manual method:

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Measurement</th>
<th>Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebellum</td>
<td>4-Slice</td>
<td>99.2% *</td>
</tr>
<tr>
<td></td>
<td>Volume</td>
<td>89.8% *</td>
</tr>
<tr>
<td>Aortic arch</td>
<td>Volume</td>
<td>76.7% *</td>
</tr>
<tr>
<td>Liver</td>
<td>Volume</td>
<td>54.7%</td>
</tr>
</tbody>
</table>
VOI Generation for Lesions

- Segmentation problem $\rightarrow$ graph-based optimization approach (Optimal Surface Segmentation)
- Graph + cost function (design is critical!)
- Integrated into 3D Slicer (www.slicer.org)
Complexity issues
Generating Data

- SUV Max, Mean
- Metabolic Tumor Volume
- Range Pixel Values
- Volume over/x SUV
- Normalization against Liver/Blood Flow
- Lowest Quartile
- Highest Quartile
- Etc…………..
PET Module Tool
Types of Analyses

P-value (log)

Feature: 1 4 6 9 11 19 22 23 25 27

- No normalization
- Liver
- Cerebellum
- Aorta

P=0.05

True positive rate

False positive rate

CV AUC = 0.875
p = 0.0049
Decision Support

- Clinical Data
- Imaging Data
- Current Patient Biomarker Profile
- "Query Biomarker Profile"
- Patients with Similar Biomarker Profile
- Select Possible Treatment
- Classify Outcome based on Treatment
- Output: Unfavorable Treatment Option
- Unfavorable Outcome Predicted
- Favorable Outcome Predicted
- Output: Favorable Treatment Option
- Output: Browse Comparable Patients
Validation Study

• Using a set of 60 cases with 230 lesions
  – 3 investigators randomly contoured each case using manual (twice) or PET module tool (twice)
  – 2760 contoured lesions
  – Compared for internal consistency and against best estimate of ground truth
**Automated tool agreement**

### Table 1

<table>
<thead>
<tr>
<th>Agreement</th>
<th>Manual N</th>
<th>Manual Mean</th>
<th>Manual 95% CI</th>
<th>Semi-automated Mean</th>
<th>Semi-automated 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within Operator</td>
<td>690</td>
<td>77.0</td>
<td>(74.7–79.4)</td>
<td>92.6</td>
<td>(90.2–94.9)</td>
</tr>
<tr>
<td>Between Operator</td>
<td>690</td>
<td>79.8</td>
<td>(78.4–81.2)</td>
<td>94.1</td>
<td>(92.7–95.6)</td>
</tr>
</tbody>
</table>

Table 1. Estimated mean dice coefficients for intra and inter-operator segmentation agreement.

### Table 2

<table>
<thead>
<tr>
<th>Method</th>
<th>Time (minutes)</th>
<th>95% CI (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual</td>
<td>8.88 ± 7.1</td>
<td>(6.47, 11.28)</td>
</tr>
<tr>
<td>Semiauto</td>
<td>3.74 ± 3.3</td>
<td>(2.40, 5.08)</td>
</tr>
</tbody>
</table>

Table 2. Estimated mean times with standard deviation and 95% confidence intervals (CI) for manual and semi-automated segmentations.
Example of intra- and inter-operator segmentation agreement for manual and semi-automated segmentation methods. (a-d) Manual slice-by-slice segmentation results. (e-h) Semi-automated full 3D segmentation results. (i) Same PET image as in images (a-h), but with a different gray-value transfer function, showing uptake peaks corresponding to individual lymph nodes in close proximity.
(Top) A virtual dose map was generated from each manual contour (Manual-C) or semi-automated segmentation (SAS).

(Bottom) Its dose gradient was -2.7 % per mm, adapted from a clinical head-and-neck intensity-modulated radiation therapy (IMRT) plan.
The differences between SAS and Manual-C in terms of EUD, Logistic TCP and Poisson TCP when using STAPLE as a ground truth tumor contour. For all cases, SAS resulted in significantly lower (i.e. “−”) both intra- and inter-observer variability standard deviations regardless of TCP modeling (p < 0.0043).

<table>
<thead>
<tr>
<th></th>
<th>Differences (SAS – Manual-C)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ΔEUD</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.3</td>
</tr>
<tr>
<td>Median</td>
<td>-0.3</td>
</tr>
<tr>
<td>Intra-Observer Variability Standard Deviations</td>
<td>0</td>
</tr>
<tr>
<td>Inter-Observer Variability Standard Deviations</td>
<td>-1.2</td>
</tr>
</tbody>
</table>
Conclusions

• Quantitative imaging represents an opportunity to improve both our ability to consistently identify targets for radiation therapy as well as improve response assessment and prognostication of cancer

• Algorithmic tools are critical components to leverage the big data source and will be combined (multiparametric) with both other imaging and radiogenomics.
Conclusions

• As the complexity increases, the ability of strictly simple principles that have commonly guided therapy decision making is likely to go away.

• Final thought on algorithms and target identification in light of TCP and NTCP.