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

John Buatti, MD

Department of Radiation Oncology

University of Iowa

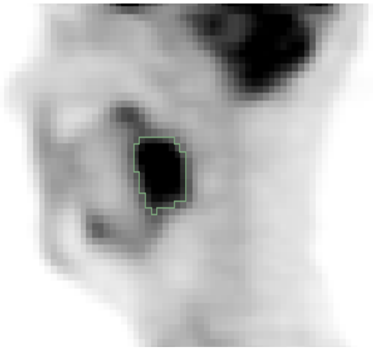
THE UNIVERSITY OF IOWA

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.

Image Analysis Output

1. Volume of interest (VOI) masks for structures with uptake:
   - Reference regions (similar to PERCIST[+])
   - Primary cancer
   - Lymph nodes
   - Etc.

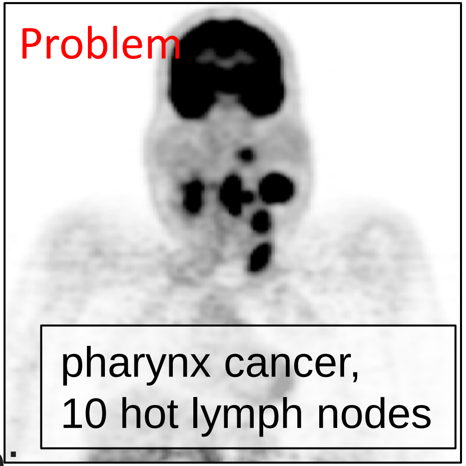


2. Quantitative indices derived from segmented VOIs:
   - SUVmax, SUVpeak, SUVaverage, Volume, Metabolic Tumor Volume (MTV), …
   - Indices may also be based on CT data

[+] *R.Wahl et al. From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. J Nucl Med. 2009.*

Advantages of Approach

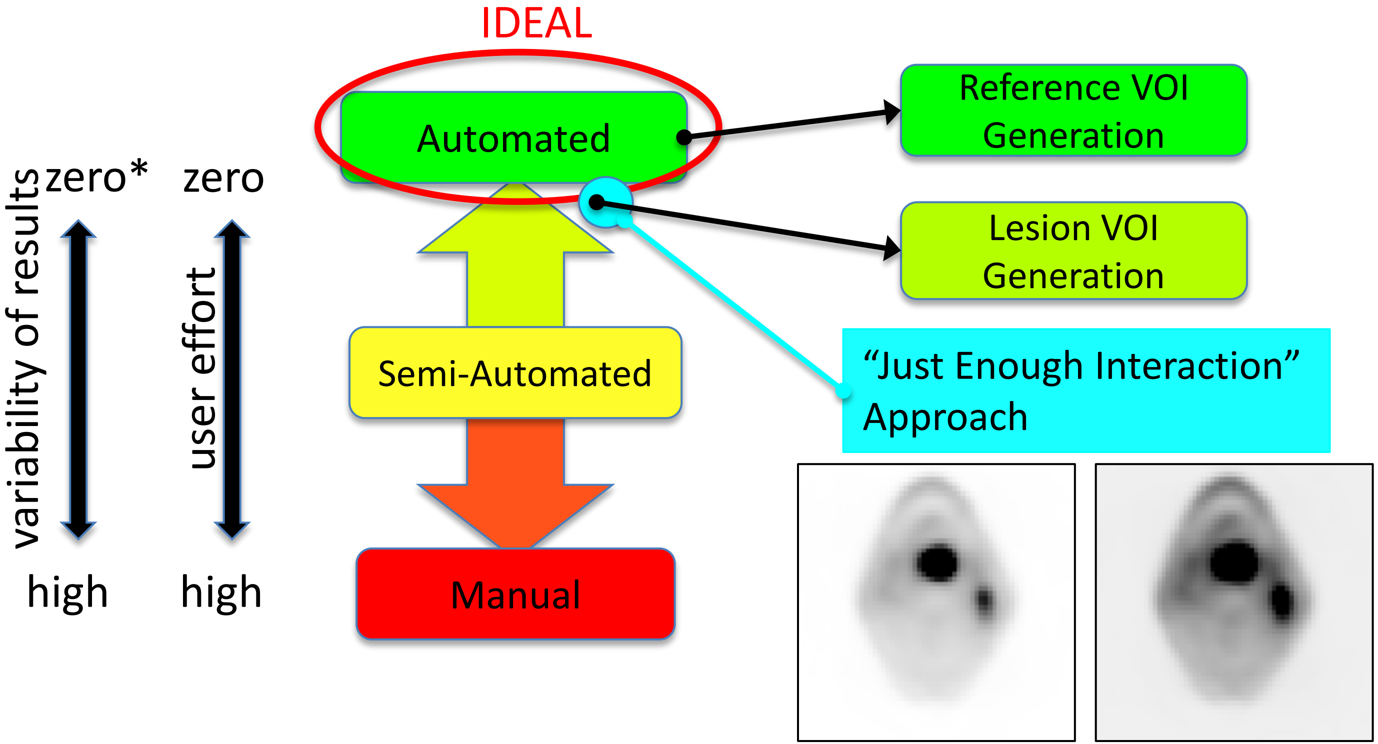
- Flexibility:
  - switch from
    - MTV → SUVpeak
    - reference region: liver → cerebellum
    - …
- New quantitative indices can be calculated retrospectively
- If outcome data or surrogate end points are available:
  - Select indices utilizing machine learning techniques
- Transparent process
- Digital volumetric models for RT targeting
- …



Q: Why not frequently utilized?

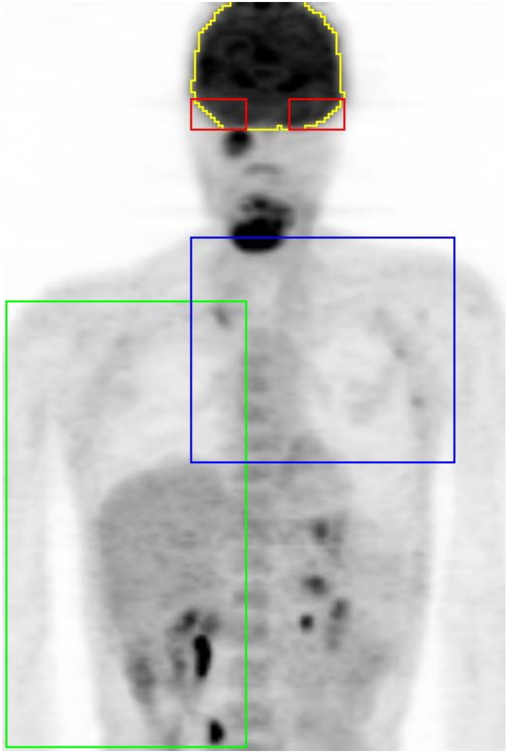
Segmentation Approach

→ key for success (usability)



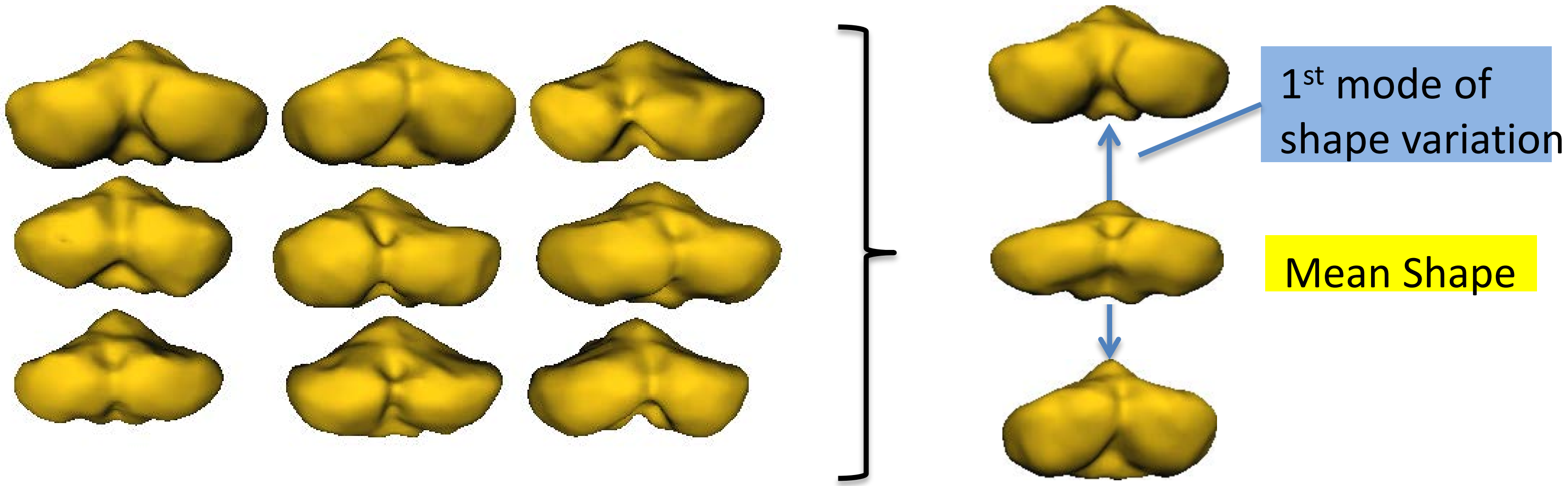
Automated Reference VOI Generation

- 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

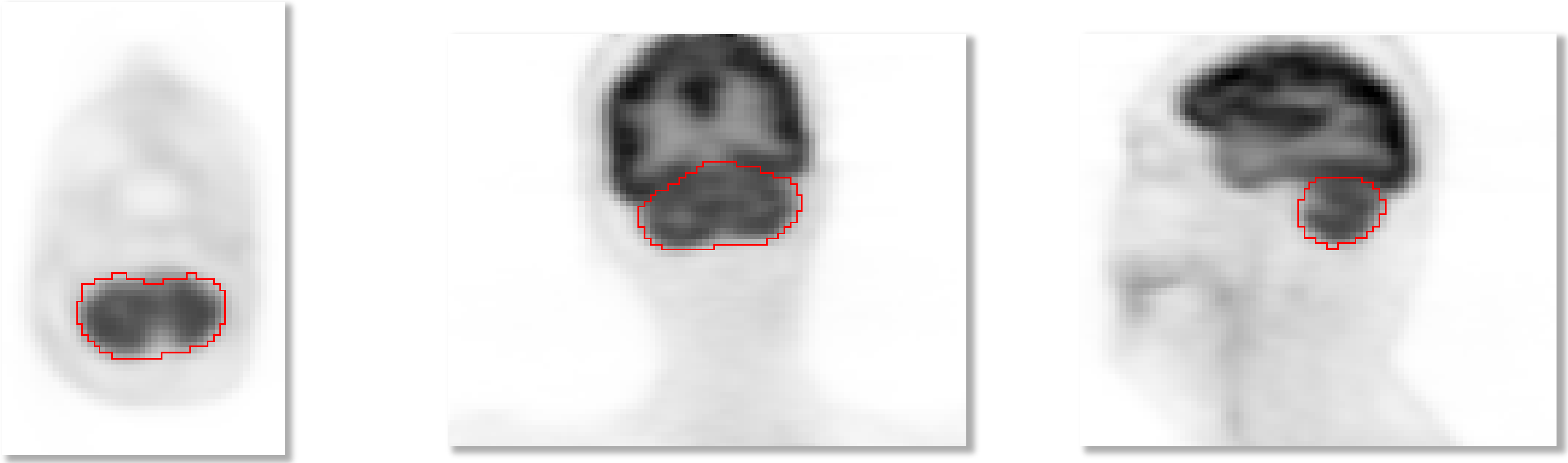
- Based on a **Robust Active Shape Model**
- Learn shapes of cerebella → model
- Match model to new image data → VOI



Training Examples

Model

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)

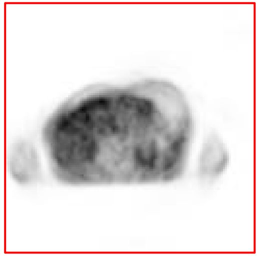
Validation - Uptake in Cerebellum

- 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

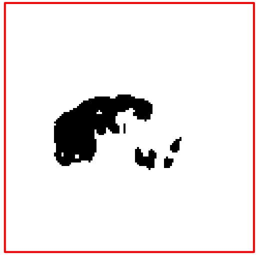
Validation - Uptake in Cerebellum

- 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

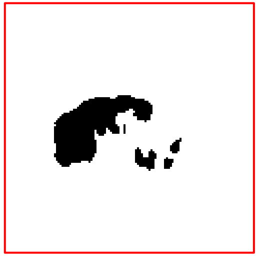
Liver VOI Generation



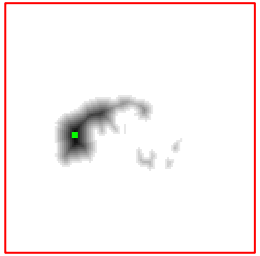
PET volume



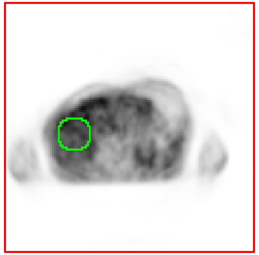
> 1 SUV



Hole closing



Distance transform

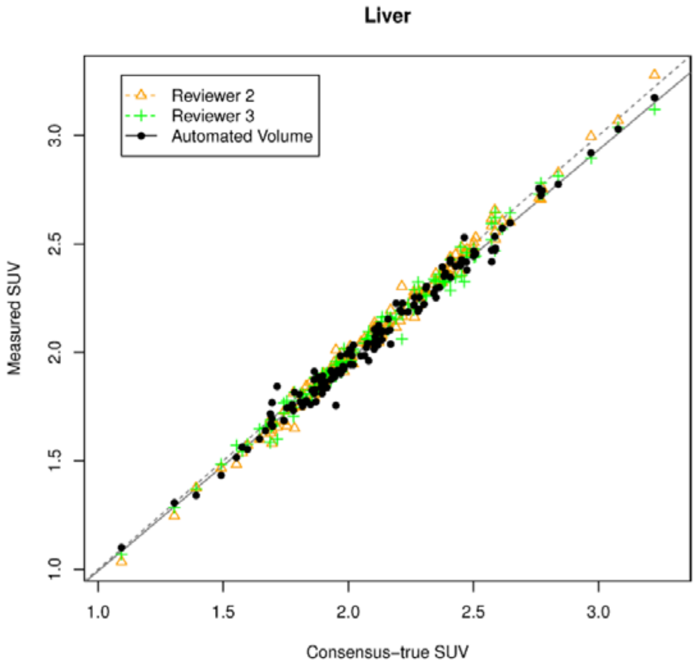


VOI

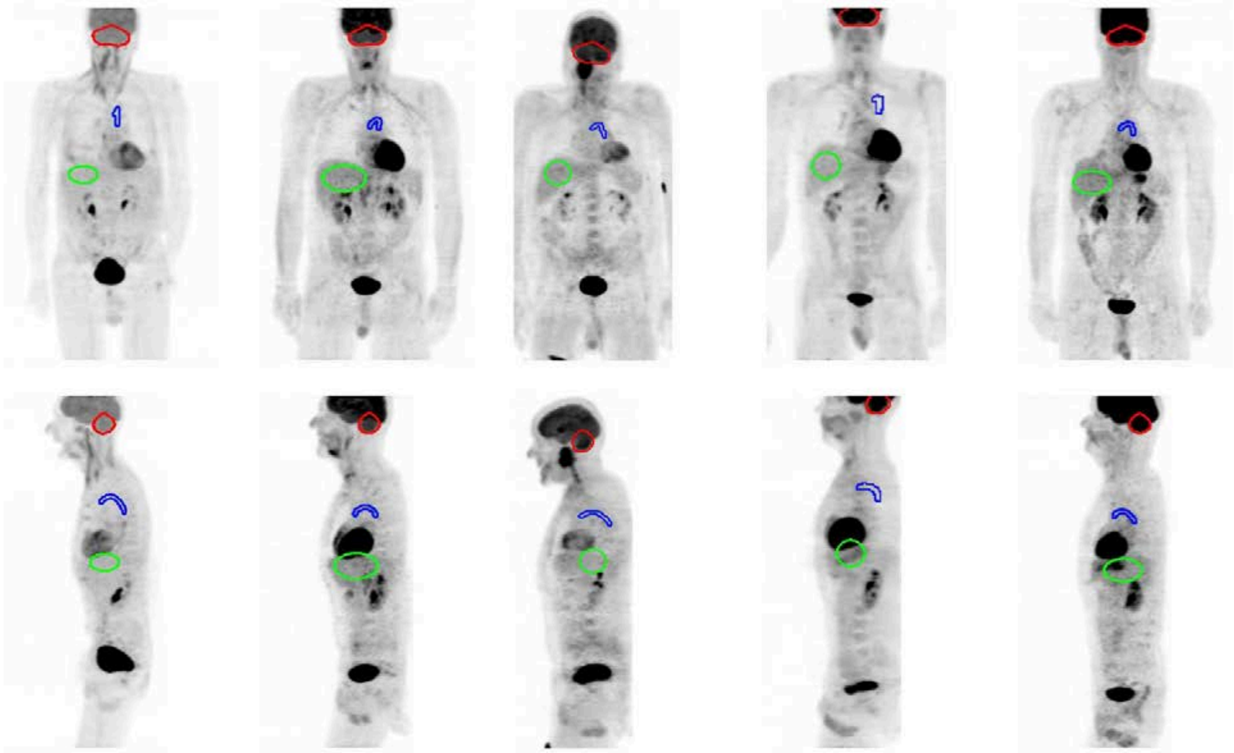
Validation - Uptake in Liver

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

Site	Auto	Intercept β_0 (SUV)	Slope β_1 (-)
Liver	Volume	0.02 (-0.23, 0.15)	0.97 (0.94, 1.00)



Examples of Automatically Defined Reference VOI (1)



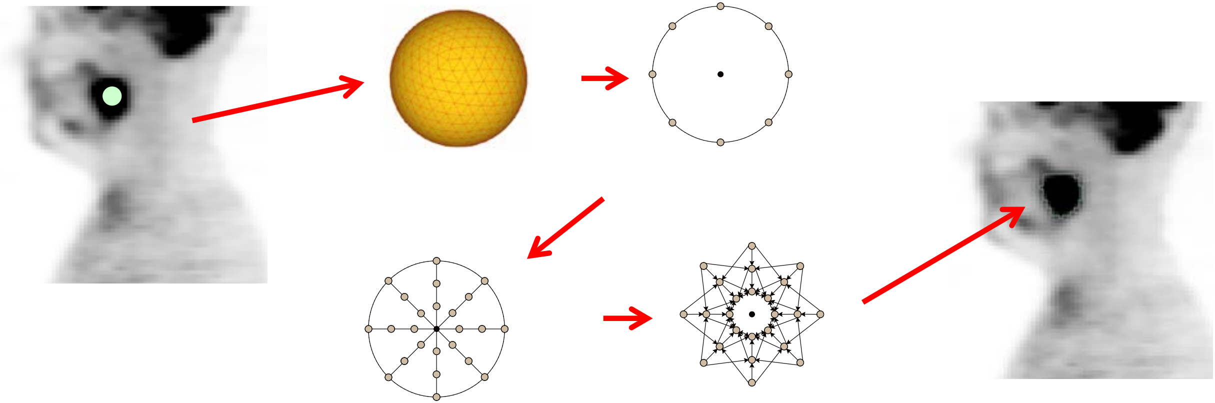
Comparison Automated vs. Manual

Decrease in total variability if the automated method was used instead of the manual method:

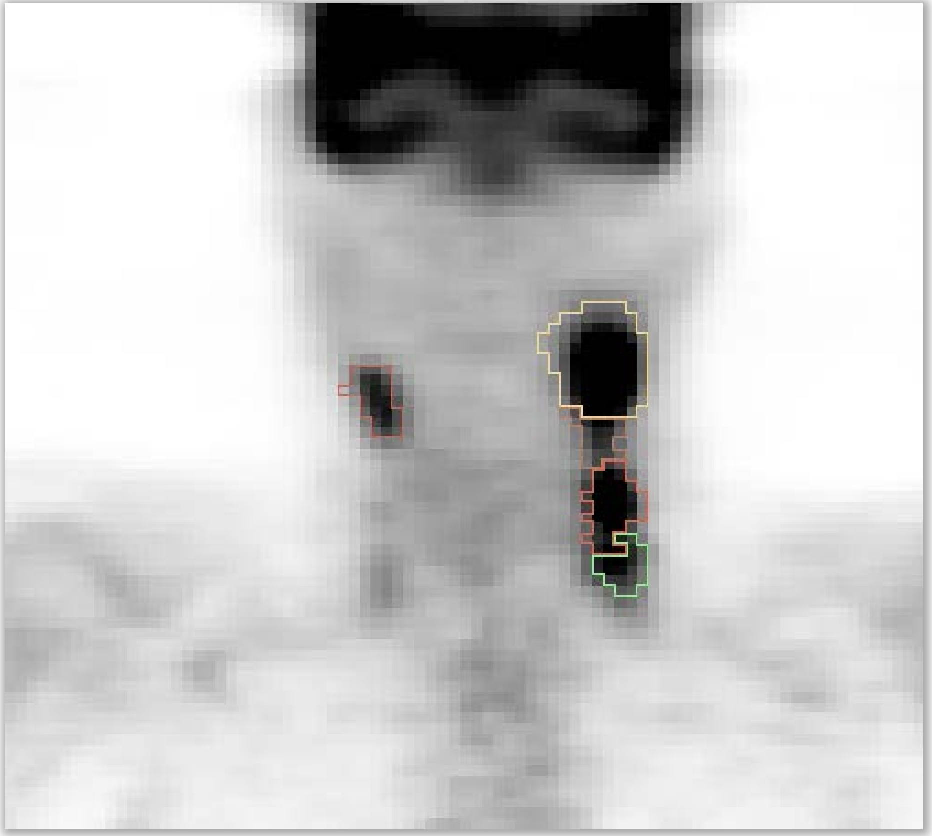
Cerebellum	**4-Slice**	**99.2% ***
	Volume	**89.8% ***
Aortic arch	**Volume**	**76.7% ***
Liver	**Volume**	**54.7%**

VOI Generation for Lesions

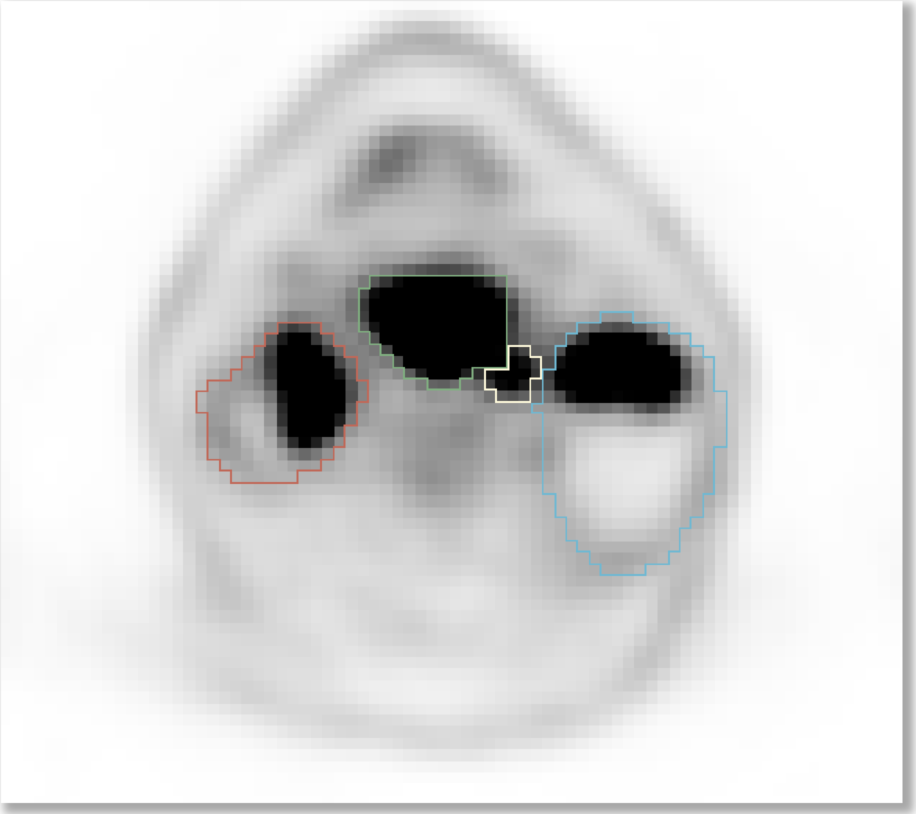
- Segmentation problem → graph-based optimization approach (Optimal Surface Segmentation)
- Graph + cost function (design is critical!)
- Integrated into 3D Slicer (www.slicer.org)



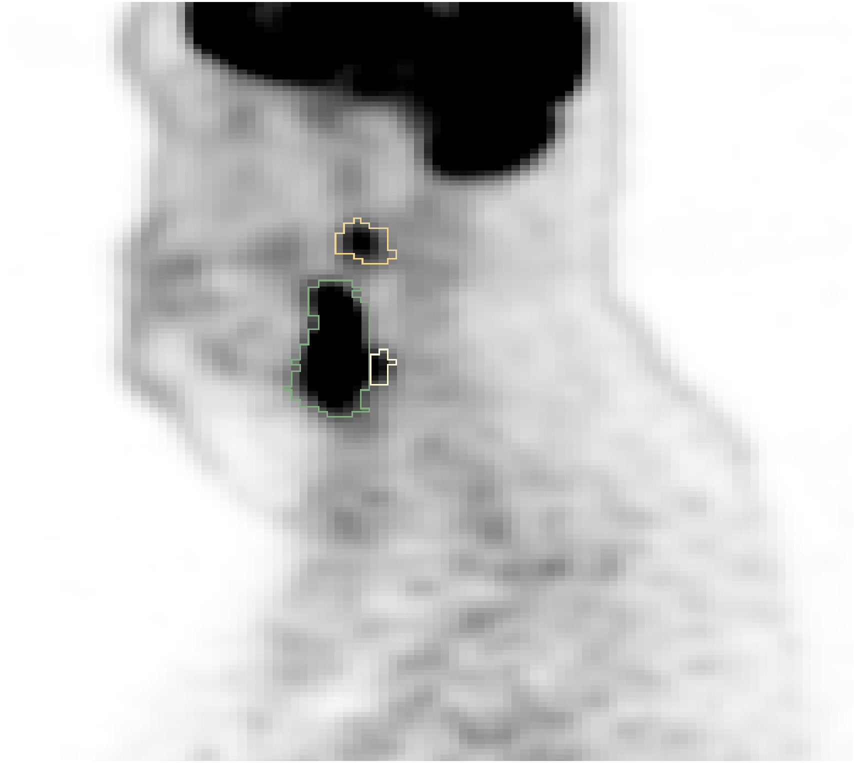
Complexity issues



Complexity issues



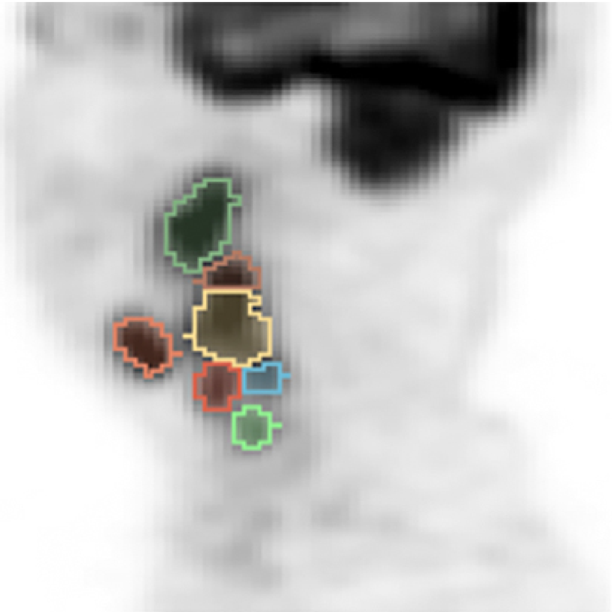
Generating Data



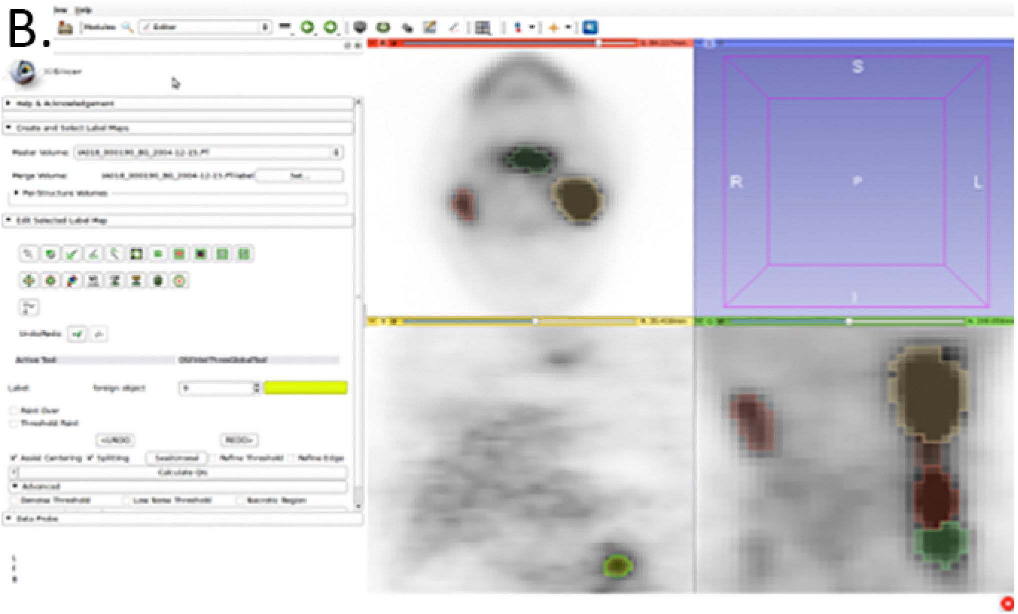
- SUV Max, Mean
- Metabolic Tumor Volume
- Range Pixel Values
- Volume over/x SUV
- Normalization against
- Liver/Blood Flow
- Lowest Quartile
- Highest Quartile
- Etc…………..

PET Module Tool

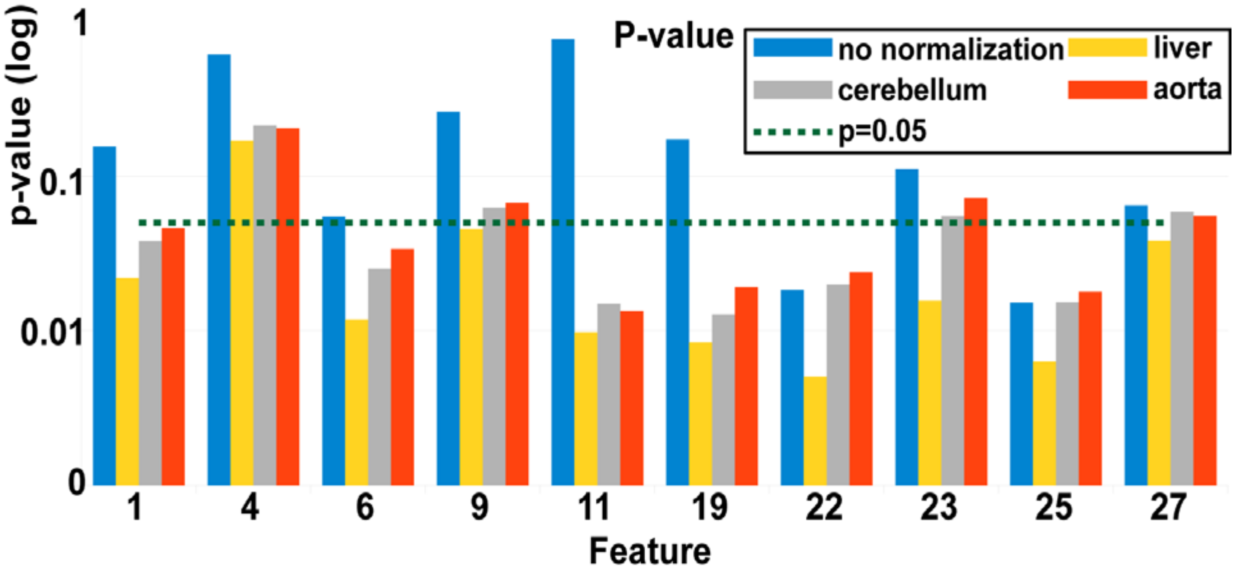
A.

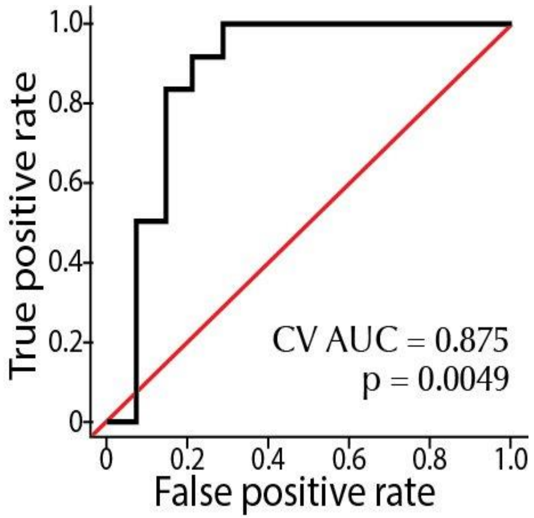


B.

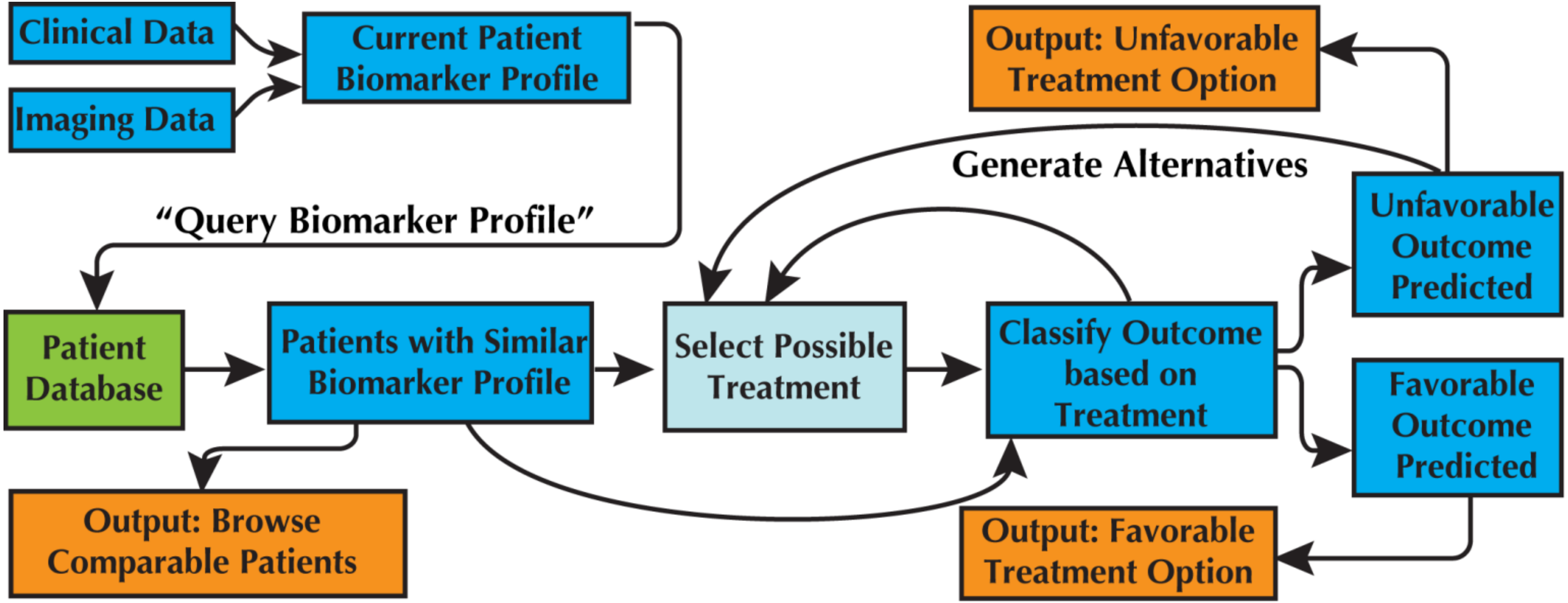


Types of Analyses





Decision Support



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

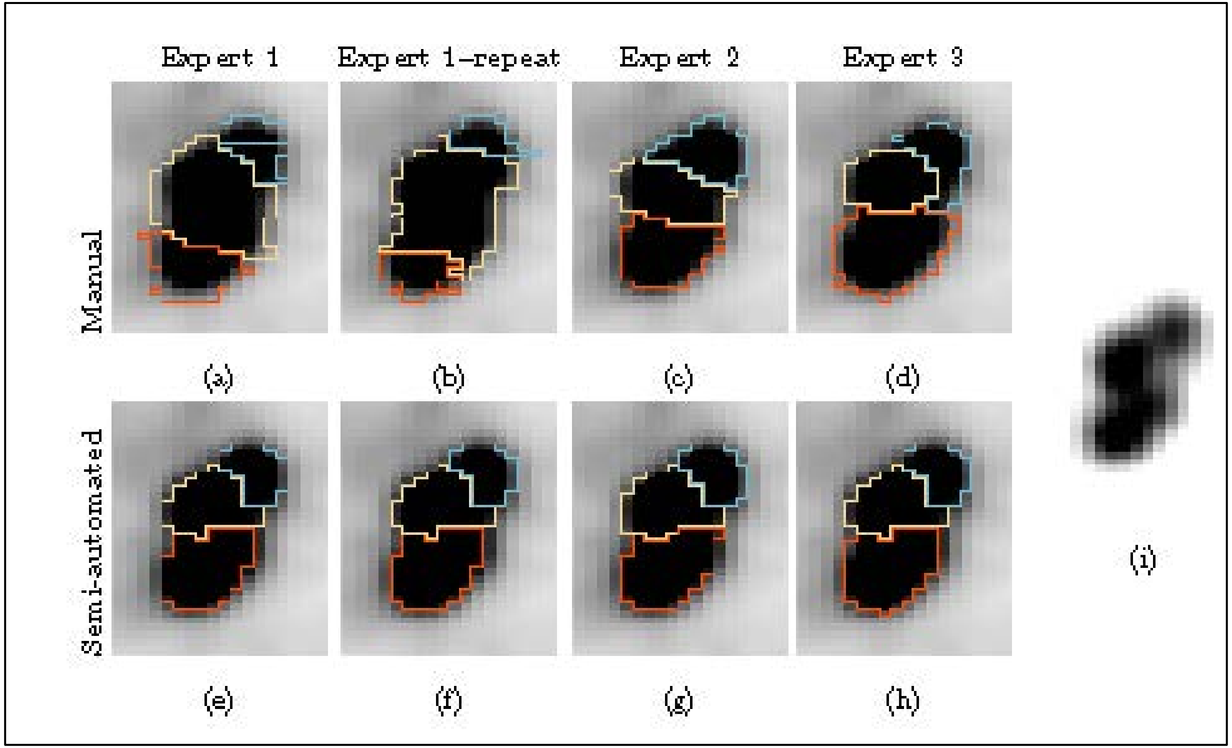
		Manual		Semi-automated	
Agreement	**N**	**Mean**	**95% CI**	**Mean**	**95% CI**
Within Operator	690	77.0	(74.7–79.4)	92.6	(90.2–94.9)
Between Operator	690	79.8	(78.4–81.2)	94.1	(92.7–95.6)

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

Table 2

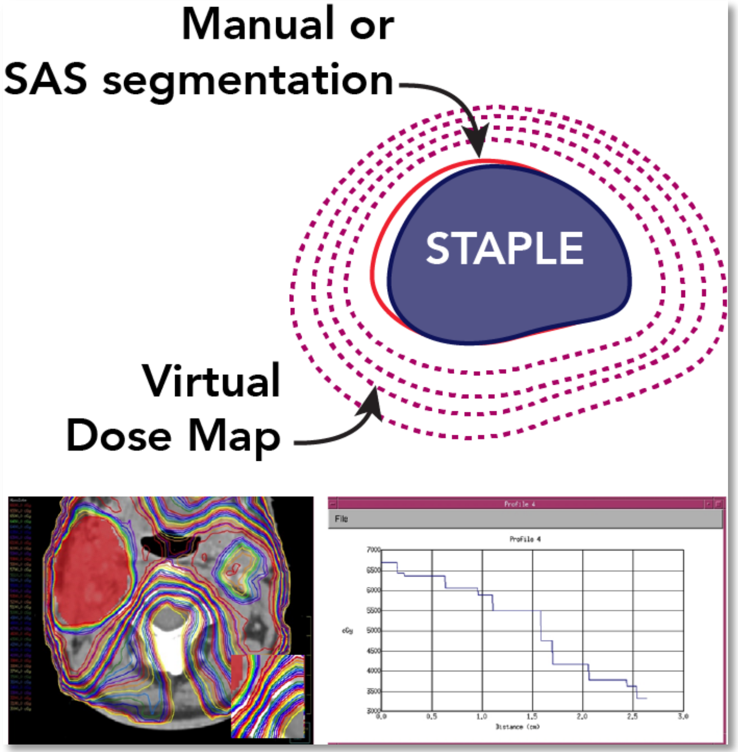
Method	Time (minutes)	95% CI (minutes)
Manual	8.88 ± 7.1	(6.47, 11.28)
Semiauto	3.74 ± 3.3	(2.40, 5.08)

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.

Tumor Control Probability



(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.

Tumor Control Probability

	Differences (SAS – Manual-C)		
	ΔEUD	ΔLogistic TCP	ΔPoisson TCP
Mean	-0.3	-1.2%	-1.1%
Median	-0.3	-1.0%	-1.0%
Intra-Observer Variability Standard Deviations	0	-0.9%	-0.8%
Inter-Observer Variability Standard Deviations	-1.2	-4.7%	-4.7%

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 ($p < 0.0043$).

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