Metabolic volume measurement (physics and methods)

Ronald Boellaard
Department of Nuclear Medicine & PET Research,
VU University Medical Center, Amsterdam
r.boellaard@vumc.nl
Metabolic volume vs tumor size

Tumor size = 1D, 2D or 3D measurement of tumor size on structural (anatomical) images (CT or MR)

Metabolic volume = 3D measurement of the metabolically most active part of the tumor

Frings, et al. JNM 2010
Metabolic volume

- "Biological" target volume – RT
- Prognostic factor (Sasanelli M, et al. 2012)
- Predictive factor (residual or change in...)
- SUV x MVOL= TLG (or TLP for $^{18}$F-FLT)

Quantitative Analysis of Response to Treatment with Erlotinib in Advanced Non–Small Cell Lung Cancer Using $^{18}$F-FDG and 3'-Deoxy-3'-$^{18}$F-Fluorothymidine PET

Deniz Kahraman1,2, Matthias Scheffler2,3, Thomas Zander2,3, Lucia Nogova2,3, Adriaan A. Lammertsma4, Ronald Boellaard5, Bernd Neumaier5, Roland T. Ullrich2,3,5, Arne Holstein1,2, Markus Dietlein1,2, Jürgen Wolf2,3, and Carsten Kobe1,2
Some automated metabolic volume methods

• Simple fixed thresholds (e.g. SUV=2.5)
  – PRO: widely available
  – CON: too simple, may fail for small lesions and low contrasts

• % thresholds (e.g. 42 or 50% of SUVmax)
  – PRO: widely available
  – CON: simple, may fail for small lesions and low contrasts

• Source-to-background or contrast oriented methods
  (e.g. Schaefer, Adaptive 42%, A50%)
  – PRO: better performance for small lesions and low contrasts
  – CON: less widely available

• Gradient(-watershed) based methods (Lee and Geets)
  – PRO: theoretically best method in case of uniform distributions
  – CON: almost not available

• Cluster based methods (e.g. fuzzy clustering, FLAB-Hatt et al.)
  – PRO: very promising results in literature, can deal with uptake heterogeneity
  – CON: not available, method hard to implement/reproduce, user interaction unclear

• All automated methods needs supervision (outliers/corrections)!
Definition of target volume with PET/CT: which method?

Results depend on segmentation method being used:

<table>
<thead>
<tr>
<th>Method</th>
<th>Volume</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT:</td>
<td>GTV - CT</td>
<td>47.5 cm³ (rood)</td>
</tr>
<tr>
<td>PET:</td>
<td>GTV - visueel</td>
<td>43.8 cm³ (groen)</td>
</tr>
<tr>
<td></td>
<td>GTV₄₀%</td>
<td>20.1 cm³ (geel)</td>
</tr>
<tr>
<td></td>
<td>GTV_{SUV}</td>
<td>32.6 cm³ (oranje)</td>
</tr>
<tr>
<td></td>
<td>GTV_{SBR}</td>
<td>15.7 cm³ (blauw)</td>
</tr>
</tbody>
</table>

**Manual**

**Semi-automated**
Theory of metabolic volume segmentation
Factors affecting metabolic volume measurements

1. Tumor characteristics
   – Tumor or metabolic volume size
   – Tumor to (local) background ratio – contrast

2. Image characteristics
   – Image resolution
   – Image noise

3. VOI method
Partial volume

constant concentration

finite resolution

Recovery

Spill-over

Courtesy of J. Nuyts
Theory of metabolic volume segmentation (1)

In this example:
SUV2.5 = 50% of max: only slight overestimation

...... SUV = 2.5

...... SUV = 2.5

...... 50% of max

Now, same metab. volume but higher uptake
SUV2.5 > 50% of max: large m. volume overestimation
Theory of metabolic volume segmentation (1)

SUV2.5 > 50% of max: large m.volume overestimation

........ SUV=2.5

........ 50% of Max

........ SBR-50%

Same uptake, smaller volume
SUV2.5 and 50% of Max overestimation metab.volume
Theory of metabolic volume segmentation (1)

SUV2.5 > 50% of max: large m.volume overestimation

Same volume, same uptake, higher background

- SUV2.5 overestimates..
- 50% of Max seems OK again....
Theory of metabolic volume segmentation (1)

SUV2.5 > 50% of max: large m.volume overestimation

Same volume, same uptake, heterogeneous background
Basically only gradient may work.....
Clinical example

Both the measured $\text{SUV}_{\text{max}}$ and tumour volume depends on image characteristic settings.

Image resolution (FWHM): 11 mm, 7 mm
Estimated volume: 4.5 mL, 1.5 mL
$\text{SUV}_{\text{max}}$: 3.3, 5.5

Clinical example: a TRT study

- Patient studies:
  - 10 NSCLC patients in dynamic FDG TRT study
    - 51±5 y, weight 76±10 kg, 388±71 MBq
    - Blood glucose level were obtained
    - All patients fasted >6 h before scanning
    - Retest scan was acquired the next day

Cheebsumon et al. JNM 2011, EJNMMI 2011
Materials and methods

- Two different contrasts were used by summing the last 3 (45-60 min p.i.) and last 6 (30-60 min p.i.) frames

- Data were reconstructed using OSEM with 2 iterations and 16 subsets followed by post-smoothing using a Hanning filter

- Additional Gaussian smoothing was performed, resulting in resolutions of 6.5, 8.3 or 10.2 mm FWHM

Frings et al. JNM 2010
VOI methods....

- **9 different tumour delineation methods were used:**
  - Absolute SUV (i.e. SUV\(^{2.5}\))
  - Fixed or adaptive threshold of the maximum pixel value\(^{(1)}\) i.e. 50% (VOI\(^{50}\)) or A50% (VOI\(^{A50}\))
  - Relative threshold level (RTL) method\(^{(2)}\) (VOI\(^{RTL}\))
  - Adaptive threshold methods\(^{(3-5)}\) (VOI\(^{Nestle}\), VOI\(^{Erdi}\), VOI\(^{Schaefer}\))
  - Iterative threshold method\(^{(6)}\) (VOI\(^{Black}\))
  - Gradient-based segmentation method that applied the Watershed transform (WT) algorithm (Grad\(^{WT}\))

Results: effect of changes in resolution

Metabolic volume depends strongly on the resolution & VOI method being used.
TRT results: effect of changes in resolution

• Volume TRT depends on the resolution & VOI method being used (up to 20%)
TRT results: effect of changes in contrast

- FDG: for most VOI methods TRT worsens with lower contrast

Cheebsumon et al. JNM 2011, EJNMMI 2011
A clinical example: validation study

This example clearly shows difference between anatomical (CT) and metabolic (PET) tumor volumes, illustrating the potential of PET to identify regions within a tumor that show different metabolic activity. In this case PET-based volume was closer to pathology-derived volume than the CT-based volume.
Materials and methods

• Patients and pathology
  – 21 whole body FDG PET/CT (Biograph, CTI/Siemens) studies were acquired for primary NSCLC patients (77±14 kg)
  – Patients fasted for >6 h before scanning
  – Mean blood glucose levels were normal (5.7±2.0 mmol·L⁻¹)
  – Data were reconstructed using OSEM (4i, 18s), having an image resolution of ~6.5 mm FWHM
  – After scanning, the primary tumour was surgically resected and the maximum diameter of this tumour was measured

Materials and methods

- 8 different automatic PET-based delineation methods were used:
  - Absolute SUV threshold (e.g. SUV^{2.5})
  - Fixed or adaptive threshold of the maximum pixel value\(^{(1)}\) i.e. 50\% (VOI\(^{50}\)) or A50\% (VOI\(^{A50}\))
  - Relative threshold level (RTL) method\(^{(2)}\) (VOI\(^{RTL}\))
  - Adaptive threshold methods (e.g. VOI\(^{Erdi}\)\(^{(3)}\) and VOI\(^{Schaefer}\)\(^{(4)}\))
  - Iterative threshold method (e.g. VOI\(^{Black}\)\(^{(5)}\))
  - Gradient-based segmentation method in combination with a Watershed algorithm (Grad\(^{WT}\))

- Manual CT-based delineation by expert physician

Materials and methods

• Data analysis
  – Comparison of PET and CT derived volumes (volume difference, slope and $R^2$)

  – Comparison of maximum tumour diameter from PET- and CT-based methods to that obtained from pathology (diameter difference, slope and $R^2$)
Results – Diameter difference: vs pathology
## Results – Slope and $R^2$ of maximum diameter

Intercept set to 0

<table>
<thead>
<tr>
<th>R²</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT-based delineation</td>
<td>0.77</td>
</tr>
</tbody>
</table>

**PET delineation methods**

| VOI $^{50}$ * | 0.82  | 1.00 |
| VOI $^{70}$   | 0.73  | 0.79 |
| VOI $^{A42}$ * | 0.82  | 1.04 |
| VOI $^{A50}$  | 0.75  | 0.95 |
| VOI $^{A70}$  | 0.81  | 0.69 |
| VOI $^{Erdi}$ | 0.71  | 0.81 |
| VOI $^{Black}$ | 0.74  | 1.00 |
| VOI $^{Schaefer}$ * | 0.75  | 0.85 |
| VOI $^{RTL}$  | 0.78  | 0.97 |
| Grad $^{WT}$ * | 0.48  | 1.17 |
| SUV $^{2.5}$ * | 0.79  | 1.16 |

Slope and $R^2$ of maximum diameter obtained from PET-based delineation methods or CT delineation against maximum diameter obtained from pathology.

* Without outliers:
- 2 outliers for VOI$^{50}$, VOI$^{A42}$, VOI$^{Schaefer}$ and Grad$^{WT}$
- 5 outliers for SUV$^{2.5}$
Results – Diameter mean difference vs pathology

Cheebsumon, EJNMMI Research (in press)
Preliminary multi-center TRT results

TRT FDG PET/CT data from 4 sites (Velasquez et al. JNM)
Advanced GI malignancies
No standardisation in place

| Table 5a - Mean & RC of relative difference in volume |
|---------------------------------|----------|-------|-------|
| **Base parameter**             | **Method / threshold** | **n** | **Mean relative difference (%)** | **RC (%)** |
| GradWT                         | 85       | 23.4  | 38.5 |
| **SUV\(_{\text{max}}\)**      | A50%     | 87    | 20.2 | 37.0 |
| Schaefer                       | 89       | 15.9  | 25.7 |
| RTL                            | 87       | 14.9  | 25.2 |
| **SUV\(_{\text{peak}}\)**     | A50%     | 87    | 16.9 | 25.5 |
| Schaefer                       | 81       | 11.9  | 24.9 |
| RTL                            | 79       | 13.2  | 23.5 |
| **SUV\(_{\text{local peak}}\)** | A50%   | 77    | 12.1 | 22.7 |
| Schaefer                       | 86       | 13.2  | 26.9 |
| RTL                            | 86       | 17.1  | 28.0 |
| **SUV\(_{\text{star}}\)**     | A50%     | 86    | 17.4 | 28.9 |
| Schaefer                       | 86       | 17.3  | 29.0 |
| RTL                            | 86       | 17.4  | 28.9 |

Use of \(\text{SUV}_{\text{peak,3D}}\) and SBR based thresholds result in improved metabolic volume measurement repeatability (SUVpeak is less sensitive to noise)
Some automated metabolic volume methods

• **Simple fixed thresholds** (e.g. SUV=2.5)
  – Many outliers, not able to provide reproducible (TRT) results for small lesions (<5mL) and at low TBR (<4)

• **% thresholds** (e.g. 42 or 50% of SUVmax)
  – May work reasonable well for NSCLC (high contrast, low background)

• **Source-to-background or contrast oriented methods**
  (e.g. Schaefer, Adaptive 42%, A50%)
  – Reasonably good performance, available in some display stations, if not then can be applied with more user interaction
  – Use of SUVpeak rather than SUVmax improves TRT performance considerably

• **Gradient(-watershed) based methods** (Lee and Geets)
  – Theoretically best method in case of uniform distributions
  – Sensitive to noise

• **Cluster based methods** (e.g. fuzzy clustering, FLAB)
  – Not tested, not easy to implement and not available

• **All automated methods needs supervision (outliers/corrections)!**
Theory of metabolic volume segmentation
Factors affecting metabolic volume measurements

- Tumor or metabolic volume size
- Tumor to (local) background ratio – contrast
- Image resolution
- Image noise
- Automated VOI method being used

Impact of $^{18}$F-FDG PET imaging parameters on automatic tumour delineation: need for improved tumour delineation methodology

Patsaree Cheebsumon · Maqsood Yaquh · Floris H. P. van Velden · Otto S. Hoekstra · Adriaan A. Lammertsma · Ronald Boellaard

- For both SUV and metabolic volume assessments standardisation is required
- With STD and optimization: good TRT repeatability
EANM STD/Guideline

- Interpretation, image quality and quantification depends on the combination of many factors (biological, technical, physics)*

- FDG PET/CT guideline* – imaging procedure
  - Feasibility of following GL shown in several trials/studies

- NB it is a harmonizing guideline/standard aiming at minimizing difference in quantitative performance between centers

- GL is optimized for use of SUVmax for quantification!

- EARL accreditation- PET/CT system calibration/perf.harmonization
  - About 70 sites across EU, likely 100 in 2013
  - Options to arrive at harmonized image quality and quantification:
    - Acquire and reconstruct data such to meet harmonizing std (preferred)
    - 2 reconstructions, one that meets std (danger of mixing up)
    - Postproces data to generate second image dataset that meets std (online or during analysis)

Uniformity of Protocols In Clinical Trials: UPICT

FDG PET/CT consensus guideline
- EANM/EARL (GL & accreditation)
- SNM & SNM-CTN
- ACR
- RSNA
- QIBA
- PET/CT Vendors

UPICT FDG PET/CT consensus GL – imaging procedure GL

UPICT GL available for external review/comment Q4/2012
Thank you for your attention