Quantitative Imaging Using PET/CT

Paul Kinahan, PhD
Director of PET/CT Physics
Imaging Research Laboratory, Department of Radiology
University of Washington, Seattle, WA

SNM Clinical Trials Network Community Workshop
February 8-9, 2009
Clearwater, FL
Summary of Issues for Quantitative PET/CT Imaging

- "The favorable experience to date is beginning to support the use of PET as a surrogate end point in trials that are aimed at testing or comparing the efficacy of new drugs or treatments" [Juweid & Cheson NEJM 2006]
  - When numbers matter in evaluating response, understanding and reducing scanner bias and variance improves the statistical power of studies.

- Evaluation of new therapies requires multicenter studies for patient recruitment
  - Pooling results between different PET/CT scanners requires knowledge of biases between scanners to improve the statistical power of studies.

- To date most of the literature using PET to evaluate response to therapy is at single sites, on single scanners, using consistent protocols and analysis.

- Until recently, there have been few systematic efforts to understand or improve quantitative accuracy, precision, and stability between multiple sites.
Sources of Error in SUV Values

\[
\text{SUV} = \frac{\text{PET}_{\text{ROI}}}{D'_{\text{INJ}} / V'}
\]

PET = measured PET activity concentration
\(D'\) = decay-corrected injected dose
\(V'\) = surrogate for volume of distribution

It is important to minimize SUV errors for **serial** (e.g. response to Rx) or **multi-center** studies

Some potential sources of error are:

- Determining that blood glucose levels are within range
- Changing dose uptake time
- Scanner clocks, calibration, and cross timing/calibration with dose calibrator
- Dose assay for each patient
- Reconstruction and other processing protocols and parameters - can there be uniform protocols?
- Analysis methods: How ROIs are drawn and whether max or mean SUV values are reported
- Weight is typically used as a surrogate for volume of distribution, but can also be further normalized for lean body mass or body surface area
Implications for Multicenter Studies

- Different sources of bias and noise have implications for estimating power of clinical trials.
- We should be able to (1) reduce bias and variance (2) know the quality of our data.
  - Number vs true SUV difference and total noise in measurements.
  - Power of 80% and $\alpha=0.05$
  - Total noise would include biological, intra- and inter-center measurement effects.

![Graph: Number of Patients Needed in Multicenter Trials](image)
Quantitative PET Initiatives

- SNM Validation Task Force
- AAPM/SNM Task Group 154 - with intersociety MOU
- RadQual/NIST Dose calibrator standard
- RSNA Quantitative Imaging Biomarkers Alliance
- NCI Cancer Imaging Program RIDER (Reference Image Database for Evaluation of Response) program
- NCI-CIP/ACRIN longitudinal stability evaluation
SNM Study: Modified NEMA NU-2 IQ Phantom

- Hot sphere diameters of 10, 13, 17, 22, 28, and 37-mm
- Target/background ratio 4:1

- Max and mean activity concentrations measured via 10-mm diameter ROIs
‘Coffee Break’ Repeat PET/CT scans with Repositioning

SUVs from 20 3D-OSEM scans with 7-mm smoothing

GE DSTE-16 PET/CT Scanner

Siemens Biograph HI-REZ-16 PET/CT Scanner

- Intra-scanner short-term variability is 3% - 4%
Sample Image Sections from Six Different Scanners

Not meant as a "Consumer's Report" evaluation, but rather to facilitate multi-center comparisons

P Kinahan, University of Washington
SNM Phantom: Key results of SUV measurements

Plots of recovery coefficient (RC) = measured in ROI/true

Variations are introduced by the scanner type, acquisition protocol, calibration differences, processing (e.g. image reconstruction method or smoothing) and ROI definition method.

averaged coefficients of variation
mean SUV: 8.6%, max SUV: 11.1%
AAPM/SNM Task Group 145: Modified ACR Phantom

one batch of Ge-68 (0.35 uCi/ml)

<table>
<thead>
<tr>
<th>Cyl</th>
<th>Diam (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>25</td>
</tr>
<tr>
<td>B</td>
<td>16</td>
</tr>
<tr>
<td>C</td>
<td>12</td>
</tr>
<tr>
<td>D</td>
<td>8</td>
</tr>
</tbody>
</table>

Reference standard 20 cm uniform cylinder (one only) 2 mCi total activity

ACR phantom w/ extra lid with Ge-68 activity in cylinders A-D only

Shipping case with:
- empty ACR phantom
- matched lid with Ge-68 in cylinders A-D
- one each of aliquots of 0.1, 0.2, 0.3 ml
- Total activity 12 uCi

well counter / dose calibrator aliquots of 0.1, 0.2, 0.3 ml

Image with FDG in main cylinder

P Kinahan, University of Washington
Multicenter study with Pediatric Brain Tumor Consortium

- 10 PET imaging centers in PBTC
- Variations SUV ~12-21%, depending on method of analysis
- Central analysis reduces variations to ~6-9% range
Reference Image Database for Evaluation of Response (RIDER)

- Database developed as a resource for evaluating methods using CT images to evaluate response to therapy (e.g. other than RECIST)
- The RIDER database was extended to include PET and PET/CT patient images of response to therapy
- Also hosts key phantom data on biases and variability
- Provides a test-bed for display/measurement software by developers/vendors
RIDER PET/CT Patient Data

PET only images
- PET image 1: Pre-therapy
- PET image 2: Post-therapy
- PET image n: Serial studies

PET and CT images
- PET DICOM image stack
- CT DICOM image stack
- PET and CT image 1: Pre-therapy
- PET and CT image 2: Post-therapy
- PET and CT image n: Serial studies

Meta-data
- De-identified Patient information: therapy and outcomes

Minimum set
- PET only images
- PET and CT images

Optional extra sets if available
- PET only images
- PET and CT images
Sample patient images for the NCIA/RIDER database

PET

CT
Scanner Calibration: Activity Correction Factor Variation in Time

- Direct effect on SUV calculation
- Anecdotal evidence suggests this happens 'often'

P Kinahan, University of Washington
3/12/2009
F-18 assays by dose calibrators
Sample of 32 dose calibrators at 3 sites using RadQual/NIST Ge-68 standard
Where does this leave us?

• All of the issues with quantitative PET imaging are well-understood
• We need more data on multi-center dose-calibrator variability and longitudinal scanner variability, but both are underway
• The current framework of clinical PET does not directly support quantitative PET imaging: Lack of covariates and DICOM fields, lack of uniform protocols, acquisition of calibration information and vendor differences
• Calibration methods are now possible and vendors want to participate, but priorities have to be set to make best use of resources and balance cost:benefit
• This applies to both the specifics of a clinical trial using imaging and improving the general framework for quantitative PET imaging (e.g. QIBA)
Summary

• When numbers matter in evaluating response with PET/CT imaging, understanding and reducing scanner bias and variance improves the statistical power of studies.

• Pooling results between different PET/CT scanners requires knowledge of biases between scanners to improve the statistical power of multicenter studies.

• The sources of bias and variance have been well understood for some time.

• What is needed is a consensus approach to developing and implementing appropriate protocols.

• Several initiatives are underway for both providing the necessary data and defining principles for such protocols.
Acknowledgements

• University of Washington
  – Robert Doot, David Mankoff, Hubert Vesselle, Michelle Wanner
• SNM Validations Task Force
  – Sandy McEwan, Paul Christian, Joel Karp
• ACRIN PET Core Lab
  – Joshua Scheuermann, Janet Saffer, Barry Siegal
• Joint AAPM/SNM Task Group 145
  – John Boone, Osama Mawlawi, Janet Saffer (ACRIN), Michael McNitt-Gray, Magnus Dahlbom, Ramsey Badawi, Bob Zimmerman, Brian Zimmerman (NIST)
• NCI Cancer Imaging Program / caBIG
  – Larry Clarke, Barbara Croft, Carl Jaffe, John Freymann, Justin Kirby, Paula Jacobs
• RSNA QIBA
  – Dan Sullivan, Richard Frank
• ACR Physics Committee
  – Peter Esser
• Pediatric Brain Tumor Consortium
  – Fred Fahey
• RadQual LLC
  – Keith Allberg

Supported by the Society of Nuclear Medicine, the American Association of Physicists in Medicine, and NCI Contract 24XS036-004