PET/CT imaging for response monitoring in multicenter studies: An update and future challenges

Paul Kinahan, PhD

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

• Research Contract, GE Healthcare
Acknowledgements

- Tom Lewellen, Robert Miyaoka, Adam Alessio, Larry Macdonald, Mark Muzi, Hannah Linden, Steve Bowen, Darrin Byrd, U Washington
- David Mankoff, Robert Doot, U Penn
- Wolfgang Weber, Memorial Sloan Kettering Cancer Center
- Robert Jeraj, U Wisconsin
- Larry Clarke, NCI-CIP
- Dan Sullivan, RSNA
- Ronald Boellaard, VUMC
- Rich Wahl, Martin Lodge, Johns Hopkins
- Osama Mawlawi, Tinsu Pan, MD Anderson Cancer Center
- Support from NCI, RSNA, SNMMI, AAPM, ACRIN
PET/CT Imaging is a powerful tool for detection, diagnosis, and staging of cancer.
Clinical Applications

IMV 2008 PET Imaging Market Summary Report
Diagnostic Accuracy of PET/CT exceeds CT or PET only

<table>
<thead>
<tr>
<th>Tumor entity</th>
<th>References</th>
<th>Purpose of the imaging studies</th>
<th>Number of patients</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PET/CT</td>
</tr>
<tr>
<td>Head and neck</td>
<td>Chen <em>et al.</em> (2006)\textsuperscript{35}</td>
<td>TNM staging</td>
<td>70</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Schoder <em>et al.</em> (2004)\textsuperscript{36}</td>
<td>Lesion detection</td>
<td>68</td>
<td>96</td>
</tr>
<tr>
<td>NSCLC</td>
<td>Lardinois <em>et al.</em> (2003)\textsuperscript{24}</td>
<td>T stage\nN stage</td>
<td>40\n37</td>
<td>98\n84</td>
</tr>
<tr>
<td></td>
<td>Shim <em>et al.</em> (2005)\textsuperscript{37}</td>
<td>T stage\nN stage</td>
<td>106\n106</td>
<td>86\n84</td>
</tr>
<tr>
<td>Colorectal</td>
<td>Kim <em>et al.</em> (2005)\textsuperscript{10}</td>
<td>Recurrence</td>
<td>51</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Votrubova <em>et al.</em> (2006)\textsuperscript{38}</td>
<td>Recurrence</td>
<td>84</td>
<td>90</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Allen-Auerbach <em>et al.</em> (2004)\textsuperscript{33}</td>
<td>(Re)staging</td>
<td>73</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>la Fougère <em>et al.</em> (2006)\textsuperscript{39}</td>
<td>(Re)staging</td>
<td>50</td>
<td>99</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Reinhardt <em>et al.</em> (2006)\textsuperscript{31}</td>
<td>(Re)staging</td>
<td>250</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Mottaghy <em>et al.</em> (2007)\textsuperscript{40}</td>
<td>(Re)staging</td>
<td>102</td>
<td>91</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Statistically significant difference when compared with PET/CT.Abbreviations: NSCLC, non-small-cell lung cancer; ND, not determined; TNM, tumor node metastasis.

*Weber et al. Nature Reviews Clinical Oncology 2008*
“Is quantitation necessary for clinical oncological PET studies interpreted by physicians with experience in interpreting PET images?” - “no.”
“Is quantitation necessary for clinical oncological PET studies interpreted by physicians with experience in interpreting PET images?” - “no.”

Image quantitation will become increasingly important in determining the effect of therapy in many malignancies.  

*R Edward Coleman (EJNM 2002)*
Quantitative PET Imaging

There is a role

- Monitoring patient response or progression
- Treatment planning
- Reporting tracer uptake (for any reason)
- Developing new therapies
- New diagnostic agents
Quantitative Assessment of Response to Therapy

Breast cancer recurrence

FDG

Baseline  s/p 4 mos. letrozole

7.4  12.4  1.9  4.5

SUVs

Quantitatively distinct

F18

71.7  11.8  20.6  9.9

Qualitatively distinct

Courtesy D Mankoff
Earlier assessment of response to therapy

Earlier SUV measurement time

Size at diagnosis

SUV measurement accuracy

Improved SUV quantitation allows earlier assessment of response to therapy

Tumor cell kill rate for effective therapy after 6 cycles of chemotherapy

Log(number of cancer cells)

Cycles of chemotherapy

Cure
Drivers for Quantitative PET

**increasing volume**

- FDG uptake is now routinely reported, and are asked for, by referring physicians
- Assessing individual response to therapy
- Treatment planning (including RT)
- New molecular diagnostic agents
- Clinical trials and Drug discovery

**short term drivers**
Isn't PET imaging already accurate?

“I’m a victim of my own success. Who should I sue?”
PET Scanning Process

Patient preparation → Scan acquisition → Image reconstruction → Image analysis → Image interpretation → uptake measure
Typical PET/CT Scan Protocol

1. Scout scan (5-10 sec)

2. Selection of scan region

3. Helical CT (30 sec)

4. Whole-body PET (15-30 min)
Standardized uptake value (SUV) in PET

- Normalize by amounts injected and weight to get the same relative distribution

A hot spot has the same SUV independent of activity injected or patient size.
Sources of Error in SUV Values

SUV = Standardized Uptake Value

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

PET = measured PET activity concentration

\[D' = \text{decay-corrected injected dose}\]

\[V' = \text{surrogate for volume of distribution}\]

Some potential sources of error are:

• High blood glucose levels
• Variations in dose uptake time
• Uncalibrated clocks (including scanner) and cross calibration of scanner with dose calibrator
• Errors in radioactive dose assay
• Variations in image reconstruction and other processing protocols and parameters
• Variations in images analysis methods: E.g. how ROIs are drawn and whether max or mean SUV values are reported
• Scanner calibration
Instrumentation Chain for FDG-PET

PET scanner

dose calibrator

9.6 mCi

pre- and post injection assays

scanner units

scanner global calibration factor

kBq/ml

decay corrected net activity

SUVs

patient weight (& height)
Error Propagation in PET Imaging

Estimate
Single-center best case: 10-12%
Single-center, typical: 10-18%
Multi-center, best case: 15-20%
Multi-center, typical: 15-50%

Source data
Minn 1999, Weber 2000, etc
Velasquez 2009
Velasquez 2009
Fahey 2009, Doot 2010, Kumar 2013

Kinahan and Fletcher, Sem US, CT, MR 2010
Impact of measurement error and sensitivity to true change on sample size

Effect size (e.g. $\Delta$SUV) = 20%
Power = 80%
Significance = 0.05

<table>
<thead>
<tr>
<th>Trial Scenario</th>
<th>Error</th>
<th># of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single site</td>
<td>10%</td>
<td>12</td>
</tr>
<tr>
<td>Multi-center (good calibration)</td>
<td>20%</td>
<td>42</td>
</tr>
<tr>
<td>Multi-center (poor calibration)</td>
<td>40%</td>
<td>158</td>
</tr>
</tbody>
</table>

Doot et al., Acad Rad 2012
Quantitative Imaging Definitions

• A biomarker is an objectively measured indicator of biological/pathobiological process or pharmacologic response to treatment

• Qualified biomarker: A disease-related biomarker linked by graded evidence to biological and clinical endpoints and dependent upon the intended use

• Imaging biomarker: a number, set of numbers, or classification derived from an image (in general imaging biomarkers are not surrogate endpoints)

• Validated assay: An assay (i.e. quantitative imaging) that has documented performance characteristics showing suitability for the intended applications

Quantitative Imaging Requirements

- Prior studies that measure bias and/or variance
- Defined protocols
- Monitoring of protocols
- Calibration and QA/QC procedures to ensure variance stays within assumed range
- Optional: Techniques and procedures that improve measurement accuracy
The Imaging Chain

• Quantitative measurements have known a measurement error, e.g. SUV = x ± y

• For quantitative imaging each component of the imaging chain requires:
  – Quality Assurance (i.e. protocol saying what to do)
  – Quality Control (checking what actually happened)

• Outline of propagation of errors through main components for all imaging methods:

  imaging physics  \[\rightarrow\] scan protocol  \[\rightarrow\] processing & reconstruction  \[\rightarrow\] analysis methods  \[\rightarrow\] final accuracy & precision

  patient status  \[\downarrow\] calibration
Recent PET Technology Innovations

- Respiratory motion compensation
- Time of flight imaging
- Advanced modeling of PET physics in image reconstruction
- Extended axial field of view
- Cost effective PET/CT scanners
- New detector systems
- PET/MR scanners
- CT dose reduction methods
Clinical PET scanners are a moving target

Modified NEMA NU-2 IQ phantom

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

Different reconstruction methods on the same PET/CT scanner

Recovery coefficient

Sphere volume (mL)

Courtesy Ronald Boellaard
Challenges with Implementing Quantitative Imaging - Industry

• Significant variability between manufacturers in scan protocols and image quality

• No tests of quantitative accuracy of images transferred between display/analysis systems

• Due to several reasons:
  – Lack of standards by which vendors can assure compliance of acquisition/processing algorithms
  – Lack of convincing (to vendors) evidence of a market for quantitative imaging
Challenges with Implementing Quantitative Imaging - Imaging Sites

- There is a tension with imaging protocols suitable for current clinical practice
- Often there is no standard clinical practice
Guidance for Industry Standards for Clinical Trial Imaging Endpoints

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the Federal Register of the notice announcing the availability of the draft guidance. Submit electronic comments to http://www.regulations.gov. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the Federal Register.

For questions regarding this draft document contact (CDER) Dr. Rafel Rieves at 301-796-2050 or (CBER) Office of Communication, Outreach, and Development at 301-827-1800 or 800-835-4709.

(FDA, August 2011)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologies Evaluation and Research (CBER)

August 2011
Clinical/Medical

“… clinical trial standard[s] for image acquisition and interpretation… exceed those typically used in medical practice.”
What do we do?

There are three main routes of action

1. Accreditation authorities
2. Standards definitions and harmonization initiatives
3. Calibration methods and/or phantoms
Quantitative Imaging Initiatives

- ACRIN Centers of Quantitative Imaging Excellence (CQIE)
- RSNA Quantitative Imaging Biomarkers Alliance (QIBA)
- NCI Quantitative Imaging Network (QIN)
- AAPM Task Group 145: Quantitative Imaging for PET
- Reconstruction Harmonization Project (ACRIN / SNM-CTN / QIN / QIBA)
- EANM and EORTC initiatives
Quantitative Imaging Network (QIN)

Laurence Clarke PhD, Science Officer
Robert Nordstrom PhD: Lead Program Director
Gary Kelloff MD: Science Officer
CIP and RRP Program Staff
QIN can deliver next generation of QI methods for data collection and analysis

- Harmonization of Data Collection & Variance Studies Across Platforms
- Advanced Data Analysis & Tool Validation
- Informatics & Data Sharing Public Resources TCIA
- Resource: NCI Clinical Trial Networks
Network teams are tasked to develop a consensus on how to compare the performance of software tools for data collection and analysis.
QIN is an early user of the public archive

http://cancerimagingarchive.net
Quantitative Imaging Biomarkers Alliance (QIBA)

• Basic premise for the RSNA:
  Extracting objective, quantitative results from imaging studies will improve the value of imaging in clinical practice

• Mission:
  Improve value and practicality of quantitative imaging biomarkers by reducing variability across devices, patients, and time.
  • Build 'measuring devices' rather than imaging devices
  • 'Industrialize' imaging biomarkers
QIBA Protocols & Profiles

• **QIBA Profile**
  Describes a specific performance **Claim** and how it can be achieved
  Establishes a written standard procedure for all parties to obtain an accurate and reproducible measurement that reflects an imaging biomarker of clinical interest

• **UPICT Protocol**
  (Uniform Protocol for Imaging in Clinical Trials)
  Consensus-derived description of a process to create quantitative medical images
FDG-PET/CT Profile Claim

If Profile criteria are met, then tumor glycolytic activity as reflected by the maximum standardized uptake value (SUVmax) should be measurable from FDG-PET/CT with a within-subject coefficient of variation of 10-12%.

Profile specified for use with: patients with malignancy, for the following indicated biology: primary or metastatic, and to serve the following purpose: therapeutic response.

If Profile criteria are met, then tumor glycolytic activity as reflected by the maximum standardized uptake value (SUVmax) should be measurable from FDG-PET/CT with a within-subject coefficient of variation of 10-12%.

Profile specified for use with: patients with malignancy, for the following indicated biology: primary or metastatic, and to serve the following purpose: therapeutic response.

QIBA Profiles

Part 1: Executive Summary

Part 2: Claim: The specific statement on measurement ability

Part 3: QIBA Acquisition Protocol: Related to UPICT protocol

Part 4: Technical Compliance Specifications
QIBA Profiles

Quantitative Imaging Biomarker Claim

Data on correlations between Quantitative Imaging Biomarker and outcomes or surrogates

QIBA Profile

Part 1: Executive Summary

Part 2: Claim: The specific statement on measurement ability

Part 3: QIBA Acquisition Protocol: Related to UPICT protocol

Part 4: Technical Compliance Specifications
Other QIBA Activities

Developing metrology standards for quantitative imaging biomarkers

Five papers submitted:
- Terminology
- Technical Performance
- Algorithm Comparisons
- Meta-analysis
- Application to Pulmonary Nodule Volume
Calibration phantoms for Quantitative PET/CT Standards and/or Accreditation

- Uniform Cylinder (used by ACRIN and many others)
- ACR PET phantom
- NEMA NU-2 Image Quality (IQ) phantom
- Modified NEMA Image Quality (IQ) phantom
- SNM CTN phantom
- Cross Calibration Phantom with NIST-traceable $^{68}$Ge standard for Dose Calibrator
- Digital reference object
PET Digital Reference Object (DRO)

- The DRO is a synthetically generated set of DICOM image files of known voxel values for PET and CT
- Intended to test computation of SUVs and ROIs
- Version 1 released 10/31/2011
- More info at depts.washington.edu/petctdro
PET Digital Reference Object (DRO)

CT (transmission)

PET (emission)

coronal section

transaxial section

ROI based analysis
Results: 13 sites, 20 different display systems

blue = okay, yellow = ?, pink = borderline, red = wrong

| ROI Information | BW | BW | BW | BW | BM | BW | BW | BW | BW | BW | BW | BW | BW | BW | BW | BW | BW | BW | BW |
|-----------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| SUV Type (BW, LBM, BSA) | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 6 | 6 | 10 | B | 6 | 6 | 10 | decimal Pl.2 |
| ROI type (2D, 3D) | 1D | 2D | 2D | 2D | 2D | 1D | 2D | Circle | 2D | 2D | 20 & 3D | 2D | 2D | 20 | 2D | |
| ROI Area or Diameter? | Diameter | area | Diameter | Area | Diameter | Area | Diameter | Area | Diameter | Area | Diameter | Area | Diameter | Area | Diameter | Area | Area/Volume |
| ROI 1 | 4.0 | 4.0 | 4.0 | 4.0 | 7.8 | 4.0 | 4.0 | 0.0000095 | 4 | 4 | 4.0000019 | 4 | 4 | 4.000004 |
| ROI 1 Std | 0.999 | 0.91 | 1.05 | 0.745834 | 4 | 0.97 | 0.97 | 1.28391 | 0.917 | 0.847 | N/A | 0.93 |
| ROI 2 | 4.0 | 4.0 | 4.0 | 4.0 | 7.8 | 4.0 | 4.0 | 0.0000095 | 4 | 4 | 4.0000019 | 4 | 4 | 4.000004 |
| ROI 2 Std | 0.999 | 0.91 | 1.05 | 0.745834 | 4 | 0.97 | 0.97 | 1.28391 | 0.917 | 0.847 | N/A | 0.93 |
| ROI 3 | 4.0 | 4.0 | 4.0 | 4.0 | 7.8 | 4.0 | 4.0 | 0.0000095 | 4 | 4 | 4.0000019 | 4 | 4 | 4.000004 |
| ROI 3 Std | 0.999 | 0.91 | 1.05 | 0.745834 | 4 | 0.97 | 0.97 | 1.28391 | 0.917 | 0.847 | N/A | 0.93 |
| ROI 4 | 4.0 | 4.0 | 4.0 | 4.0 | 7.8 | 4.0 | 4.0 | 0.0000095 | 4 | 4 | 4.0000019 | 4 | 4 | 4.000004 |
| ROI 4 Std | 0.999 | 0.91 | 1.05 | 0.745834 | 4 | 0.97 | 0.97 | 1.28391 | 0.917 | 0.847 | N/A | 0.93 |
| ROI 5 | 4.0 | 4.0 | 4.0 | 4.0 | 7.8 | 4.0 | 4.0 | 0.0000095 | 4 | 4 | 4.0000019 | 4 | 4 | 4.000004 |
| ROI 5 Std | 0.999 | 0.91 | 1.05 | 0.745834 | 4 | 0.97 | 0.97 | 1.28391 | 0.917 | 0.847 | N/A | 0.93 |
| ROI 6 | 4.0 | 4.0 | 4.0 | 4.0 | 7.8 | 4.0 | 4.0 | 0.0000095 | 4 | 4 | 4.0000019 | 4 | 4 | 4.000004 |
| ROI 6 Std | 0.999 | 0.91 | 1.05 | 0.745834 | 4 | 0.97 | 0.97 | 1.28391 | 0.917 | 0.847 | N/A | 0.93 |
| ROI 7 | 4.0 | 4.0 | 4.0 | 4.0 | 7.8 | 4.0 | 4.0 | 0.0000095 | 4 | 4 | 4.0000019 | 4 | 4 | 4.000004 |
| ROI 7 Std | 0.999 | 0.91 | 1.05 | 0.745834 | 4 | 0.97 | 0.97 | 1.28391 | 0.917 | 0.847 | N/A | 0.93 |

ROI Measurements
results for each of the 6 ROIs
CONCLUSION
State of the art for FDG-PET/CT: Quantitative imaging requirements

• Test-retest studies in the literature demonstrate that quantitative image acquisition protocols are definable and possible
• To enable quantitative image acquisition protocols we need
  – Standards by which users can assure compliance, e.g. QIBA Profile
  – Methods to collectively agree on data transfer and analysis, e.g. QIN/ACRIN methods
  – Education for (and adoption by) radiologists, if they are to remain in the image processing chain
PET image reconstruction harmonization

- Harmonized and optimized reconstruction
- Range for harmonized reconstruction
- Current range for PET scanners