Measurement Issues in Imaging

Richard L. Wahl, M.D.
Division of Nuclear Medicine
Johns Hopkins Medical Institutes
Baltimore, MD

Why Include Imaging in a Clinical Trial?

Primary evaluation of the imaging test for:
- Diagnostic accuracy
- Prognostic accuracy
- Predictive accuracy
- As a biomarker

Evaluation as an endpoint for a therapy:
- for example, response or non response to a cancer therapy

Amyloid burden in the brain for a trial of an AD drug

Secondary correlative marker, Eg. Survival a primary endpoint, tumor shrinkage a secondary endpoint.

What are our Imaging Tools?

- X-ray based imaging: Chest X ray, Bone radiograph, CT scan, angiogram
- Sound based—Ultrasound (including Doppler US)
- Nuclear Medicine—Gamma camera imaging, PET, SPECT imaging
- Magnetic Resonance Imaging— including varying pulse sequences, DCE MRI,
- Optical Imaging: Human eye, bioluminescence, NIR (best in animals)

Why Include Imaging in a Trial?

- Varying pharmacokinetics in whole body, organs and tumors
- Varying receptor status in tumors
- Varying proliferative/apoptotic rates in tumors
- Varying in many characteristics
- Current cancer treatments are substantially based on the “Average Patient” not the individual patient.
- Goal is to individualize treatments to optimize responses and minimize toxicity.
- Imaging shows the phenotype in the physiological milieu.

X-Ray Based Imaging

- Strengths:
  - Excellent anatomic resolution
  - Quick and widely available
  - Provides size measurement used in RECIST

- Weakness:
  - Ionizing radiation
  - Information content somewhat limited
  - Tumor size changes are “last thing” to occur with successful or unsuccessful rx.
  - Lack of specific contrast agents
FDG-PET/CT imaging (FOV 25 cm)

Ultrasound Based Imaging
- Strengths
  - No ionizing radiation
  - Reasonably good resolution
  - Relatively Inexpensive
  - Widely available
- Weaknesses
  - Limited depth of penetration
  - Limited quantitative information
  - Lack of specific contrast

Nuclear Medicine PET and Single Photon Imaging
- Strengths
  - Low mass quantities required (tracer)
  - Quantitative
  - Can evaluate deep structures
  - Wide range of tracers available (or not)
  - Can combine with CT
  - Reproducibility understood
- Weaknesses
  - Limited resolution
  - Ionizing radiation
  - Radiotracers may not be widely available
  - Cost
  - Short half lives make logistically challenging

Magnetic Resonance Imaging
- Strengths
  - Relatively widely available
  - Resolution is good and can be excellent
  - Many pulse sequences available
  - Can provide physiological information
  - Some contrast agents emerging (nodal/vascular)
- Weaknesses
  - Predominantly anatomic in many applications
  - Pulse sequences are not well-standardized across manufacturers or institutions
  - Reproducibility is not yet well characterized
  - Cost
  - NSF (severe complication in patients with poor renal function and repeated Gadolinium dosing)
Breast Applications

Baseline

1st cycle (7 days)

4th cycle (5 days)

Before surgery

Optical Imaging

• Strengths
  - Very sensitive to small numbers of molecules
  - Semi quantitative (attenuation effects)
  - Some use in OR and superficial organ assessments
  - Excellent pre-clinical tool

• Weaknesses
  - Not easily scalable to humans, at least for deep structures due to light absorption
  - Limited specific contrast agents (but they can be made)

Results

d0d4
d8
d11
d18
t

Treated (35 uCi/mouse) Untreated

Is imaging tool “Fit for the Task?”

• Must understand question one wishes to address
• Must understand limitations and strengths of imaging tool
• Consultation with a subject expert is recommended if imaging has a major role in a clinical trial.

Qualitative Analysis vs Quantitative

• Variability in the Interpretation of Screening
• Mammograms by US Radiologists
• Findings From a National Sample
• Craig A. Beam, PhD; Peter M. Layde, MD, MSc; Daniel C. Sullivan, MD
• Arch Int. Med 1996; 156:209-213

Qualitative reads are Qualitative

• Not all readers are created equally
• Performance can and does vary among readers
• For test/re test studies, it seems apparent that the same reader should be used for both the baseline and the follow up
Results of Multi Reader Study

• There is a range of at least 40% among US radiologists in their screening sensitivity.
• There is a range of at least 45% in the rates at which women without breast cancer are recommended for biopsy.
• As indicated by receiver operating characteristic curve areas, the ability of radiologists to detect cancer mammograms varies by as much as 11%.

Biases in Imaging

• Reader experience
• Variation in equipment or methods in a center or across centers
• Reader blinding to truth data
• Stage migration
• Verification bias
• Dichotomous of continuous recording of results?
Structure and Affinity of the Neutral Thioflavin-T Analogue [11C]PIB

\[
\begin{align*}
\text{thioflavin-T} \\
\text{[11C]6-OH-BTA-1 or [11C]PIB}
\end{align*}
\]

\[A(\beta(1-40)) K_i = 580 \text{ nM} \quad A(\beta(1-40)) K_d = 2 \text{ nM}\]

Impact of PET/CT in Cardiac Disease

- Single photon imaging is established throughout the world as a tool to individualize patient treatments
- While very effective, is less accurate in RCA distribution and in larger patients such as in the US
- Limitations for viability
- Limitations for “balanced” disease
- PET/CT and PET are more accurate in many studies. When should they be used? All/some/No patients?

Cardiac Imaging: Biological Considerations

- Myocardium:
  - Flow
  - Viability
  - Innervation
  - Necrosis/Apoptosis
- Blood vessels: plaque, inflammation

Ru-82 PET Stress/Rest Polar Map
Dynamic Rb-82 PET-CT for Quantification of Myocardial Blood Flow

Cardiac PET/CT
- Obtain Myocardial Perfusion images (PET)
- Obtain Structural and anatomical abnormalities within the field of view (CT)
- Obtain Calcium score (CT)
- Obtain Coronary structure and detection of lesions (CTA)
- Obtain area(s) of viable tissue (PET)

Efficacy of Current Cancer Rx
- 100% of patients get Rx.
- 10-15% in metastatic lung, esophagus, pancreatic cancers
- Paradigm of treating 10 to benefit 1 is not economically sustainable
- Need PERSONALIZED approaches to enrich response rates
- Benefit to those who will benefit most and avoid harm to those who will not

PET/CT and SPECT/CT are Quantitative Methods
- Most applications to date have been Qualitative
- In treatment response assessment, especially if looking for small treatment induced changes, Quantitation will be needed
- Quantitation requires greater attention to technical details than qualitative imaging
- Standardization of methods is required

Potential Roles of Advanced Imaging in Drug Development
- Pre-Clinical
  - In vitro
  - Animal Models
- Clinical
  - Phase I
  - Phase II
  - Phase III

Imaging What?
- Drug Itself?
- Surrogate for the drug? (eg. In-111 for Y90)
- Tumor characteristics best suited to drug
  - Receptor or hypoxia, proliferation rate, blood flow
- Tumor response to the drug
  - FDG, FLT, size, vascular permeability or flow
- Size:
Roles of PET/CT and SPECT/CT in Cancer Management

- Prior to treatment
  - Diagnosis (cancer or not?) — including directing biopsy
  - Staging before therapy
    - Pre surgical
    - Pre Radiation therapy (planning)
  - Pre Radionuclide therapy (dosimetry)
    - Predicting Response*
- After treatment initiated
  - Assessment of response after completing therapy
  - Intra-therapy monitoring
  - Response Adaptive Approaches

Imaging What?

- Drug Itself?
- Surrogate for the drug? (e.g., In-111 for Y90)
- Tumor characteristics best suited to drug
  - Receptor or hypoxia, proliferation rate, blood flow
- Tumor response to the drug
  - FDG, FLT, size, vascular permeability or flow
- Size:

I-131 Tositumomab Treatment Regimen

<table>
<thead>
<tr>
<th>Day 0</th>
<th>Day 7-14</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dosimetric Dose</strong></td>
<td><strong>Therapeutic Dose</strong></td>
</tr>
<tr>
<td>450 mg unlabeled tositumomab</td>
<td>450 mg unlabeled tositumomab</td>
</tr>
<tr>
<td>35 mg tositumomab radiolabeled I 131 (5 mCi)</td>
<td>35 mg tositumomab radiolabeled I 131 to deliver specific cGy TBD (variable mCi)</td>
</tr>
<tr>
<td>• Unlabeled dose infused over 1 hour</td>
<td>• Unlabeled dose infused over 1 hour</td>
</tr>
<tr>
<td>• Radiolabeled tracer dose infused over 20 minutes</td>
<td>• Radiolabeled therapeutic dose infused over 20 minutes</td>
</tr>
</tbody>
</table>

Graphic Estimate of Total Body Residence Time

Effect of Clearance Rate on Radiation Exposure (mCi)

- Individuals with a rapid clearance rate require a higher dose of radiation (in mCi)
  - Rapid Clearance
  - Days: 1, 2, 3, 4
  - 75 cGy
- Individuals with a slow clearance rate require a lower dose of radiation (in mCi)
  - Slow Clearance
  - Days: 2, 3, 4
  - 75 cGy
Imaging What?

- Drug Itself?
- Surrogate for the drug? (eg. In-111 for Y90)
- Tumor characteristics best suited to drug
  - Receptor or hypoxia, proliferation rate, blood flow
- Tumor response to the drug
  - Size: RECIST 1.0, 1.1; WHO, Volume
  - FDG, FLT, size, vascular permeability or flow

Organ Dosimetry Dose Escalation Study of Y90 Ibritumomab Tiuxetan and ASCT: Study Schema

Roles of PET/CT in Cancer Management

- Prior to treatment
  - Diagnosis (cancer or not?)—including directing biopsy
  - Staging before therapy
    - Pre surgical
    - Pre Radiation therapy (planning)
    - Pre Radionuclide therapy (dosimetry)
  - Predicting Response*
- After treatment initiated
  - Assessment of response after completing therapy
  - Intra-therapy monitoring
  - Response Adaptive Approaches
**Imaging What?**

- Drug Itself?
- Surrogate for the drug? (eg. In-111 for Y90)
- Tumor characteristics best suited to drug
  - Receptor or hypoxia, proliferation rate, blood flow
- Tumor response to the drug
  - Size: RECIST 1.0, 1.1; WHO, Volume
  - FDG, FLT, size, vascular permeability or flow

**TABLE 6: Comparison of WHO Response Criteria and RECIST (6th ed).*

<table>
<thead>
<tr>
<th>Observation</th>
<th>WHO</th>
<th>RECIST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement</td>
<td>1. Measurable, dimensioned ≥ 20 mm</td>
<td>1. Measurable, dimensioned ≥ 20 mm, with equal CO = 15 mm.</td>
</tr>
<tr>
<td></td>
<td>2. Nonmeasurable, all other lesions, excluding small lesions, evaluable is not recommended</td>
<td>2. Nonmeasurable: all other lesions, excluding small lesions, evaluable is not recommended</td>
</tr>
<tr>
<td>Objective response</td>
<td>1. Measurable disease change ≥ 50% decrease in sum of products of all tumors</td>
<td>1. Target lesion change ≥ 30% decrease in sum of products of all target lesions</td>
</tr>
<tr>
<td></td>
<td>2. Target lesion change &lt; 50% decrease in sum of products of all target lesions</td>
<td>2. Target lesion change &lt; 30% decrease in sum of products of all target lesions</td>
</tr>
</tbody>
</table>

**Anatomic Response II**

- Slow
- Not sensitive to early response in some tumor types (GIST, Sarcoma)
- Variability in measurement
- Tumor Size a late downstream marker
- Difficult to detect some tumors
- Hard to separate tumor from treatment response or other pathology
Molecular and Functional Alterations in Cancer I.

- Increased glucose metabolism
- Increased Amino acid transport
- Increased Protein and membrane synthesis
- Increased DNA synthesis
- Overexpression of receptors/antigens
- Increased blood flow or vessel density
- Decreased oxygen tension in lesions

FDG: A Tracer of Early Steps of Glucose Metabolism

FDG \xrightarrow{K1 \text{ Extracellular}} \xrightarrow{K2 \text{ Intracellular}} \xrightarrow{K3 \text{ Intracellular}} \xrightarrow{K4 \text{ Intracellular}} \text{FDG-6-P}

FDG Signal

- “Heartbeat of the Cell”
- Downstream Marker of Many Processes

Why do we need PET in Treatment Response Assessment?

- PET can detect tumors not detectable by CT
- PET responses often occur earlier than those seen on CT
- PET has higher predictive ability at end of treatment than CT
- PET may be more reproducible than CT for measuring tumor response to therapy
Post-therapy PET

- Multiple studies demonstrating superior prognostic value of PET vs CT at therapy completion
  - Can differentiate viable tumor from fibrosis
  - Can detect residual disease in context of radiographic CR
- Relapse up to 100% if PET+, 5-17% if PET-

Jerusalem et al, Blood 1999; 94:429-433
Spaepen K et al, JCO 2001; 19: 414-419

Bone Marrow Involvement
IWG response criteria

<table>
<thead>
<tr>
<th>Original</th>
<th>Revised (includes Hodgkin's)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR</td>
<td>Normal size</td>
</tr>
<tr>
<td></td>
<td>FDG-avid tumor: mass of any size, as long as PET &quot;negative&quot;</td>
</tr>
<tr>
<td></td>
<td>Variably FDG-avid or unknown: normal size</td>
</tr>
<tr>
<td>CRu</td>
<td>If nodes &gt; 1.5 cm, &gt; 75% decrease</td>
</tr>
<tr>
<td>PR</td>
<td>≥ 50% decrease</td>
</tr>
<tr>
<td></td>
<td>FDG-avid tumor: ≥ 50% decrease, but at least one PET positive focus</td>
</tr>
<tr>
<td></td>
<td>Variably FDG-avid or unknown: regression in size</td>
</tr>
</tbody>
</table>


“International Harmonization Project (IHP) Recommendations for FDG-PET in Patients with Lymphoma (abst 24, EANM, Juweid et al)

- Recommendations for PET, not PET-CT
- PET mandatory for post Rx HD and DLBCL (pre Rx PET recommended)
- PET pre Rx strongly recommended for variably FDG avid NHL and post Rx (if clinical trial)
- “Mid Treatment” PET only should be done in Clinical Trial
- “No justification” for routine post-Rx surveillance with PET.

Roles of PET/CT and SPECT/CT in Cancer Management

- Prior to treatment
  - Diagnosis (cancer or not?)—including directing biopsy
  - Staging before therapy
    - Pre surgical
    - Pre Radiation therapy (planning)
  - Pre Radiomimetic therapy (dosimetry)
  - Predicting Response
- After treatment initiated
  - Assessment of response after completing therapy
  - Intra-therapy monitoring
  - Response Adaptive Approaches
When is Response Assessed?

- End of a Treatment Course: perhaps most relevant to have a complete response at this time
- Mid treatment course
- Shortly after treatment is initiated

Why might mid-treatment PET be superior to post-treatment?

Early PET result implies a certain rate of tumor kill.
Practical Considerations:

- Consistent Protocol Essential
- Same time from injection to imaging
- Same fasting
- Same Imaging sequence (caudal to cranial)
- Variability in estimates of SUV of 10-20%
- Changes in SUV under 20% may be chance
- Normalization to liver can help

Mid-treatment PET strongly predicts PFS in aggressive lymphoma

Roles of PET/CT and SPECT/CT in Cancer Management

- Prior to treatment
  - Diagnosis (cancer or not?)—including directing biopsy
  - Staging before therapy
    - Pre surgical
    - Pre Radiation therapy (planning)
  - Pre Radionuclide therapy (dosimetry)
  - Predicting Response
- After treatment initiated
  - Assessment of response after completing therapy
  - Intra-therapy monitoring
  - Response Adaptive Therapy

J0348: rationale

- Early \([^{18}F] FDG-PET\) (after 2-4 cycles) is highly predictive and more individualized than the IPI
  - Event rates of 78-100% if PET+ midtreatment, versus 8-16% if PET-, consistently seen
- Benefit of early BMT most pronounced in poor-risk pts
- HYPOTHESIS: Early PET can identify those pts who most stand to benefit from early treatment intensification
- PRIMARY ENDPOINT: 25% absolute improvement in 2-yr EFS (from 20% to 45%) of PET+ pts, compared with historical outcomes, through early BMT
**Study design**

Aggressive NHL, any stage, any IPI

- (R)CHOP for 2 or 3 cycles
- PET - complete conventional therapy
- PET + if no disease progression
- (R)ESHAP or (R)ICE x 2
- High dose therapy and ABMT

---

**J0348**

- **Primary endpoint:** 25% absolute improvement in 2-yr EFS (from 20% to 45%) of PET+ pts, compared with historical outcomes, through early BMT
- **Goal:** 55 pts, expect 50% PET + and 19 transplants

---

**PET assessment**

**NEGATIVE**
- 0 no abnormal activity (tumor cold compared with background)
- 1+ minimal activity (tumor less than background)
- 2+ equivocal (tumor = background)

**POSITIVE**
- 3+ moderate activity (tumor greater than background)
- 4+ strong activity (tumor much greater than background)

---

**Response Adaptive Therapy**

- Fifty-nine newly diagnosed patients, 98% with B cell lymphoma, had PET/CT performed after 2 or 3 cycles of first-line chemotherapy.
- Mid-treatment PET was positive in 33 (56%);
- 28 received ASCT with an actuarial 2-year EFS of 75% (65% confidence interval, 60%-93%).
- On intention-to-treat analysis, 2-year EFS was 67% (53%-86%) in all PET-positive patients and 85% (77%-100%) in PET-negative patients (overlapping CI).

---

**E3404: Phase II Study of Response-Adapted Therapy**

- Newly diagnosed large B-cell lymphoma
  - (Stage III, IV, or Bulky II)
- PET + R-ICE x 4
- PET during C3
- PET - R-CHOP x 2

**Central review of mid-treatment PET, designated + or -**

---

**Conclusions:**

- PET/CT after initial therapy is complete has prognostic value in lymphomas > CT alone.
- PET/CT after 1-2 cycles appears predictive of effectiveness of treatment in several tumor types. Will allow for response adaptive therapy.
- PET/CT with non FDG ligands, eg. FES, may be predictive before treatment is begun.
FLT Lung Cancer Response

- FLT-PET images of the thorax were obtained before and 7 days after the start of gefitinib (250 mg/d) therapy in non smokers with new or recurrent NSCLC.
- SUV max and % decline in SUV max were assessed in tumor at baseline and 7 days. Compared with CT after 6 weeks of therapy.
- 31 patients of whom 28 had complete data.
- CT at 6 weeks showed PR in 14 (50%), SD in 4 (14%), and PD in 10 (36%) after 6 weeks of treatment.
- Pretreatment SUV max of the tumors did not differ between responders and nonresponders.


...Figure 2...

...Figure 4...

Qualitative vs Quantitative or Both?

- Visual assessments can be used to assess response
- Strengths: No special instrumentation, integration of all data.
- Weaknesses: Tendency to be binary, not reproducible in several series
- Quantitative: Ignores qualitative data, can be erroneous due to technical factors, "Standardized" is not so standard after all.

Practical Considerations:

- Consistent Protocol Essential
- Same time from injection to imaging
- Same fasting
- Same Imaging sequence (caudal to cranial)
- Variability in estimates of SUV of 10-20%
- Changes in SUV under 20% may be chance

EORTC 1999

EORTC Response Criteria

<table>
<thead>
<tr>
<th>CR</th>
<th>Complete Disappearance of all Metabolically Active Tumor (i.e. decreased to background levels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR</td>
<td>&gt;15% decline in SUV after 1 cycle, &gt;25% decline after 2 or more cycles. Reduction in extent (size) of FDG uptake is not required</td>
</tr>
<tr>
<td>SD</td>
<td>Increase in FDG SUV of &lt;25% or decrease of &lt;15% in SUV and no increase in extent of uptake (&lt;20% in longest dimension)</td>
</tr>
<tr>
<td>PMD</td>
<td>Increase in SUV of over 25%, increase in extent of FDG uptake by &gt;20%, New FDG positive metastases</td>
</tr>
</tbody>
</table>

If there were RECIST Criteria for PET, It Would be Defined as...

“PERCIST”

- Positron
- Emission
- Response
- Criteria in
- Solid
- Tumors
From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumors

What is Measured?

- **SUL Peak**
- **In Hottest Tumor Focus**

- **PERCIST** (4): Measurable target lesion is hottest single tumor lesion; SUL of "maximal 1.2 cm diameter volume ROI in tumor" (SUL peak), SUL peak is at least 1.5 fold greater than (w) SUL mean + 2SD (in 3 cm spherical ROI in normal right lobe of liver).

- If liver is abnormal, primary tumor should have uptake >2.0 times SUL mean of blood pool in a 1 cm diameter ROI in descending thoracic aorta extended over 2 cm Z axis.

- The tumor with the maximum SUL peak is measured post-treatment. While typically this is in the same region or the tumor with the highest SUL peak at baseline, it need not be.

- Uptake measurements should be made for the peak and maximum single voxel tumor SUL.

- Other SUV metrics including SUL mean at 50 or 70% of SUV peak can be collected as exploratory data. TLG can be collected similarly based on voxels more intense than 2SD above liver mean SUL (see below). These parameters can be for 1 measurable target lesion.
Introduction

- A number of different ROI definitions have been employed including:
  - Mean within an irregular ROI defined by isocontours.
  - Mean within a fixed size ROI centered on the most metabolically active region.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{SUV}_{\text{mean}}$</td>
<td>SUV for activity in largest diameter of tumor and 2 adjacent slices, representing largest cross-section of tumor</td>
</tr>
<tr>
<td>$\text{SUV}_{\text{peak}}$</td>
<td>SUV of 1-cm ROI (0.75–1.25 cm) placed in region of highest $^{18}$F-FDG uptake</td>
</tr>
<tr>
<td>$\text{SUV}_{\text{max}}$</td>
<td>SUV for single pixel with highest activity in tumor</td>
</tr>
<tr>
<td>$\text{SUV}_{\text{70%}}$</td>
<td>SUV generated using 70% threshold of maximum tumor SUV and isocontour adapted for local background</td>
</tr>
</tbody>
</table>

Results: Statistical Quality

- Statistical quality of the images deteriorates with decreasing scan duration for both 2D and 3D.

Results: $\text{ROI}_{\text{max}}$

- $\text{ROI}_{\text{max}}$ increasingly overestimates recovery as noise increases (shorter scan durations).

Results: $\text{ROI}$

- Insert has an SUV of 2.5 (2.5:1 insert-to-background ratio).
Normalization for Quality Control

- Normal liver SUL must be within 20% (and <0.3 SUL mean units) for baseline and follow-up study to be assessable. If liver is abnormal, blood pool SUL must be within 20% (and <0.3 SUL mean units) for baseline and follow-up study to be assessable. 
- Uptake time of baseline study and follow-up study 2 must be within 15 minutes of one another to be assessable. Typically, these are at a mean of 60 minutes post injection, but not less than 50 min post injection.
- Same scanner, or same scanner model at same site, injected dose, acquisition protocol (2D vs. 3D), and software for reconstruction should be used. Scanners should provide reproducible data and be properly calibrated.

Complete Metabolic Response

- Complete metabolic response (CMR) complete resolution of [18F]-FDG uptake within the measurable target lesion so that it is less than mean liver activity and indistinguishable from surrounding background blood pool levels.
- Disappearance of all other lesions to background blood pool levels. % decline in SUL should be recorded from measurable region as well as (ideally) time in weeks after treatment was begun (i.e. CMR -90, 4).
- No new FDG avid lesions in a pattern typical of cancer. If progression by RECIST must verify with follow up.

Partial metabolic response (PMR)

- Reduction of a minimum of 30% in target measurable tumor FDG SUL peak.
- Absolute drop in SUL must be at least 0.8 SUL units, as well.
- Measurement is commonly in the same lesion as the baseline, but can be another lesion if that lesion was previously present and is most active lesion after treatment. ROI does not have to be in precisely the same area as baseline scan, though typically it is.
- No increase >30% in SUL or size of target or non-target lesions (i.e. no PD by RECIST or IWC) (if PD anatomically, must verify with follow up). A reduction in the extent of the tumor FDG uptake is not a requirement for partial metabolic response. % decline in SUL should be recorded as well as (ideally) time in weeks after treatment was begun (i.e. PMR -40, 3). No new lesions.
Stable Metabolic Disease

- Stable metabolic disease (SMD) Not CMR, PMR nor PMD. Note, the SUL peak in metabolic target lesion should be recorded as well as (ideally) time from start of most recent therapy in weeks (i.e. SMD - 15,7). No new lesions

Progressive metabolic disease (PMD)

- >30% increase in FDG SUL peak, with >0.8 SUL unit increase in tumor SUV peak from the baseline scan in pattern typical of tumor and not of infection/treatment effect.
- OR - Visible increase in the extent of [18F]-FDG tumor uptake (75% in TLG volume with no decline in SUL).
- OR - new [18F]-FDG avid lesions which are typical of cancer and not related to treatment effect or infection. PMD other than new visceral lesions should be confirmed on follow up study within 1 month unless PMD also is clearly associated with progressive disease by RECIST 1.1. PMD should be reported to include % change in SUV peak, (ideally time post treatment in weeks) and whether new lesions are present/absent and their number (i.e PMD, +35, 4, New-5).

PERCIST Continuous Response Scale

- Because SUL is a continuous variable, dividing response criteria into a limited number of somewhat arbitrary response categories loses much data.
- For this reason PERCIST preserves percent declines in the SUV peak in each reported category. Because the rapidity with which a scan normalizes is important (faster appears better), PERCIST asks for the time from start of treatment as part of the reporting.
- For example, a CMR 90, 1 is probably superior to a CMR 90, 10, especially if the latter patient were SMD 20,1. More than one measurement of PET response may be needed at differing times and it may be treatment type dependent.

Number of Lesions

- PERCIST 1.0 only evaluates the SUL peak of the hottest tumor. This is a possible limitation of the approach, but lesions and their responses are highly correlated in general.
- Additional data are required to determine how many lesions should be assessed over 1.
- A suggested option is to include the 5 hottest lesions, or the 5 observed on RECIST 1.1 which are the most measurable. % change in SUL can be reported for the single lesion with the largest increase in uptake or the smallest decline in uptake. Additional studies will be needed to define how many lesions are optimal for assessment.
Inter-Observer Variability of SUV

Same tumor data set measured multiple times by independent observers
100% agreement in SUV determination
- Minn et al, Radiology, 1995
- 10 tumors each measured twice by 2 independent observers
- Semi-automated image analysis software

“Good” inter-observer agreement
- Marom et al J Thorac Imaging 2006
- 5 readers measured 20 primary tumors four times

Untreated primary lung cancers on average > 2 cm

Assessment of Inter-observer Reproducibility in Quantitative FDG PET and CT Measurements of Tumor Response to Therapy

HA Jacene, S Leboulleux, S Baba, D Chatzifotiadis, B Goudarzi, O Teytelbaum, Horton, T Kamel, K Macura, H Tsai, J Kowalski and RL Wahl

Objectives

To directly compare inter-observer reproducibility of
1) SUV & CT size measurements in malignant tumors pre- and post-therapy
2) % change in SUV & CT size measurements in response to therapy

Percent change SUV<sub>bw max</sub>

ICC – 0.94

Percent change Longest CT size

ICC – 0.70

Percent change 2D CT size

ICC – 0.33
Dynamic contrast enhanced MRI (DCE MRI) for Phase I anti-angiogenic trial:
Comparison of the transfer constant (Ktrans) to blood flow and permeability derived by a distributed parameter model.

C. H. Thng¹, G. I. Hartono², T. C. Koh³, H. Rumpel⁴, A. B. Ong⁵, N. Sukrit⁶, B. C. Tai⁷, R. A. Soo⁶, R. Humerickhouse⁶, B. C. Goh⁶

National Cancer Centre, Singapore
Nanyang Technological University, Singapore
Singapore General Hospital
National University Hospital, Singapore
¹Yang Le Li Cancer Institute, National University Hospital, Singapore
²Medlab Laboratories

Background

- DCE MRI is commonly used for assessment of angiogenesis

Tracer pharmacokinetic modeling

**Generalized Kinetic Model (Ktrans, ve)**

\[
\frac{dC}{dt} = K^\text{trans} \left( C - C_0 \right)
\]

**Distributed Parameters Model (F, PS, v1, v2)**

\[
\frac{dC}{dt} = F^\text{trans} \left( C - C_0 \right)
\]

Modeling equations derive Model IRFs

Glossary

<table>
<thead>
<tr>
<th>GK model</th>
<th>DP model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ktrans</td>
<td>F (Flow)</td>
</tr>
<tr>
<td>Ktrans</td>
<td>PS (Permeability-surface area product)</td>
</tr>
<tr>
<td>ve</td>
<td>v2 (% Intravascular volume)</td>
</tr>
</tbody>
</table>


Results

- **Prediction of late progression**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>% drop at D19</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ve</td>
<td>-4.6%</td>
<td>88.2%</td>
<td>87.5%</td>
</tr>
<tr>
<td>PS</td>
<td>-27.4%</td>
<td>64.7%</td>
<td>87.5%</td>
</tr>
<tr>
<td>v2</td>
<td>-17.7%</td>
<td>64.7%</td>
<td>87.5%</td>
</tr>
<tr>
<td>MAUC-N</td>
<td>-63.2%</td>
<td>58.8%</td>
<td>87.5%</td>
</tr>
<tr>
<td>Ktrans</td>
<td>-27.4%</td>
<td>35.5%</td>
<td>50%</td>
</tr>
</tbody>
</table>

- **Prediction of early progression**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>% Inc</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ve</td>
<td>+37.5%</td>
<td>12.5%</td>
<td>94.1%</td>
</tr>
<tr>
<td>PS</td>
<td>+15.8%</td>
<td>37.3%</td>
<td>94.1%</td>
</tr>
</tbody>
</table>
Measurement Issues: Summary

- Rigor of imaging approach depends on if imaging is a primary or secondary/tertiary aim
- Rigorous use of imaging requires attention to details including reader selection, methods of acquisition and methods of analysis
- Clinical reads, especially qualitative reads, are highly variant depending on skill sets of readers
- Major national push to quantitative imaging biomarkers applicable across disease types
- Quantitative imaging should add rigor to clinical studies

Acknowledgements

- Heather Jacene, M.D.
- Eric Frey, Ph.D.
- Bin He, Ph.D.
- George Sgouros, Ph.D.
- Martin Lodge, Ph.D.
- Wayne Kasecamp, CNMT
- Ken Zasadny, Ph.D.
- James Engles, M.S., MBA.
- Ian Flinn, M.D.