Quantitative Imaging In Clinical Trials Using PET/CT: Update

Paul Kinahan, Robert Doot

Imaging Research Laboratory

Department of Radiology

University of Washington, Seattle, WA

Supported by

RSNA Quantitative Imaging Biomarkers Alliance

American Association of Physicists in Medicine

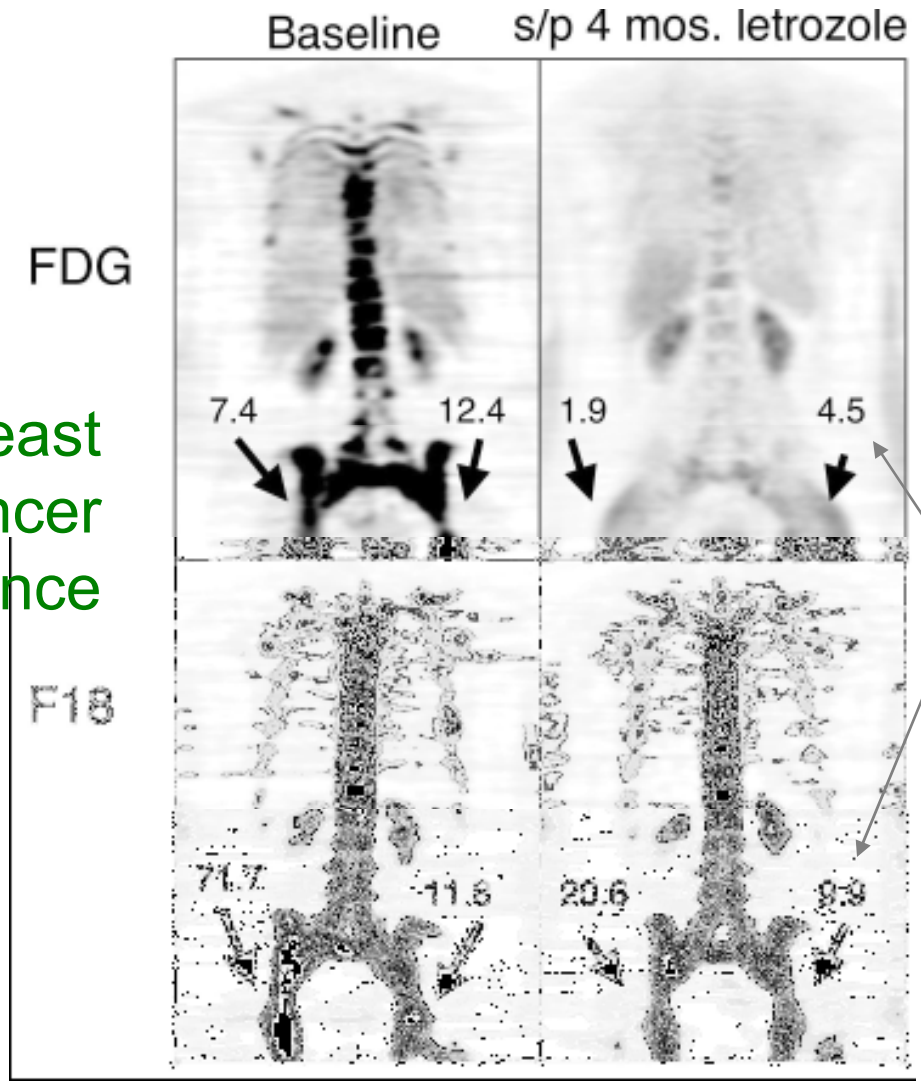
Society of Nuclear Medicine

NCI Cancer Imaging Program contract 24XS036-004 (RIDER)

Seattle Cancer Care Alliance Network

NIH Grant U01-CA148131 (QIN)

Quantitative Assessment of Response to Therapy



Breast cancer recurrence

Qualitatively distinct

SUVs

Quantitatively distinct

Courtesy D Mankoff

Errors in Numbers in PET/CT

- Technical factors
  - Relative calibration between PET scanner and dose calibrator (10%)
  - Time-varying scanner calibration (5%)
  - Residual activity in syringe (5%)
  - Incorrect synchronization of clocks (10%)
  - Injection vs calibration time (10%)
  - Quality of administration (50%)
- Physical factors
  - Scan acquisition parameters (15%)
  - Image reconstruction parameters (30%)
  - Use of contrast agents (10%)
  - ROI (50%)
- Biologic factors
  - Uptake period (15%)
  - Patient motion and breathing (30%)
  - Blood glucose levels (15%)
  - Other factors (5%)

Modified from R. Boellaard and R. Jeraj

PET/CT Quantitation Initiatives

- European Organization for Research and Treatment of Cancer (EORTC)
- American College of Radiology Imaging Network (ACRIN) PET Core Lab
- NIH/NCI
  - Imaging Response Assessment Teams (IRATs)
  - Reference Image Database for Evaluation of Response (RIDER)
- American Association of Physicists in Medicine (AAPM)
  - Quantitative Imaging Initiative Task Group 145 (joint with SNM) PET/CT
- Radiological Society of North America (RSNA):
  - Quantitative Imaging Biomarkers Alliance (QIBA)
- European Association for Nuclear Medicine (EANM)
- Cancer and Leukemia Group B (CALGB) PET Core lab
- Society of Nuclear Medicine (SNM)
  - Validation Task Force
  - Clinical Trials Network

Quantitative PET/CT Accreditation Bodies

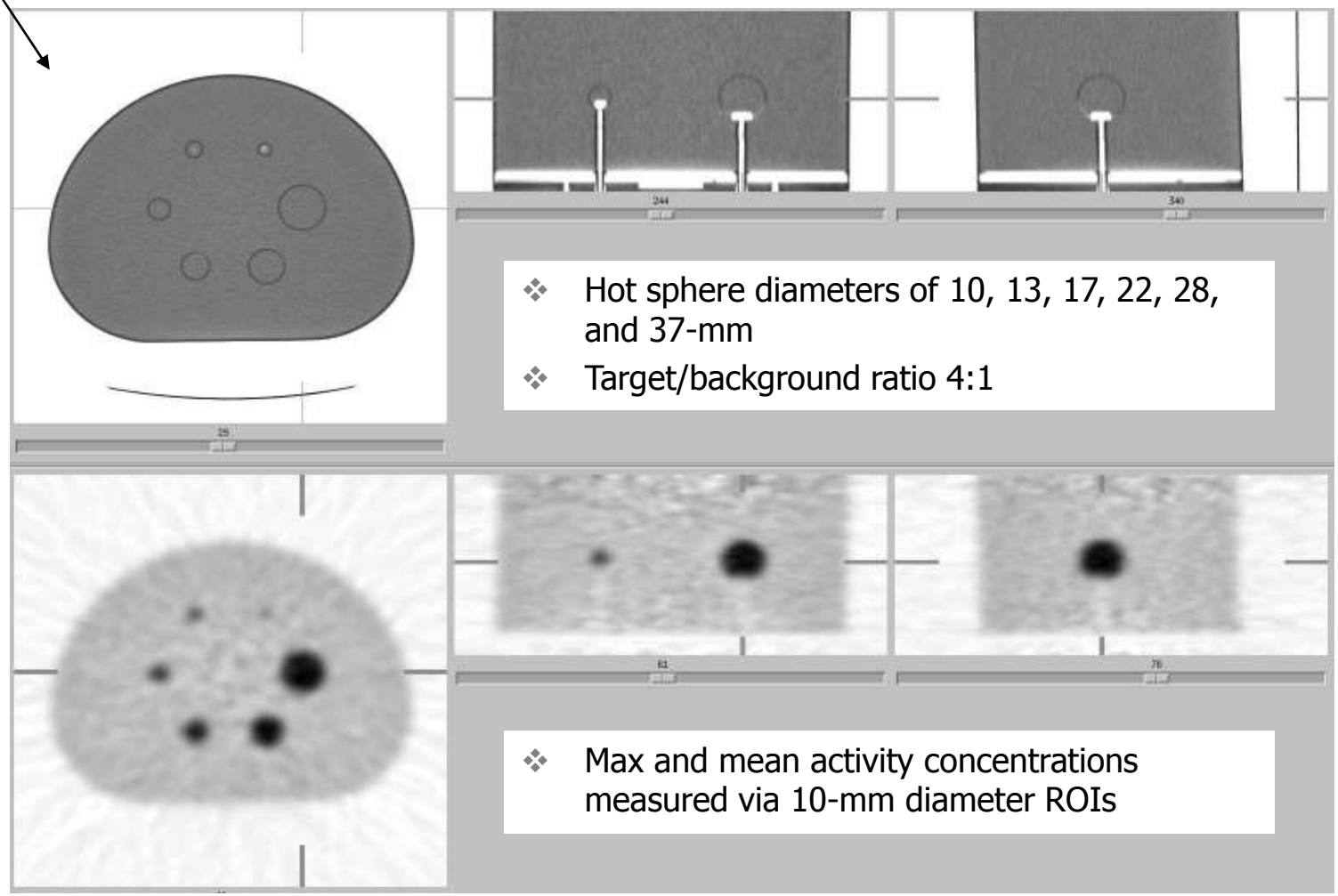
- Medical Imaging Technology Assessment (MITA)
- several Clinical Research Organizations (CROs)
- American College of Radiology (ACR) - 1000 sites!
- PET Core Labs (ACRIN, CALGB, …)
- EANM
- Cancer UK
- SNM
  - Clinical Trials Network

Calibration Phantoms / Methods

- Main Phantoms
  - Uniform cylinder (should get SUV = 1.0 ± 0.1)
  - MITA (NEMA) NU-2 Image Quality
  - ACR PET accreditation
  - AAPM/RIDER modified ACR
  - SNM Clinical Trials Network
  - Cross-calibration kit using NIST ^{68}Ge standard (v2 of modified ACR)
- Dose Calibrator
  - NIST ^{68}Ge standard for ^{18}F dose calibration

Modified NEMA NU-2 IQ Phantom

similar to abdominal x-section



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

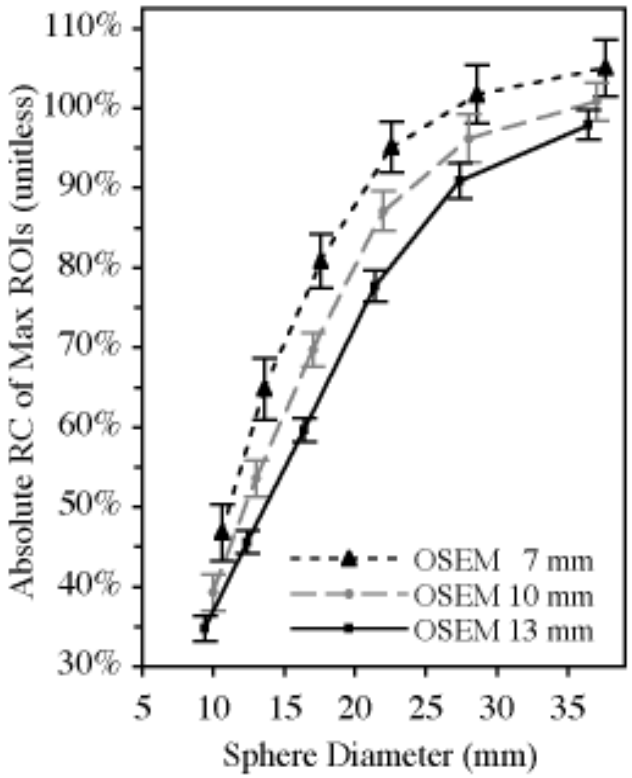
- Max and mean activity concentrations measured via 10-mm diameter ROIs

Used ^{68}Ge in epoxy to remove filling variations at sites

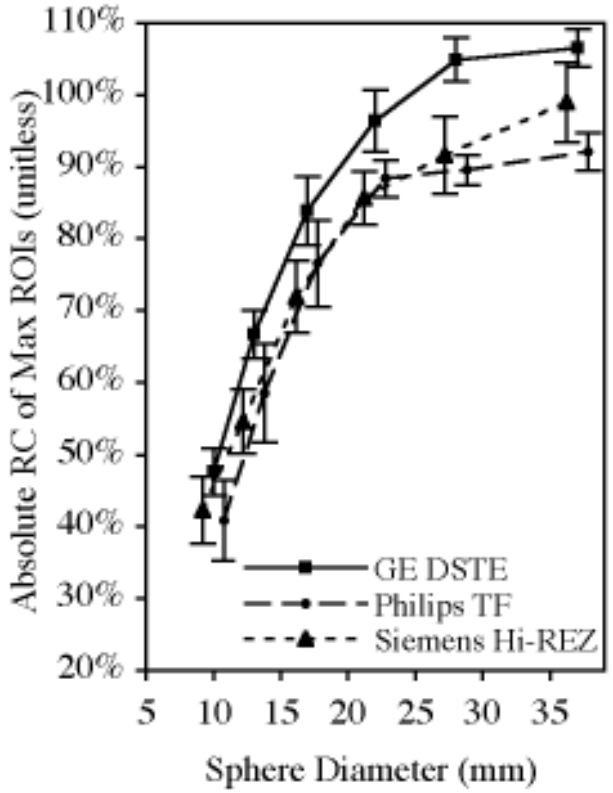
Single-site repeat PET/CT scans

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

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



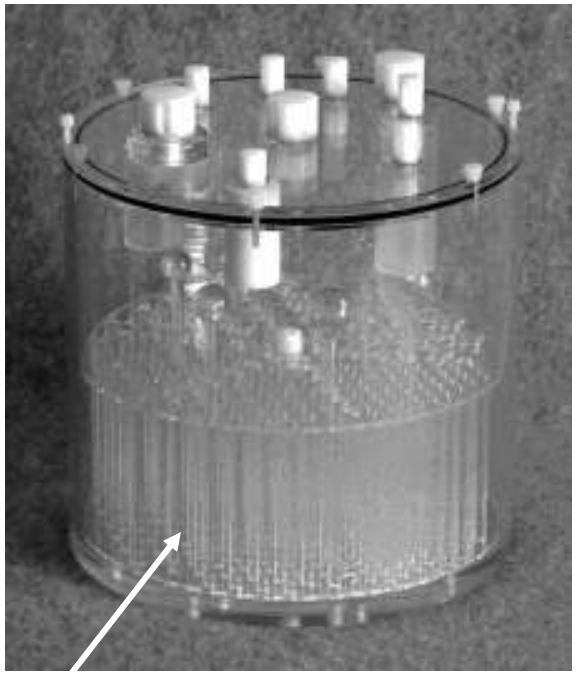
Absolute recovery coefficients from 3D-OSEM reconstructions using 7, 10, and 13 mm smoothing.



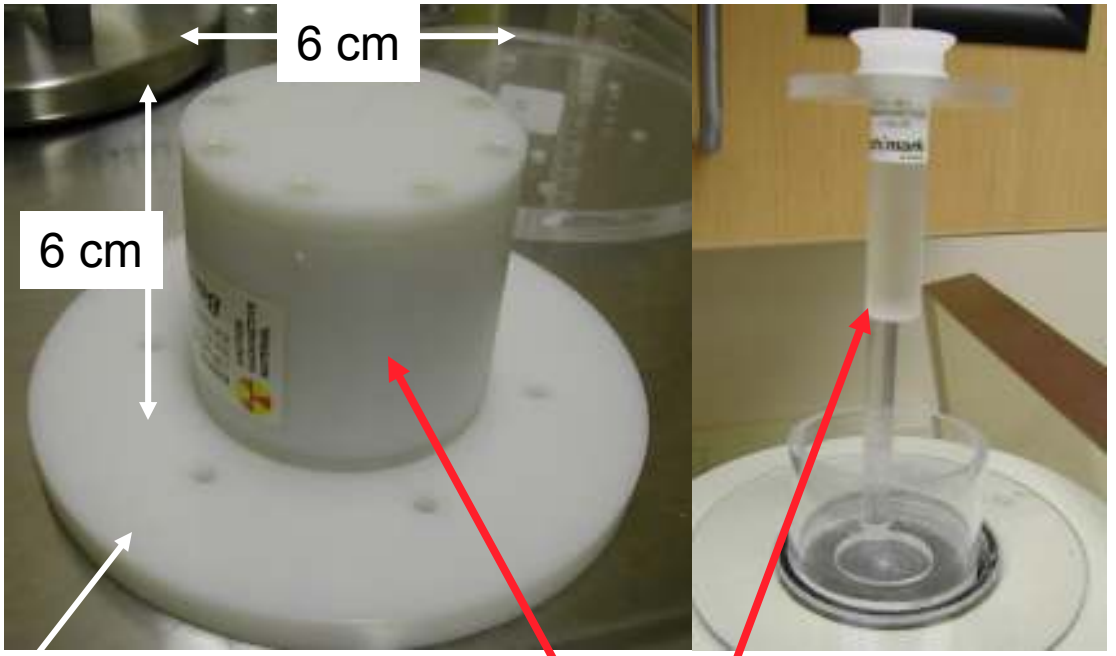
Maximum ROI recovery coefficients versus sphere diameter for the same phantom repositioned and imaged 20 times using PET/CTs from three vendors

Doot et al. 2010

Version 2 Modified ACR phantom with long half-life source matched to NIST standard



removable resolution insert



adapter base plate

^{68}Ge in epoxy sources from same batch using NIST traceable methods (1.3% error)

Preliminary results: Multi-site repeated scans

- All units in kBq/ml
- (number) in brackets is value from repeat scan after > 3 months
- True activity 217 kBq/ml (All activity measures decay corrected to 9/2/09)

			Measure errors	
Site	Dose calibrator	PET mean	Dose calibrator	PET (mean)
1	237	213	9.2%	- 1.8%
2	236 (235)	256 (219)	8.6% (8.2%)	18.0% (1.2%)
3	235 (236)	204 (231)	8.3% (9.0%)	- 5.8% (6.4%)
4	216 (212)	200 (185)	- 0.6% (-2.3%)	- 7.7% (-14.8%)
5	217 (209)	200 (182)	- 0.1% (-3.6%)	- 7.6% (-16.0%)
	1st 2nd	1st 2nd	1st 2nd	1st 2nd

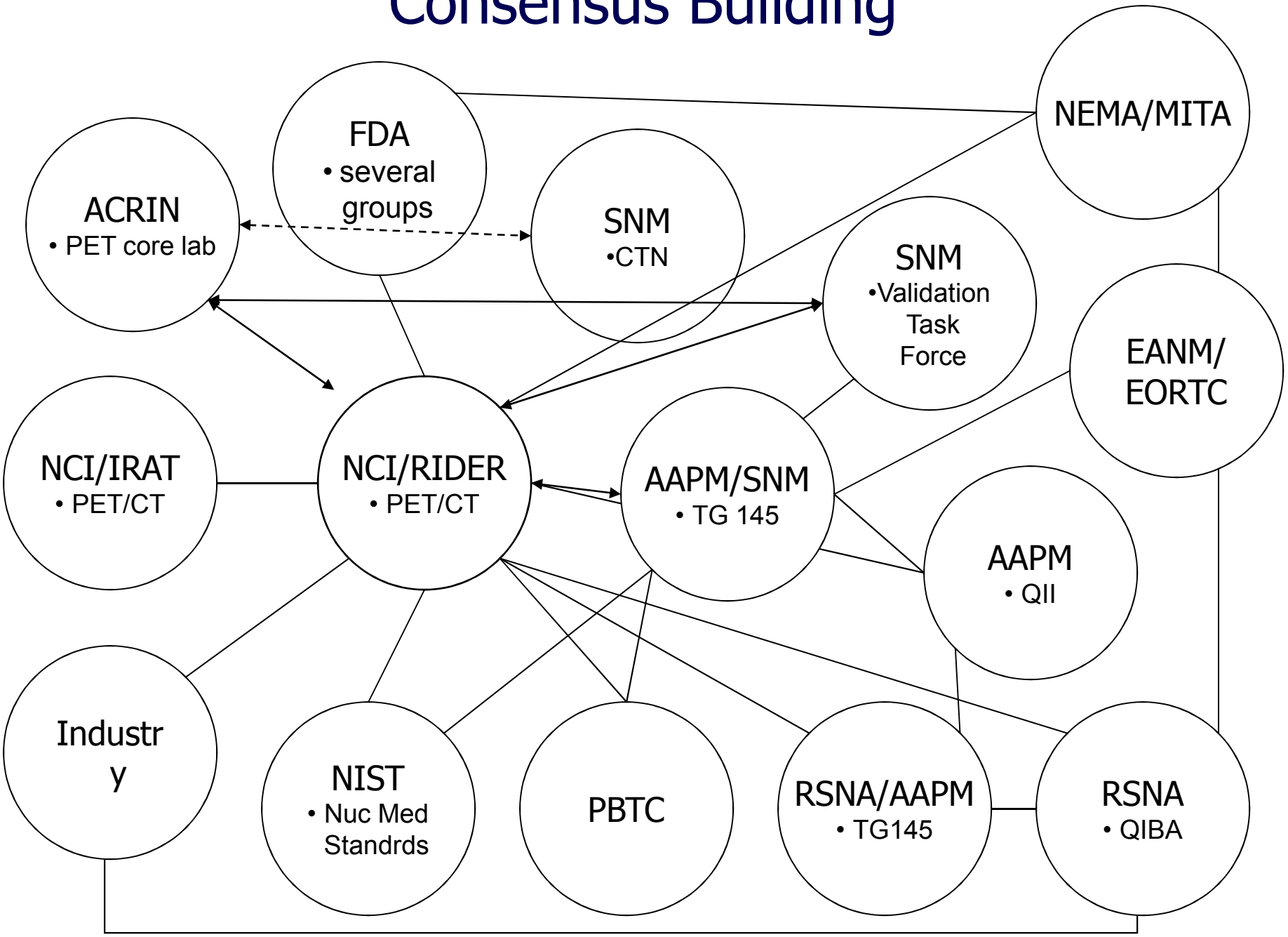
Next steps

- PET/CT is evolving from a valid *qualitative* clinical tool with excellent image fidelity to a *quantitative* clinical research tool
- Imaging results are quantitative if we pay attention to all aspects of image acquisition processing and analysis (scanner QA/QC ≠ quantitation)
- Reducing/controlling variability may be more important (and feasible) than reducing bias
- There are, however, simple changes that can reduce bias
- Paying attention means reporting *what was done*, not just what was specified, for the protocol: acquisition, processing, and analysis
- Through collaboration we can:
  - Determine impact of image bias/variance on clinical trials
  - Minimize impact across multiple sites by adhering to standards
- Quantitative imaging results can be used as disease response or stratification markers if:
  - We adhere to standards to *quantitatively validate* imaging
  - Acquire sufficient number of quantitatively validated studies with outcomes

Numbers Do Matter

- "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]*
- 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
- Until recently, there have been few systematic efforts to understand or improve quantitative accuracy, precision, and stability between multiple sites.

Consensus Building

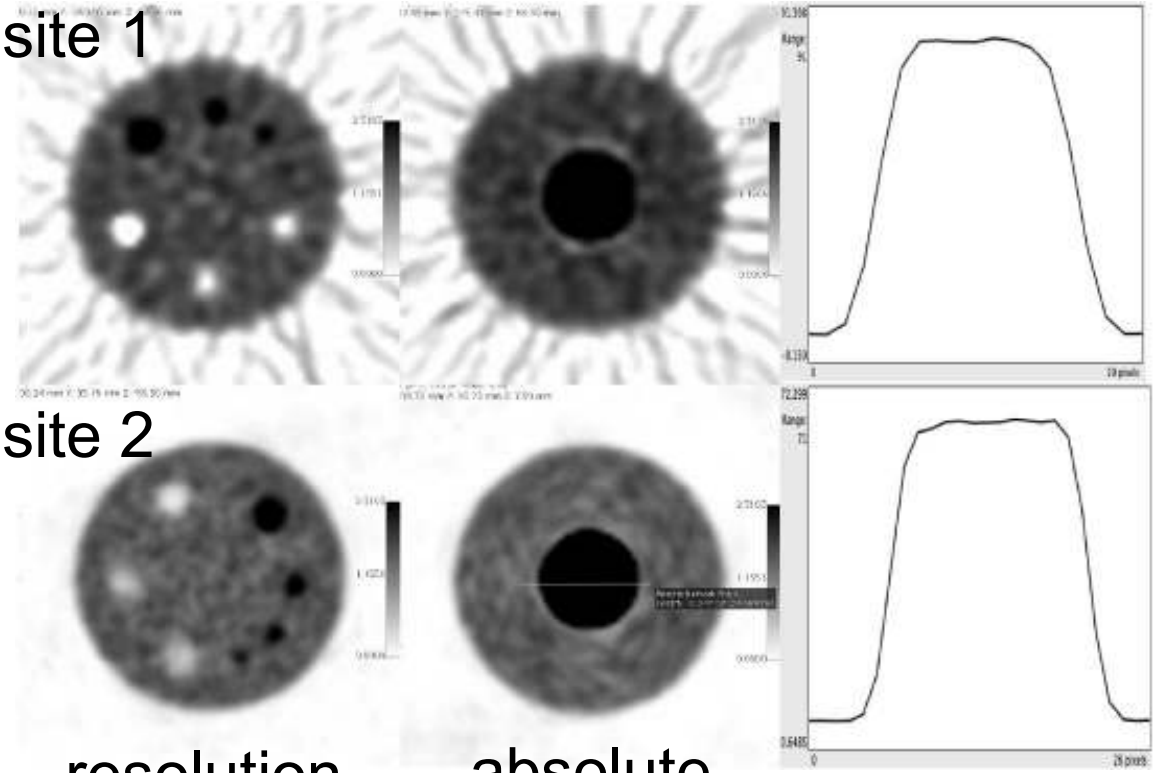


Version 2 Modified ACR phantom with long half-life source using NIST standard

assembled

site 1

site 2



resolution effects

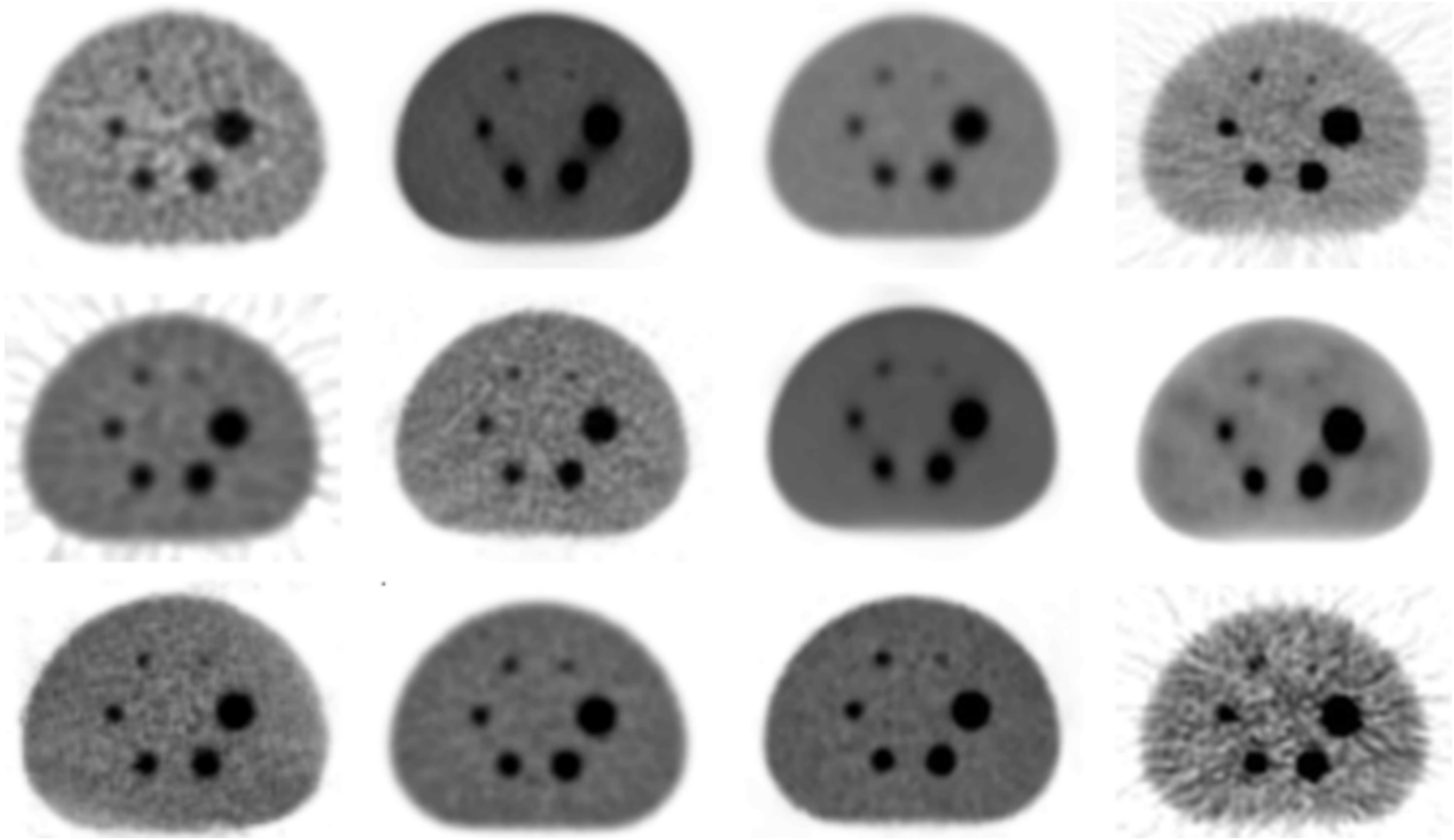
absolute quantitation

(also measure uniform region in between to check for SUV=1)

Profiles:
No partial volume effects in center

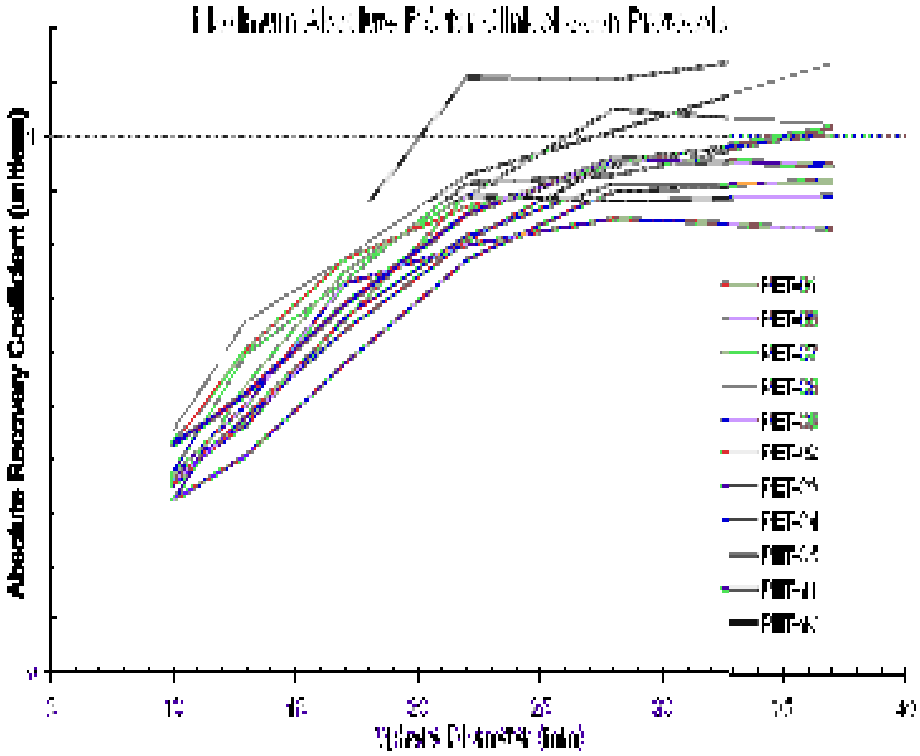
SNM Validation Phantom Study

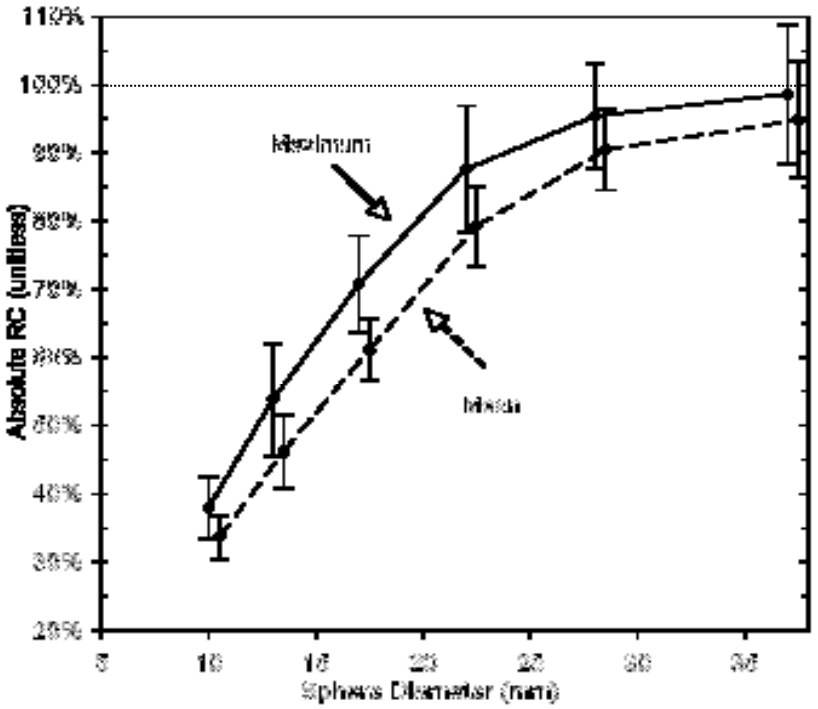
- Sample images of the IDENTICAL object from 12 different PET and PET/CT scanners



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

Multi-center repeated PET/CT scans





- Values for 11 scanners at at 8 academic imaging centers.
- Results should be independent of sphere diameter.

averaged coefficients of variation

mean SUV: 8.6%, max SUV: 11.1%

Doot PhD Thesis 2008, Kinahan et al 2008 SNM

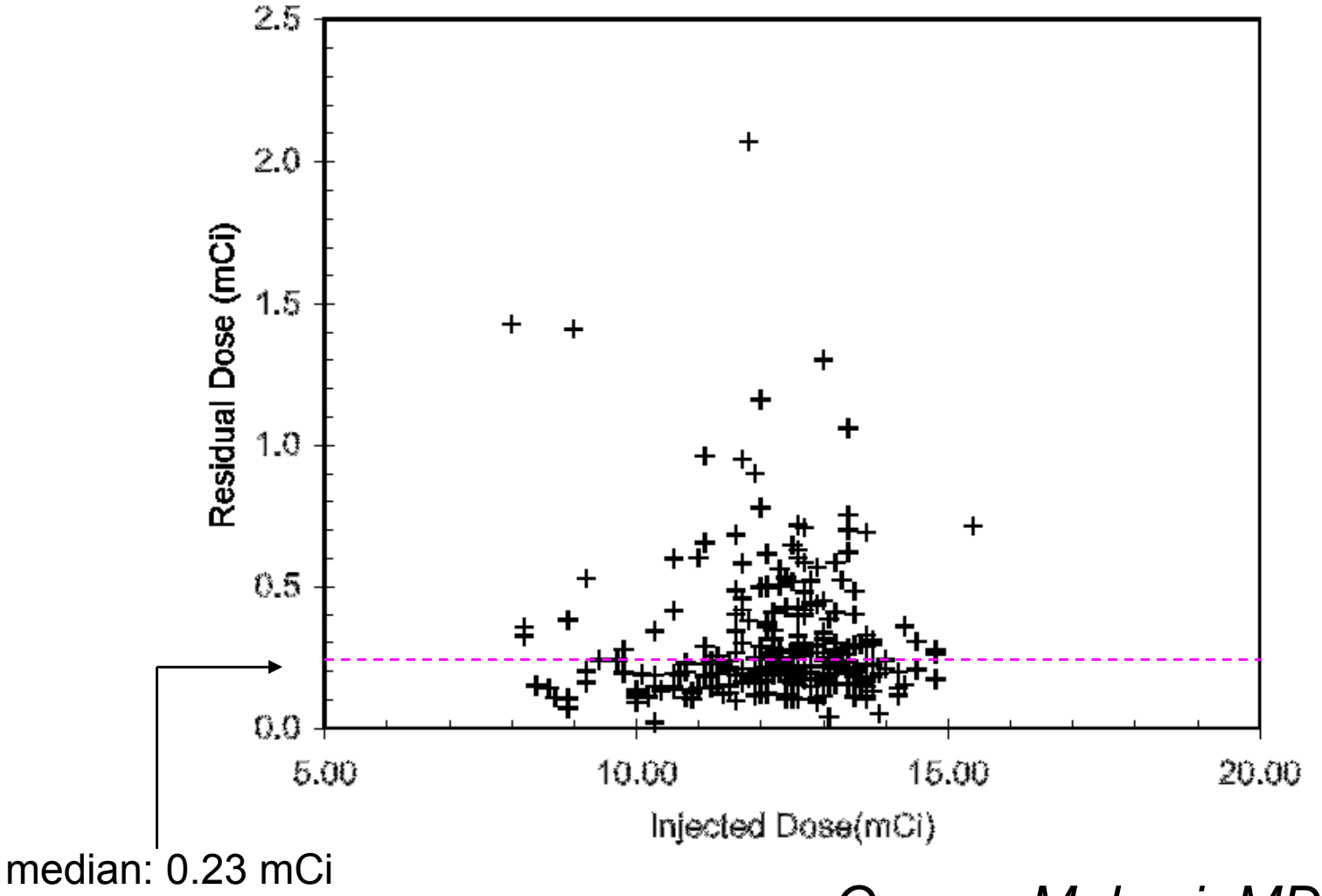
With Current Clinical Practice, do Numbers Matter in PET Images?

- *R Edward Coleman Eur J Nucl Med (2002)*
  - The answer to the question "Is quantitation necessary for clinical oncological PET studies interpreted by physicians with experience in interpreting PET images?" is "no."
  - Image quantitation will become increasingly important in determining the effect of therapy in many malignancies

- What do we need accurate SUVs for?
  - Clinical research, Clinical trails, and Drug discovery
  - Individual response to therapy
  - SUVs are now routinely reported, and asked for by referring physicians

Residual dose

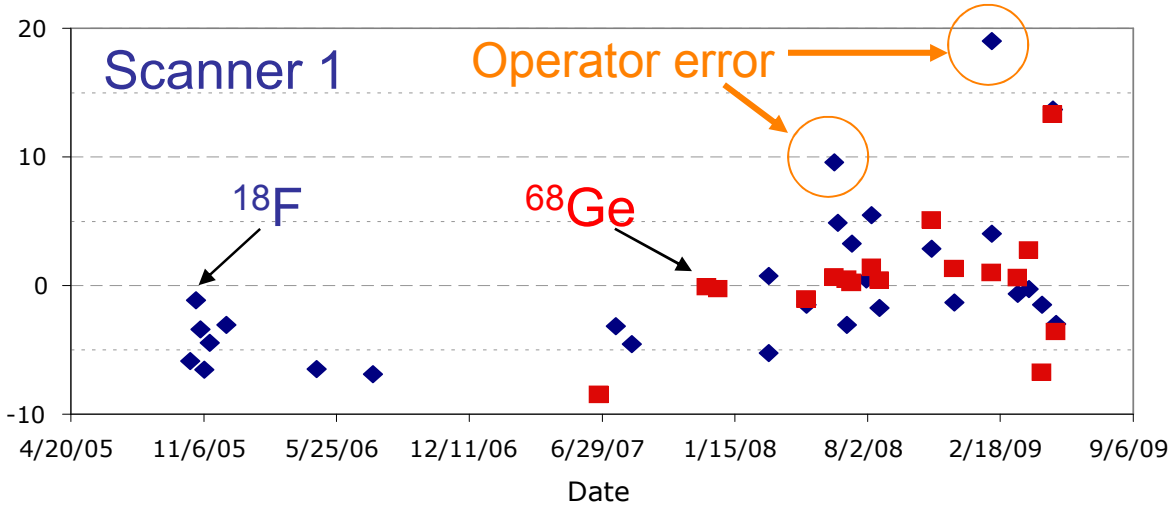
- *n* = 250 patients



Osama Malawi, MD Anderson

We know short term variability, but not long term variability

Activity Correction Factors



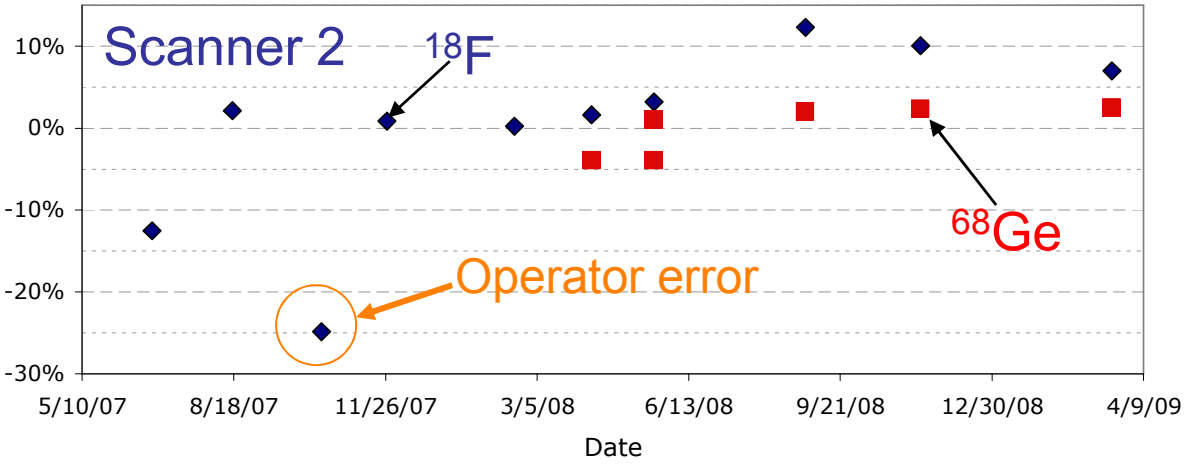
Standard Deviation

^{68}Ge= 3.2%

^{18}F = 6.1%

Outliers Removed

^{18}F = 4.1%



Standard Deviation

^{68}Ge= 3.1%

^{18}F = 11.0%

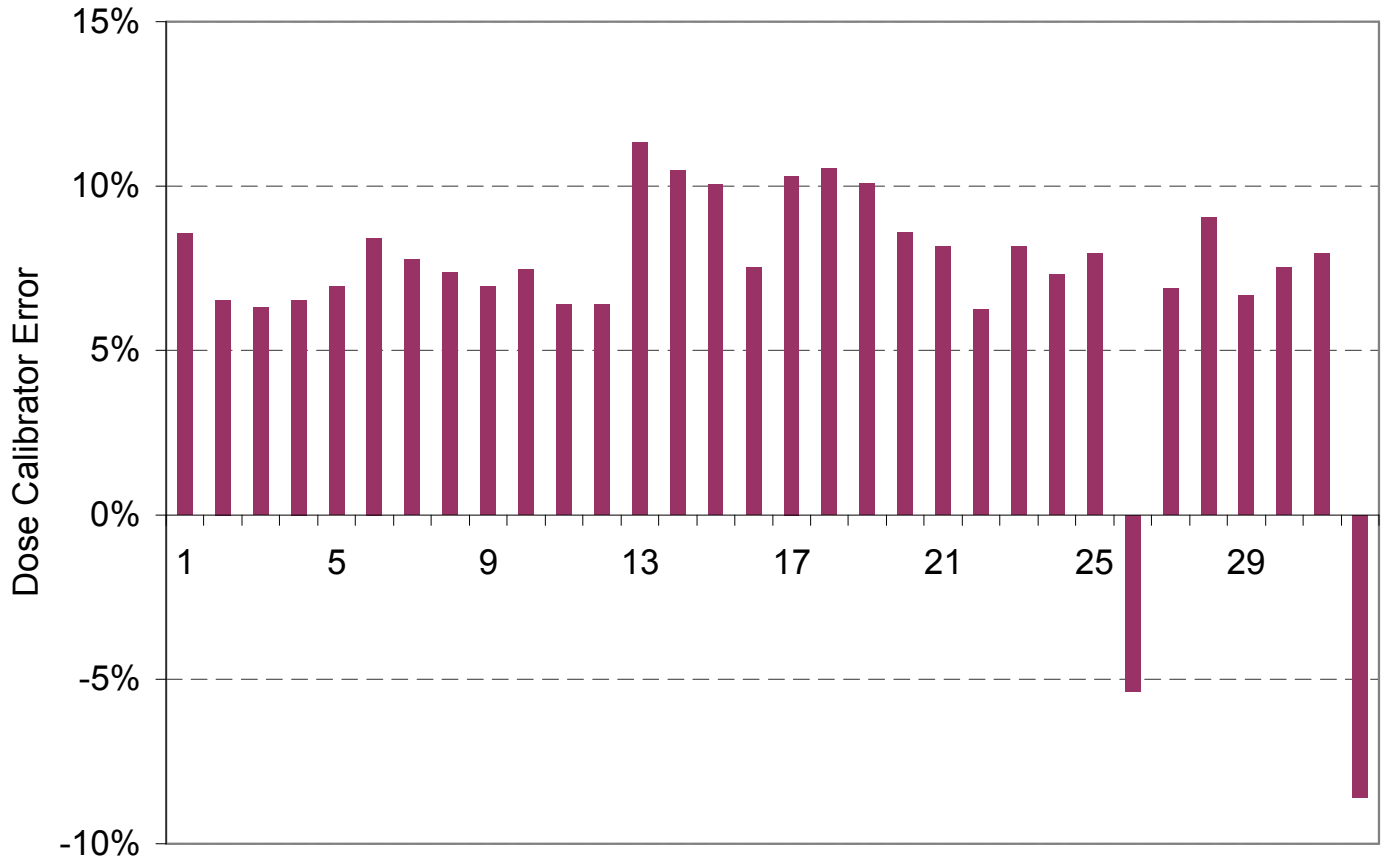
Outliers Removed

^{18}F = 7.0%

Lockhart SNM 2009

Dose Calibrators have significant variability, and not all scanners calibrate against a dose calibrator

Sample of 32 dose calibrators at 3 sites using RadQual/NIST Ge-68 standard



Zimmerman SNM 2009