Quantitative molecular imaging biomarkers and impact on patient safety

Robert Jeraj
Professor of Medical Physics, Human Oncology, Radiology and Biomedical Engineering
Fakulteta za matematiko in fiziko, Univerza v Ljubljani
University of Wisconsin Carbone Cancer Center, Madison, WI
robert.jeraj@fmf.uni-lj.si; rjeraj@wisc.edu
Type of imaging (biomarkers)

Qualitative imaging (Diagnostics)

Quantitative imaging (Quantitative Imaging Biomarkers)

BEFORE

DURING

AFTER THERAPY

DIAGNOSIS and STAGING

TARGET DEFINITION

EARLY

LATE

TREATMENT ASSESSMENT
Qualitative imaging chain

1. Scanning protocol
2. Data acquisition
3. Image reconstruction
4. Image interpretation

Imaging physics

Patient status
Imaging biomarkers require quantification of the whole imaging chain!
Main issues for Quantitative Imaging Biomarkers (QIB)

- **Imaging Equipment ≠ Measurement Device**

- **Measurement Device:**
  - Specific measurand(s) with known bias and variance (confidence intervals)
  - Specific requirements for reproducible quantitative results
  - Example: a pulse oximeter

- **Imaging Equipment:**
  - Historically: best image quality in shortest time (qualitative)
  - No specific requirements for reproducible quantitative results (with few exceptions)
QIB challenges

- **General QIB challenges:**
  - Lack of detailed assessment of sources of bias and variance
  - Lack of standards (acquisition and analysis)
  - Highly variable quality control procedures
  - QC programs / phantoms, if any, typically not specific for quantitative imaging
  - Little support (historically) from imaging equipment vendors
  - No documented competitive advantage of QIB (regulatory or payer)

- All lead to **varying measurement results** across vendors, centers, and/or time
QIB challenges

- **Other QIB challenges:**
  - Cost of QIB studies (comparative effectiveness) / reimbursement
  - Radiologist acceptance
    - Limited number of use cases for QIBs vs. conventional practice
    - QIBs are not part of radiologist education & training
    - The software and workstations needed to calculate and interpret QIBs are often not integrated into the radiologist’s workflow
    - Clinical demand on radiologists is high --- “time is money”
Consumer expectations of QIB

- **Oncologists** (94%) expect some or all tumors to be measured at the time of standard initial clinical imaging. (Jaffe T, AJR 2010)

- **Pulmonologists** desire CT-derived quantitative measures in COPD and asthma patients. (ATS/ERS Policy statement, Am J Resp Crit Care Med 2010)

- **Hepatologists** desire quantitative measures of liver fat infiltration (Fitzpatrick E, World J Gastro 2014)


- **Neurologists and psychiatrists** desire quantitative measures of brain disorders (IOM Workshop, August 2013).

- **Regulatory agencies** desire more objectivity in interpretations.
Problem: QIB uncertainties

Sources of Variance
- Patient Handling
- Acq. Protocols
- Reconstruction
- Segmentation

Measure = 7 ±6
Reducing QIB uncertainties

Problem

Measure = 7 ±6

Goal

Measure = 7 ±1

Cause

Sources of Variance

Differences in:
- Patient Handling
- Acq. Protocols
- Reconstruction
- Segmentation
...

Solution

When all participating actors conform...

- Requirements for: Acquisition Params, Recon Params, Resolution, Noise Reqs, Processing Params, Patient Prep & Operation, Segmentation, Calibration

Courtesy of Kevin O'Donnell
Harmonization

- **Harmonization of acquisition**
  - Minimize limitations due to different scanner hardware and software

- **Harmonization of scanning protocols**
  - Creating harmonized imaging protocols, which need to be tuned to specific scanners

- **Harmonization of image analysis**
  - Unifying image analysis protocols, which often means centralized analysis

- **Harmonization of reporting**
  - Standardized reporting, otherwise not comparable data
How much variability is there?

SNMMI’s Clinical Trials Network (CTN) sent the same phantom to 170 sites, and collected and analyzed the PET/CT images.
How much variability is there?

SNMMI’s Clinical Trials Network (CTN) sent the same phantom to 170 sites, and collected and analyzed the PET/CT images.

<table>
<thead>
<tr>
<th>Type</th>
<th>SUV&lt;sub&gt;max&lt;/sub&gt; bins for 10 mm right lung lesion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early era</td>
<td>1.5</td>
</tr>
<tr>
<td>Mid-range</td>
<td>1.6</td>
</tr>
<tr>
<td>High end TOF</td>
<td>2.2</td>
</tr>
<tr>
<td>PRF</td>
<td>2.8</td>
</tr>
</tbody>
</table>

### Typical academic site (UW example)

<table>
<thead>
<tr>
<th>ACR Phantom</th>
<th>Discovery VCT</th>
<th>Discovery 710</th>
<th>Signa PET/MR</th>
<th>Discovery IQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>2.7</td>
<td>2.8</td>
<td>2.7</td>
<td></td>
</tr>
</tbody>
</table>

Note: scanners have already been tuned to fall within ACR’s guidelines.
ACR phantom scanned on DVCT and D710

Scanner harmonization (phantom)

DVCT

- 25 mm RC = 0.99
- StDev = 3.53

Original D710

- 25 mm RC = 1.15
- StDev = 4.32

Harmonized D710

- 25 mm RC = 1.03
- StDev = 3.71
Harmonization changes values!

Baseline scan (on GE Discovery 710)

SUV\text{max} = 15.1

Harmonized to GE Discovery VCT

Follow-up scan (on GE Discovery VCT)

SUV\text{max} = 6.9

SUV\text{max} = 8.4

\[ H = \frac{\Delta\text{SUV}^{\text{orig.}}_{\text{max}} - \Delta\text{SUV}^{\text{harm.}}_{\text{max}}}{\Delta\text{SUV}^{\text{orig.}}_{\text{max}}} = 66\% \]

10 weeks post-treatment
Harmonization changes values!

Baseline scan (on GE Discovery 710)

Follow-up scan (on GE Discovery VCT)

10 weeks post-treatment

Reference regions

Original

$\text{SUV}_{\text{max}} = 15.1$

$\text{SUV}_{\text{liver mean}} = 2.65$

$\text{SUV}_{\text{aorta mean}} = 1.89$

Change in normalized values:

\[
\frac{\text{SUV}_{\text{max}}}{\text{SUV}_{\text{reference mean}}}
\]
Example: lung cancer patient

Lung Lesion

<table>
<thead>
<tr>
<th>Months</th>
<th>Maximum SUV</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-39%</td>
</tr>
<tr>
<td>6</td>
<td>+89%</td>
</tr>
<tr>
<td>13</td>
<td>-76%</td>
</tr>
<tr>
<td>DVCT</td>
<td>+74%</td>
</tr>
<tr>
<td>D710</td>
<td></td>
</tr>
</tbody>
</table>

Original
Example: lung cancer patient

Lung Lesion

Maximum SUV

Months

- D710
- DVCT

Harmonized
Original

+13%
+19%
+74%
-39%
+89%
-62%
-76%
+3%
Response classification

±30% changes in $\text{SUV}_{\text{max}}$ used for classification based on PERCIST (Wahl et al, JNM, 2009)
Response classification

$\Delta SUV_{\text{max}}$ (%)

±30% changes in $SUV_{\text{max}}$ used for classification based on PERCIST (Wahl et al, JNM, 2009)
All lesions (sorted by H) → Lesions changed SUV max response classification in 35/90 cases → Classification changes occurred in 10/20 patients

Is normalization able to capture the same changes that harmonization does?
Harmonization vs normalization

<table>
<thead>
<tr>
<th>Method</th>
<th>Changed Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harmonization</td>
<td>35</td>
</tr>
<tr>
<td>Liver Normalization</td>
<td>17</td>
</tr>
<tr>
<td>Aorta Normalization</td>
<td>17</td>
</tr>
</tbody>
</table>

Harmonization

Liver Normalization

Aorta Normalization
Conclusions

- **Quantitative Image Biomarkers (QIB)** are needed for assessment of treatment response.

- **Harmonization** is necessary for decreasing uncertainties of QIB (e.g., QIBA profiles).

- Harmonization directly impacts clinical outcome evaluation.

- **QIBs directly impact patient safety!**