Image Quality Assessment in Nuclear Medicine

Paul Marsden, KCL

Mayneord-Phillips Trust Workshop
‘Multi-modality Routine Image Quality Assessment: QC/QA and Standards’
14 October 2014
Overview of talk

• Radionuclide imaging
• What does image quality depends on?
• Performance measurement
• Quality control
• Image quality assessment
• Multi-centre trials
• Conclusions
## Functional imaging with radiopharmaceuticals

1. Obtain/produce gamma or positron emitting radionuclide
2. Label an interesting pharmaceutical with it
3. Administer to the patient/subject
4. Wait for radiopharmaceutical to accumulate
5. Scan the patient to detect the emitted radiation
6. View/analyze images of the radiotracer distribution in-vivo

Radionuclide imaging is EXTREMELY SENSITIVE - pmolar radiotracer concentrations are detectible
Two different flavours of radionuclide imaging

**Single photon / conventional Nuclear Medicine**

- Single-photon emitting radionuclides
- Radionuclides usually metals, eg $^{99m}$Tc
- Straightforward tracer production using radionuclide generator and radiopharmaceutical ‘kits’
- Widely used – many current applications

**Positron Emission Tomography (PET)**

- Positron emitting radionuclides
- Biogenic radionuclides: C,N,O,F
- Difficult and expensive radionuclide production and radiopharmaceutical synthesis
- Fewer compounds – expanding use
- Physics of imaging process results in superior image quality
Collimated single photon system

Absolute sensitivity
~ 0.05 %

Positron Emission Tomography

Absolute sensitivity
~ 0.5 – 5%
> Better images
Hardware set-up (~3 monthly – many hours)
Adjust gains of all PMTs in the scanner to optimize…

Tube operating range

Coincidence timing
Energy discrimination
Crystal maps
Normalisation (~3 monthly – many hours)

• The problem
  – many 10,000s of detector elements
  – Complex systematic and random variations
  – Hardware setup will not remove all these
  – Detectors will drift over time

• Normalisation
  – acquire sinogram by uniformly illuminating all lines of response (rod source or uniform cylinder)
  – Divide all acquired sinograms by this ‘normalisation’ sinogram

• Results of not doing normalization
  – One detector out > fan in image
  – One element in sinogram > streaks
  – Lower SNR
  – Drift in overall calibration
NEMA Tests

- Manufacturers quote camera performance in terms of the NEMA specification
  - National Electrical Manufacturers Association
- Compare NEMA test results of different cameras
- Disadvantages of NEMA tests
  - Complex and time consuming
  - Test equipment not always readily available
  - Tests do not lend themselves to regular in-house checks
Acceptance Testing

Ideal time to get involved with Accept. Testing as valuable inside-knowledge of the system can be gained at this point from interaction with engineers & apps specialists.

After camera is installed in hospital, acceptance tests are performed to ensure it is...

- working to its optimum performance
- working to the specification outlined by manufacturer
Scanner spatial resolution

• ‘System resolution’
  measured for point or line source
  not count limited
  varies over field of view
  typically 4-6 mm FWHM

• ‘Image resolution’
  measured in final image
  smoothing to reduce noise
  typically 5-10 mm FWHM

measure transaxial and axial spatial resolution at various position in field of view by profile or by stepping source
Sensitivity measurement

20cm cylinder
Cps/Bq/ml
Correct for scatter
Low count rate

1994

2001

Metal source surround
Absolute sensitivity
Noise Equivalent Count Rate (NEC)

- NEC = ‘True coincidence count rate that would give you the same SNR as the one you actually get’

Slide courtesy Prof Paul Marsden, KCL
Is NEC really an index of image quality?

- The image quality depends primarily on both the SNR and the spatial resolution of the scanner.

Herzog et al. IEEE TNS 51:5;2662;2004

*Slide courtesy Prof Paul Marsden, KCL*
Phantom to assess image quality and accuracy of attenuation & scatter corrections (NEMA 2001)
Performance Characteristics

Frequently assessed as part of routine QC procedures

- Spatial resolution
- Image non-linearity
- Image non-uniformity
- Count rate characteristics
- Energy resolution
- Sensitivity
Integral Uniformity
Measure of the difference between the max and min counts per pixel in the image

\[
\text{Integral Uniformity} = \frac{C_{\text{max}} - C_{\text{min}}}{C_{\text{max}} + C_{\text{min}}} \times 100 \%
\]

Differential Uniformity
Max difference between the highest (H) and lowest (L) counts in 2 pixels over a set of 5 pixels

\[
\text{Differential Uniformity} = \frac{H - L}{H + L} \times 100\%
\]
Daily PET QC procedures (daily - <30mins)

- All the gain/normalisation settings are stored
- Daily check just compares how much values have changed since last normalisation/setup....
Annual Checks: PET/CT

• CT annual checks
  – More complex measurements eg computed tomography dose index (CTDI), spatial resolution, mechanical checks

• PET/CT scanner alignment checks
  – performed quarterly to annually or after scanner repair

Image supplied by: Dr. Mike Partridge, Joint Department of Physics, The Institute of Cancer Research / The Royal Marsden NHS Foundation Trust
Additional daily checks – uniform emission phantom

• Cylindrical uniform Ge-68 phantom
• 5-10 minute PET scan acquired
• Whole body imaging parameters used
• Images reconstructed
• Check PET; stability; CT; CTAC; alignment; software, faults etc

10% limit on SUV variation
Objective image quality assessment
<table>
<thead>
<tr>
<th>Task</th>
<th>Extracting required information</th>
<th>Figure of merit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor detection (classification)</td>
<td>for accurate statistics</td>
<td></td>
</tr>
<tr>
<td>Numerical observer (e.g. channelized Hotelling)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task</th>
<th>Measure tracer uptake by placing ROI over the tumor</th>
</tr>
</thead>
</table>

| Tracer uptake quantitation (estimation): | MSE = \text{bias}^2 + \text{variance} |

**Receiver operating characteristic (ROC) curve:**
- Accuracy \approx \text{Area under curve}

**Ideal test**
- No useful information
- False Positive Rate
- True Positive Rate

![Image showing ROC curve](image-url)
<table>
<thead>
<tr>
<th>Category</th>
<th>Factor</th>
<th>Typical range</th>
<th>Maximum effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical errors</td>
<td>Relative calibration between PET scanner and dose calibrator</td>
<td>-10% - +10%</td>
<td>±50%</td>
</tr>
<tr>
<td>Technical errors</td>
<td>Residual activity in syringe or administration system</td>
<td>0% - 5% (typically &lt;15%, but can be much greater in worst-case situations)</td>
<td></td>
</tr>
<tr>
<td>Technical errors</td>
<td>Incorrect synchronization of clocks of PET/CT camera and dose calibrator</td>
<td>0% - 10%</td>
<td>21% observed</td>
</tr>
<tr>
<td>Technical errors</td>
<td>Injection vs. calibration time</td>
<td>0% - 10%</td>
<td></td>
</tr>
<tr>
<td>Technical errors</td>
<td>Paravenous administration of 18F-FDG</td>
<td>0% - 50% or more, strongly depending on quality of administration</td>
<td></td>
</tr>
<tr>
<td>Biologic factors</td>
<td>Blood glucose level</td>
<td>-15% - +15%</td>
<td>±75%</td>
</tr>
<tr>
<td>Biologic factors</td>
<td>Uptake period</td>
<td>0% - 15% at 60–90 min</td>
<td>±30%</td>
</tr>
<tr>
<td>Biologic factors</td>
<td>Patient motion or breathing</td>
<td>0% - 30%</td>
<td>±60%</td>
</tr>
<tr>
<td>Biologic factors</td>
<td>Patient comfort</td>
<td>NaN, mainly giving rise to false-positive results (SUV&lt;sub&gt;BW&lt;/sub&gt; = 2–12) and possibly incorrect SUV</td>
<td></td>
</tr>
<tr>
<td>Biologic factors</td>
<td>Inflammation</td>
<td>NaN, mainly giving rise to false-positive results and possibly incorrect SUV</td>
<td></td>
</tr>
<tr>
<td>Physical factors</td>
<td>Scan acquisition parameters</td>
<td>0% - 15%</td>
<td>±15%</td>
</tr>
<tr>
<td>Physical factors</td>
<td>Image reconstruction parameters</td>
<td>-30% - 0%</td>
<td>±30%</td>
</tr>
<tr>
<td>Physical factors</td>
<td>ROI</td>
<td>0% - 55%</td>
<td>±55%</td>
</tr>
<tr>
<td>Physical factors</td>
<td>Normalization factor for SUV</td>
<td>Trivial</td>
<td></td>
</tr>
<tr>
<td>Physical factors</td>
<td>Blood glucose level correction</td>
<td>-15% - 15%</td>
<td>±75%</td>
</tr>
<tr>
<td>Physical factors</td>
<td>Use of contrast agents during CT-AC</td>
<td>0% - 15%</td>
<td>±50%</td>
</tr>
</tbody>
</table>
FDG protocols

Tumour SUV

~ 60 mins

1 week

Time

SUV

Time

~ 60 mins

1 week

Time

FDG protocols
NCRI PET Clinical Trials Network

- 25 UK PET centres accredited (29 scanners)
  - List of accredited sites available to researchers

- Accreditation of sites

- ‘Core lab’

- Standards to be extended to more complex trials
  - Input from ‘PET Scanning Methodology Expert Panel’
  - Smaller number of experienced sites to perform mechanistic studies

- Collaborate with European and US groups (ACRIN, SNM, EANM etc.) to harmonise standards across countries
Error Propagation in PET Imaging

Estimate

Single-center best case: 10%
Single-center, typical: 10-18%
Multi-center, best case: 15-20%
Multi-center, typical: 15-40+%

Source data

e.g. Minn 1999, Weber 2000
Velasquez 2009
Velasquez 2009
Fahey 2009, Doot 2010

Note – these issues have not yet been addressed to the same extent for more complex studies...

Paul Kinahan
Conclusions

- Overall system performance and how it is affected by basic hardware is well understood
- Effect of performance and hardware errors on image quality is much less well understood
- Clinical image quality still evaluated on a very ad-hoc basis
- This is because task-based ROC type studies are very difficult to perform