

核医学検査をバイオマーカーとするために求められること

九州大学大学院医学研究院 保健学部門 医用量子線科学分野
佐々木雅之、三輪建太、赤松剛、松友紀和

1. はじめに

核医学検査は機能や代謝が測定できることから機能診断や代謝診断とも呼ばれており、近年は生体の分子機能を反映した画像として分子イメージングの中心的な手法とされている。さらに、得られた結果は診断や治療の目安として利用されるようになり、バイオマーカーとして扱われるようになった。核医学検査は画像が表示できるがゆえに画像診断のひとつとして扱われる一方で、他の画像診断に比して空間分解能や組織分解能が劣るために、画像診断としては低く扱われがちである。つまり、核医学検査は生体の代謝や機能を画像化する検査であることから、他の画像診断とは異なる情報を得ることができるという特徴や有用性が評価されていない場合が少なくない。

核医学検査の中で機能代謝診断としての有用性が高く認められてきた検査に脳血流 SPECT 検査がある。認知症などでは血流が反映する脳機能の評価に用いられるとともに、脳血管障害では脳血流の定量評価が治療方針の決定に用いられている。脳血流 SPECT の定量的評価としての役割は日本の研究者たちの功績によって確立されてきたが、この中で定量化には放射性医薬品と集積原理、使用放射性核種、コリメータ、空間分解能、減弱補正、散乱線補正、などが重要な因子となるこ

とがわかってきた。

一方、腫瘍診療では局所的な治療である外科的治療と放射線治療と全身的な治療である殺細胞的な化学療法が重要な役割を担ってきた(図1)。従来は腫瘍が消失したか、またはいかに小さくなったかなどの形態的变化が治療の効果の目安であった。近年、分子標的薬などの静細胞的な化学療法が登場したことで、治療効果が形態的变化ではとらえることができない場合がでてきた。さらに、殺細胞的な化学療法でも形態的变化が現れる前の治療途中で効果を予測することが求められるようになった。これらの従来の効果判定法では対応できない場合に、代謝的变化が目安として期待されている。特に複数の病巣を有する腫瘍の場合は病巣ごとの評価が必要な場合があり、核医学検査は画像で病巣ごとの代謝の違いを評価ができるバイオマーカーとして注目されている。

核医学検査がバイオマーカーとして確立するには、いつ、どこで、だれが、何のために、どう撮影しても、同様に判定できる結果が得られなければならない。すなわち、生体内の放射能分布を正確に測定する定量性が最も重要となる。本稿では、腫瘍 PET 検査を例に定量性の向上について概説する。

2. 定量性に影響する因子

核医学検査の定量性には様々な要因が影響することが知られている(図2)。まず、データを収集するPET装置では、クリスタル、光電子増倍管、電気回路などの違いが影響する。また、収集したデータから画像を再構成するためのソフトウェア、データの劣化を補正するための減弱補正、散乱線補正などのアルゴリズム、近年になって臨床機に導入された飛行時間差情報(time-of-flight)の利用なども影響する。検査の方法でも、投与放射能、データ収集の開始時刻、データ収集時間、投与装置とドーズキャリブレーションやオートウェルカウンタなどの測定装置間のキャリブレーションの精度などが重要である。さらに、被検者側の要因として、前処置の状態、検査時血糖値、糖尿病の有無、呼吸や心拍に起因する移動、体動さらには病状なども影響を与える。

これらの要因を施設ごとに、また施設を超えて統一することで定量性の向上を図る取り組みがなされているが、われわれはこれまでに、腫瘍PET検査の定量性に影響を与える、1)病変検出能と部分容積効果、2)呼吸性移動の制御、3)再現性と繰り返し性、について取り組んできた。また、世界的には検査の標準化と調和化にむけた取り組みが進んでいる。

3. 病変検出能と部分容積効果

腫瘍PET検査は多くの場合はバックグラウンド集積の中のホット高集積を検出する。バックグラウンド集積とホット高集積との差がシグナルであり、このシグナルをいかに検出するかが診断能に相当する(図3)。

高いシグナルを得るにはバックグラウンドと病巣の間で大きな集積差があるトレーサを使用しなければならない。しかし、トレーサの種類は限られており、自由にトレーサを変更できるわけではない。

バックグラウンド集積には統計変動がある。シグナルを明瞭に検出するには、シグナルはバックグラウンドの変動よりも大きくなければならない。バックグラウンドの変動は小さければ小さいほどシグナルを検出しやすい。バックグラウンド変動が少ない画像を得るには、十分に大きい計数を得なければならない。現在の装置を使用する場合は収集時間を長くする必要はあるが、臨床の現場では必ずしも容易ではない。このためには、高感度装置を開発すること、装置の時間分解能を向上させることが重要である。

小病変のホット高集積は装置性能やデータ収集法の影響などで過小評価される。装置の空間分解能に依存する過小評価は部分容積効果と呼ばれる(図4)¹⁾。対象が小さい場合には、検出器と光子の相互作用の特性によって計数が低くなり、濃度境界が不明瞭になる現象はfinite resolution effectと呼ばれる。また、デジタル表示の画素サイズが濃度境界や濃度範囲と一致しないための計数低下はtissue fraction effectと呼ばれる。核医学画像の部分容積効果補正にはいくつかの試みがなされてきたが、いまだ決定的な補正方法は確立されていない。我々は、対象のサイズ、シグナル、バックグラウンドに応じて部分容積効果を補正する方法を開発した²⁾。この方法を利用することで小病変の良悪性の鑑別能が向上

した。しかし、画像そのものを補正する実用的な方法は確立していない。

4. 呼吸性移動の制御

PET 検査では収集時間が比較的長いいため、データ収集中の被検者の体動、心拍による動き、呼吸による動きによって過小評価が生じる。データ収集中に集積が動くと、集積が広範囲に分散するため集積範囲が広くなるとともに集積程度が低くなる。さらに、CT 画像との位置ズレも生じると各種補正が不適切となるとともに病変部位の誤認も生じうる (図 5)。

高感度装置では比較的短時間でのデータ収集が可能であり、息止め撮像が利用できる。我々は約 20 秒間の息止め撮像を 5 回行い、3 回以上を加算することで読影に十分な画質を得ることが明らかとした³⁾。この結果、自由呼吸下の画像と比較して、集積程度を改善できた。

呼吸同期検査は何らかの方法で呼吸の状態をモニターし、必要と思われる部分の PET データのみを用いて画像を再構成する方法である。各呼吸をそれぞれ一定の位相に分割したのちに特定の位相のデータを加算する方法を位相同期法または時間同期法と呼ぶ。我々はファントムを用いた検討により、PET の呼吸同期には CT も呼吸同期を行い位相をそろえる必要であることを明らかとした⁴⁾。また、呼吸の振幅を目安として、特定の振幅のデータのみを加算する方法を振幅同期法と呼ぶ。我々はファントムを用いて位相同期と呼吸同期を比較し、規則的な呼吸の場合は両者ともに有効であるが、振幅が変動したり基線が変動するような不規

則な呼吸の場合は振幅同期がより有用であることを明らかとした⁵⁾。

5. 繰り返し性と再現性

Boellaard らは、繰り返し性 repeatability と再現性 reproducibility が定量評価/バイオマーカーとしての根本的な必要事項である、と述べている (図 6)⁶⁾。さらに彼らは、繰り返し性とは同一被検者に対して同じ装置で複数回撮像した時の一致性であり、test-retest reproducibility と同義であると定義している。また、再現性は同一被検者に対して施設や装置を変えて撮像したときの一致性、と定義している。すなわち、漠然とイメージする再現性(広義の再現性)には繰り返し性と再現性を区別して考える必要がある。これまでに同一被検者を 2-3 週間以内に 2 回検査をした場合の SUV の一致性は多数検討されており、単一施設内での 2 回検査の差は 10-18%、複数施設間での 2 回検査の差は 15-50%とされている (図 7)。

6. 標準化と調和化

これまでに述べたように様々な要因で結果の差異が生じることが知られている。異なる施設間のもとより、装置が異なれば結果が異なり、同一装置であっても撮影方法や画像再構成方法が異なっても結果は異なる。ひとつひとつの要因が、それぞれの程度の変動をもたらすかについてもすでに報告がある (図 8)⁷⁾。

装置の性能や精度を評価する方法は製造企業などの工業会や各国の学会が取り組んできている。PET 装置の性能評価方

法として広く用いられているのは National Electric Manufacturers Association (NEMA) の nu-2 であり、国内では JESRA の方法が用いられている。また、PET 装置や関連装置の QAQC の指針も出されている。これらを踏まえて、検査結果の繰り返し性と再現性を改善する目的で検査の標準化 standardization の取り組みが世界的になされてきた (図 9)。欧米の核医学会をはじめ日本核医学会でも各種のガイドラインを出してきた。ガイドラインの多くは、被検者に関すること、放射性医薬品と投与方法に関すること、撮像装置の設定とデータ収集プロトコールに関すること、画像再構成と処理に関すること、結果の判定法に関することなどの標準的な方法が記載されており、検査の方法を統一することで同一の結果が出るようにとの指標とされている。すなわち標準化は、ある一定の基準を設定し、得られた結果が基準値に一致するように正確性 accuracy に焦点を当てた考え方である (図 10)。しかし、これまでなされてきた標準化の努力によっても依然として結果が同一とならない状況は続いている。

これらの状況を踏まえて欧米の学会等は、分子イメージングの定量的安定性を確立するために調和化 harmonization と呼ばれる取り組みを始めている。調和化とは異なる施設や装置の測定結果をある柔軟な基準範囲内 upper and lower limits に収まるに合わせこむ取り組みをいう。これは結果の精度 precision に焦点をあてた考え方である。異なるメーカーの異なる性能の装置、異なる世代の

異なる性能の装置、異なるソフトウェアでの画像再構成などの結果を同一基準で判定するために有用である。特に、診療経過の中で異なる施設で検査を受けた場合、同一施設であっても前回と今回で検査装置が異なった場合、治験や臨床研究などで多施設での検査結果を総合して解析しなければならない場合などでは重要である (図 11)。一方で、多くの検査結果を調和させるには、性能の低い装置や施設の結果を包含できる基準を設定する必要があり、最先端技術の活用とは逆行する場合があります。今後の解決すべき問題である。

7. おわりに

核医学検査は古くから機能診断・代謝診断と呼ばれてきたにも関わらず、臨床の現場ではその定量的意義は十分に信頼されて活用されてきたとは言い難い。また、分子イメージングやバイオマーカーという呼び方も臨床の現場ではまだ定着していない。核医学検査がバイオマーカーとして確立するには、いつ、どこで、だれが、何のために、どう撮影しても、同様に判定できる結果が得られなければならない。すなわち、生体内の放射能分布を正確に測定する定量性が最も重要となる。

8. 引用文献

- 1) Soret M, et al. J Nucl Med 2007; 48: 932-945.
- 2) Sakaguchi Y, et al. Ann Nucl Med 2010; 24: 585-591.
- 3) Mitsumoto K, et al. Nucl Med Commun.

2011; 32: 1148-1154.

4) Sakaguchi Y, et al. Ann Nucl Med
2010; 24: 585-591.

5) Tsutsui Y, et al. Ann Nucl Med 2014;
28: 770-779.

6) Boellaard R, et al. Eur J Nucl Med
Mol Imaging 2015; 42: 328-354.

7) Boellaard R. J Nucl Med 2009; Suppl
50:11S.

図1 癌の治療と治療効果判定

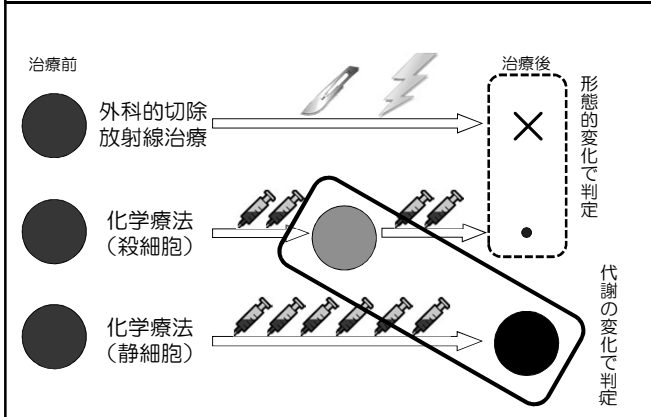


図2 核医学検査の定量性に影響する要因

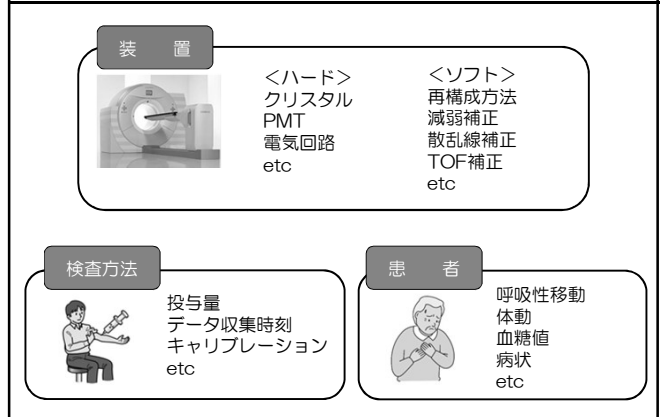


図3 Lesion detectabilityの改善には？

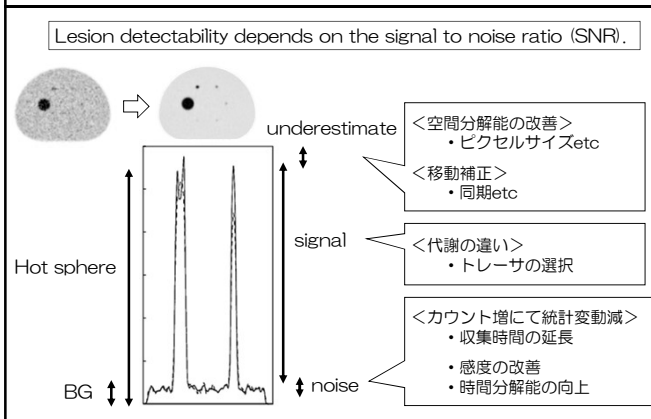


図4 核医学検査の部分容積効果(PVE)

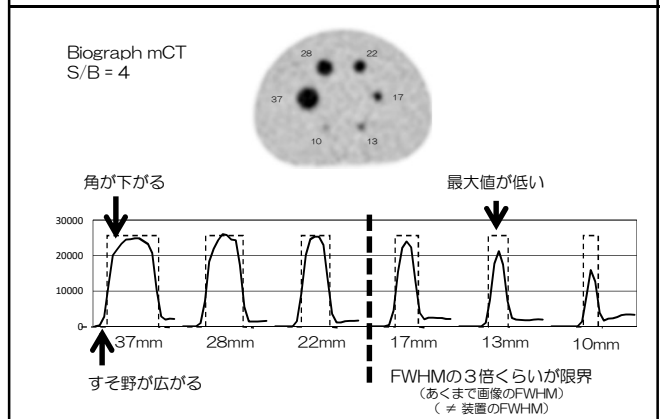


図5 呼吸性移動対策

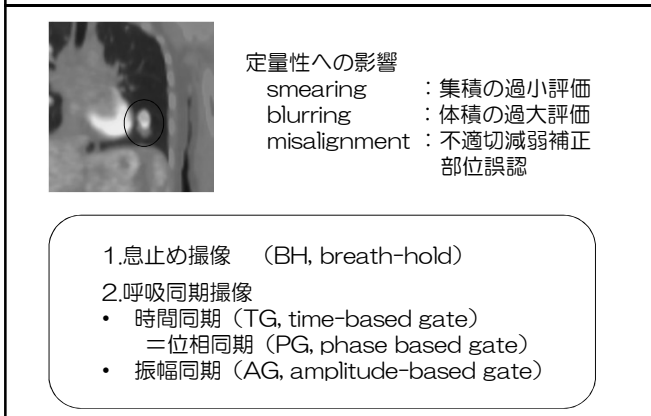


図6 繰り返し性と再現性

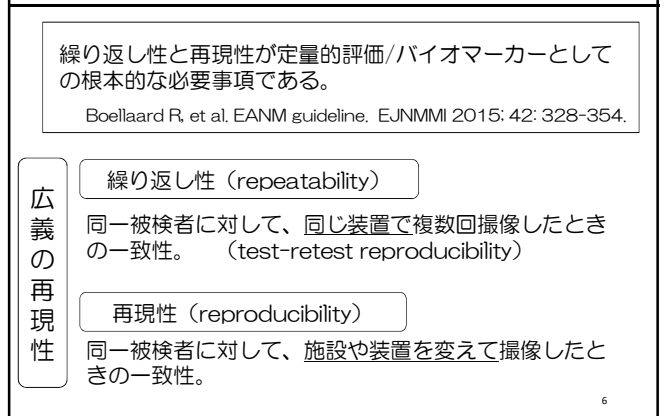


図7 繰り返し性：臨床の検討

test-retest reproducibility

・同一患者を2回検査（2-3週間以内）したときの
悪性腫瘍SUV_{max}のrelative difference (%)

単一施設	10-18%	Minn H, Radiology 1995. Weber WA, J Nucl Med 1999. Valasquez LM, J Nucl Med 2009. Nahmias C, J Nucl Med 2008.
複数施設	15-50%	Valasquez LM, J Nucl Med 2009. Weber WA, J Nucl Med 2015. Kumar V, Clin Nucl Med 2013. Doot RL, Med Phys 2010.

図8 定量に影響する因子とSUVの変動

技術的因子

Cross-calibration	(10%)
投与精度（残存放射能など）	(5%)
時刻設定のずれ	(10%)
静脈投与テクニック（注射漏れなど）	(50%)

物理的因子

収集条件	(15%)
再構成条件	(30%)
造影剤の利用	(10%)
ROIの種類と測定方法	(50%)

生物学的因子

血糖値	(15%)
待機時間（uptake time）	(15%)
被検者の動きと呼吸	(30%)
被検者の状態（心理状態を含む）	(5%)

Boellaard R, J Nucl Med 2009, Suppl. 50: 11S

図9 標準化と調和化

関連装置のQA/QC

PET装置の性能評価

NEMA nu-2 standards JESRA

標準化（standardization） ガイドライン

1998 SNM	2007 日本核医学会
1999 EORTC	2008 Netherland
2005 ACR, SNM, SCBTMR	2010 EANM
2006 SNM, NCI	2014 日本核医学会

それでも揃わない

調和化（harmonization）

図10 標準化と調和化

○ 標準化（standardization）

ある厳格な一定の基準（Lower threshold）を設定する。
accuracy（正確度）に焦点。

○ 調和化（harmonization）

ある柔軟な基準範囲内（Lower and upper limits）に収める。
precision（精度）に焦点。

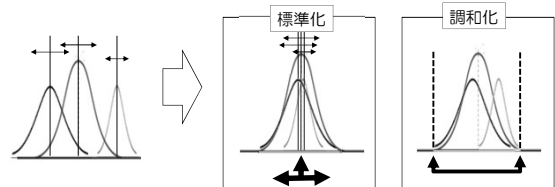


図11 標準化と調和化の取り組み例

標準化：Clinical Practice

- ・ ACR (American College of Radiology) PET/CT Accreditation Program

標準化：Clinical Trial

- ・ UPICT (Universal Protocols for Imaging in Clinical Trials)
- ・ RSNA QIBA (Quantitative Imaging Biomarkers Alliance)
- ・ SNMMI CTN (Clinical Trial Network)

標準化と調和化

- ・ NCI QIN (Quantitative Imaging Network)
- ・ EANM and EORTC initiatives