The Physics of PET/CT scanners

Ruth E. Schmitz, Adam M. Alessio, and Paul E. Kinahan

Imaging Research Laboratory
Department of Radiology
University of Washington

What Makes PET Useful?

Positron emission tomography (PET) offers a number of unique advantages compared to other imaging modalities. PET measures the two annihilation photons that are produced back-to-back after positron emission from a radionuclide tagged tracer molecule, which is chosen to mark a specific function in the body on a biochemistry level (Fig. 1). Hence PET provides molecular imaging of biological function instead of anatomy. The detection of both annihilation photons in coincidence yields increased sensitivity over all other forms of medical imaging. The time-coincident imaging of the two high-energy annihilation photons (discussed below) allows accurate attenuation correction from either a dedicated transmission scan or from CT information. This allows extraction of accurate information from PET images. Only minute amounts of imaging substrate need to be injected (tracer principle) because of the high sensitivity of PET. In addition, positron emitting isotopes that are used in medical imaging (C-11, N-13, O-15, F-18, etc.) are relatively short-lived, which enables optimal use of imaging photons while keeping patient radiation dose low. Furthermore, many of these isotopes can be incorporated into biological substrates (glucose, H_2O, NH_3, CO_2, O_2, etc.) and pharmaceuticals, without altering their biological activity.

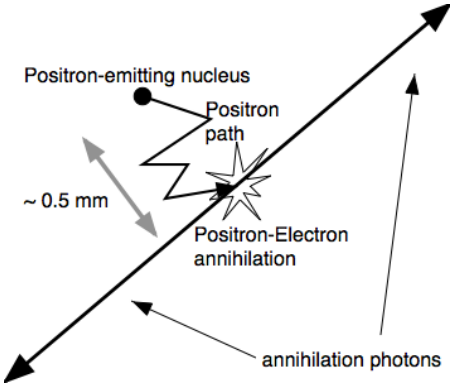

Fig. 1. General principle of PET imaging: decay of radionuclide, positron (β^+) emission, multiple scatter in tissue, annihilation with electron, and production of two back-to-back 511 keV annihilation photons. (Not to scale.)

Compared to CT and MR images, PET images appear much blurrier and or noisier, due to the relatively limited number of photons that can be collected during an imaging study. In addition, detector resolution is poorer due to the detector physics. X-ray CT scanners can easily resolve points less than 1 mm in size, while PET scanners cannot reliably resolve point sources smaller than 4-5 mm at best, and closer to 10 mm in practice. However, this does not impair their high sensitivity to focal tracer concentrations or their usefulness in accurate quantitative functional imaging.

In this chapter we give a very basic introduction to the physics of PET imaging. Several textbooks provide a more in-depth treatment [1-3].

1. Radioactive Decay

- General Principles

Radioactive isotopes are atoms whose inner core, their nucleus, is unstable, i.e. in a state with too much energy. Nuclei consist of a densely packed arrangement of protons and neutrons. By undergoing decay, the nuclei change their composition and properties to arrive in a less energetic and more stable state

The decay process follows an exponential law, i.e. the number of decays per second is always proportional to the number of un-decayed nuclei present. The same is true for the rate of decay, also called activity, which is determined by the half-life of the particular nuclide – the time it takes for half of the original nuclei to decay. Most common in PET is ^{18}F, which has a half-life of 109 minutes. After some time t, the activity left, $A(t)$, is proportional to the initial number, $A(0)$, and an exponential term involving the half-life, τ, of the nuclide:

$$A(t) = A(0)e^{-t(\ln 2 / \tau)}.$$

Radioactive rates or activity are measured in units of Becquerel (1 Bq = 1 decay/second) in the SI system or the traditional Curie (1 Ci = $3.7 \cdot 10^{10}$ decays/sec). A common scale factor used in the clinic is 1 mCi = 37 MBq.

- Positron Emission and Annihilation

In β^+ (positron) decay (Fig. 1), a nuclide transforms one of its core protons (p) into a neutron (n) and emits a positron (β^+), essentially a positively charged electron, and a neutrino (ν): $p \rightarrow n + \beta^+ + \nu$. The average positron range in matter depends on the positron's energy and material characteristics, such as the density and the atomic number. For ^{18}F-FDG, positron ranges are rather short, typically less than a millimeter.

At the end of its path, the positron, being anti-matter to electrons, will annihilate (recombine) with an atomic electron. In the annihilation, electron and positron convert

their mass into energy and produce a pair of 511 keV *annihilation photons* traveling in opposite directions. The 511 keV photon energy (*E*) comes from Einstein's famous equation $E = mc^2$, where *m* is the mass of the electron or positron (a very small number) and *c* is the speed of light (a very large number - squared) This annihilation radiation is what is detected in PET and what is used to form images of tracer concentration in the body.

- Interaction of Photons with Matter

The dominant annihilation photon interaction in human tissue is *Compton scatter*. The photon interacts with an electron, ejecting it from its atomic shell. The photon experiences a loss of energy and an associated change of direction, typically out of the detector, and so is unavailable for image formation.

Compton scatter and other interactions lead to an attenuation of the annihilation photons. The number of photons that are observed in a straight line from where they were produced decreases exponentially with increasing length of the material traversed. The thickness of soft tissue required to reduce the intensity of a beam by one half is approximately 7 cm, as opposed to 3-4 cm for x-rays. This after ~14 cm of soft tissue the 511 keV annihilation photon flux would be reduced to ¼ of its original intensity, and through the abdomen the photon flux can be reduced to 1/50 of its original intensity. Thus attenuation is often the dominant factor in PET image quality, especially for thicker patients.

2. Data Acquisition

Photon detection and Scintillation detectors

The general goal of photon detection is to measure the total energy deposited by the photon when it traverses the detector. For highest sensitivity and accuracy all of the photon's energy should be deposited, but in practice this is not always possible.

In most PET scanners today, scintillation detectors are used as detection elements. They couple inorganic scintillation crystals that emit visible or near ultraviolet light after interaction with an incident high-energy (511 keV) photon, to photo detectors that detect and measure the scintillation photons.

In scintillation crystals, the incident annihilation photon (nominally 511,000 eV energy) interacts and creates tens of thousands visible wavelength photons (about 1 eV energy each) in a very short flash or 'scintillation'. The number of scintillation photons produced in the crystal is proportional to the energy deposited by the annihilation photon.

Scintillators for PET photon detection can be rated on four of their characteristic properties:

The *stopping power* is the inverse of the mean distance traveled by photons before they deposit energy in the crystal. This length depends on density and effective atomic number (Z) of the material. A short travel distance is favorable, because it will yield more interactions with the 511 keV photons and a better efficiency for detecting them in crystal of fixed size.

The *decay constant* describes how long the scintillation flash lasts in the crystal. Shorter decay constants are desirable, because they allow for counting higher photon rates and lower background rates.

A good *energy resolution*, i.e. a small ratio of energy variance over energy, means that there are only small fluctuations in the energy measurement. This gives a means to distinguish against PET photons that have Compton scattered (and lost energy) before being measured. The energy resolution depends on the light output and the intrinsic energy resolution of the crystal.

The *light output* as the name indicates is the number of scintillation photons produced by each incident photon. Again this should be as high as possible, allowing the best spatial and energy resolution.

The most commonly used PET scintillators are listed in Table 1. Other materials are being evaluated (e.g. LaBr). Manufacturers are divided on the choice of material: Currently, BGO and LYSO is favored by GE, LSO by Siemens, and LYSO by Philips. Recently developed time-of-flight PET scanners (TOF-PET) use the scintillator LYSO, which has properties that are very similar to LSO.

Table 1. Scintillators used in PET Scanners.

Material	Cost	Light Output[1]	Effective Density[2]	Light Decay Time[3]	Comments
NaI(Tl)	cheap (relatively)	highest	lowest	long	Hygroscopic No longer used
BGO	expensive	lowest	highest	long	Does not support TOF PET
LSO (or LYSO)	more expensive	high	high	very short	Some patent disputes
GSO	more expensive	very high	somewhat lower than LSO	very short	No longer used

[1] determines energy and spatial resolution
[2] determines scanner sensitivity
[3] determines scanner deadtime and random coincidences rate as well as ability to be used with time-of-flight (TOF) PET imaging

Abbreviations: BGO = bismuth germinate, NaI(Tl) = thalium-doped sodium iodide, LSO = lutetium oxyorthosilicate, LYSO = lutetium yttrium orthosilicate, GSO = gadolinium orthosilicate

The most commonly used photo detectors for PET are photo-multiplier tubes (PMTs). PMTs are vacuum tubes with a photo cathode where incoming light photons produce electrons that are accelerated and amplified. The resulting electrical current is proportional to the number of initial scintillation photons and therefore to the energy deposited in the scintillation crystal by the PET photon.

By segmenting the scintillator blocks, using many small PMTs, or exploiting the properties of position sensitive PMTs, the location of the photon detection can be determined. The most commonly used setup today is the *block detector* (Fig. 2). Here, small individual scintillation crystals, a few millimeters in size where they face the patient, are tightly packed into blocks, which are typically coupled to four or more small photo-multiplier tubes. To determine the interaction position of the annihilation photon from the spread-out scintillation photon signals, the relative outputs from the PMT signals are compared. The calculated location then determines the crystal element that the photon is assigned to.

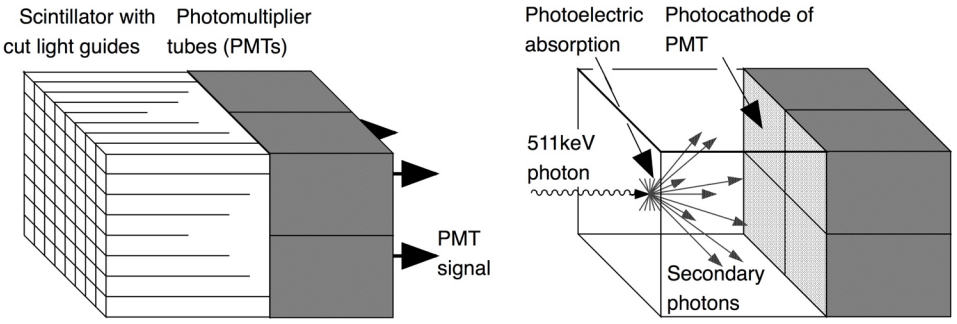

Fig. 2. Schematic of a block detector with finely segmented scintillator crystals read out by four photo-multiplier tubes.

Spatial resolution (in the detector) of a few millimeters is possible with this scheme, as it is determined by the size of the crystal cross-section.

A full PET scanner is constructed as a cylindrical assembly of block detectors in a ring structure several blocks deep. The sensitive volume inside the detector cylinder that a patient can occupy is called the field-of-view (FOV), which in human scanners is typically 70 cm in diameter and 16 – 18 cm in axial length (Fig. 3).

- Coincident Photon Events

An important advantage of PET imaging is the fact that due to the positron annihilation we expect to observe two photons at roughly the same time (in coincidence) in the detector ring. The annihilation event, i.e. the radioactive tracer, will then be located somewhere on the line connecting the two photon-detection points, as on the left side of Figure 3. This knowledge of the photon direction is a huge advantage over single photon emission tomography (SPECT) where collimators have to be used to restrict possible photon directions at the detectors at the cost of a large reduction in sensitivity.

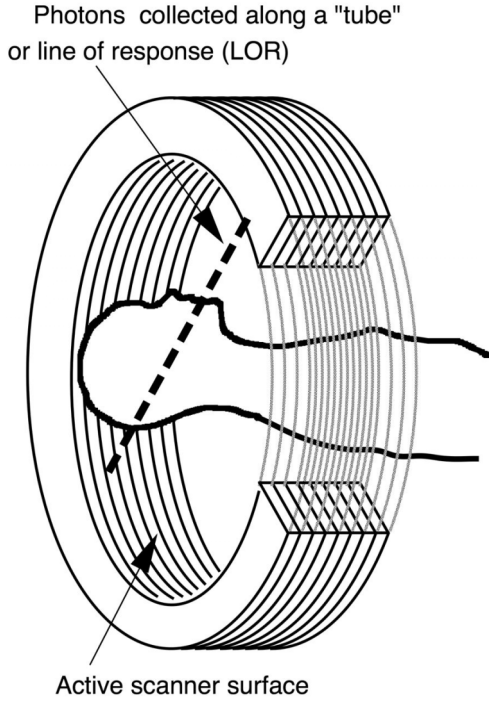

Fig. 3. PET scanner schematic with a possible line-of-response.

There are several factors that lead to the photon detections not occurring at the exact same time: The annihilation may occur closer to one detector surface than the other, which will result in a slight but measurable delay of one photon, where the photons travel at the speed of light, or 1 m in 3.3 ns. Most important for temporal mismatches is the finite timing resolution of the detector, i.e. its timing uncertainty, which arises from the decay time of the scintillation in the crystal and the processing time of the PMT signals. These effects lead to the use of a coincidence time window on the order of 6-10 ns.

If two photons are detected within each other's coincidence window they are assumed to be arising from the same annihilation and an event is attributed to the line-of-response (LOR) that connects the two detection points in the sensitive imaging volume (Fig. 4a).

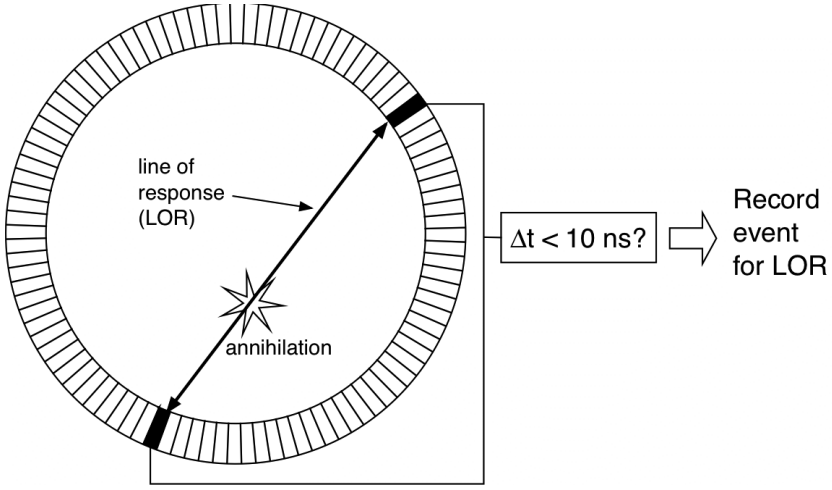

Fig. 4a. Coincidence processing in PET data acquisition.

With *time-of-flight* PET imaging the relative time difference (Δt) between the detection of the two annihilation photons is used to determine the most likely location (d) of the annihilation event along the LOR (Fig. 4b) where c is the speed of light. Time-of-flight PET imaging was previously investigated, but was never adopted. The recent development of scintillators suitable for time-of-flight PET (i.e. LYSO) combined with advances in timing resolution and timing stability of detector electronics have lead to a resurgence of interest in time-of-flight PET scanners, with commercial models being introduced. The advantage of time-of-flight PET, however, remains to be evaluated over the next few years.

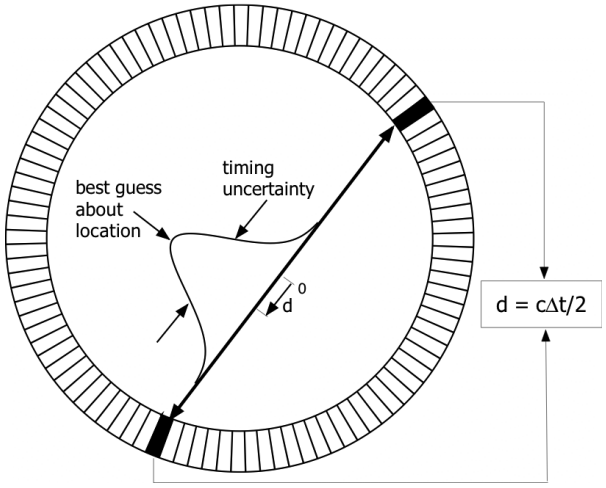

Fig. 4b. Coincidence processing in time-of-flight (TOF) PET data acquisition.

The detected coincidence events (called *coincidences*) can be classified into true coincidences and background events (Fig. 5). The latter are distinguished as either accidental (or *random*) coincidences where the two photons did not arise from the same annihilation event – or coincidences that did originate from the same annihilation, but where the true annihilation position does not lie on the line connecting the two photon positions, because one photon has experienced Compton scatter within the patient and therefore has had a change of direction (*scattered coincidence*s).

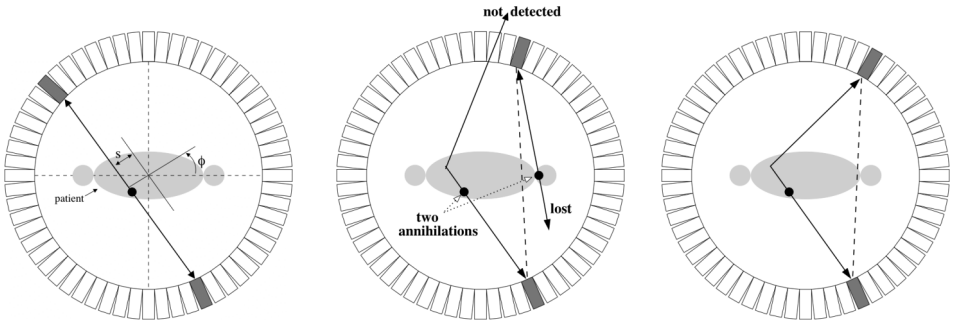

Fig. 5. Types of coincident events. From left to right: True coincidence, random (accidental) coincidence and scattered coincidence. In the last two types, the annihilation event (marked with a black circle) does not lie on the apparent line of response between the two photon detections.

- Sinograms

In the scanner, coincidence events are observed and identified along their lines of response (LORs) between pairs of detector elements (Fig. 5, left). To organize these raw data as they are acquired, the LORs are stored in a in such a way that all the LORs passing through a single point in the patient trace a sinusoid curve in the raw data histogram, hence the term *sinogram* for the raw data format. The formation of sinograms is an important middle step in the PET data acquisition process, since necessary corrections are often applied at this level.

- Data Corrections

The PET data acquisition process is not a perfect one. Interactions in the patient attenuate the emitted photons, detector elements vary in their detection efficiency, and random and scattered coincidences are recorded along with the true coincidence events. These effects need to be corrected to obtain clinically useful images and accurate quantitative information from PET studies.

The most important of the corrections is *attenuation correction* (AC): Photons that encounter more or denser material on their path from the annihilation site to the detectors are more likely to be absorbed or scattered (i.e. attenuated) than photons that travel through sparser parts of the body. If images are reconstructed from sinograms without AC, this can lead to less dense areas, like the lungs, appearing darker (emitting more photons) than surrounding denser tissue, like the mediastinum (Fig. 6). This is clearly an artifact that arises from the fact that lung tissue exhibits lower attenuation, not higher uptake. It not only impairs the visual appearance of the image, but also leads to wildly inaccurate quantitation of tracer uptake. To apply attenuation correction, it is necessary to determine the attenuation through the patient for all LORs. On stand-alone PET scanners, this is done with a *transmission scan* where an external positron source is rotated around the patient and the attenuation of the transmitted photons is determined. In PET/CT scanners, the acquired CT image is used for PET attenuation correction.

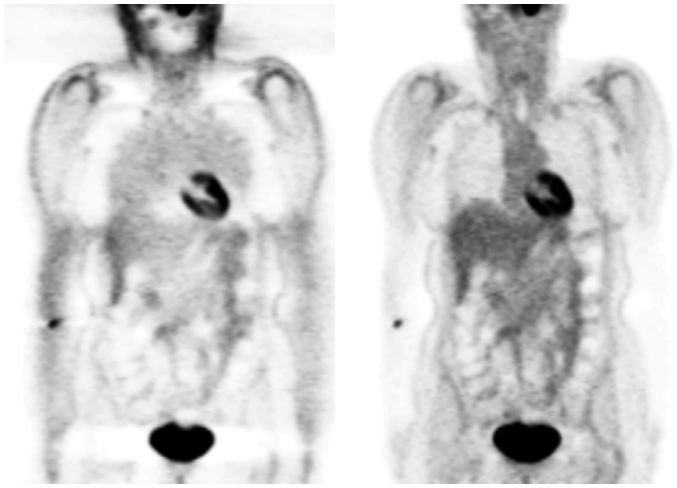

Fig. 6. Whole-body PET image without (left) and with (right) attenuation correction. Artifacts from not performing attenuation correction include the lungs and skin showing higher tracer uptake than muscle. In these images, darker regions represent higher tracer uptake using the common inverse-grey color table.

- 2D versus 3D acquisitions

Axially, PET scanners consist of several rings of detector elements that may or may not be separated by thin annular rings or *septa* of photon-absorptive material, typically tungsten, that provide *collimation*. With collimation, all data is acquired in 2-dimensional slices between the septa. This type of acquisition is therefore called *2D*, even though the reconstructed stack of images gives 3-dimensional information about the tracer uptake throughout the patient. When a scanner is operated without collimation (i.e. no septa), coincidences from all axial angles in the FOV will be accepted, making this a *fully-3D*

acquisition protocol. Data storage, correction, and image reconstruction is considerably more complex in the fully-3D case. Current PET scanners operate in either 2D-only or 3D-only mode, or in a 2D/3D mode for those with retractable septa.

Figure 7 shows the effect that collimation has on the acquisition of coincidence counts: The septa block a fairly large number of true coincidences from ever reaching the detector surface, decreasing sensitivity. However, they also reduce scattered and random coincidences, thus improving contrast. Of special importance here are accidental counts that partly originate from outside of the area between the detector surfaces (true FOV), because without collimation, the scanner is sensitive to activity from a very large area outside the true FOV. The decision on whether to use 2D or fully 3D acquisitions is still under debate, weighing the reduction of background counts against the loss of sensitivity. Brain imaging – typically with small activity concentrations outside the true FOV – is a clear indication for 3D imaging with increased sensitivity, while whole-body imaging – with usually much more activity directly surrounding the true FOV – does not show a clear preference and is usually done in 2D mode on scanners with 2D/3D capabilities.

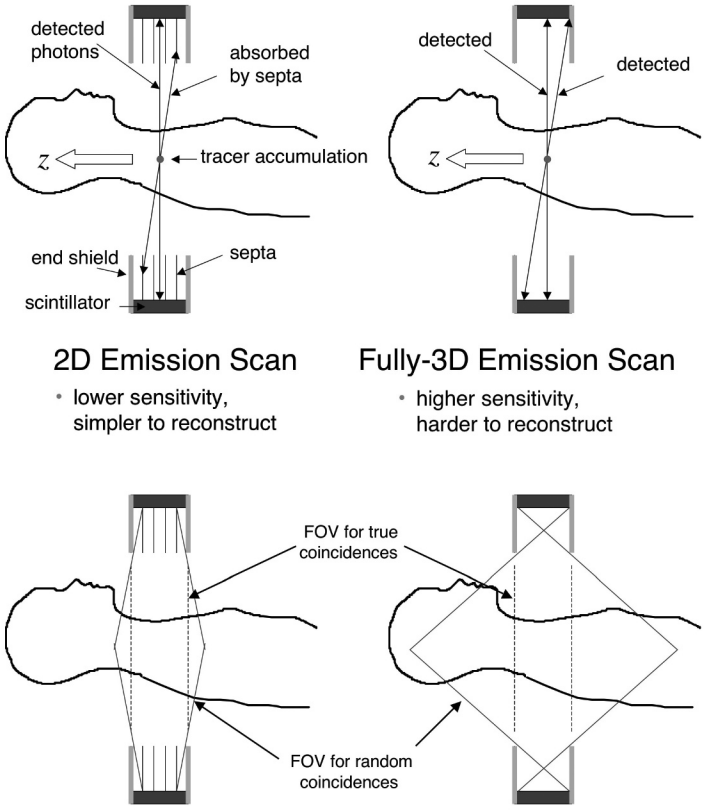

Fig. 7. Effect of 2D and 3D acquisition protocols on true and background coincidence counts.

3. Image Reconstruction

After the acquisition of PET data in sinograms and their corrections for attenuation and other effects as described above, the next stage in the PET processing chain is to reconstruct an estimate of the in vivo tracer distribution. This process of *image reconstruction* is the most mathematically complex step and is well described elsewhere [1-3]. Here we point out the differences between the two most common methods: filtered-backprojection (FBP), which is a well-established method, and ordered-subsets expectation maximization (OSEM), which is a more recent iterative approach that allows for a more accurate model of the PET acquisition process. Figure 8 shows a visual comparison of FBP and OSEM images, reconstructed from the same patient sinogram. FBP is well-understood and robust (it is also used in CT), but it does not account for noise.

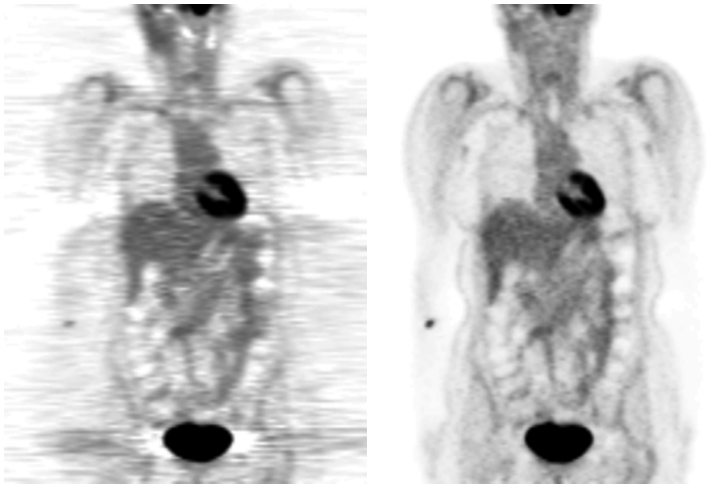

Fig. 8. Comparison of coronal sections of PET images reconstructed with FBP (left) and OSEM (right, same as Fig. 6 - right). The FBP image shows characteristic streak artifacts and appears overall noisier than the OSEM image.

Modeling the noise in the data sets generates much more complicated sets of equations that can only be solved iteratively, such as with the expectation maximization (EM) algorithm. This process is too slow for clinical needs, but with the advent of the ordered-subsets acceleration of the EM algorithm (OSEM) and faster processors, iterative methods are becoming more common. In many PET centers OSEM is now the reconstruction method of choice.

- Noise/Resolution Tradeoffs and Image Quality

If images appear very noisy after reconstruction, they may be *smoothed* in a further step to give the eye an easier task, especially for localizing disease. However, since smoothing

averages together neighboring image pixels, it is connected to a loss of resolution, so small structures may not be distinguishable anymore. Figures 9 and 10 illustrate this connection. It is task-dependent, observer-dependent, and non-trivial to define an optimal region in the noise/resolution space. Currently there are no standards for adjusting this trade-off.

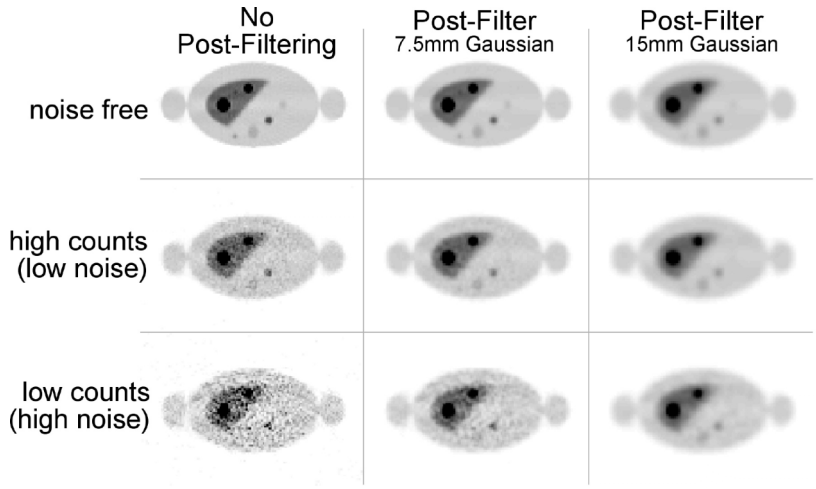

Fig. 9. OSEM images showing the tradeoffs between noise (increasing downwards) and smoothing (increasing to the right).

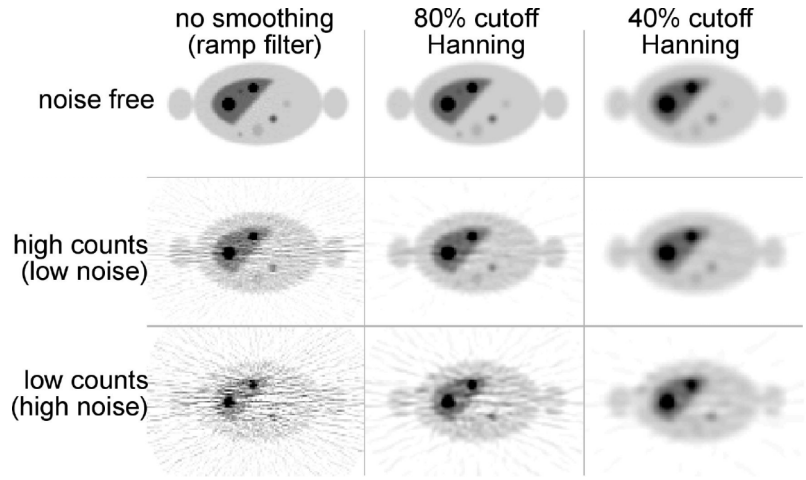

Fig. 10. FBP images showing the tradeoffs between noise (increasing downwards) and smoothing (increasing to the right).

4. PET/CT Scanner Design

The primary purpose of combining CT and PET systems in a single scanner is the precise anatomical localization of regions identified on the PET tracer uptake images. Although

it is possible to use non-rigid image registration to align separately acquired whole-body PET and CT images, challenges remain in the practical implementation and validation of software-based methods. In recent years, the advent of combined PET/CT systems has pushed dedicated pure PET scanners almost completely off the market due to the convenience and ease of creating co-registered PET and CT images for oncology, radiation oncology, and cardiology applications.

In this section we give a brief overview of this new technology; for more details, the reader is referred to [4].

- Basic Components

PET/CT systems are combinations of stand-alone CT and PET scanners in one gantry with a shared patient bed. They must be able to be very accurately aligned. The patient bed is an important and non-trivial component, because there should be no differential bending in the bed between the PET and CT scans. Figure 11 shows a schematic of a PET/CT scanner.

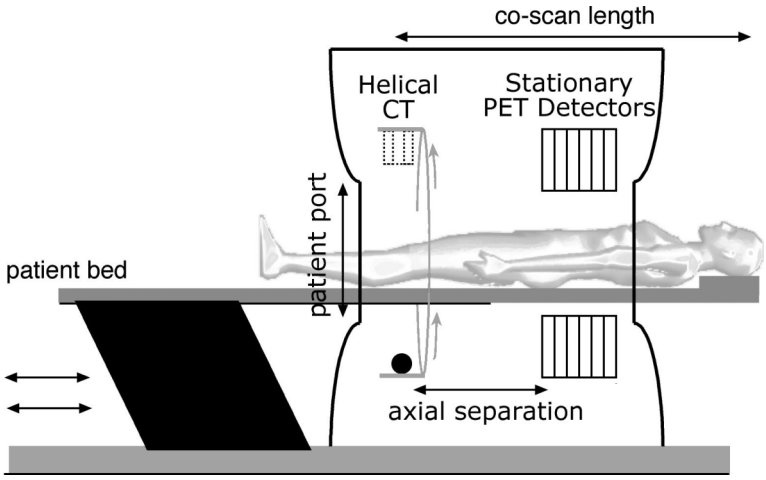

Fig. 11. Illustration of the main components of a PET/CT scanner.

The typical PET/CT protocol begins with a CT scout scan to define the scan area, followed by a helical CT scan, and finally the PET scan. The entire acquisition sequence is typically controlled from the CT console. However, each sub-system has its separate acquisition system with 3 or more databases overall (CT, PET, and PET/CT display), making the process somewhat complicated with the current generation of scanners.

The data flow in the combined PET/CT acquisition is outlined in the schematic in Fig. 12: The x-ray CT scan provides anatomical images that after some processing can also be used for attenuation correction in PET, and the PET/CT software can display both images side-by-side or overlaid (fused), as in Fig. 13. It is to be noted that there are no 'fused'

images in PET/CT – the PET and CT image always remain separate entities. Displaying them together is an overlay process rather than creating a new type of image.

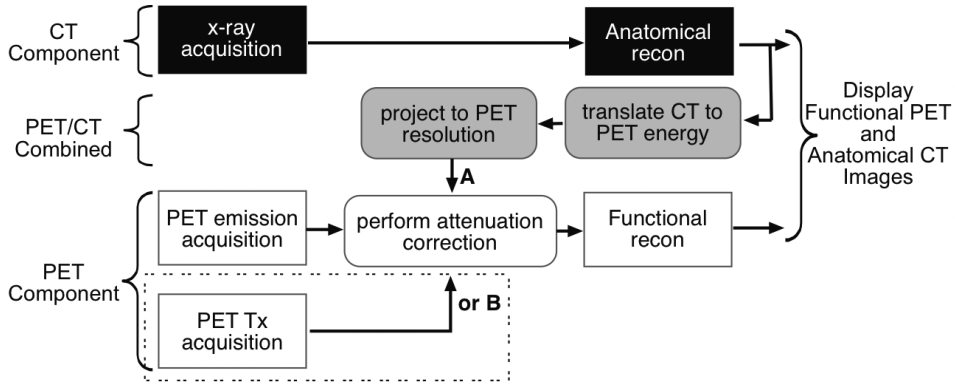

Fig. 12. Data flow in a PET/CT scanner.

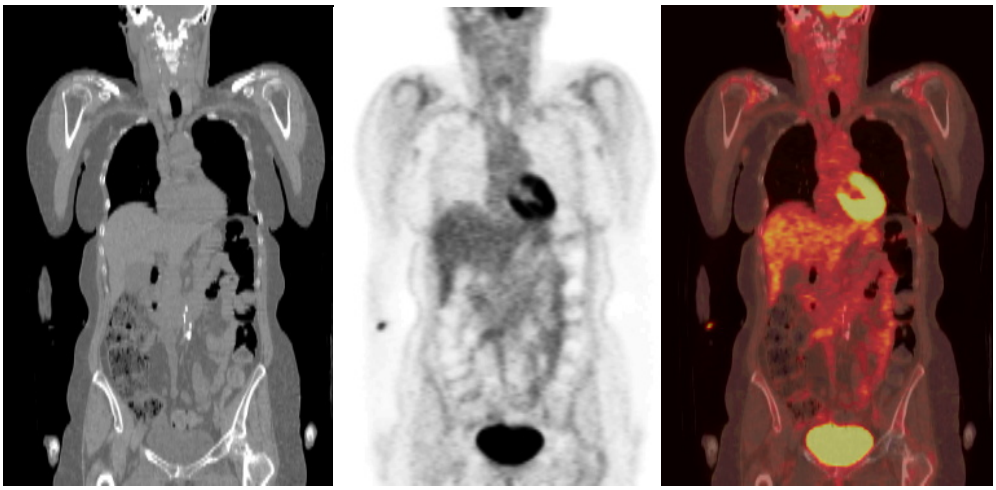

Fig. 13. Images from a PET/CT scanner: anatomical CT image (left), functional PET (middle, right, same as Fig. 8 - right), and overlaid images (right) of a whole-body scan.

- CT-based Attenuation Correction (CTAC)

An important synergy of PET/CT scanners is the use of the CT images for attenuation correction of the PET emission data. All manufacturers of PET/CT scanners incorporate x-ray CT-based attenuation correction algorithms in their systems, and for newer PET/CT scanners it is the only option offered. CT-based attenuation correction offers the significant advantage that the CT data has much lower statistical noise and can be acquired in a shorter time than a standard PET transmission scan. CT transmission scans

can also be acquired after the PET tracer is injected, giving the ability to collect unbiased post-injection transmission scans. This shortens the time spent by a patient on the scanner bed and provides more efficient use of scanner time.

To be used for attenuation correction, the CT data must be transformed to an estimate of the attenuation coefficients at 511 keV. In the bilinear scaling method [4], the attenuation map at 511 keV is estimated by using separate scaling factors for bone and non-bone components based on the CT image values as shown in Figure 14.

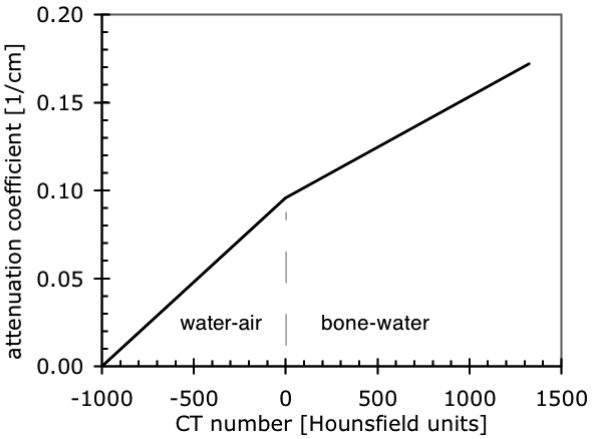

Fig. 14. Bilinear scaling transform used to convert CT image values to PET attenuation coefficients.

There is, however, no unique transformation from CT energies (~30 to 140 keV) to 511 keV due to the possibility of independent variations in density and atomic number Z, which can allow two materials with different atomic numbers to have similar CT values and different attenuation coefficients at 511 keV. Conversely, it is possible for two distinct materials with the same value of attenuation coefficient at 511 keV to yield different CT numbers, and errors in the CTAC image will propagate to errors in the PET image in the same location. This case arises if contrast agent or metallic or high-density implants are present in the CT image. Also, if there are positional mismatches between the PET and CT images due to, for example, respiratory motion, there will also be errors introduced into the PET image. Thus, while CT-based attenuation correction can lead to significant improvements in PET image quality, artifacts can arise from the presence of contrast agent, metallic or high-density implants, and respiratory motion.

References

[1] Cherry SR, Sorensen JA, Phelps ME. *Physics in Nuclear Medicine*. Orlando, FL: Grune & Stratton, Inc.; (2003).

[2] Valk PE, Bailey DE, Townsend DW, Maisey MN. *Positron Emission Tomography: Basic Science and Clinical Practice*. London: Springer-Verlag; 2003.

[3] Wernick MN, Aarsvold JN. *Emission Tomography*. San Diego, CA: Elsevier Academic Press; 2004.

[4] Alessio AM, Kinahan PE, Cheng PM, Vesselle H, Karp JS. PET/CT scanner instrumentation, challenges, and solutions. *Radiol Clin N Am*. 2004;42(6):1017-32.