

Chapter One

PET Background

An understanding of positron emission tomography and the various image processing techniques involved in the preparation of images of blood flow is essential to their analysis.

In this chapter we outline the fundamentals of positron emission tomography to provide the background to the data we consider in subsequent chapters. We review the statistical aspects in detail. The overall aim is to provide a general background to PET, at a level of exposition suitable for the newcomer.

Those familiar with the concepts and terminology of functional mapping may wish to skip most of this chapter, perhaps reading only the description of the “V5” data set (§1.7.) before proceeding to chapter 2. Similarly, the reader with little time may prefer to accept that images whose intensity is indicative of regional neuronal activity can be obtained with PET, note the description of the “V5” study, and also proceed to chapter 2.

1.1. Additional Reading

Overviews of PET

There exist various reviews of PET and Neuroimaging that may be of interest. Ter-Pogossian *et al.* (1980) review PET in its early days. They describe the mechanics of PET: the preparation of isotopes, physics of electron annihilation and event detection, and reconstruction. The description of filtered back-projection is especially clear. Of particular interest is their discussion of tracer kinetics, and the problems of obtaining quantitative images from tracer density. They look forward to the “future” of PET, concentrating on the study of the heart, and their discussion provides an introduction to heart-PET.

The paper by Raichle (1983b) briefly covers the main ingredients of PET. This paper also contains an excellent review of tracer kinetics, and the assumptions inherent in obtaining quantitative measurements of metabolic processes, and ends with a section detailing possible clinical applications of PET.

Brownell *et al.* (1982) gives an even briefer review of PET. Despite the age of this paper, it does provide a useful overview of nuclear Magnetic Resonance Imaging (MRI), and contains extensive references for the theory of PET and MRI.

State of the art

“Positron Emission Tomography and Autoradiography”, edited by Phelps *et al.* (1986) is a collection of chapters written by experts in the various disciplines associated with positron emission tomography. This book represented the state of the art in the mid-eighties, and still constitutes the single most thorough exposition of PET methodology.

For more recent information, the published proceedings of recent conferences are useful. The “Brain PET” bi-annual conference was first held during June 1993 in Akita, Japan, with proceedings “Quantification of Brain Function: Tracer Kinetics and Image Analysis in Brain PET” edited by Uemura *et al.* (1993).

Journals

The main journals for the publication of PET methodology are the *Journal of Cerebral Blood Flow and Metabolism* and the *Journal of Computer Assisted Tomography*. Recently *Human Brain Mapping* has emerged, a journal dedicated to functional neuroimaging to which many researchers are now sending their papers, particularly those working in functional MRI (fMRI).

1.2. Data Acquisition

1.2.1. The Tomograph

The positron tomograph essentially consists of a cylinder of bismuth germanate ($\text{Bi}_4\text{Ge}_3\text{O}_{12}$) detectors arranged in rings. A recent dedicated brain scanner (ECAT 953B, Siemens/CTI, Knoxville, Tennessee, described by Spinks *et al.*, 1991) has 16 rings, each of 384 detectors of size $5.6 \times 56.45 \times 30 \text{mm}^3$. The detectors are grouped into blocks of 8×8 , mounted on a set of four photomultiplier tubes. The detector hit in an 8×8 block is obtained by combining the readings from the four photomultipliers, using Anger logic. The ring diameter is 76cm, with an axial length of 10.8cm. The patient port size is 35cm.

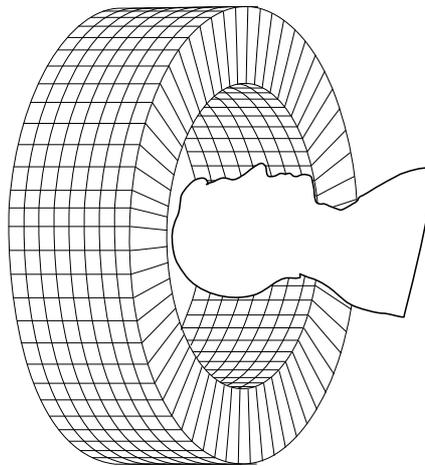


Figure 1

Schematic of the arrangement of detectors in a PET tomograph

1.2.2. Data acquisition

Subject preparation, isotope decay, electron annihilation

The subject is positioned in the tomograph, using external landmarks (eyes, ears) for rough alignment. A chemical labelled with a positron emitting isotope is introduced into the subject, either intravenously or by inhalation of a gas.

The radioisotope decays, each decay event resulting in the emission of a positron. The positron loses kinetic energy as it travels through the surrounding tissue. After about 1–6mm the positron is highly susceptible to interaction with an electron. This results in the annihilation of both particles, their mass being conserved as energy by the emission of two photons of gamma radiation, each of energy 511keV. Conservation of momentum dictates that the two photons will be emitted in almost opposite directions (fig.2), the colliding positron having little momentum. The axis of the pair of emitted gamma photons is completely randomly oriented.

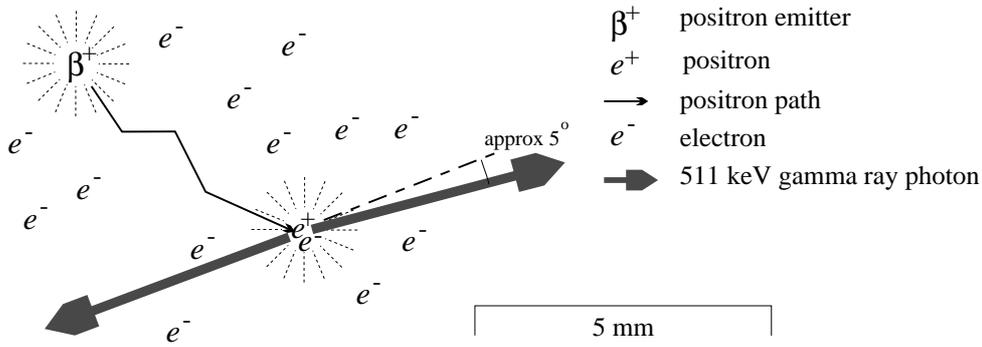


Figure 2
Positron emission, electron annihilation and gamma ray emission

Detection of annihilation events, VORs

These two photons of gamma radiation then travel through the tissue and air. If the two photons are detected by the tomograph, then two crystals will be hit at approximately the same time. These two crystals then define a volume of response (VOR) in which the electron annihilation must have taken place (fig.3). The tomograph computer counts the number of times the two detectors defining each VOR are hit (almost) simultaneously during the scanning period.

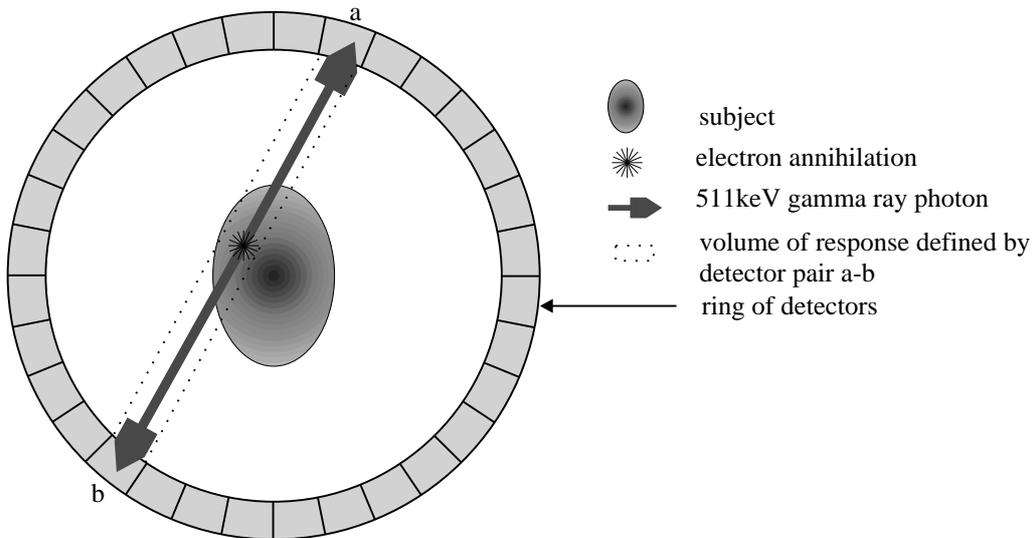


Figure 3
Schematic of detection in a PET tomograph

1.2.3. Sinograms

These VOR counts are stored as sinograms. A sinogram is an image of the VOR counts for a particular ring of detectors (direct plane sinograms), or for a pair of rings (cross plane sinograms). For a 16 ring tomograph, 256 sinograms are acquired, 16 direct plane sinograms and 16×15 cross plane sinograms.

A direct plane sinogram is constructed as follows: Let the number of parallel VORs possible be p ; and the number of detectors in the ring $2n$. Then from a given starting

point in the ring, consider the p parallel VORs defined by that detector, and the $p-1$ detectors round from it, with the p opposing detectors (fig.4). The counts for these p VOR constitute the pixel values for the first row of the sinogram. Subsequent rows of the sinogram are obtained by considering the p parallel VORs defined by the p detectors one along from the previous set. There are n sets of p parallel VORs for a single ring, giving an $n \times p$ sinogram. Cross ring sinograms are similarly formed: Each row consists of the counts for the p parallel VOR defined by p detectors in one ring and the opposing detectors in the other ring. There are $2n$ sets of p parallel VOR, giving a $2n \times p$ sinogram, which is split up into two of dimension $n \times p$.

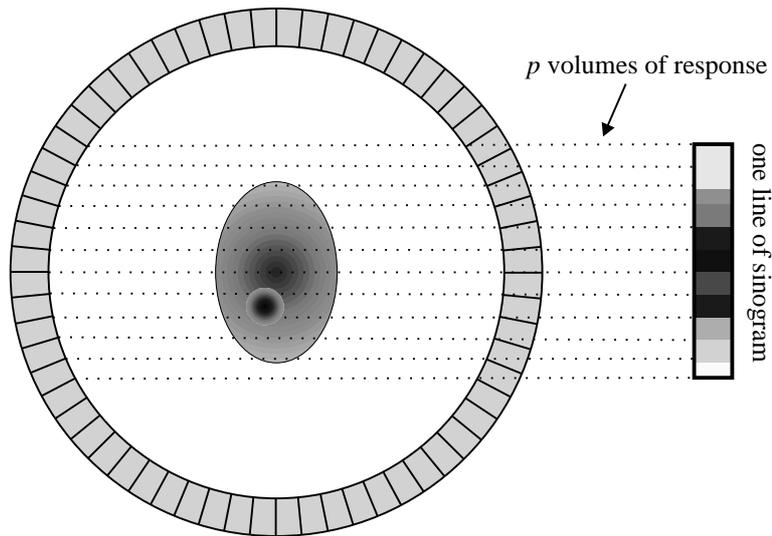


Figure 4
Construction of one line of a sinogram

A direct plane sinogram for a non-central point source will be a sine wave (fig.5). Hence the name sinogram. For the ECAT 953B n is 192 and p is 160.

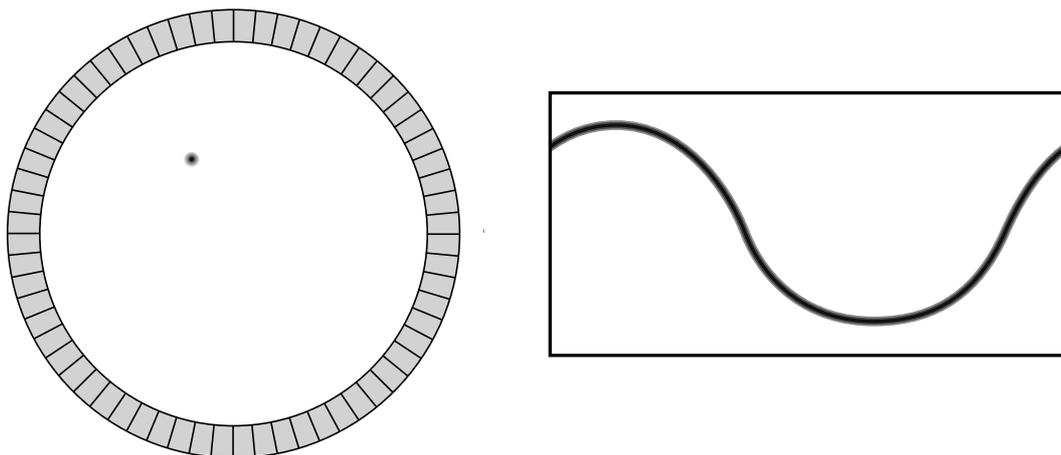


Figure 5
Sinogram for a non-central point source



Figure 6

Direct plane sinogram from a scan of the authors brain. Data frame energy window sinogram 64 of run 12 on subject n278. Acquired on an ECAT-953B dedicated brain scanner, February 1992, during a visit to the Medical Research Council Cyclotron Unit at the Hammersmith Hospital, London.

1.2.4. Additional data

In addition to the VOR counts for the scan of the subject, a number of additional measurements have to be taken. These are to allow corrections for detector efficiency, attenuation, background radiation, scatter and randoms. We briefly describe each in turn, to give some insight into some of the more problematic physical aspects of PET scanning. A comprehensive discussion of quantification for positron emission tomography can be found in Hoffman and Phelps (1986).

Normalisation–Blank scan

It is impossible to manufacture the detector crystals and associated electronics such that all the detectors have the same sensitivity. Discrepancies in detector sensitivities show up on sinograms as light sinusoidal “streaks” (fig.6). To correct for the relative differences in sensitivity a *blank scan* is acquired at the beginning of each day.

During the blank scan radioactivity is provided by three positron emitting rods aligned parallel to the axis of the tomograph, which are rotated about the axis of the tomograph to give the appearance of a uniform source to all the detectors. Acquisition is continued over a long period, so that the number of recorded events for each VOR is

great. The relative sensitivities of the detector pairs for each VOR, can then be accurately estimated by the differences in detection rates.

These relative sensitivities can be built into the statistical model for PET, or can be used to correct observed VOR counts by a simple scaling. The latter is usually used in practice.

Transmission Scan

Although the photons of gamma radiation emitted following the annihilation of positron and electron have a relatively high energy of 511keV, some are absorbed by the tissue through which the photons have to pass in order to reach the detector. The probability of *attenuation* is related to the density to gamma rays of the tissue along the line of flight. To estimate this density, a transmission scan is taken during the scanning session, usually at the beginning or the end.

The subject is positioned in the scanner as usual, but no tracer is introduced. The rotating rod sources provide a uniform source of gamma rays. Data is acquired for about three minutes, such that the detection rates for each VOR can again be regarded as accurately estimated by the observed detection rates. The ratio of this rate to the blank scan rate for each VOR is then an estimate of the probability of non-absorption along the VOR. These attenuation probabilities can then be built into the model for PET reconstruction, or can be used to correct the observed VOR counts by a simple scaling. Again the latter is usually used in practice.²

Background radiation

For a simple scan, two sets of sinograms are acquired. The scan is split into two time frames, the *background frame* and the *data frame*. The radiotracer is introduced at the beginning of the data frame. The sinograms for the background frame indicate the structure of background radiation and any residual activity from a previous scan.

The background frame may be used to estimate and remove background and residual activity. However, due to the low activity, the number of detected events is low, and estimates are highly variable. This is evidenced in the noisy appearance of reconstructed background frame images. Any linear correction of a data frame therefore adds this noise to the image.

²It is possible to reconstruct the three-dimensional linear attenuation function. This is known as *transmission tomography* and is considered again in the filtered back-projection subsection of the reconstruction section (§1.4.3.1., p.22).

Scatter

Scatter occurs when one (or both) of the emitted gamma photons are deflected during their passage to the detectors, possibly leading to detection in a VOR that does not contain the point of annihilation (fig.7).

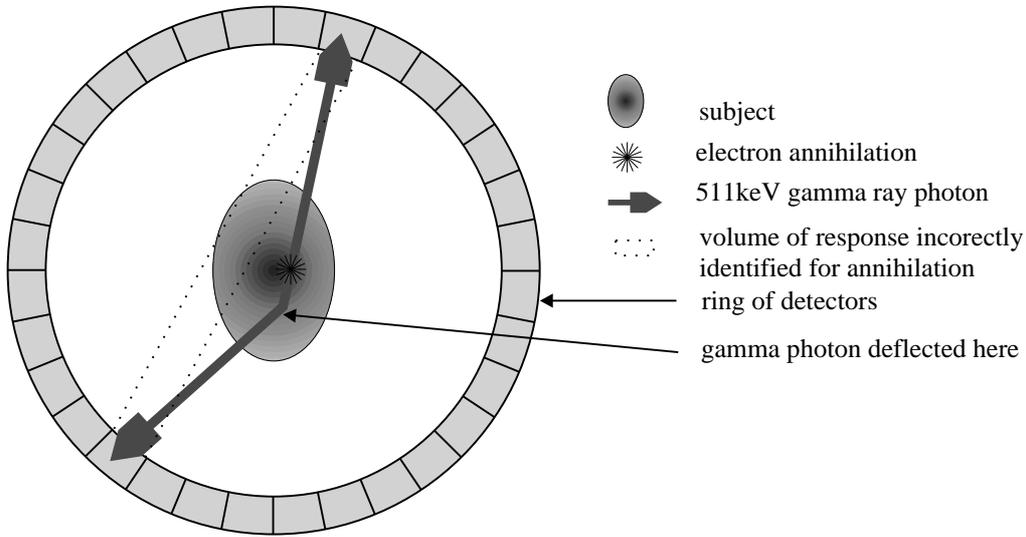


Figure 7
Scatter

To counter the effects of scatter, the energies of the detected photons can be used, as a photon will lose some of its 511keV of energy when it is scattered. If only photons with energies lying in a predefined energy window containing 511keV are considered, and a VOR count only recorded for almost simultaneous detection of two photons with energies lying in this window, then some scatter should be eliminated.

However, there will still be some scattered photons with a high enough energy to be in the energy window, as illustrated in fig.8. To enable a correction for this, two sets of VOR counts are collected. One set for coincident detection of photons with energies in the energy window, and one set for coincident detection of photons one or both of which have energies in the lower window.

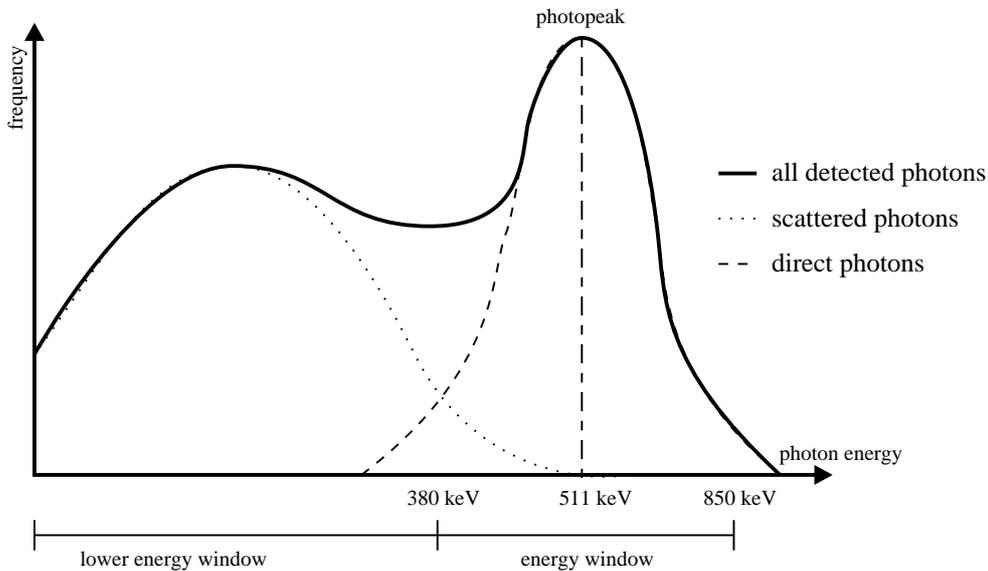


Figure 8
Energy distribution of detected gamma photons

The simplest correction proceeds by subtracting a fraction of the lower energy window sinograms from the upper energy window sinograms. (That is, a fraction of the VOR counts for the lower window are subtracted from the corresponding VOR counts for the upper energy window, for each VOR.) The fraction used is ideally the fraction of the scattered photon energy distribution appearing in the upper energy window, the *scatter fraction*. This can be estimated using phantom studies, in which sources of known radioactivity are scanned once when surrounded by water, and once in a vacuum, when there will be no scatter. The assumption implicit in such a correction is that the structure of the events in the lower energy window (predominantly scattered events) is shared by the scattered events of the upper energy window. Scatter is an artefact of the radiotracer density and the tissue density (to gamma rays) of the subject, so the assumption is not unfounded.

The low count rates for the lower energy window results in estimates of scatter that are very variable, so scatter correction adds noise. More sophisticated methods for scatter correction are the subject of current research.

Multiple incidences

A *multiple incidence* occurs when three or more detectors are hit (almost) simultaneously. In this situation, a VOR can not be identified, and the events are ignored.

Randoms

Randoms are multiple incidences, where two (nearly) simultaneous annihilations have exactly one photon detected. The VOR defined by the two detectors has its count falsely incremented.

1.3. 2D & 3D PET, SPECT

1.3.1. 2D & 3D PET

Until recently, algorithms for reconstructing three dimensional images from all possible VOR counts were not in common use. Reconstruction was only possible in two dimensions, reconstructing a slice of the subject from each direct plane sinogram. The slices were then stacked to obtain a three dimensional image. Only VORs corresponding to detector pairs in the same ring could be considered. Some limited cross ring VORs were allowed to enable planes between the rings to be considered, with *induced plane* sinograms formed by averaging counts between adjacent rings. This type of PET is known as 2D PET. For a 16 ring tomograph there would be only 31 sinograms, 16 direct plane sinograms, and 15 induced plane sinograms.

As 2D PET only needs VOR counts for direct plane and limited cross ring VOR counts, older scanners have fixed tungsten collimating septa around the detectors, reducing the view of each detector to detectors in the same and adjacent rings. This has the effect of reducing the detection of randoms, scatter, multiple incidences and background radiation. Scanners produced in the early nineties have removable septa, enabling data acquisition for 2D or 3D PET (fig.9), and new scanners operate solely in 3D mode.

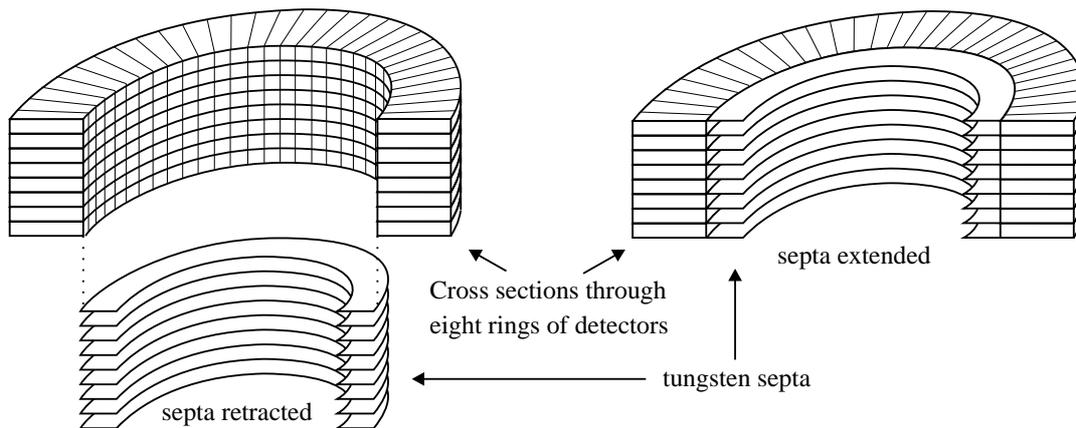


Figure 9
Septa

Sensitivity of 3D PET

In 2D mode, less than 1% of the annihilation photon pairs are detected (Cherry *et al.*, 1991, p655). Removal of the septa greatly enhances the sensitivity of the tomograph: The probability of both gamma photons from an annihilation event hitting detector blocks is on average about six times greater (Spinks *et al.*, 1991), with larger increases for annihilations occurring near the centre of the scanner.

For a subject with radioactivity suitable for a 2D study, the rate of photons hitting the detectors in 3D mode is so high that many are not detected because of the detector

dead time.³ This saturation of the tomograph is avoided by using lower doses of radioactivity, typically a quarter of the dose administered in 2D (Cherry *et al.*, 1991).⁴

Removal of the septa and acquisition of all lines of response makes the scanner more likely to detect scattered photons, multiple incidences, randoms and background radiation. So, raw count rates can not be used to compare the sensitivity of 2D and 3D modes of PET acquisition. Noise Equivalent Counts (NEC), due to Strother *et al.* (1990), measures count rate relative to image signal-to-noise ratios by accounting for scatter and randoms. Using this measure, Bailey *et al.* (1991) found that for activity concentrations suitable for a 3D acquisition ($< 1\mu\text{Ci/cc}$, corresponding to an injected dose of approximately 28mCi), the NEC was around five times that for 2D mode. Furthermore, this NEC was slightly better than that of a 2D acquisition with four times the activity. Cherry *et al.* (1993) present similar findings, using (grey matter) standard deviation in single difference images rather than NEC as the measure of scanner sensitivity.

Summary: 3D better than 2D

Thus 3D acquisition provides images of comparable quality to 2D mode, but requires the administration of less radiation to the subject per scan. This allows more scans to be taken on an individual for the same total injected dose, giving more data, and hence improving the power of statistical procedures to analyse data sets of images from such experiments. This has been evaluated in blood flow studies of the human brain by Bailey *et al.* (1993) in terms of NEC, and by Cherry *et al.* (1993) in terms of estimated standard deviation in mean difference images⁵ from phantom⁶ data. Both these papers report a two-fold increase in sensitivity of the 3D method using four times as many scans as the 2D method, for the same administered activity. Further fractionalisation of the dose gives further improvements, but leads to an unacceptable number of scans.

1.3.2. SPECT

There are radio-chemicals available that emit gamma rays directly on decay. A single photon is emitted, so a special tomograph is required, commonly referred to as a *gamma camera*. The radio-chemicals in question are less radioactive than positron emitters, do not require a high power cyclotron to produce them, and can be handled easily. The isotopes involved have long half lives, so tracers can be made remotely. This makes Single Photon Emission Computed Tomography (SPECT) a financially viable clinical tool, and most large hospitals possess a SPECT tomograph.

³When gamma photon hits the detector crystal, a scintillation event takes place causing the block to flash. The flash is detected by the photomultipliers which generate an electric signal. A scintillation event takes a small time to take place, during which incident gamma photons go undetected. This period is called the *dead time*

⁴About 50–80 mCi H_2^{15}O is injected into subjects to obtain a satisfactory 2D acquisition.

⁵Mean difference images are introduced in §2.3.1.1. where the simple paired *t*-test approach for multiple subject activation studies is discussed.

⁶A *phantom*, in radiography, is a shaped container containing a known concentration of activity. The most famous is the Hoffman brain phantom (Hoffman *et al.*, 1990).

The SPECT tomograph

Since only one gamma photon is emitted by the decaying isotope, the detectors of a SPECT tomograph need to be shielded by lead collimators so that each has a narrow field of view, enabling the identification of a VOR for a detected emission (fig.10).

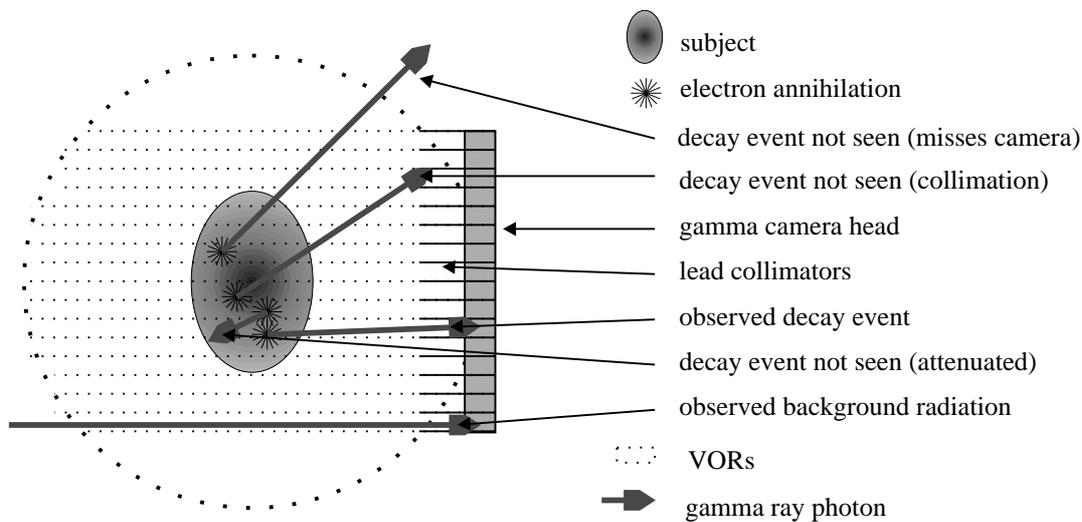


Figure 10

Schematic of a simple SPECT gamma camera. The camera head is positioned at various places around the subject to enable data to be collected for many VOR.

Resolution of SPECT scans

The resolution of SPECT imaging is much lower than that of PET for a number of reasons. The low energy gamma rays (140 KeV for technetium-99m, ^{99m}Tc) are very prone to attenuation, resulting in decay events not being observed. The restricted field of view of the detectors also reduces the probability of observing the gamma radiation emitted by a decaying isotope. As a result of this, and the low activity of SPECT tracers, a SPECT scan is usually acquired over quite a long period, typically in the region of twenty minutes. Even then, the number of observed decay events is still quite poor combined with PET, being a couple of orders of magnitude smaller. These low counts lead to highly variable estimates of tracer concentration, apparent in reconstructed images as noise. Background radiation is of comparable energy to the gamma rays of SPECT, and its detection adds to the noise in SPECT images. To obtain VOR counts of sufficient size for reasonable reconstruction, gamma cameras use larger detectors than a positron tomograph, resulting in fewer VORs and hence lower reconstructed resolution. For a comprehensive overview of SPECT, see English and Brown (1986).

Advantages of PET over SPECT

The poor quality, long scan times, and long half lives of the tracers (precluding multiple scans on the same individual) in SPECT make the expense of PET justifiable, at least for research. In addition the array of gamma emitting tracers is somewhat limited, and many biochemicals can only be labelled with positron emitters. PET has wider biological scope, and is a more powerful research tool.

1.4. Reconstruction

The raw data acquired by the tomograph are the VOR counts. These are the numbers of annihilations on a particular VOR resulting in detected gamma radiation along that VOR. From this data it is desired to estimate the regional decay rates, or regional activity (rA), as this indicates the distribution of the tracer.

1.4.1. A statistical model for PET

Let $\mathbf{n}^* = (n_1^*, \dots, n_r^*, \dots, n_R^*)$ be the vector of counts for the R VORs. These are clearly realisations of random variables, the form of which we now derive. The model follows that of Vardi *et al.* (1985).

Average emission intensity

The radioactive decay of the positron emitting isotopes in the subject forms an inhomogeneous Poisson process in both space and time. Let the intensity of this process be $\lambda(\mathbf{x}, t)$; for position $\mathbf{x} \in \Xi$ and time t , measured from the start of the scan, which we take to be of length T . Here Ξ is the subset of Euclidean 3-space \mathfrak{R}^3 corresponding to the image space of the tomograph, for suitable Cartesian axes.

If we assume a constant tracer density over the duration of the scan, then this can be expressed as $\lambda(\mathbf{x}, t) = \lambda(\mathbf{x}, 0) \exp(-\beta t)$ by the law of radioactive decay, where $\lambda(\mathbf{x}, 0)$ is the emission rate at the start of the scan. β is related to the half-life of the tracer, $t_{1/2}$, by $\beta = \ln(2)/t_{1/2}$, and is hence known.

Over the scan period, $t \in [0, T]$, the emission intensity at $\mathbf{x} \in \Xi$ is:

$$\begin{aligned} \int_0^T \lambda(\mathbf{x}, t) dt &= \int_0^T \lambda(\mathbf{x}, 0) \exp(-\beta t) dt \\ &= \frac{\lambda(\mathbf{x}, 0)(1 - e^{-\beta T})}{\beta} \\ &= \lambda(\mathbf{x}) \quad \text{say} \end{aligned} \tag{1}$$

so it suffices to estimate $\lambda(\mathbf{x})$, the *average emission intensity*.

Model for average emission intensity

Let $c(\mathbf{x}, r)$ be the probability that an emission from \mathbf{x} results in an electron annihilation detected in VOR r . Then, the emissions from point \mathbf{x} detected in VOR r , form a thinned Poisson process with thinning probability $c(\mathbf{x}, r)$, independently for all $\mathbf{x} \in \Xi$ and $r = 1, \dots, R$. Thus the number of detected emissions for the VORs, n_r^* , are independent Poisson random variables with means

$$E[n_r^*] = \lambda_r^* = \int_{\mathbf{x} \in \Xi} \lambda(\mathbf{x}) c(\mathbf{x}, r) d\mathbf{x} \tag{2}$$

and we want to estimate $\lambda(\mathbf{x})$ given n_r^* .

Discretisation of the problem

To simplify the mathematics, and to avoid the estimation of a continuous function, it is usual to consider Ξ divided up into K voxels⁷ (volume elements), usually cuboid in shape. Within each voxel $\lambda(\mathbf{x})$ is assumed to be constant.

Let $\{V_k\}_{k=1}^K$ be a partition of Ξ into K voxels. That is $V_k \subset \Xi$, $k = 1, \dots, K$; $V_k \cap V_{k'} = \emptyset$ for $k \neq k'$; and $\bigcup_{k=1}^K V_k = \Xi$. We shall refer to voxels directly by their index, referring to the voxel with index k (the volume element V_k), simply as “voxel k ”.

Discrete model for average emission intensity

The number of emissions within voxel k over the period of the scan is a Poisson random variable, with mean

$$\begin{aligned} \lambda_k &= \int_{V_k} \int_0^T \lambda(\mathbf{x}, t) dt d\mathbf{x} = \int_{V_k} \lambda(\mathbf{x}) d\mathbf{x} \\ &\approx |V_k| \lambda(\mathbf{x}) \end{aligned}$$

for some $\mathbf{x} \in V_k$, if $\lambda(\mathbf{x})$ is fairly constant within V_k , where $|V_k|$ is the volume of voxel k . Thus $\lambda_k / |V_k|$ is a step function approximating $\lambda(\mathbf{x})$, and we proceed to estimate $\lambda = (\lambda_1, \dots, \lambda_K)$. Substituting this approximation into eqn.2 gives:

$$E[n_r^*] = \lambda_r^* = \sum_{k=1}^K \lambda_k p(k, r) \quad r = 1, \dots, R \quad (3)$$

$$\text{where } p(k, r) = \frac{1}{|V_k|} \int_{V_k} c(\mathbf{x}, r) d\mathbf{x}$$

Eqn.3 defines a Poisson regression equation expressing the observed VOR counts n_r^* in terms of λ , which it is desired to estimate, subject to the positivity constraints $\lambda_k \geq 0$, $k = 1, \dots, K$. The average emission intensity for voxel k , λ_k , is referred to simply as the *regional activity* (rA) at voxel k . Since λ_k is expressed in counts per unit volume, reconstructed images are known as *counts images*.

Computation of detection probabilities

The $p(k, r)$ are the probabilities that an emission occurring in voxel k is detected in VOR r , the *detection probabilities*. Ignoring positron flight, attenuation, scatter, detector inefficiencies and assuming that the two gamma photons are emitted in opposite directions, this probability is (up to a normalising constant) the angle of view of VOR r into voxel k , and can be computed from a knowledge of the scanner geometry.

Positron flight, attenuation, detector inefficiencies and the angle of gamma photons can all be built in to these probabilities, see Vardi *et al.* (1985) and Shepp *et al.* (1982) for further details. These refinements make computation of the $p(k, r)$ more lengthy, too lengthy for computation on the fly to be efficient, and the probabilities must be computed and stored, requiring considerable storage space.

⁷Voxel, is the 3D analogue of a *pixel*, which is short for “picture element”.

This computation is made notably simpler (especially in 2D PET imaging) by the consideration of an annular array of voxels, with the number of voxels per annulus some multiple of the number of detector elements per ring. With this arrangement, proposed by Kearfott (1985), Each detector element would view exactly the same arrangement of voxels and opposing detector elements as any other detector element (in that ring, for the 3D case). This greatly reduces the number of distinct $p(k,r)$. Kaufman (1987) discusses an implementation of such a method. If an attenuation correction more complicated than that for a rotationally symmetric attenuating medium (with axis that of the tomograph) is being considered, then there is no symmetry, and the advantages of an annular voxelation are lost, a point noted by Green (1985).

1.4.2. Maximum Likelihood Reconstruction

Vardi *et al.* (1985) discuss the maximum likelihood estimation of λ . The likelihood of the observed data is:

$$L(\lambda) = P(\mathbf{n}^* | \lambda) = \prod_{r=1}^R e^{-\lambda^*(r)} \frac{\lambda^*(r)^{n^*(r)}}{n^*(r)!}$$

giving log-likelihood

$$\begin{aligned} l(\lambda) &= \ln(L(\lambda)) = \sum_{r=1}^R (-\lambda_r^* + n_r^* \ln(\lambda_r^*) - \ln(n_r^*!)) \\ &= \sum_{r=1}^R \left\{ -\sum_{k'=1}^K \lambda_{k'} p(k', r) + n_r^* \ln \left[\sum_{k'=1}^K \lambda_{k'} p(k', r) \right] - \ln(n_r^*!) \right\} \end{aligned}$$

with first partial derivative with respect to λ_k as

$$\frac{\partial l(\lambda)}{\partial \lambda_k} = -\lambda_k p(k, \bullet) + \lambda_k \sum_{r=1}^R \frac{n_r^* p(k, r)}{\sum_{k'=1}^K \lambda_{k'} p(k', r)} \quad (4)$$

where $p(k, \bullet) = \sum_{r=1}^R p(k, r)$ is the probability that an emission in voxel k is detected.

Vardi *et al.* (1985) show that the matrix of second partial derivatives of $l(\lambda)$ is negative semi-definite, and hence that $l(\lambda)$ is concave. Thus the single local maximum (if it exists) is a global maximum over $\lambda \in \mathfrak{R}^3$, and can be found by setting the first partial derivative to zero.

Bearing in mind the positivity constraint on λ ($\lambda \in \{\mathfrak{R}^+ \times \dots \times \mathfrak{R}^+\}$), sufficient conditions for $\hat{\lambda}$ to be the maximum likelihood estimator of λ are the Kuhn Tucker conditions:

$$\frac{\partial l(\lambda)}{\partial \lambda_k} \Big|_{\hat{\lambda}} = 0 \quad \text{and} \quad \frac{\partial l(\lambda)}{\partial \lambda_k} \Big|_{\hat{\lambda}} \leq 0 \quad \text{if } \hat{\lambda}_k = 0$$

There are many possible schemes for solving to obtain $\hat{\lambda}$. A particularly simple iterative scheme suggested by the form of the partial derivatives (eqn.4) is:

$$\lambda_k^{(i+1)} = \frac{\lambda_k^{(i)}}{p(k, \bullet)} \sum_{r=1}^R \frac{n_r^* p(k, r)}{\sum_{k'=1}^K \lambda_{k'}^{(i)} p(k', r)} \quad (5)$$

giving a sequence $\lambda^{(1)}, \dots, \lambda^{(i)}, \dots$ converging to the maximum likelihood estimate $\hat{\lambda}$. This can be thought of as a preconditioned steepest ascent method, having properties similar to steepest ascent in many situations and considerably improving on it in others (Kaufman, 1987). This iterative scheme is an instance of the EM algorithm, as we shall now illustrate.

EM reconstruction

Here we have a classic example for the application of the EM algorithm of Dempster *et al.* (1977). The ideal data would be the number, n_k , of detected emissions in each voxel ($k=1, \dots, K$). The available data are the VOR counts. Both are marginal sums of the quantities n_{kr} , the number of emissions in voxel k detected in VOR r ($k=1, \dots, K; r=1, \dots, R$). The VOR counts, n_r^* , can be viewed as incomplete data, with the n_{kr} as complete data.

With this in mind, Vardi *et al.* (1985) re-formulate the problem as follows: Emissions in voxel k subsequently detected in VOR r form a Poisson process with mean $\lambda_{kr} \stackrel{\text{def}}{=} \lambda_k p(k, r)$, since it is a thinning of the emission process, independently for all voxels k and VOR r . Thus the number of detected emissions in voxel k , n_k , is a Poisson random variable with mean

$$E[n_k] = \sum_{r=1}^R \lambda_{kr} = \lambda_k p(k, \bullet) \quad (6)$$

The estimation, or E-step, is to estimate n_k as \hat{n}_k , given $\lambda^{(i)}$ and the data n_r^* .

$$\begin{aligned} \hat{n}_k &= E\left[n_k \mid \lambda^{(i)}, n^*\right] \\ &= E\left[\sum_{r=1}^R n_{kr} \mid \lambda^{(i)}, n^*\right] \\ &= \sum_{r=1}^R E\left[n_{kr} \mid \lambda^{(i)}, n^*\right] \\ &= \sum_{r=1}^R E\left[n_{kr} \mid \lambda^{(i)}, n_r^*\right] \quad \left| \text{by independence of } n_r^* \text{'s} \right. \\ &= \sum_{r=1}^R \frac{n_r^* \lambda_{kr}^{(i)}}{\sum_{k'=1}^K \lambda_{k'r}^{(i)}} \quad \left| \begin{array}{l} \text{since if } X_i \sim \text{Po}(a_i), \\ X_i \mid \left(\sum X_i = x^*\right) \sim \text{Bin}\left(x^*, a_i / \sum a_i\right) \end{array} \right. \\ &= \lambda_k^{(i)} \sum_{r=1}^R \frac{n_r^* p(k, r)}{\sum_{k'=1}^K \lambda_{k'}^{(i)} p(k, r)} \end{aligned} \quad (7)$$

The maximisation step, or M-step, is to obtain $\lambda^{(i+1)}$, the maximum likelihood estimate of λ , given the current estimated detected emissions from voxel k . Since n_k is a Poisson random variable with mean $\lambda_k p(k, \bullet)$, (eqn.6), $\lambda_k^{(i)} = \hat{n}_k / p(k, \bullet)$.

Combining the E and M steps gives the iterative scheme of eqn.5. It then follows from Theorem 1 of Dempster *et al.* (1977) that $l(\lambda^{(i)}) < l(\lambda^{(i+1)})$. Since $l(\lambda)$ is convex, convergence to the maximum likelihood estimate (on this voxellation, see below) is guaranteed. See Vardi *et al.* (1985) for a proof and discussion.

There are two interesting properties of this EM solution. Firstly, the total counts are preserved:

$$\sum_{k=1}^K \lambda_k^{new} p(k, \bullet) = \sum_{r=1}^R n_r^*$$

Secondly, if $\lambda_k^{old} > 0$ then $\lambda_k^{new} > 0$ also, for all voxels $k=1, \dots, K$. That is, the positivity condition is satisfied for positive starting values.

Comments on the maximum likelihood approach

Shepp & Vardi (1982) tested their EM approach for 2D PET against a filtered back-projection method, and (in Vardi *et al.*, 1985) against a least squares method, using simulated data. Shepp *et al.* (1984) test the EM approach with real phantom data, comparing the reconstructed images with those from filtered back-projection. In all the papers the EM reconstruction was preferred on visual grounds for lack of reconstruction artefacts and for aesthetic appeal. Mintun *et al.* (1985) also reported an improvement in resolution in EM reconstructions over the standard filtered back-projection for a low resolution 2D camera.

Lange & Carson (1984) also developed EM algorithms for emission and transmission tomography, aiming their exposition at a mathematically literate audience familiar with tomography.

Kaufman (1987) discusses the practicalities of implementing the approach. Consideration of positron flight, detector inefficiencies and the non-linearity of the two gamma photons (but not attenuation or scatter) in the computation of the detection probabilities $p(k, r)$ improved 2D reconstructions very little. Indeed, vast simplifications were made in the computation of these probabilities, without substantial degradation of the reconstruction.

The maximum likelihood solution is not necessarily unique, indeed it cannot be unique if the number of VORS (R) is less than the number of voxels (K). The algorithm converges to one of the reconstructions maximising the likelihood, depending on the choice of initial values. Thus for cubic voxels, a unique solution for the EM algorithm imposes a lower bound on the voxels size, this size dependent on the geometry of the scanner. For voxel sizes smaller than this critical size the estimated value is not unique, but it is argued (Shepp & Vanderbei ([23] in Green, 1990)) that for typical data sets the positivity constraint renders the maximum essentially unique.

Convergence to the maximum likelihood function estimate $\hat{\lambda}$ (for this voxellation of Ξ) is guaranteed, but may take a long time. Knowing when to stop the iterations is a source of indeterminacy. Stop too soon and a vague image is recovered. After many iterations the estimate $\lambda^{(i)}$ becomes very noisy, a common feature of maximum likelihood estimates for images. Silverman *et al.* (1990) illustrate this vividly. Vardi *et al.* (1985) suggest using a small degree of smoothing on the maximum likelihood estimate, or a constant starting estimate $\lambda^{(0)}$ and fixed (small) number of iterations (for example 64). This latter scheme, starting with a level playing field and iterating until an ‘‘appealing’’

estimate is found, is commonly called “Maximum Likelihood” reconstruction by PET practitioners, though the estimates obtained are far from the actual MLE.

Silverman *et al.* (1990) considered the above points and applied their smoothed EM algorithm, the EMS algorithm, which introduces an S step after each EM iteration, spatially smoothing the current estimate $\lambda^{(i)}$. This gives aesthetically pleasing reconstructed images which are quite smooth and free from the obvious reconstruction artefacts that plague filtered back-projection. The smoothing step reduces the spectral radius of the EM iteration, and thus accelerates convergence. Experience suggests that convergence is to a unique solution, although only approximate theory exists to back up this assertion. However, there is little justification for the smoothing in this form, and these images were criticised by Ford (1990) for being too smooth, as interest is in more than the “picture” of the annihilation events.

Although the EM algorithm has been successfully applied to reconstruction it isn’t used in practice. This is partly because of the noisiness and indeterminate stopping criteria for the iterations, but mainly because computationally it is much more intensive than filtered back-projection, especially for \mathfrak{D} reconstruction. Current interest in statistical reconstruction is concentrated on the use of prior information in an empirical Bayesian context, a review of which is beyond the scope of this chapter.

1.4.3. Filtered Back-projection

Filtered Back-projection (FBP) is the reconstruction method routinely used in most circumstances. Essentially the counts are “projected” back across the image space, interfering to give an image of estimated counts per voxel. The technique was developed to reconstruct images from projections in fields as diverse as gravitation theory, radio astronomy and electron microscopy. More recently the technique has been used in transmission tomography, particularly X-ray computed tomography (X-ray CT).⁸ Shepp & Kruskal (1978) review the history of CT, and look at the various methods of reconstruction from a mathematical perspective. Emission tomography developed from transmission tomography, and the existing reconstruction methods were adapted for the new modality. Thus to review FBP we first look at transmission tomography and how it relates to the PET problem.

1.4.3.1. Transmission tomography

In transmission tomography, parallel beams of monochromatic⁹ X-rays of constant intensity¹⁰ I_{in} (in counts per second) are directed through a thin slice of the subject, and the number of X-ray photons transmitted through the subject from each beam is measured. This is done for a number of orientations, resulting in a data structure similar to that of PET (Fig.11). The quantity of interest here is the tissue density to X-rays, the *linear attenuation function* $\mu(\mathbf{x})$.

⁸X-ray CT is commonly referred to a CAT, an acronym for Computerised Axial Tomography, the word “axial” emphasising that the X-rays are administered perpendicular to the axis of conventional X-ray imaging.

⁹The source is assumed to be monochromatic, that is only X-ray photons of a given energy are emitted. Alternatively, one can assume that the physical attenuation is independent of the energies of the X-ray photons.

¹⁰That is the source has a half-life so long that the rate of decay can be considered constant for the time periods in question.

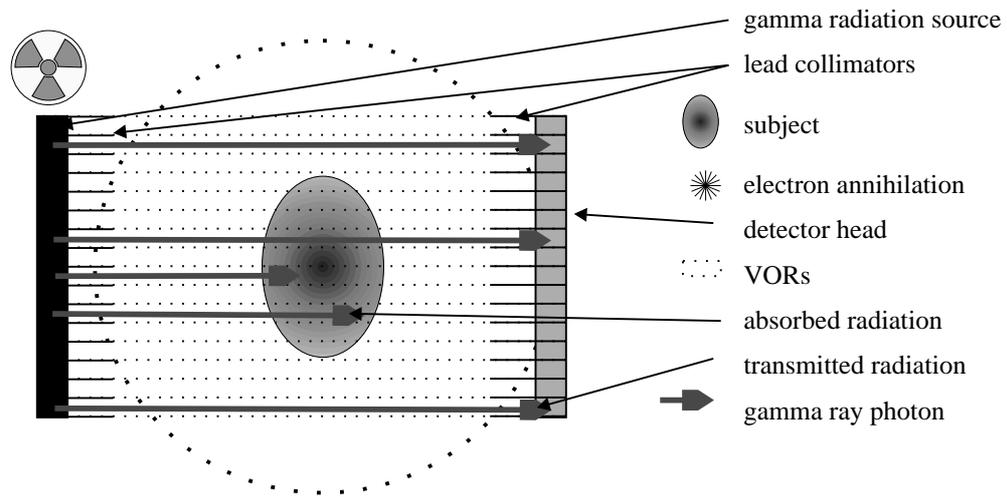


Figure 11
Schematic of transmission tomography

For a line segment $\delta \subset \mathfrak{R}^3$ through a medium with linear attenuation function $\mu(\mathbf{x})$, the total attenuation along the line is μ_δ , given by the Radon transform:

$$\mu_\delta = \int_{\delta} \mu(\mathbf{x}) d\mathbf{x} \quad (8)$$

Estimation of μ_δ in transmission tomography

For any X-ray photon travelling along δ the probability of non-absorption, is proportional to $\exp(-\mu_\delta)$, so the process of X-ray photons arriving at the detector is a thinning of the emission process. The thinning probability is estimated by $n_\delta^{in} \div n_\delta^{out}$, where n_δ^{in} and n_δ^{out} are the numbers of photons emitted by the source, and the number of those transmitted, respectively. Clearly it is impossible to count both. However, since intense X-rays are used, $n_\delta^{in} \div T \approx I_{in}$, where T is the length of the acquisition for line δ . This gives $I_{in} \times T \div n_\delta^{out}$ as a good estimate of the thinning probability, and hence of μ_δ .

1.4.3.2. Approximation of emission tomography problem to fit transmission tomography framework

Emission tomography can be thought of in this framework, by assuming that the probability of a positron emission from point $\mathbf{x} \in X$ being detected by the pair of detectors defining VOR r , $c(\mathbf{x}, r)$, is constant for all points \mathbf{x} within VOR r , and zero for all other points. Then, eqn.2 simplifies to

$$E[n_r^*] = \lambda_r^* = c(\bullet, r) \int_{V_r} \lambda(\mathbf{x}) d\mathbf{x} \quad (9)$$

Here $V_r \subset X$ is the set of points in VOR r , and $c(\bullet, r)$ is the probability an emission occurring inside VOR r is detected by the detectors defining that VOR.

Eqn.9 is now in the same form as for transmission tomography. The observed data n^* are taken as estimates of the VOR intensities λ^* , and an estimate of $\lambda(\mathbf{x})$ is sought whose integrals along the r^{th} VOR is n_r^* .

Inadequacies of the approximation

There are two fundamental problems with this approach to emission tomography. Firstly, even with the simplistic modelling of $c(\mathbf{x},r)$ as the angle of view from \mathbf{x} to the detectors defining VOR r , it is clear that $c(\mathbf{x},r)$ must vary considerably within the VOR, disappearing for points on the edge of the VOR. The simplifying assumption precludes consideration of positron flight, attenuation, scatter and detector inefficiencies. Corrections for these phenomena must be made to sinograms, or by post-processing reconstructed images. Within this framework it is impossible to account for the fact that the two gamma photons are emitted in only near opposite directions.

Secondly, and more importantly, the count-rates in PET are many orders of magnitude less than those in transmission tomography. Thus the observed count-rates give poor estimates of the actual VOR intensities λ^* . The large variation of n^* about its mean is seen as noise in the sinograms, noise which is magnified by the filtered back-projection procedure.

1.4.3.3. Reconstruction from projections

Considering eqn.8, Radon (1917) showed that if the linear attenuation function $\mu(\mathbf{x})$ is continuous and has compact support then $\mu(\mathbf{x})$ is uniquely determined by the values of μ_δ for all line segments δ intersecting the support of $\mu(\mathbf{x})$. Moreover, he obtained a fairly simple inversion formula for the computation of $\mu(\mathbf{x})$ from its projections. However, with the finite number of “projections” examined in CT, the necessary discrete approximations to Radon’s inversion formula give poor reconstructions. (Shepp & Kruskal, 1978, dwell on this in more detail.)

Other methods of reconstruction from projections have been developed, and are discussed in detail by Herman (1980). Of the available methods, Filtered Back-projection, also known as *Convolution Back-Projection* (CBP), has emerged as the most frequently used. In all methods the “*fundamental law of tomography*” applies: The greater the number of projections obtained, the better the resolution of the reconstructed image.

A complete description of filtered back-projection technique for 3D PET is well beyond the scope of this background chapter. We shall take a simplistic pictorial approach, describing the fundamentals of \mathbb{D} reconstruction by FBP, and then commenting on the 3D method. For the remainder of this section we use the terminology for reconstruction of functions from projections. A “profile” is the value of a line integral through the image space, and a “projection” is the set of all profiles through the image space in the same direction.

Back-projection

Basically, back-projection reconstruction is a summation method. Essentially, the density at a point is estimated by adding all the obtained profiles through that point, effectively back-projecting each profile (i.e. sinogram pixel) across the image space and summing. For a simple 2D object observed in two orthogonal directions, fig.12 illustrates the approach.

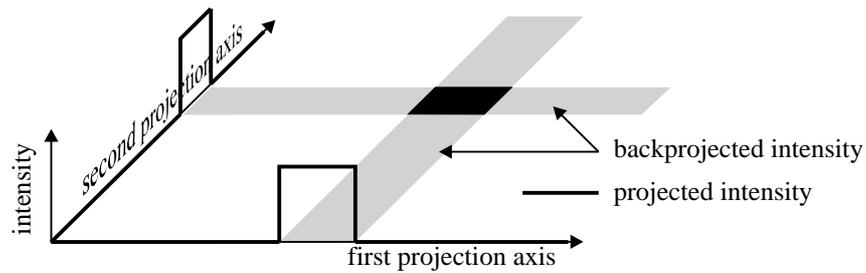


Figure 12
Back-projection of a point source

As is clear from fig.12, the reconstructed density is only a crude approximation to the original density. In particular, points outside the object receive positive contributions from some of the projections, leading to star artefacts, often seen in reconstructed images, (fig.13).

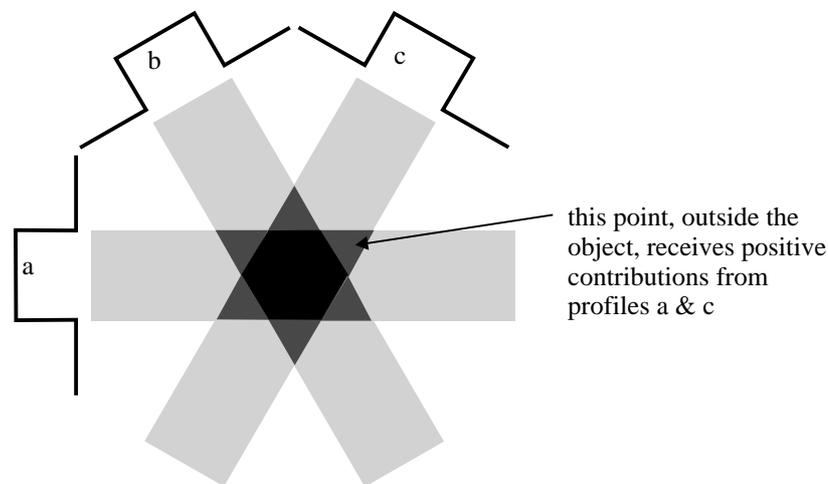


Figure 13
Star artefact in back-projection of a point source

Filtered Back-projection

To eliminate the star artefacts surrounding dense points, back-projection can be modified by convolving the projected data with a *reconstruction filter*. The filter kernel is constructed with a positive central lobe and negative side lobes, such that points just outside an object receive both positive and negative contributions from the filtered projections when back-projected. Fig.14 illustrates the general idea, and fig.15 illustrates how artefacts are avoided, at least in the vicinity of a dense point. After appropriate normalisation, the image constructed is then an estimate of the number of emission events per voxel.

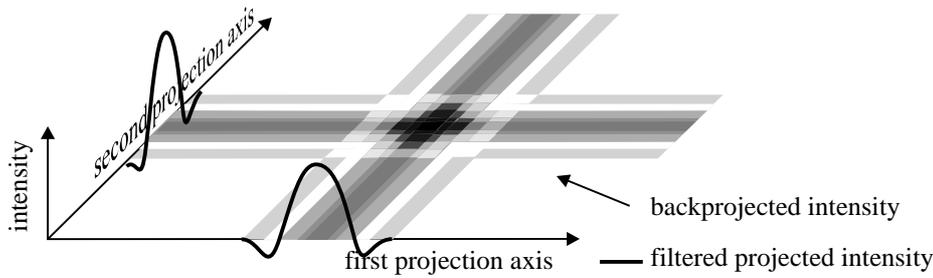


Figure 14
Filtered back-projection of a point source

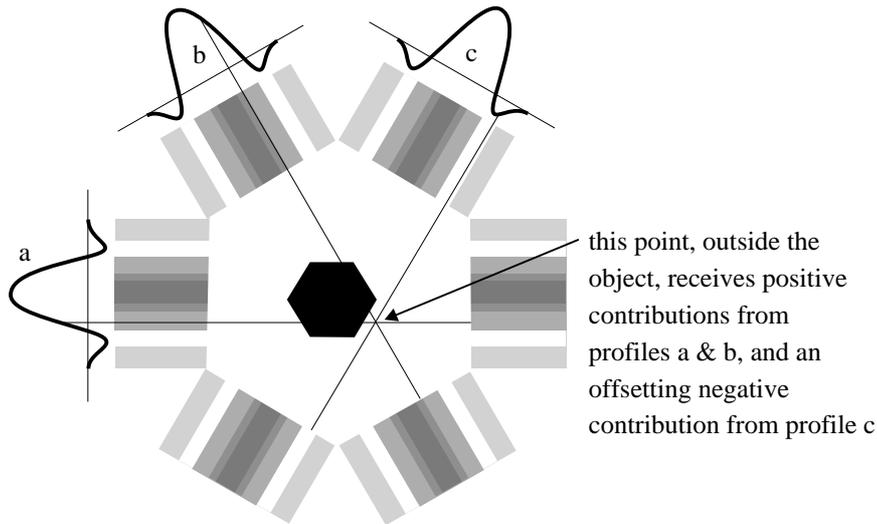


Figure 15
Artefacts in filtered back-projection of a point source

For a more mathematical description, the reader is referred to the paper by Shepp & Kruskal in *American Mathematical Monthly* (1978). Most papers reviewing PET give similar descriptions of filtered back-projection (Ter-Pogossian *et al.*, 1980; Raichle, 1983b).

3D filtered backprojection

In theory, 3D FBP is a simple extension of the 2D case. In practice there are a number of problems in implementing FBP for PET data acquired in 3D mode.

A projection is *complete* if profiles are available for all lines that pass through the region of interest (brain) in the direction of projection (given the discreteness of the sampling). If all projections are filtered with the same kernel, then to realise the advantages of FBP it is necessary to have a full set of complete projections, that is complete projections for all possible angles (given the discreteness of the sampling). For 2D reconstruction, the necessary information to reconstruct a plane is contained in the corresponding direct-plane (or induced-plane) sinogram. In 3D PET many of the possible projection angles are unobservable because of the tomograph geometry. Furthermore, many of the projections that are possible are incomplete, since oblique projections become increasingly truncated as the direction of projection nears the axis of the tomograph.

Careful choice of filters for each projection (compared by Defrise *et al.*, 1989) enables artefact-free FBP reconstruction for (incomplete) sets of complete projections. Defrise *et al.* (1990) discuss filtered back-projection for such data.

To obtain complete projections it is necessary to estimate the missing profiles in partially observed projections. This is done by reconstructing the density in 2D mode, using only the direct (and induced) plane sinograms, and obtaining the missing profiles from this rough estimate of the actual density. The effect is to artificially extend the axial extent of the scanner by adding extra rings of detectors. The estimated sinograms for these artificial rings are termed *synthetic* sinograms. Typically the number of synthetic sinograms is comparable with the number actually observed.

Townsend *et al.* (1991) discuss the implementation of filtered back-projection for 3D PET data acquired on the CTI-953B, a tomograph with retractable septa. The procedure requires the following steps: (Taken from Townsend *et al.*, 1991, to which the reader is referred for additional details.)

- 1) Acquire 2D, septa-extended blank and transmission scans, and a 3D, septa-retracted emission scan.
- 2) Normalise the 256 emission sinograms using coefficients obtained from the high statistics, septa-extended blank scan.
- 3) Reconstruct the 2D attenuation map and forward project to obtain oblique correction factors.
- 4) Subtract the estimated scatter from the emission sinograms.
- 5) Correct the scatter corrected emission sinograms for attenuation.
- 6) Reconstruct (in 2D) the emission sinograms for ring differences of 0 & ± 2 (direct plane) and ± 1 & ± 3 (cross plane).
- 7) Forward project the sinogram corresponding to the measured sinogram with the maximum activity, and compare (fit) with the measurement to obtain the scaling factors between the forward projected and measured sinograms.
- 8) Forward project the unmeasured sinograms for all ring index differences required for the reconstruction.
- 9) Scale the forward projected sinograms using the coefficients obtained in step (7).
- 10) Form complete 2D projections and apply the appropriate filter.
- 11) Back-project all VORS, both measured and synthetic.
- 12) Re-scale the reconstructed values to counts/second/voxel.

1.4.3.4. Comments on filtered back-projection

There are two appealing properties of the filtered back-projection approach for estimating the average emission intensity for PET data. Firstly, the method is linear, so additive corrections to sinograms can equivalently be applied to reconstructed images, after the corrections have themselves been reconstructed. Secondly, the method is analytic, and doesn't rely on iterative implementations. Even so, a 3D reconstruction of one scan from the CTI-953B takes 3.3 hours on a SUN SPARCStation2 (Townsend *et al.*, 1991).

We have already noted two inadequacies of the filtered back-projection approach, namely the simplification of the emission tomography problem to the transmission

framework, and the use of VOR counts as good estimates of their means without recourse to any probability modelling.

1.4.4. Reconstructed images

Co-ordinate systems: Tomograph axes

The Cartesian co-ordinate system of reconstructed PET images is fixed relative to the tomograph. Looking into the patient port of the tomograph; the X-axis is horizontal, running from right to left; the Y-axis is vertical, running from below to above; and the Z-axis is horizontal, running directly away from you along the axis of the tomograph. Thus for a subject placed in the tomograph head first and on their back, the X-axis runs (roughly) from their left to their right, the Y-axis from their back to their front, and the Z-axis from their toe to their head. The origin is placed arbitrarily on the tomograph axis, and the axes are scaled in millimetres. Fig.16 illustrates the positioning and orientation of the axes.

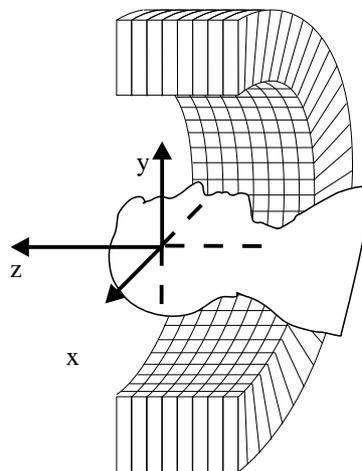


Figure 16
Positioning and orientation of tomograph axes

Planes of PET images

3D images are usually viewed as a set of 2D images which form the 3D image when stacked one on top of the other. The 2D images most frequently presented are *transaxial* slices, planes parallel to the XY plane. *Sagittal* and *coronal* sections are planes parallel to the YZ plane and XZ plane respectively.

Transverse sections are presented viewed from above and behind the subject: The posterior of the brain is at the bottom of the image, the subjects left at the left side of the image, and the X and Y axes are in their usual positions. Sagittal sections are presented viewed from the subjects right with the brain upright: The posterior of the brain at the left of the image and the top of the brain at the top of the image. (Posterior) Coronal sections are viewed from behind the subject, with the subjects left on the left of the image.

FBP reconstructed image

At the MRC Cyclotron Unit at the Hammersmith Hospital in London, scans are routinely reconstructed using three-dimensional filtered back-projection with a Hanning filter cut off at 0.5 cycles per voxel. The discretisation of Ξ for reconstruction is of

128×128×31 cuboid voxels of dimensions 2.09×2.09×3.5mm³ (x×y×z), with voxel edges parallel to the tomograph axes. The resulting images are then re-sampled axially to give 128×128×43 approximately cubic voxels of dimension 2.09×2.09×2.09mm³.

Orthogonal sections through such an image are shown in fig.17. The subject (the author) was injected with [¹⁵O]water and emissions recorded according to the protocol described below (§1.6.1.2.). The full 43 transverse planes are shown in Fig.18. Reconstruction artefacts can be seen in the lower planes.

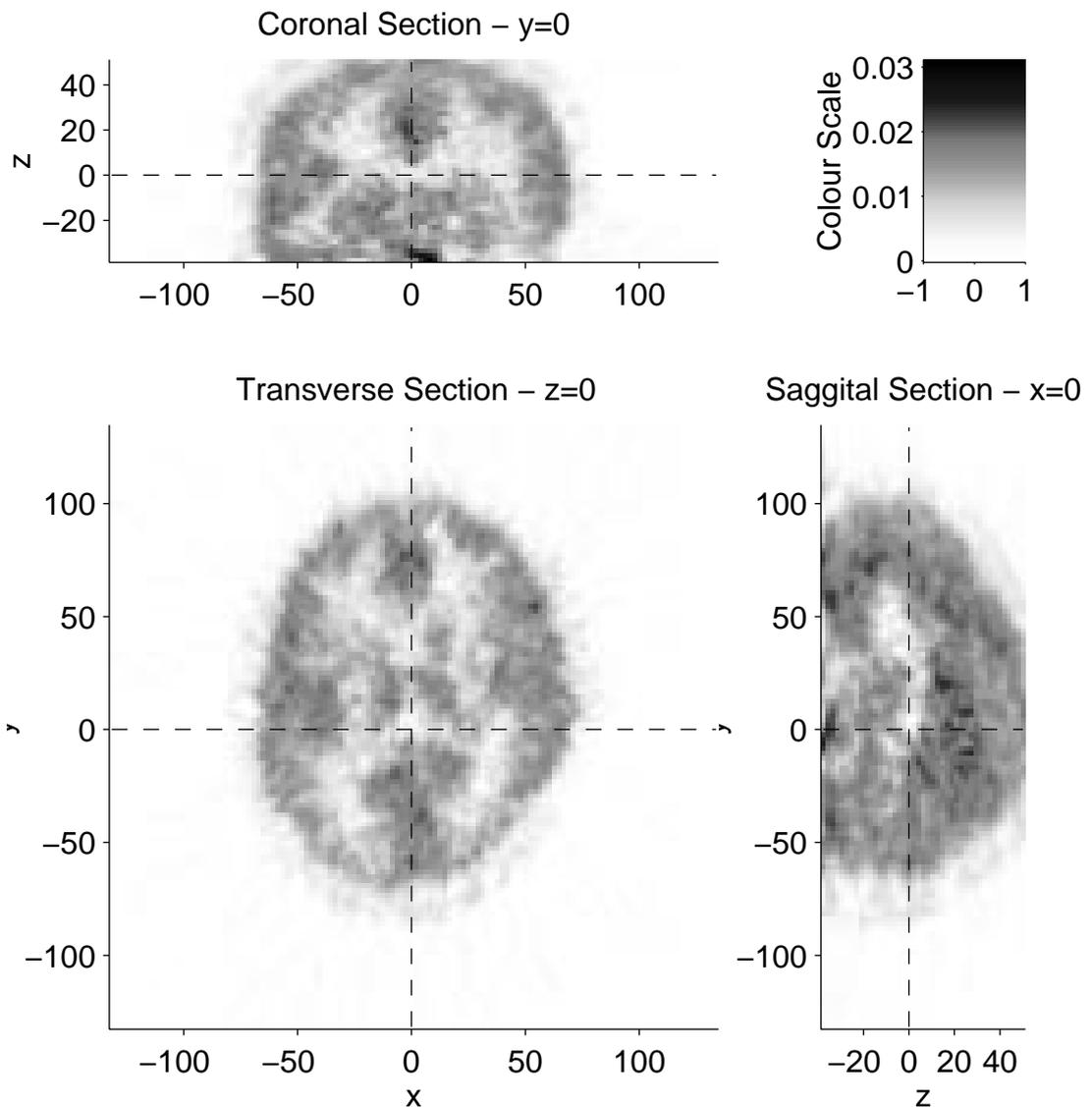


Figure 17
Orthogonal sections of the author's brain.
Colour scale in (estimated) annihilations per second.
Tomograph axes, graduated in mm.

Resolution

The flight of the positrons from the radioisotopes before electron annihilation imposes a limit on the resolution of these images. Resolution for PET scanners is measured in terms of Full Width at Half Maximum (FWHM), the width of the point spread function at half its maximum. This measures the resolving power of the tomograph and reconstruction method since the FWHM is the smallest displacement at which two point sources can be distinguished. The point spread function is measured by scanning a point source of radioactivity located centrally in the scanner. Clearly the point spread function depends on the reconstruction used, and on position in the scanner. Modern scanners typically have a FWHM of 8–15mm in the X and Y directions.

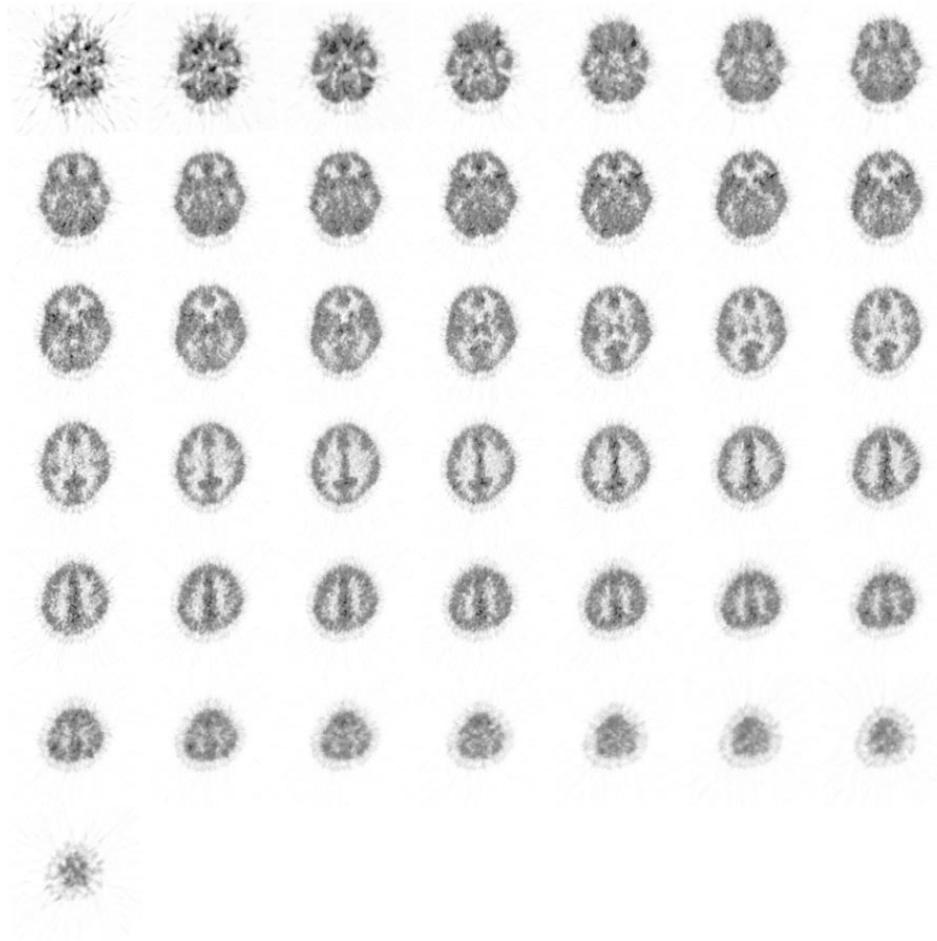


Figure 18

The 43 transverse planes of the authors brain. The planes are 2.09mm apart, and are presented from bottom to top (increasing Z direction), row-major order.

1.5. Tracer and Modelling Issues

1.5.1. Radioactive isotopes and labelling

The radiotracers used in Positron Emission Tomography are chosen to enable the study of particular metabolic and physiological processes in the living body. Three positron emitting radioactive isotopes of particular use are oxygen 15 (^{15}O), nitrogen 13 (^{13}N) and carbon 11 (^{11}C) (Ter-Pogossian *et al.*, 1990), with half lives of 123, 598 and 1220 seconds respectively (Tennent, 1971). These isotopes are particularly valuable because biological systems consist mainly of compounds of oxygen, nitrogen and carbon. Labelling a biochemical, by replacing one of its atoms with a radioactive isotope, enables imaging of the density of the labelled compound, and hence non-invasive measurement of metabolic processes in the living body.

Observable metabolic functions

Using these four radioactive isotopes in different biological compounds, the quantitative study of many facets of metabolism is possible. Among the quantities it is possible to measure are: blood flow, blood volume, organ perfusion, brain permeability to various amino acids, tumour drug levels, tissue pH, tissue lipid content, heart metabolism, tissue hematocrit (ratio of cells to plasma), mapping of monoamine oxidase receptors (lung), neuroreceptor pharmacology (brain) and tissue drug levels in epilepsy. See Raichle (1983b) for a more detailed summary of Positron Emission Tomography and its research and clinical applications.

1.5.2. Short half lives and low activities of the isotopes

Implications for application

The short half lives of the radio isotopes used limit the use of PET to imaging fairly fast metabolic processes. Fortunately the overwhelming majority of brain metabolic processes have proved to be fast relative to the radionuclide half-lives. The low activity of commonly used tracers forces each scan to be acquired over a period of a few minutes. This precludes the use of PET to study very fast metabolic processes, or processes that are quickly reversed (in that the labelled metabolite is cleared away).

Multiple scans on healthy subjects

The low dose of radioactivity required for a single PET scan is well below the maximum allowable for volunteers in medical research¹¹. Thus healthy subjects can be examined using PET, allowing investigation of normal function. This is in contrast to many areas of medical research, where knowledge is acquired from patients who are ill or are undergoing other forms of treatment. More importantly, the low doses permit repeat scanning of an individual. Short half-lives allow repeat scanning within a short time, making it feasible to obtain multiple scans on a single subject in a single session, so

¹¹In the European community, the maximum activity that can be administered to volunteers in medical science is 5 milli-Sieverts per year, roughly equivalent to two years background activity. For a typical (3D) PET scan using the tracer H_2^{15}O only about a twelfth of this amount need be injected per scan. (Source: RSJ Frackowiak, MRC Cyclotron Unit.) In the USA, the maximum allowable activity for administration to volunteers in medical science is much higher than in the UK. Fourteen run experiments are commonplace there, but more could be performed. The limiting factor on the number of possible scans is the length of time a subject can spend in the scanner.

avoiding long term time differences. The multiple scans of a subject under similar or different conditions permit the investigation of clinical questions within an individual. In a statistical sense, subjects can act as their own controls.

1.5.3. Cost of PET

The development of PET was hampered by the short half lives of the positron emitting isotopes most useful for labelling. Fast methods for creating the isotopes and then binding them into the relevant biochemicals are needed. Production must be on-site, requiring a cyclotron, and a resident team of radio-chemists. Cyclotrons for this purpose typically accelerate protons up to an energy of 15 million electron volts (MeV) and deuterons (heavy hydrogen nuclei) up to half that energy. Such apparatus weighs some twenty tons, is shielded by several feet of concrete. With the cost of a tomograph lying in the region of a million pounds, PET is a very expensive imaging modality. Currently (1992) there are only about 40 centres world-wide.

1.5.4. Modelling

So far we have only considered using PET to estimate the regional activity (rA) of the radioactive isotope administered to the subject. A study that seeks only to identify the site of a focal abnormality in an individual, for example a tumour or infarct characterised by abnormally low activity, might be complete at this stage with no need for further processing.

To quantitatively obtain measurements of metabolic function it is necessary to model the relationship between the tracer activity, and the metabolic parameter of interest. Even in comparative studies it is usually necessary to convert "counts" images to appropriate biological parameters, because of global differences in tracer uptake from person to person, or even from scan to scan in the same individual. The principles of tracer kinetic modelling are presented in depth by Huang & Phelps (1986). A brief review from a statistical perspective is given by McColl *et al.* (1994, §3).

Compartmental models

The link between regional tracer concentration (estimated from the measurements of radioactivity) and biological parameters is usually provided by *compartmental modelling*. These tracer-kinetic models represent each stage in the relevant biological process as a *compartment*. Differential equations describe the flow of the metabolite of interest (and its labelled analogue) between the compartments. The parameters of these differential equations are those of the metabolic process of interest, and can be estimated given the tracer concentration and additional physiological parameters. Typically the concentration of the tracer in the blood pool is needed, requiring sampling of the subjects blood during the scanning session via an arterial catheter. Such a model is applied for each voxel of the reconstructed PET image, giving *parametric* images, each voxel having value the estimated metabolic parameter of interest.

1.5.5. Imaging cerebral blood flow

Particularly important is the measurement of regional Cerebral Blood Flow (rCBF). The method is due to Kety (1951), and was adapted for PET by Herscovitch *et al.* (1983) and Raichle *et al.* (1983a). A simple one compartment model describes the relationship between regional cerebral blood flow, the arterial (plasma) concentration of a freely

diffusible blood flow tracer (measured by arterial sampling), and the total tissue/plasma activity (measured by PET).

Freely diffusible blood flow tracers

[¹⁵O]Water is commonly used as a freely diffusible blood flow tracer. Although H₂¹⁵O has only limited diffusibility, it is usually used by most PET centres in preference to truly freely diffusible blood flow tracers such as [¹¹C]butanol. The water method gives slight underestimates of rCBF when compared to Butanol, as noted by Herscovitch *et al.* (1987). However, [¹⁵O]water is easy to make to a specified activity, using a relatively cheap, low energy cyclotron. Further, [¹⁵O]water has a short enough half life (123 seconds) that a subject can be repeatedly scanned with only a five minute wash-out period, without cumbersome schemes to account for residual activity.

Use of “counts” images.

Conversion from local radiotracer density to local cerebral blood flow requires arterial sampling of the blood radioactivity levels at given intervals, so that the *time activity curve* for the tracer in the cerebral plasma can be estimated. This requires the insertion of a cannula into an artery. The cannula can become quite painful by the end of a three hour scanning session, and the pain may dominate the subjects thoughts, compromising experiments where rCBF is being measured as an indicator of mental activity.

However, if the amount of tracer administered is carefully controlled, then over the normal range of tracer uptake by the plasma, and for a fixed global CBF, the relationship between regional activity (measured by PET) and rCBF is fairly linear. Thus the modelling and conversion of “counts”, or *regional activity* (rA) images to quantitative rCBF images is frequently not carried out when rCBF images are to be compared for relative changes (Mazziotta *et al.*, 1985; Fox *et al.*, 1989).

In fact, Mazziotta *et al.* (1985) report that the relationship between flow and counts is slightly convex, so changes in regional activity are slightly conservative indicators of changes in rCBF.

1.6. Functional Mapping with PET

The investigation of the functional organisation of the brain is termed *functional mapping*. The use of PET in functional mapping experiments has matured over the past two decades to the point where it is now regarded as a mature research tool.

1.6.1. Imaging brain activity

Early research into the workings of the brain showed that the brain is very well organised, specific different tasks being dealt with by specific physical areas of the brain. With Positron Emission Tomography the physical locations of function in the human brain can be identified by using a radiotracer whose density reflects local brain activity.

Electrical neuronal activity

The brain works by communication between neurones using small electrical currents, generated and transmitted by chemical reactions (Fishbach, 1993). Local neuronal electrical activity is thus the absolute measure of local brain activity. Early methods measured this directly, using electrodes placed on the scalp (electroencephalogram) or even in the brain itself.

1.6.1.1. Tracers for functional mapping

Metabolic activity induced by neuronal activity

Increased neuronal (electrical) activity causes an increase in metabolic activity for maintenance of the neuronal membrane ionic gradients (Yarowsky *et al.*, 1981). Under normal conditions the brain uses glucose and oxygen as energy substrates (Phelps *et al.*, 1986, ch1)¹². By examining the metabolism of these substrates using PET, the activity of the human brain can be observed. The regional cerebral metabolic rate of oxygen (rMRO₂) is the most direct measure of regional energy metabolism, but this is hard to measure. The regional cerebral metabolic rate of glucose (rCMRG) can be measured by labelling an analogue of glucose, deoxyglucose, which follows the same metabolic pathways until the blood-brain barrier has been crossed, and is then effectively trapped in the brain (for a short time). Experiments in rats and monkeys using autoradiography¹³ dramatically illustrate the relationship between physiological function and energy metabolism in the brain (Sokoloff, 1979). The deoxyglucose method has been used successfully in man with PET, using ¹⁸F as a label. However, scanning requires a 30 minute equilibration period, during which time patterns of neuronal activity may change substantially. Also, repeat scanning of a subject on the same day isn't possible, because of the long half-life of ¹⁸F (110 minutes). Usually fluoro-deoxyglucose (FDG) is used in *dynamic* scans, where a sequence of sequential scans are acquired after the introduction of the tracer, and its progress through the brain examined over the time course of the experiment. A more efficient method to study mental activity is via rCBF.

¹²The brain is one of the most active organs metabolically in the body. Although comprising only c2% of the total body weight, it accounts for c20% of the total body oxygen consumption (at rest), and maintains a power output of at least 20 watts (Phelps *et al.*, 1986, ch1).

¹³ In autoradiography (AR) a laboratory animal is injected with a radiotracer, decapitated, and their frozen brain sliced and placed on photographic film to obtain precise images of radiotracer density.

rCBF as indicator of regional neuronal activity

The brain has almost negligible stores of energy substrates when compared their rates of utilisation, and is hence dependent on blood circulation to supply its needs. Complete interruption of cerebral circulation leads to loss of consciousness within 10 seconds, and irreversible brain damage within a few minutes. So great is the demand for energy substrates that an increase in neuronal activity induces a dilation of the veins in the vicinity, in turn increasing the blood flow in the region (Kety, 1991). This induced increase in rCBF lags behind the neuronal activity by mere seconds.

Thus, PET measurements of rCBF, or rA as an indicator of rCBF, are indicative of the average neuronal activity over the period of the acquisition. The “average” here being weighted by the activity of the tracer, and hence emphasising the neuronal activity at the beginning of the acquisition period.

1.6.1.2. H₂¹⁵O infusion protocol

There are many different protocols for acquiring scanning using [¹⁵O]water. For routine studies at the MRC Cyclotron Unit (Hammersmith Hospital, London) on the 953B dedicated brain positron tomograph, the following protocol has been adopted (Bailey *et al.*, 1993)

Scanning protocol

The subject is positioned in the scanner, using external landmarks to obtain rough alignment, and a 2D transmission scan acquired for attenuation correction procedures. (A blank scan having been made at the beginning of the day.) A thin indwelling cannula for tracer administration is inserted into a superficial vein in the right forearm. After the transmission scan the inter-plane septa are removed to allow 3D acquisition of the emission scans. For each scan, H₂¹⁵O is infused via the cannula at a constant rate of 10ml/min. The task is initiated at the commencement of data acquisition, which consists of a 60s background frame, at the end of which tracer administration is started. The data is acquired in a single 3D frame for 150s. The infusion is terminated after 120s, the acquisition continuing for a further 30s. In each run approximately 1040 MBq (28 mCi) of [¹⁵O]water is administered. Emission scans are acquired every 12 minutes, to allow for tracer washout and decay, and 12 scans are acquired in the session.

Infusion system

The infusion system consists of a small furnace containing a palladium catalyst over which ¹⁵O₂ (from the cyclotron) is passed as a dilute (1:99) mixture in nitrogen. The hot H₂¹⁵O vapour produced is passed across one side of a cellulose acetate semi-permeable membrane. Isotonic saline, driven by a conventional medical infusion pump, is passed across the other side of the membrane and the H₂¹⁵O transfers into the saline. The device is compact and sits under the tomograph patient bed. A calibrated Geiger counter is incorporated into the infusion system to monitor accurately the amount of radioactivity infused. When not switched to the subject, H₂¹⁵O is drained into a lead encased reservoir. The constant production of H₂¹⁵O helps achieve a stable, reproducible delivery of tracer as the system settles into a constant pattern.

1.6.2. Types of study

Most PET functional mapping experiments take one of a few forms.

1.6.2.1. Common factors

All functional mapping experiments using PET have common factors dictated by the constraints of the technique.

Comparative

Distinct parts of the brain naturally have different activity and rCBF, so functional mapping experiments must proceed by comparing rCBF (rA). The 3D PET H₂¹⁵O method allows for each subject to be scanned numerous times in a three hour scanning session. Subjects can therefore be scanned repeatedly under the mental conditions of interest, and act as their own controls.

Multi-scan

Changes in rCBF induced by neuronal activity are very subtle, typically at most 10%, and are dominated by the noise in the reconstructed images. Increases in rCBF between a “rest” task and an “activation” task are not readily apparent in reconstructed images taken under the two conditions. It is necessary to obtain repeat observations and apply statistical tests. For some basic functions the changes in rCBF can be seen in a single individual scanned 12 times using the 3D PET H₂¹⁵O method. With less sensitive 2D acquisition, more subtle activations or more complicated protocols it is necessary to pool rCBF data across individuals to reliably detect rCBF changes. When it is desired to make deductions about a population rather than an individual, it is necessary to scan a sample of individuals from the population.

Small subject numbers

PET scanning is expensive. A three hour, 12 scan PET session costs in the region of two thousand pounds (Sterling). For this reason, sample sizes are very small. The largest studies completed to date have only 12 to 14 subjects. Typically only 6 subjects will take part in a study.

Single subject studies

With the increasing sensitivity of scanners, better reconstruction and more sensitive statistical procedures, significant results can be obtained on a single subject. This has led to move to “single subject” studies. Often interest lies in the individual: “Where is the brain gyrus in this individual which initiates motion in the right index finger?” However, in many single subject experiments the subject is assumed to be representative of the population and the results accepted for that population.

No prior information

Knowledge of the functional specialisation of the human brain is still limited, and even basic tasks have only recently been mapped. This lack of prior knowledge regarding the location of function necessitates the consideration of the whole brain volume, or the use of hypothesis generating experiments before a full study examining condition induced changes in rCBF at hypothesised locations. The expense of PET leads researchers to adopt the former approach, trying to get the most out of the data they have, placing emphasis on good statistical procedures to analyse these data sets.

1.6.2.2. *Activation studies*

The simplest and commonest paradigm is the *activation study*, designed to accurately locate the brain region responsible for performing a particular cognitive task. Subjects are scanned under two *conditions*, a baseline or “rest” condition, and “activation” condition, carefully chosen by the experimenter on psychological grounds such that the mental processing required by the two conditions differs only in the cognitive task it is wished to locate. Localised changes in measured rCBF between the two scans are due to natural variation, scanner noise, or are induced by different patterns of neuronal engagement. The problem is to identify changes that fall into the latter category. Usually arterial sampling is not carried out, and images of regional activity are analysed as indicators of regional neuronal activity.

Most experiments are concerned with mapping simple tasks. The somatosensory cortex is readily mapped by activation studies, using a rest baseline, and a simple sensory stimulation as the activation condition. Similarly the motor cortex is easily mapped. Contrasting a rest baseline condition with an activation condition where the subject repeatedly extends a finger identifies the brain region responsible for control of the finger.

In a *simple activation experiment*, only two conditions are used. The aim is to reliably locate a particular task. In a *compound activation study*, several conditions are used. Interest may lie in the difference in neuronal engagement between many conditions, or many activation conditions might be compared with a single baseline condition in a hypothesis generating experiment.

Typical experiment

The typical simple activation study will have four, six, or twelve right handed male volunteers between the ages of 25 and 50 years. Each subject is scanned twelve times in a single session, each scan under either the baseline or the activation condition. The conditions are assigned to the scan times for each subject before the scanning session. Usually the allocation is balanced, but not necessarily random. Typically subjects will be scanned alternately under the two conditions, with half the subjects starting with the baseline condition for scan one, the remaining half starting with the activation condition. This gives a one-way block design with replication, where the blocks are the subjects, and the data points are images of rA or rCBF.

Single subject activation studies are not usually performed explicitly. Rather activation studies are performed and the subjects considered individually as well as a group.

1.6.2.3. *Correlation with an external variable*

It is not always possible to design conditions inducing mental states differing only in the cognitive component it is wished to locate. Sometimes it is wished to locate a brain region which is increasingly engaged as the severity of some task increases. In many of these situations it is possible to obtain a *score*, indicating the severity of the task. Assuming a linear relationship between rCBF and the score, the regions increasingly engaged by the tasks can be identified as those with rCBF and score highly positively correlated. These studies are referred to as *correlational* studies. There are two variants, *across subjects*, and *within and across subjects*.

Across subjects

This paradigm is typically used to locate the brain region whose function is impaired by a particular disease. A number of sufferers with varying degrees of

debilitation are scanned whilst performing the task the disease impairs. The scores are clinical scores reflecting the severity of the disease. The location of the function which worsens with the onset of the disease may be found as those regions where rCBF correlates negatively with the score. These studies are particularly prone to changes in global CBF, and different normalisation schemes can result in different results, a point illustrated by McCrory & Ford (1991).

Within & across subjects

Alternatively, interest may lie in locating the brain regions increasingly engaged by progressively more difficult tasks. Subjects are scanned a number of times whilst undertaking a task, each subject performing the task at various levels of difficulty. The score for each scan is then simply either a measure of the task difficulty, or a measure of the subjects performance. The regions increasingly engaged by the task are located as those whose rCBF correlates positively with the scan score.

An example of this approach is a study where 12 subjects were scanned whilst attempting to remember lists of words. During each scan, word lists of given length were presented verbally by the experimenter, for immediate recall. Over the 12 scans each subject faced list lengths from 3 to 14 words in length, the actual order of allocation of list lengths to scans for each subject having been made randomly. The score for each scan is simply the length of the word lists presented during that scan. Regions increasingly engaged in longer term memory tasks should have rCBF highly correlated with score.

1.6.2.4. Cross group comparison

Other studies seek differences in function between groups. Typically a group of normal volunteers are compared with a group of patients suffering under some condition affecting their mental performance. Each group undergoes an activation study, and differences in response between the groups are sought.

1.6.2.5. Flow correlation, Functional connectivity

So far we have only considered experimental paradigms for locating particular functional centres in the brain. The brain can be thought of as a large and complex network of interconnected neurons. There are known neuronal projections between brain regions which are physically far apart. These connections should manifest themselves in PET data as correlations in the rCBF measurements of different regions. Recent efforts in functional brain mapping have sought to locate these connections, examining sequences of rCBF scans for correlation. These experiments are known as *functional connectivity* experiments .

Occasionally it is desired to locate the brain loci functionally connected to a known region. In this situation correlations between the rCBF of the hypothesised region and all other regions can easily be computed, and subsequently examined. Frequently there is no hypothesised primary brain loci. In these situations factor analytic approaches are being developed to elucidate the correlation structure.

1.6.3. Region of interest methods

There are two broad classes of analysis methods employed for PET data, *Region of Interest* (ROI) approaches and *voxel-by-voxel* approaches.

1.6.3.1. ROI methods

In ROI approaches, anatomical regions which are of interest are defined. These anatomical regions are identified on the PET scans. The average rCBF in each of the regions is calculated, usually by taking the mean value of the voxels inside the region. This results in a large reduction in the dimensionality of the data set. A typical rCBF image consists of approximately 77000 voxels, whereas typically only tens of ROI are defined.

ROI identification–Hypothesis generation

Usually the ROI are chosen with a particular hypothesis in mind. The functional anatomy of monkeys has been extensively studied using autoradiography, so regions of the human brain corresponding to those of the monkey whose function is known may be chosen in an experiment hoping to extend monkey results to humans. Alternatively, pilot studies or previous experiments may have suggested that a particular region be implicated in a certain task. Occasionally, the whole brain is partitioned into many small regions of interest, identified on anatomical grounds (Mazoyer *et al.*, 1993).

Analysis

When there are a selection of regions of interest, standard multivariate procedures have been applied. However, there are many problems, as indicated by Ford *et al.* (1991). The small sample sizes of functional mapping experiments results in data where often there are more ROI than there are replications. This results in a singular sample correlation matrix, precluding the use of many multivariate procedures.

When interest is in a particular region, a simple univariate statistical procedure can be used. With many ROI statistical testing can be performed on a region-by-region basis, using univariate techniques at each region. This approach is possibly more robust than the multivariate methods, though may be excessively conservative once adjustments for the many multiple comparisons have been made. An advantage of applying simple univariate statistical procedures is that medical researchers are more likely to be familiar with them.

Discussion

The review of Ford *et al.* (1991) discusses many of the statistical problems with ROI analyses, illustrated with real data sets from a variety of experiments. Section 4 of McColl *et al.* (1994) brings a review of ROI analyses up to date.

Many of the issues in ROI analysis are of concern in voxel-by-voxel methods, to which we now turn.

1.6.4. Voxel-by-voxel methods of analysis

1.6.4.1. Overview

The smallest brain region that can be defined for an image is that corresponding to a single voxel. Defining each voxel as a ROI also gives the highest resolution possible for locating a functional component. In the absence of any prior hypothesis regarding the expected location of the particular function under study, voxel ROI provide the framework for the most general analysis. These approaches are termed *voxel-by-voxel* methods.

Statistic images

Voxel-by-voxel approaches proceed by analysing the data for each voxel across the data using univariate techniques. This results in the computation of a statistic for each voxel, giving *statistic images*. The significance of these statistic images must then be assessed. Several schemes to account for the multiplicity of testing in this special spatial setting have been derived.

The formation and analysis of statistic images from activation experiments are the prime topics of this thesis. In chapters 2 and 3 we review current methods for formulating and assessing the significance of statistic images for PET functional mapping experiments. In the remainder of the thesis we consider new methods for assessing the significance of statistic images.

1.6.4.2. Re-alignment

Clearly for a voxel-by-voxel analysis it is imperative that the images obtained be in alignment, such that any particular voxel location corresponds to the same physical brain gyrus for all images.

During the course of a three hour scanning session, the subject is certain to move about on the scanner bed. This motion can be minimised by the use of custom made head holders and other restraints. These are rather uncomfortable if the head is held too rigidly, and may be distressing to the subject (fig.19). As it is the subjects thoughts that are of prime interest in activation studies, restraint must be as unobtrusive as possible, and most head restraints allow some movement. Subjects are encouraged to remain motionless for the duration of each scan, to prevent blurring the acquisition. Laser beams are often mounted on the tomograph gantry, and are aligned with lines drawn on the subjects head, enabling repositioning to approximately the same position at the start of each scan. This precaution helps minimise changes in head position and orientation relative to the tomograph between scans.

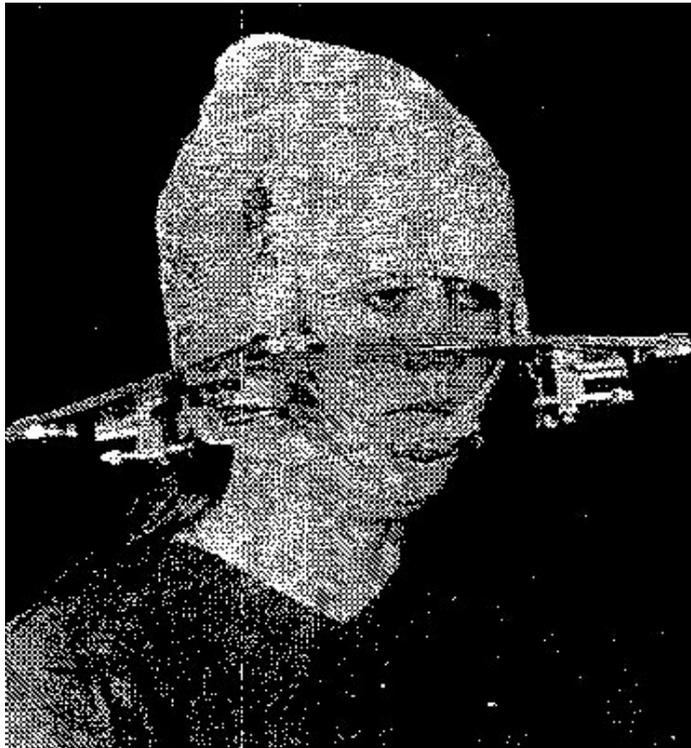


Figure 19
A (rather severe) head holder

Despite these precautions, motion between scans is likely, leading to mis-registration of the reconstructed images. To correct these translational and rotational motions various realignment procedures have been developed.

Image alignment

To align two images the co-ordinate axes of one must be rotated and translated to match the anatomical position and orientation of the axes in the other image. This transformation is parameterised by three rotational and three translational parameters. The new co-ordinate axes can be obtained by a rotation of θ about the z-axis, a rotation of ϕ about the y-axis, and a rotation of ψ about the x-axis, followed by a translation of the origin to point $\mathbf{o}' = (o'_x, o'_y, o'_z)^T$.

If the new axes are O'X', O'Y' and O'Z', then the co-ordinates of point $(x, y, z)^T$ relative to these new axes are $(x', y', z')^T$ relative to the old axes, where:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\phi\cos\theta & \cos\phi\sin\theta & -\sin\theta \\ \sin\psi\sin\phi\cos\theta - \cos\psi\sin\theta & \cos\psi\cos\theta + \sin\psi\sin\phi\sin\theta & \sin\psi\cos\phi \\ \sin\psi\sin\theta + \cos\psi\sin\phi\cos\theta & \cos\psi\sin\phi\sin\theta - \sin\psi\cos\theta & \cos\psi\cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} o'_x \\ o'_y \\ o'_z \end{pmatrix}$$

$$\mathbf{x}' = \mathbf{R}_{XYZ} \mathbf{x} + \mathbf{o}' \quad \text{Eqn.(10)}$$

Eqn.10 is a function of six parameters, three rotation angles and three translation distances. These parameters are known as the *reslice* parameters (Woods *et al.*, 1992).

Once the reslice parameters have been estimated for an image $A=(A_1, \dots, A_K)$, the image can be aligned to the new axes O'X' O'Y' O'Z' by re-sampling the image, using tri-linear interpolation¹⁴ to give new image $A'=(A'_1, \dots, A'_K)$ with cuboid voxels aligned with the (new) axes. Point \mathbf{x}' in the reoriented co-ordinates corresponds to $\mathbf{R}_{XYZ}^{-1}(\mathbf{x}' - \mathbf{o}')$ in the original co-ordinates.

Once two images from the same subject have been aligned, the same voxel from each image corresponds to the same volume of brain.

Estimation of reslice parameters: Fiducial Markers

Various methods have been devised to estimate the reslice parameters. Usually, one image is chosen as a reference image, and other images (*reslice images*) are aligned to it.

The simplest methods for realigning involve the use of *fiducial markers* whose positions are fixed relative to the subjects head. If radioactive markers are used then the markers show up on reconstructed images, and aligning the markers aligns the brain images. The reslice parameters satisfy $\mathbf{x}' = \mathbf{R}_{XYZ} \mathbf{x} + \mathbf{t}$ (eqn.10) with \mathbf{x}' the co-ordinates of a marker in the reference image and \mathbf{x} the co-ordinates of the same marker in the reslice image. With three (non-collinear) markers the resulting simultaneous equations can be solved for the reslice parameters $(\theta, \phi, \psi, t_x, t_y, t_z)$. These methods are rather crude, must be applied prospectively, and involve additional radiation exposure for the subject. Thus mathematical image alignment methods that can be applied retrospectively have been investigated.

Estimation of reslice parameters: Mis-alignment score methods

These methods for realigning two images proceed by minimising some mis-alignment cost function with respect to the reslice parameters. The reslice image is re-sampled with the reslice parameters that minimise the mis-alignment cost.

¹⁴Tri-linear interpolation is described fully in appendix A:2.

The cost functions proposed rely on the observation that two rCBF images on the same subject are similar, up to a scale factor (due to global changes), noise, and some localised (small) condition specific change. Mintun & Lee (1990) first applied this approach, using the stochastic sign change criterion (SSC) to provide a cost function. The misalignment cost function is computed as the number of sign changes (zero crossings) in the subtraction image formed by subtracting from the reference image the re-sampled reslice image for the current reslice parameters. When the images are aligned, the subtraction image should be mainly random noise and the number of sign changes large. The reslice parameters are sought as those maximising the number of sign changes. Minoshima *et al.* (1992) found this method sensitive to small global changes.

Eberl *et al.* (1993) used the downhill Simplex multidimensional minimisation technique to find reslice parameters minimising a misalignment cost function formed by summing the absolute difference in voxel values between the two images. This was found to be substantially more accurate than the SSC method, and robust to changes in gCBF.

Woods *et al.* (1992) developed a similar method that has become the standard realignment procedure, because of the availability of good software that has received extensive empirical validation. A proportional model is used, the idealised assumption being that when the two images are aligned, the value of any voxel in the reference image is roughly related to the value of the corresponding voxel in the reslice image by a single multiplicative factor. The misalignment cost function is the sample variance (computed over all voxels) of the ratio of the two scans.

With the emergence of single subject studies, there has been a move towards the use of superimposed anatomical MRI images of the same subject. To co-register the PET and MRI images, modifications of the various alignment techniques are used, the most successful being that of Woods *et al.* (1993b).

1.6.4.3. Anatomical normalisation

No two brains are physically identical, but the basic *gyral anatomy* (topology) is identical, and functions are tied to a particular gyrus. Since corresponding voxels in images of two different subjects will not correspond to the same brain gyrus, some form of anatomical normalisation is needed rectify this before analysis of images across subjects on a voxel-by-voxel basis. This normalisation is termed *stereotactic normalisation*.

Stereotactic atlas, reference brains

To standardise the study of neuroanatomy, various brain atlas have been proposed. Talairach & Tournoux (1988) dissected a human brain, photographing transverse slices onto which a co-ordinate system was overlaid and the brain gyri and lobes labelled. This brain has become a reference brain for much neuroanatomy, and the *Talairach co-ordinate system* has become a standard for describing brain locations. Hence, even with studies of an individual, anatomical normalisation is necessary to enable locations to be reported in standard Talairach co-ordinates.

The Talairach co-ordinate system is based on the *inter-commissural* or AC-PC line, the imaginary line joining the centres of the anterior and posterior commissures. The midpoint of the AC-PC line is the origin of the Talairach co-ordinate system (fig.20). The Y-axis runs along the AC-PC line from posterior to anterior, the X-axis from (subject) left to right, and the Z-axis from the bottom to the top of the brain. The YZ plane bisects the brain into its two hemispheres. Co-ordinates are given in millimetres from the origin in the axial directions. The *AC-PC plane* is the plane $z = 0$.

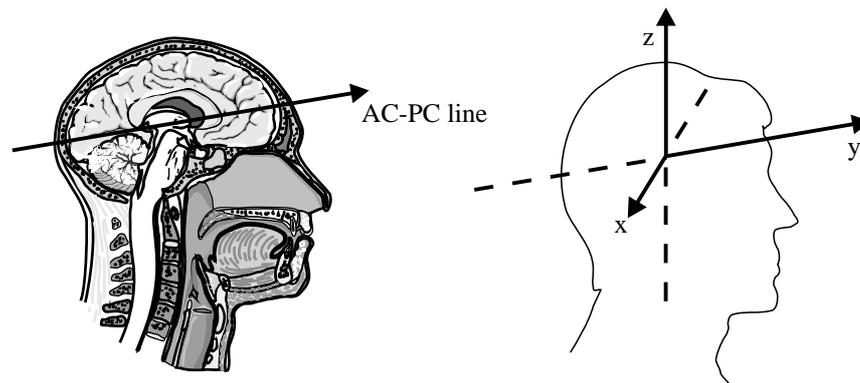


Figure 20
The intercommisural line, and Talairach axes

Estimation of AC-PC line

Most methods for stereotactic normalisation to Talairach co-ordinates begin by estimating the position of the AC-PC line. Early methods relied on external landmarks, head fixation and a secondary imaging modality showing enough anatomical detail to locate the AC-PC line by eye (Fox *et al.*, 1985). The improvement in resolution of PET images enabled the identification of anatomical landmarks such as the edges of the ventricles.¹⁵ The positions of certain landmarks have been shown to lie near the AC-PC line for many brains studied by MRI (Friston *et al.*, 1989). The AC-PC line can then be estimated by regression. Once the AC-PC line is fitted, the Talairach origin is found, and the X and Z axes located by fitting the YZ plane to bisect the brain into its two hemispheres.

Most centres now have templates for aligned PET images¹⁶ which can be compared to the study image to estimate the AC-PC line automatically. Typical of early approaches is that of Friston *et al.* (1991a), as implemented in early versions of the SPM software:¹⁷ The YZ plane is located by eye as the plane bisecting the brain into its two hemispheres, a Y-axis is chosen in this plane in the rough vicinity of the AC-PC line, and an origin arbitrarily chosen. The image is re-sampled to these new axes. The true AC-PC line lies in the YZ plane. To locate it, 15 coronal sections ($y = \text{constant}$) from the image are taken, and the height (z) of the AC-PC line in each is estimated.¹⁸ These heights are regressed on the y values of the coronal sections, giving the estimate of the AC-PC line.

¹⁵The brain contains several hollow pockets of spinal fluid, called *ventricles*. The spinal fluid is virtually free of tracer, and shows up black on PET scans of rCBF.

¹⁶Template images are usually formed as the average of many PET scans that have been stereotactically normalised.

¹⁷For a recent approach, see Friston *et al.* (1994). This is the method implemented in the SPM94 software.

¹⁸The height of the AC-PC line for each coronal section is estimated by comparison with the template image: Each section is aligned (by shift in Z-direction) with a selection of coronal sections from the template, and the best fit found. The height of the AC-PC line in the template sections is known, so adding this to the Z-shift for the best fit gives the estimated height of the AC-PC line.

Stereotactic normalisation

Once the Talairach origin and axes have been located, the PET image can be aligned to these new axes. There remains the problem of stereotactically transforming the PET brain to the atlas brain, deforming the image about the origin so that it has the same morphology as the Talairach brain.

The simplest methods simply scale the image in the axial directions so that the overall brain size matches the atlas (Fox *et al.*, 1985). After scaling the image is re-sampled to a standard resolution, typically with $2 \times 2 \times 4 \text{mm}^3$ voxels. This gives a stereotactical normalisation procedure with 9 degrees of freedom; the three rotations and three translation parameters of the alignment, and three scaling parameters.

The most general linear transformation from tomograph axes to Talairach co-ordinates is effected by three-dimensional *affine transformations*, a general linear transformation (except reflection or central inversion) followed by a translation (Woods *et al.*, 1993a). If \mathbf{x} is the position of a point on the tomograph axes then the co-ordinates after affine transformation are $\mathbf{x}' = \mathbf{L}\mathbf{x} + \mathbf{o}'$ where \mathbf{L} is a positive definite 3×3 matrix and \mathbf{t} the translation vector, giving a stereotactical normalisation procedure with 12 degrees of freedom. Once \mathbf{L} and \mathbf{o}' have been estimated the image can be re-sampled into Talairach co-ordinates with the desired voxel dimensions.

Reference brains

Woods *et al.* (1993a) have extended their alignment method (Woods *et al.*, 1992) to affine transformations, allowing anatomical normalisation to a reference brain for PET and MRI images. This has led to studies where all images are aligned within subjects, and then all subjects images anatomically normalised to a reference brain, either that of the first subject of a previously studied subject. Results are presented superimposed on the MRI of the reference brain, or on an average of the anatomically normalised subject MRIs, the blur in the latter giving some indication of variability in anatomy after normalisation.

Non-linear transformations

Friston *et al.* (1991c) proposed a general plastic transformation to reshape an image to a template formed as the average of previously stereotactically normalised images. A continuous general re-sampling function $r(x)$ was proposed such that point x in a one-dimensional continuous "image" $a(x)$ corresponds to point $r(x)$ in the standard space. Given a template image $t(x)$, $r(\bullet)$ was estimated to satisfy $t(x) = a(r(x))$ subject to $r'(x) \approx 1$, the latter restriction reflecting the fact that the distortion should be slowly varying. This was applied (Friston *et al.*, 1991a) to images previously oriented to the Talairach axes. A quadratic re-sampling function relating the heights of the image planes to the corresponding height in the template was estimated using polynomial regression, and the image volume re-sampled for this quadratically deformed z-axis. Each transverse plane was then re-sampled to a polar co-ordinate system and deformed to match the corresponding transverse plane from the template by estimating the general \mathbf{D} re-sampling function $r(\bullet)$ along each radii, and re-sampling. This approach gives good results and has been widely used, mainly due to the availability of software as part of older versions of the SPM package.

1.6.4.4. Primary smoothing

Images reconstructed by filtered back-projection are quite noisy. The highest (spatial) frequency of the noise is constrained only by the reconstruction filter used (typically a Hanning filter, with cut off frequency arbitrarily chosen at 0.5 cycles per

voxel). At this point in the analysis, it is usual to smooth the images. This *primary smoothing*, as it is known is applied for a number of reasons.

Motivation

In single subject experiments there are only 6–12 scans available for inference. Changes in rCBF induced by changes in neuronal activity are quite subtle, may be dominated by the noise in any one scan, and therefore difficult to detect. In the terminology of signal processing, the noise has a high spatial frequency, and blood flow a low frequency. Smoothing the image with a low-pass filter increases the “signal-to-noise”, making changes in rCBF more easily detectable, at the expense of a loss in resolution.

In multiple subject studies there are usually enough replications that subtle changes in rCBF between conditions are detectable, provided the activation has a coherent location in the stereotactically normalised images. Since the alignment and anatomical normalisation procedures are not perfect, and since humans exhibit limited biological variation in functional location, the same stimulus in different subjects will not necessarily cause rCBF changes at the same location in the anatomically normalised images. To allow for this residual functional variability, images are smoothed by running a moving average filter over the voxellated image. This extends spatially the influence of the actual activity at the brain location corresponding to a particular voxel, and hence allows slightly separated changes in rCBF from several individuals to “constructively interfere” in statistic images.

For basic brain functions (the senses, motor control) the anatomical localisation of function is particularly precise. Thus for the study of basic functions smoothing should not be used on the grounds of functional variability alone. Further, the changes in rCBF induced by neuronal activity in these areas quite marked, so smoothing may unnecessarily degrade resolution. The mechanics of smoothing are discussed in appendix B.

1.6.4.5. Adjusted images

After reconstruction, alignment within subject, anatomical normalisation (if required) and primary smoothing, the images are ready for a voxel-by-voxel analysis. Scans at this stage of the analysis are known as *adjusted images*. The same voxel from all the images has value reflecting neuronal activity in the same functional area in each of the subjects, subject to some errors in realignment, anatomical normalisation and natural biological variability in functional specialisation.

Example

Fig.21 depicts the rCBF image of fig.17 (p29), after stereotactical normalisation using the methods of Friston *et al.* (1989 & 1991a&b), and primary smoothing using a Gaussian filter of width 20mm×20mm×12mm.

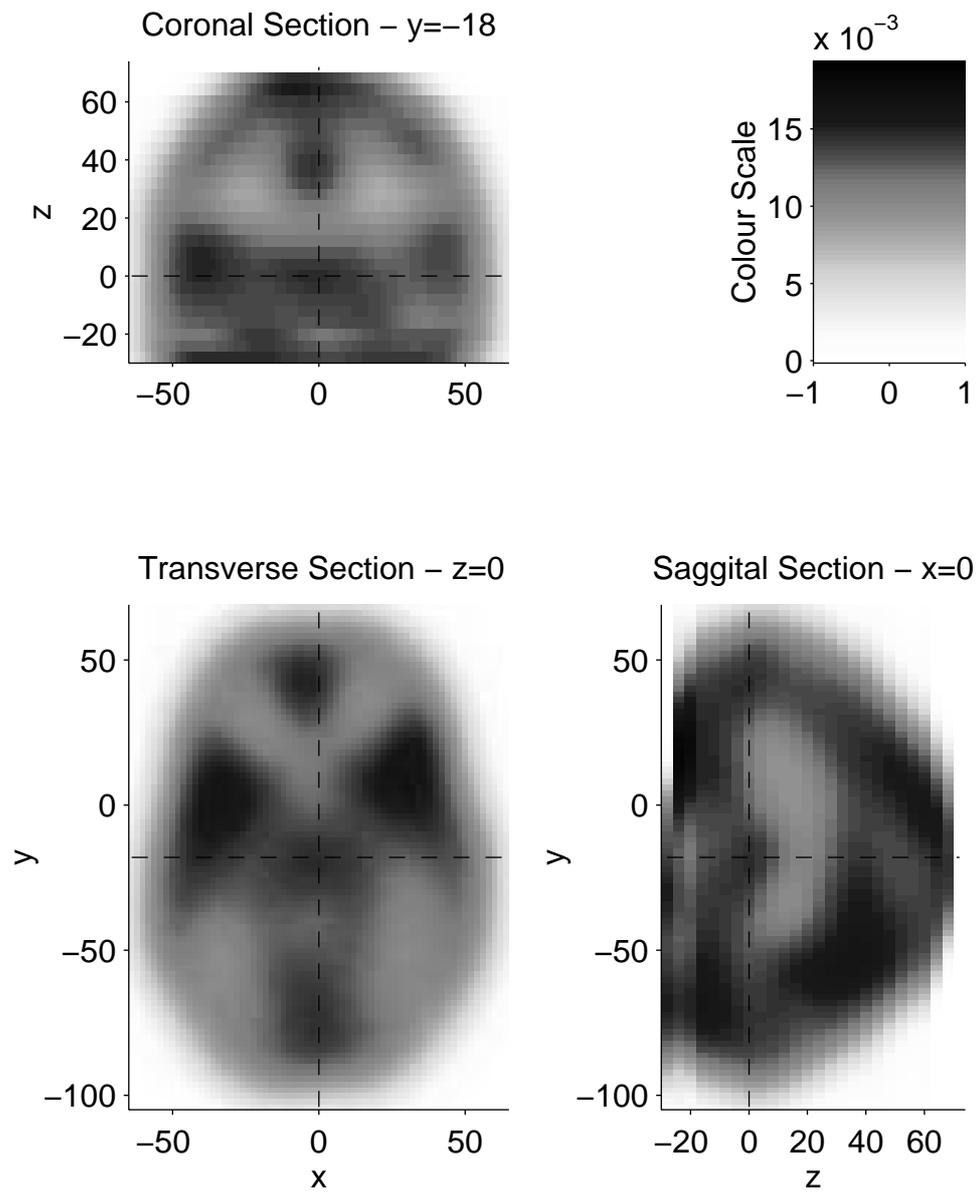


Figure 21

Orthogonal sections of the author's brain, after anatomical normalisation and smoothing. Colour scale in (estimated) annihilations per second per voxel. Co-ordinates are millimetres in Talairach space.

1.7. Data Set–The “V5” Study

Throughout this thesis we shall be applying the methods discussed to a real PET data set. The functional mapping experiment we shall use is a simple activation study, the “V5” study, recently published by Watson *et al.* (1993), where the reader is referred for full details of the experiment.

Data acquisition

In brief, there were twelve normal volunteers in the study group, nine males and 3 females, with ages ranging from 21 to 70 years. Each subject was scanned twelve times in a single 3hr scanning session, using the 3D acquisition mode of the CTI 953B PET tomograph. ^{15}O -Water was used as a freely diffusible blood flow tracer, infused according to the protocol of §1.6.1.2. The tomograph VOR data were corrected for detector efficiency and attenuation, and reconstructed by three-dimensional filtered back-projection with a Hanning filter cut off at 0.5 cycles per voxel. The discretisation of Ξ for reconstruction was of $128 \times 128 \times 31$ cuboid voxels of size $2.09 \times 2.09 \times 3.5 \text{mm}^3$ ($x \times y \times z$). The resulting estimated activity images had resolution $8.5 \times 8.5 \times 4.3$ mm at Full Width at Half Maxima (FWHM). These regional activity (r_A) images were used as r_{CBF} indicators, as arterial sampling and the subsequent quantitation to r_{CBF} was not carried out. Images were realigned within each subject using software provided by Woods *et al.* (1992). The brain was identified by reference to coregistered MRI images, and other areas removed from the images. Thus, all voxels with $r_A > 0$ correspond to regions within the brain tissue. The resulting images were transformed to the stereotactic atlas of Talairach and Tournoux using the methods of Friston *et al.* (1989, 1991), with re-sampling to a $65 \times 87 \times 26$ volume of $2 \times 2 \times 4 \text{mm}^3$ cuboid voxels. Finally images were smoothed by convolution with a Gaussian kernel of FWHM $10 \text{mm} \times 10 \text{mm} \times 12 \text{mm}$ to overcome functional variability and residual anatomical variability.

Scan conditions

There were two conditions under which scans were acquired. During condition A scans subjects viewed a random stationary pattern of small squares. For condition B the squares moved smoothly in random directions. The conditions were presented alternately, the series commencing with A in six randomly chosen subjects and B for the remaining six. The aim of the study was to identify visual area V5, the part of the visual cortex responsible for processing motion.¹⁹

Acknowledgement

The author wishes to thank JDG Watson (formerly a research fellow with the Medical Research Council Cyclotron Unit clinical sciences section) for assistance with the data, the Wellcome Trust for supporting the research, and S Zeki (Department of Anatomy, University College London) and RSJ Frackowiak (MRC-CU clinical sciences section) for permission to use the data.

¹⁹Previous autoradiographic studies of the Macaque monkey had demonstrated the specialisation of various parts of the visual cortex (which takes up the whole posterior of the brain). A 2D PET study (Zeki *et al.*, 1991) had indicated this functional specialisation was present in man, and found the approximate location of areas V5 and V4 (the colour centre of the visual cortex). The present study was undertaken to accurately locate area V5 in relation to the sulcal and gyral pattern of the cerebral cortex.

