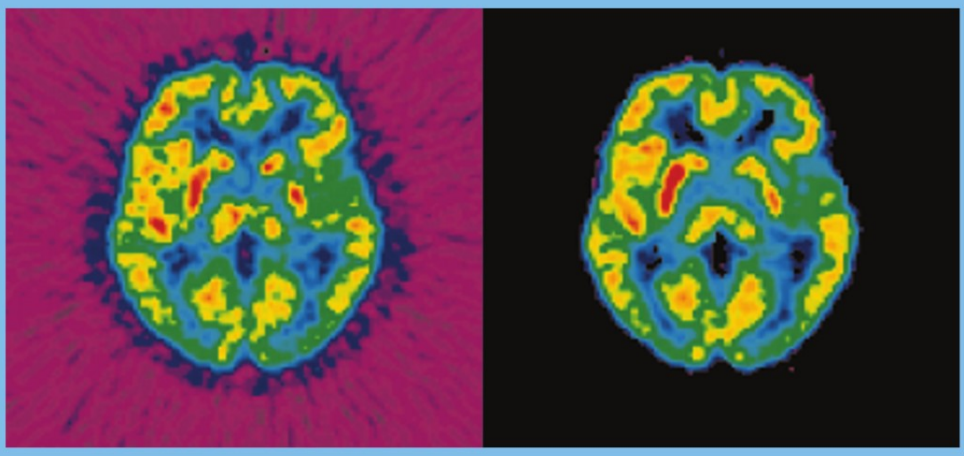
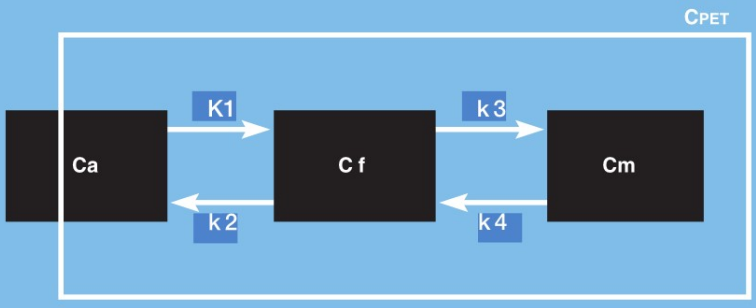
Quantitative Analysis in Nuclear Medicine Imaging

Edited by

Habib Zaidi





Springer

Quantitative Analysis in Nuclear Medicine Imaging

Habib Zaidi, Ph.D.
Editor

Quantitative Analysis in Nuclear Medicine Imaging

With 110 Illustrations

Springer

Habib Zaidi
Division of Nuclear Medicine
Geneva University Hospital
Geneva, CH-1211
Switzerland
habib.zaidi@hcuge.ch

Library of Congress Control Number: 2005931370

ISBN-10: 0-387-23854-9 e-ISBN 0-387-25444-7
ISBN-13: 978-0387-23854-8

Printed on acid-free paper.

Printed in Singapore. (SPI/KYO)

9 8 7 6 5 4 3 2 1

springeronline.com

Contents

List of Contributors

Habib Zaidi, Ph.D, PD Geneva University Hospital, Division of Nuclear Medicine, CH-1211 Geneva, Switzerland

M'hamed Bentourkia, Ph.D Dept of Nuclear Medicine and Radiobiology, University of Sherbrooke, Sherbrooke, Quebec J1H 5N4 Canada

Abdel Ouahab Boudraa, Ph.D Ecole Navale, Département Signal, Lanvéoc Poulmic BP600, F 29240 Brest, France

Michael Braun, Ph.D Department of Applied Physics, University of Technology, Sydney, NSW 2007, Australia

A Bertrand Brill, Ph.D Department of Radiology and Radiological Sciences, Vanderbilt University, Nashville, TN 37232-2675, USA

Eric C. Frey, Ph.D Division of Medical Imaging Physics, The Russell H. Morgan Department of Radiology and Radiological Sciences, Johns Hopkins University, Baltimore, MD 21287, USA

Guido Germano, Ph.D Department of Medicine, Artificial Intelligence in Medicine (AIM) Program, Cedars Sinai Medical Center, Los Angeles, CA 90048, USA

Bruce H. Hasegawa, Ph.D Department of Radiology, University of California, San Francisco, San Francisco, CA 94143-0946, USA

Brian F. Hutton, Ph.D Institute of Nuclear Medicine, University College London, London W1E 6BT, UK; also Centre for Medical Radiation Physics, University of Wollongong, NSW 2522, Australia

Kenneth F. Koral, Ph.D Dept. of Radiology, University of Michigan Medical Center, Ann Arbor MI 48109-0552, USA

David A. Mankoff, MD, Ph.D Department of Radiology, University of Washington Medical Center, Seattle, WA 98195, USA

Mark Muzi, Ph.D Department of Radiology, University of Washington Medical Center, Seattle, WA 98195, USA

Johan Nuyts, Ph.D Katholieke Universiteit Leuven, Department of Nuclear Medicine, B3000 Leuven, Belgium

Olivier G. Rousset, Ph.D Department of Radiology, Johns Hopkins University, Baltimore, MD 21287, USA

Piotr Slomka, Ph.D Department of Medicine, Artificial Intelligence in Medicine (AIM) Program, Cedars Sinai Medical Center, Los Angeles, CA 90048, USA

Michael G. Stabin, Ph.D Department of Radiology and Radiological Sciences, Vanderbilt University, Nashville, TN 37232-2675, USA

Benjamin M. W. Tsui, Ph.D Division of Medical Imaging Physics, The Russell H. Morgan Department of Radiology and Radiological Sciences, Johns Hopkins University, Baltimore, MD 21287 USA

Koenraad Van Laere, MD, Ph.D, DrSc Katholieke Universiteit Leuven, Division of Nuclear Medicine, B3000 Leuven, Belgium

Preface

This book provides a review of image analysis techniques as they are applied in the field of diagnostic and therapeutic nuclear medicine. Driven in part by the remarkable sophistication of nuclear medicine instrumentation and increase in computing power and its ready and inexpensive availability, this is a relatively new yet rapidly expanding field. Likewise, although the use of nuclear imaging for diagnosis and therapy has origins dating back almost to the pioneering work of Dr G. de Hevesy, quantitative imaging has only recently emerged as a promising approach for diagnosis and therapy of many diseases. An effort has, therefore, been made to place the reviews provided in this book in a broader context. The effort to do this is reflected by the inclusion of introductory chapters that address basic principles of nuclear medicine instrumentation and dual-modality imaging, followed by overview of issues that are closely related to quantitative nuclear imaging and its potential role in diagnostic and therapeutic applications. A brief overview of each chapter is provided below.

Chapter 1 presents a general overview of nuclear medicine imaging physics and instrumentation including planar scintigraphy, single-photon emission computed tomography (SPECT) and positron emission tomography (PET). Nowadays, patients' diagnosis and therapy is rarely done without the use of imaging technology. As such, imaging considerations are incorporated in almost every chapter of the book. The development of dual-modality imaging systems is an emerging research field, which is addressed in chapter 2. Different designs of combined tomographs were developed for diagnostic purposes in clinical oncology and are now commercially available. Various methods have been proposed to solve the problem of recovering the image from the measured projection data sets. There are two major classes of image reconstruction algorithms used: direct analytical methods and iterative methods. Still at present, the most widely used methods of image reconstruction are direct analytical methods because they are relatively quick. However, the images tend to be 'streaky' and display interference between regions of low and high tracer concentration. Images produced by iterative methods are computationally much more intensive; however, with the development of

parallel architectures and current generation PC clusters in the GHz range, the potential for using these in conjunction with iterative techniques, the problem becomes less awful. Chapters 3 and 4 describe the techniques and associated algorithms used in the reconstruction of nuclear medicine tomographic images.

Nuclear medicine imaging offers the possibility of quantitative measurements of tracer concentration *in vivo*. However, there are several issues that must be considered in order to fully realize this potential. In practice, the measured line integrals must be corrected for a number of background and physical effects, the most significant being the blurring introduced by the collimator response function (in SPECT), the limited spatial resolution and associated partial volume effect, photon attenuation and the contribution in the images of events arising from photons scattered in the object and the collimator/gantry assembly. Chapters 5-8 review the fundamental aspects related to the physics of these image degrading factors and algorithms and methods used to compensate for their effect.

There is increasing interest in being able to automatically register medical images from either the same or different modalities. Registered images proved to be useful in a range of applications, not only by providing more correlative information to aid in diagnosis, but also by assisting with the planning and monitoring of therapy, both surgery and radiotherapy. The objective of Chapter 9 is to present a general overview of medical image registration with emphasis on the application and issues relevant to nuclear medicine. Image segmentation plays a crucial role in many medical imaging applications by automating or facilitating the delineation of structures or organs and other regions of interest (e.g tumors). Chapter 10 examines specific problems and implementation issues related to segmenting noisy nuclear medicine images including transmission data obtained from external radionuclide scanning sources. A critical appraisal of the current status of semi-automated and automated methods for the segmentation of functional medical images is presented. Current segmentation approaches are reviewed with an emphasis placed on revealing the advantages and disadvantages of these methods for nuclear medical imaging applications. The use of image segmentation in other imaging modalities is also briefly discussed along with the difficulties encountered in each modality.

Because of their importance in image quantification, chapter 11 provides a comprehensive review of Monte Carlo modeling techniques relevant to nuclear medicine imaging and describes Monte Carlo software packages available for carrying out these complex and lengthy calculations. The kinetics and biodistribution of internally administered radionuclides must be measured for both diagnostic and therapeutic applications. The previous chapters are essential in understanding the mathematical aspects critical to the design of successful models for tracer kinetics described in chapter 12. The use of whole-body imaging for the determination of the amount of activity present in the body is of interest for the estimation of whole-body

radiation dose. For total activity within an isolated organ or tumour, one can use an imaging method that is simpler than 3D imaging with SPECT or PET. That is, one can get the total activity using only two opposed views of the patient, the so-called conjugate views, by sacrificing the knowledge about the 3D spatial distribution of the activity. The different methods used for quantification of planar nuclear medicine images are reviewed in chapter 13. Chapters 14, 15 and 16 examine specific implementations of quantitative functional imaging approaches in the special cases of brain, cardiac and oncologic imaging, respectively. The role of quantitative imaging in providing a means for estimating and guiding the overall scheme of imaging-based, patient-specific absorbed dose calculations is reviewed in chapter 17. This chapter provides also a review of the traditional dose schema, developed by the Medical Internal Radionuclide Dose (MIRD) Committee that is most widely implemented for radionuclide dosimetry.

Finally, I would like to thank all of the contributors for their invaluable hard work and for keeping to a very tight schedule. The topic of this book is rapidly evolving and the editor felt it important to minimise the time required to get this book into press. That this problem was so very well avoided speaks to the commitment and dedication of the contributors. I found compilation of this book to be a rewarding and educational experience and hope that the reader is left with the same experience.

H. Zaidi

1
Overview of Nuclear Medical Imaging: Physics and Instrumentation

H. Zaidi* and B.H. Hasegawa†

1. Introduction

Radionuclide imaging, including planar scintigraphy, single-photon emission computed tomography (SPECT) and positron emission tomography (PET), relies on the tracer principle, in which a minute quantity of a radiopharmaceutical in introduced into the body to monitor the patient's physiological function. In a clinical environment, radionuclide images are interpreted visually to assess the physiological function of tissues, organs, and organ systems, or can be evaluated quantitatively to measure biochemical and physiological processes of importance in both research and clinical applications. Nuclear medicine relies on non-invasive measurements performed with external (rather than internal) radiation sources and detectors in a way that does not allow the radionuclide measurement to be isolated from surrounding body tissues or cross-talk from radionuclide uptake in non-target regions.

Within the spectrum of macroscopic medical imaging (Figure 1), sensitivity ranges from the detection of millimolar concentrations of contrast media with CT and MRI, to picomolar concentrations in PET: a 10^9 difference.[1] With CT and MRI, contrast is produced by detecting differences in tissue density and water content; however, with radionuclide imaging, contrast is conferred by detection of a clearly identified molecule labelled with a radioactive isotope of one of its natural constituent elements. Signal sensitivity is a prerequisite for studies of biological pathways and binding sites which function at less than the micromolar level. It also is important to avoid the pharmacological effects of administering a labelled molecule to study its inherent biodistribution. The sensitivity of *in vivo* tracer studies is achieved *par excellence* with PET, which uses electronic collimation and thereby operates with a wide acceptance angle

*PD Dr H. Zaidi, Geneva University Hospital, Division of Nuclear Medicine, CH–1211 Geneva, Switzerland
†Prof. B.H. Hasegawa, Department of Radiology, University of California, San Francisco, CA, USA

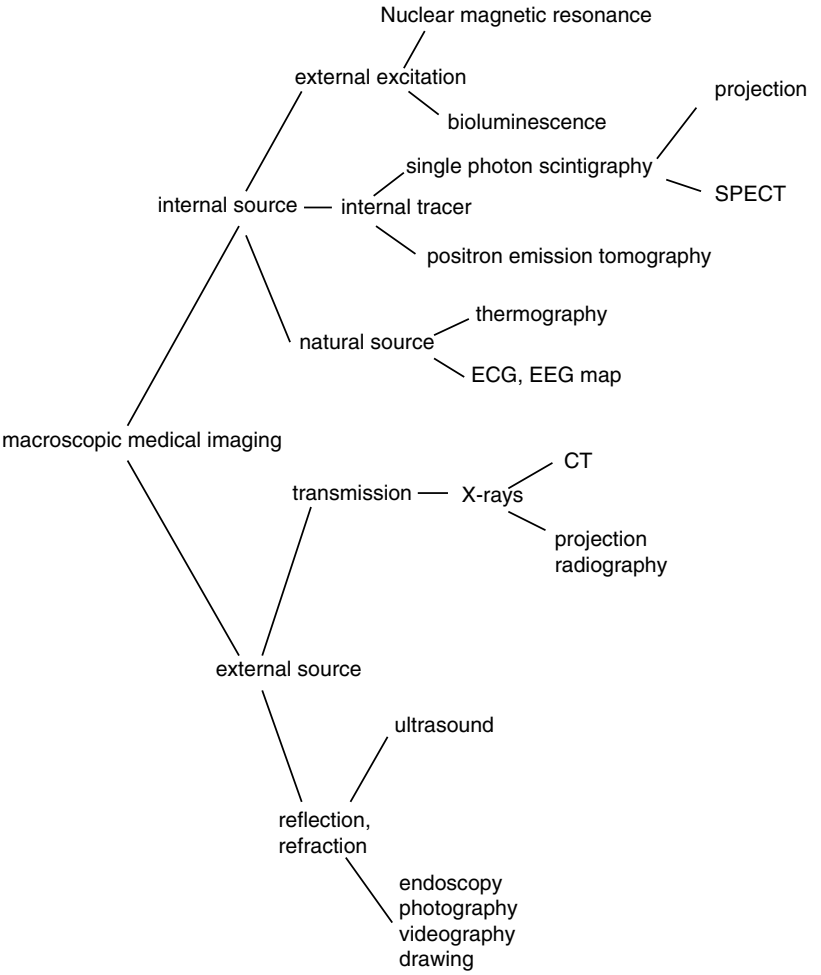


FIGURE 1. The wide spectrum of macroscopic medical imaging techniques (courtesy of Prof. F. Deconinck, Vrije University, Belgium).

for detecting emitted annihilation photons. Consequently, the sensitivity of PET per disintegration, with comparable axial fields of view, is two orders of magnitude greater than that of SPECT. PET also benefits by detecting radiopharmaceuticals that have short physical half-lives and high specific activities, which enable clinical and research studies to be performed at low radiation doses and with low molecular concentrations of the tracer.

Both the academic community and the nuclear medicine industry maintain a continuous learning cycle and assessment of products quality to advance the technology and the operational capabilities of both SPECT and PET cameras. As PET has become integrated into clinical practice, several design trends have developed; with systems now available with a

spectrum of features, from those designed for "low cost" clinical applications to others designed specifically for very high-resolution research applications. There also is a continual upward revision and refinement in both hardware and software components for all of these systems. The development of dual-modality imaging systems is an emerging research field (see chapter 2) and now offers unique capabilities for the medical imaging community and biomedical researchers. In this chapter, the physical principles, basic features and performance parameters of nuclear medicine instrumentation are outlined, and some of the practical issues involved in optimizing the design aspects discussed.

2. Planar Projection Imaging

2.1 The Anger Scintillation Camera

Virtually all commercial scintillation cameras used for imaging gamma-ray emitting radiopharmaceuticals are based on the original design proposed by Anger about 50 years ago, which is considered the working horse of contemporary nuclear medicine.[2] For these reasons, the scintillation camera often is called an "Anger camera" or "gamma camera". Figure 2 illustrates the principle and basic components of the Anger scintillation camera which incorporates a large scintillation sodium iodide crystal doped with thallium (NaI(Tl)), equipped with a parallel-hole collimator that limits the acceptance angle and defines the spatial distribution of gamma radiation viewed by the scintillator. Behind the crystal, a light guide is optically coupled to an array

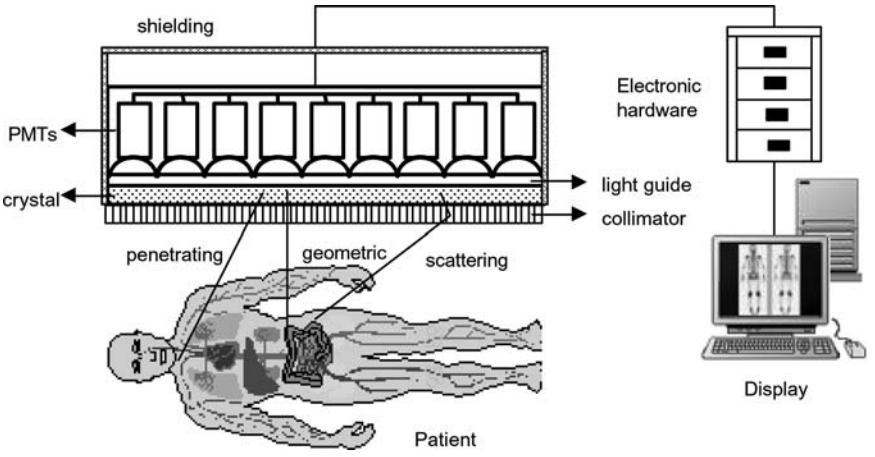


FIGURE 2. Schematic description of the principles and basic components of an Anger scintillation camera.

of light sensitive photomultiplier tubes (PMT's) that proportionately convert the distribution of scintillation light into electronic signals. The PMT's outputs then are processed using "Anger logic" electronics that generate output signals representing the (x,*y*) spatial position and energy of the individually detected gamma-rays on an event-by-event basis. Images displayed on the console represent the accumulation of the individual events recorded during an imaging study. Depending on the size of the scintillation camera, whole organs such as brain, heart and liver can be imaged. In many cases, large scintillation cameras are capable of imaging the entire body and are used, for example, in whole-body skeletal imaging.

2.2 Collimators

The parallel-hole collimator is configured as a honeycomb with lead or tungsten septa that form thousands of straight parallel channels so that the scintillation camera records only those photons that strike the detector surface from an approximately perpendicular direction. A typical parallel-hole collimator might have a bore diameter of roughly 2 mm, and a bore length of approximately 4 cm giving a high degree of angular selectivity. In this way, the collimator design is the most important factor in determining the geometric efficiency or sensitivity of the imaging system. For example, a typical Anger camera equipped with a low-energy parallel-hole collimator detects roughly one in every ten thousand gamma rays emitted isotropically by a point source in air. Even fewer gamma rays will be detected if the point source is located within the body or other object, due to the effects of photon attenuation by the material surrounding the source. Therefore, the system sensitivity (obtained with the collimator) is always worse than intrinsic sensitivity of the detector operated without the collimator. Similarly, the collimator also largely determines the spatial resolution of the imaging system. For example, a typical modern Anger camera has an "intrinsic" spatial resolution of 3 to 4 millimetres, representing the precision of spatial localization obtained for a point source placed directly on the camera surface without a collimator. In comparison, the system spatial resolution obtained with a collimator might be 7 to 12 millimetres at a detector-object distance of 10 cm, with the spatial resolution degrading linearly with distance. The system spatial resolution depends heavily on the type of collimator used and to a less extent the intrinsic resolution of the detector. For example, a scintillation camera equipped with an ultra-high resolution parallel-hole collimator can achieve a system spatial resolution in the range of 4 to 8 millimetres.

Fortunately, the design of the collimator can be changed to improve geometric sensitivity; however, this produces a corresponding loss of geometric resolution. Alternatively, changing the collimator design to improve collimator resolution will decrease collimator sensitivity. Therefore, researchers must always consider this trade-off when working on new colli-

mator designs since high resolution and great sensitivity are two paramount (but often competing) goals of gamma camera imaging.[3] Obviously, sensitivity has seen an overall improvement by the introduction of multi-detector SPECT systems which now are common in clinical use.

There have been several collimator designs in the past fifteen years, which optimised the trade-off between resolution and sensitivity for a given detector design.[4] Converging-hole collimators, for example those configured with fan-beam and cone-beam geometries[5] have been built to improve both resolution and sensitivity by increasing the amount of the Anger camera that is exposed to the radionuclide source. This increases the number of counts, which improves sensitivity, but limits the field-of-view. Other types of collimators with only one or a few channels, called pinhole collimators, have been designed to image small organs and human extremities, such as the wrist and thyroid gland, in addition to small animals.[6,7] The latter application has received considerable attention during the last decade with the aim of achieving spatial resolutions better than 1 mm.[8-10] More recent collimator designs, such as planar-concave,[11] half fan- and cone-beam[12] and astigmatic,[13] have also been conceived. Other types of collimators have also been reported, including rotating slit/slat collimators to increase the geometric efficiency resulting from the enlarged solid angle of acceptance afforded by these designs.[14-17]

2.3 *Scintillation and Solid-State Detectors for Gamma Cameras*

The first systems for imaging radionuclides that emit gamma-rays and positrons were designed with single large NaI(Tl) scintillation crystals. NaI(Tl) is the most common detector for scintillation cameras owing to its relatively high atomic number and corresponding high probability of photoelectric absorption for the energies involved in conventional nuclear medicine imaging. The development of dedicated small field of view pixelated cameras has received considerable attention during the last decade. The pixelated crystal limits the degree to which the scintillation light spreads laterally, and thereby can improve spatial resolution in comparison to cameras that use continuous crystals.[18] A typical design used 4 mm thick CsI(Tl) crystal with a 1.13 mm pixel pitch readout by position sensitive PMT's (PSPMT's).[19] It should be emphasized that such a design improves spatial resolution at the expense of deteriorating the energy resolution resulting from light losses in the pixelated crystal compared to that of a single crystal. However, cameras with pixelated scintillators also can have the scintillator segments coupled to individual photomultiplier tubes, allowing them to be operated somewhat independently of one another to increase count-rate capabilities for first-pass and other high count-rate imaging applications.

The development of strip solid-state detectors for medical imaging is motivated by their excellent energy resolution and direct gamma radiation conversion, thus removing the need for using PMT's, and thereby allowing the system to be more compact, lighter in weight, and more rugged. The better energy resolution of semiconductor detectors is the consequence of the lower energy needed to create an electron-hole pair (3-6 eV) compared to the energy required to create a scintillation photon (~ 30 eV) in a conventional NaI(Tl) scintillation crystal. In addition, solid-state detectors obviously do not suffer the light losses that occur between a scintillation crystal and the PMT's[15] which also improves the signal generation process in the camera. Among available semiconductor detectors, high-purity germanium (HPGe) offers the best energy resolution and thereby allows efficient rejection of Compton scattering in the patient.[20] One such a system designed for clinical applications had 154 (250 mm long, 1.62 mm wide, 12 mm thick) detector strips collimated by parallel tungsten sheets perpendicular to the detector face.[21] However, due to its narrow energy band-gap, HPGe detectors must be operated at cryogenic temperatures. As a consequence, the need for cooling prevented their widespread applicability and encouraged the investigation of the potential of semiconductor detectors operating at room temperature including mercuric iodide (HgI_2), cadmium telluride (CdTe) and cadmium zinc telluride (CdZnTe or CZT) either in the form of single detectors or as segmented monolithic detectors.[22] Both CdTe and CZT currently are regarded as especially promising candidates for nuclear medicine imaging applications with the aim of replacing the conventional NaI(Tl) scintillator for clinical practice, especially for small-field applications such as sentinel node detection and radiopharmaceutical-guided surgery.[23]

One interesting design, the SOLid STate Imager with Compact Electronics (SOLSTICE), has the potential of producing a quantum advance in the performance of nuclear medicine instrumentation. This system uses a room-temperature semiconductor detector (CdZnTe) to offer the excellent signal characteristics obtained with direct gamma ray conversion. The system also uses a novel rotating slat collimator (Figure 3) which improves both spatial resolution and detection efficiency in a way that relaxes the performance limitations imposed on scintillation cameras by conventional parallel-hole collimators.[24] While the instrument is in a development phase, some promising results have been achieved during the last few years, including studies confirming the significant potential of this design for high resolution small animal imaging.[15] The practicality, clinical utility, and cost-effectiveness of the system still needs to be demonstrated.

2.4 Photodetectors

In nuclear medicine, photodetectors are used to measure the light emanating from the scintillation crystal, and must be used in a way that maintains the spatial resolution, energy resolution, and detection efficiency of the entire

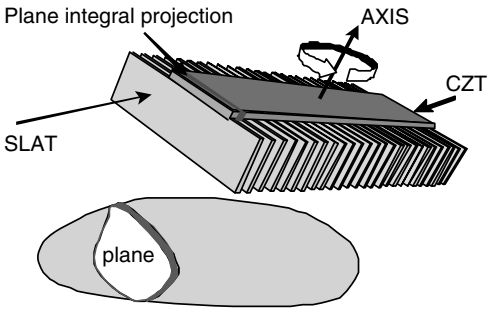


FIGURE 3. Schematic illustration of the SOLSTICE prototype system. Plane integrals are measured by the SOLSTICE imaging system for each projection angle in contrast to line integrals measured in the conventional approach. A specially designed reconstruction algorithm which handles this acquisition geometry has been developed for this purpose. (Courtesy of Dr D. Gagnon, Philips Medical Systems).

imaging system. When the light output of the scintillator is small, the noise of the photodetector can significantly contribute to the loss of energy resolution. The emission wavelength of the scintillator light must be carefully matched with the spectral response of the photodetector. Virtually all commercial gamma camera imaging devices available operate by coupling the scintillation crystal to an array of single channel PMT's. However, special cameras for imaging small organs (*e.g.*, female breast) and small animals use pixelated photodetectors for radionuclide imaging applications. In the near future, it also is possible that multi-anode PSPMT's, silicon p-i-n photodiodes (PDs) or avalanche photodiodes (APDs) will be incorporated into scintillation cameras.

Solid-state photodiodes exhibit many advantages compared to conventional PMT's. They are relatively small, operate at much smaller voltage, and more importantly exhibit higher quantum efficiencies. Furthermore, photodiodes are insensitive to axial and transversal strong magnetic fields and therefore have the potential to be operated within magnetic resonance imaging systems. By using this technology, the sensitive area of the detector could be read-out more efficiently taking advantage of recent developments of monolithic pixelated photodetectors. Thanks to the progress made in the microelectronic industry, high-purity detector grade silicon now is available commercially and can be used to produce low-noise silicon planar photodiodes, avalanche photodiodes, and silicon drift photodetectors.[25] Commercially available APDs typically are not pixelated and can have a sensitive area of $\sim 20\,\mathrm{mm}^2$. PDs and APDs also can be segmented to form a linear or two-dimensional array of pixels on a single device.[26] Their geometry can be modified to suit specific applications, with the configuration of the scintillator matrix adapted to the available read-out pixelation of the

photodetector. However, given the current cost of solid-state photodiodes, they need to offer very significant advantages to replace the currently adopted technology.

Although not previously used in medical imaging, hybrid photodetectors (HPDs) represent an interesting technology with potential for high resolution photon detection.[27] An HPD consists of a phototube with two key components: a semi-transparent photocathode (deposited by vacuum evaporation on the entrance window) with high sensitivity for photons in the visible and near UV range, and a silicon (Si) sensor which serves as the anode.[28] The photocathode receives light from a scintillator and converts the incident photons to electrons. These photoelectrons are accelerated by an electrostatic field (10-20 kV between cathode and anode), which is shaped by a set of ring electrodes, onto the segmented Si sensor; the anode thereby creates a signal proportional to the kinetic energy of the incident photoelectrons. In this way, the HPD combines the sensitivity to single photons of a conventional PMT with the spatial and energy resolution of a solid-state sensor. In addition, this design overcomes the intrinsic limitations of a classical PMT with respect to the statistical fluctuations in the number of electrons at the first dynodes. Proximity focused HPDs can be operated in strong magnetic fields, as long as the field direction is aligned with the tube axis. Axial magnetic fields have even the beneficial effect of reducing the intrinsic spatial resolution of the device, which is a consequence of the angular and energy distribution of the photoelectrons at emission from the photocathode.[29] HPDs have been incorporated as major components of a PET camera design discussed in section 6.[30] Nuclear medicine instrumentation continues to benefit from the ongoing advances in microelectronic technologies, which have led to highly integrated CMOS front-end circuits, fast data acquisition processors and ultra rapid field programmable gate arrays (FPGAs).

3. Single-Photon Emission Computed Tomography

Planar radionuclide imaging (also called planar scintigraphy) is used extensively in clinical routine practice and offers the advantage of fast acquisitions over a large area at a relatively low cost. The drawback with this technique is, however, the lack of information regarding the three-dimensional spatial distribution of radioactivity in the body (see chapter 13). Furthermore, the acquired images can exhibit limited contrast since the radioactivity signal from the target often is combined with that from overlapping structures. In contrast, SPECT produces images that represent the three-dimensional distribution of radioactivity. This improves both image quality as well as the potential for the quantification of the radioactivity distribution *in vivo*. Transverse tomographic images can be reconstructed from projection data acquired at discrete angles around the object

using one of the algorithms described in chapters 3 and 4. There are, however, multiple factors that must be considered in quantitative tomographic imaging. These factors include the system sensitivity and spatial resolution, dead-time and pulse pile-up effects, the linear and angular sampling intervals of the projections, the size of the object, and photon attenuation, scatter, and partial volume effects, voluntary and involuntary patient motion, kinetic effects, and filling of the bladder (and other normal excretory routes) by the radiopharmaceutical. Since image quality in nuclear medicine is limited by the photon statistics of the acquired data, the administered dose and the imaging time are extremely important. In practice, the limited count statistics in most clinical studies affect the accuracy and precision of quantitative SPECT.

SPECT has become a major tool for the *in vivo* localisation of radiopharmaceuticals in nuclear medicine and now is performed routinely with commercially available radiopharmaceuticals to answer important clinical questions including those in cardiology, neurology, psychiatry, and oncology. In conjunction with new and existing radiopharmaceuticals, quantitative SPECT may be used noninvasively to measure blood-flow, metabolic function, receptor density, and drug delivery. In oncology, quantitative SPECT is important in radiation dosimetry and treatment planning for internal radionuclide therapy and specifically for radioimmunotherapy.[31]

Advances in dedicated SPECT instrumentation may advance the use of clinical high resolution imaging of the brain. These systems include the recently marketed NeuroFocus™ multi-conebeam imager (Neurophysics Corporation, Shirley, MA), which produces radionuclide images with an intrinsic resolution of ~ 3 mm. The operation of the NeuroFocus™ scanner follows the same principles as scanning optical microscopes to obtain high-resolution, three-dimensional images of biological tissue. A highly focused point of light is scanned mechanically in three dimensions to uniformly sample the volume under observation. Since the energetic gamma rays emitted by single-photon tracers cannot be focused by conventional optics, NeuroFocus™ uses proprietary "gamma-lenses™" known as scanning focal-point technology.

3.1 Compton Cameras

Except for those that use coded apertures, all collimated imaging systems exhibit a limiting detection sensitivity that is inversely proportional to the system's spatial resolution. This fundamental trade-off has motivated the development of Compton cameras (see Figure 4), which provide information about the incoming photon direction electronically without any restriction with respect to the solid detection angle.[32] The mechanical collimation of the Anger camera is thus replaced by "electronic collimation" hereby removing the coupling between sensitivity and spatial resolution. This is achieved by

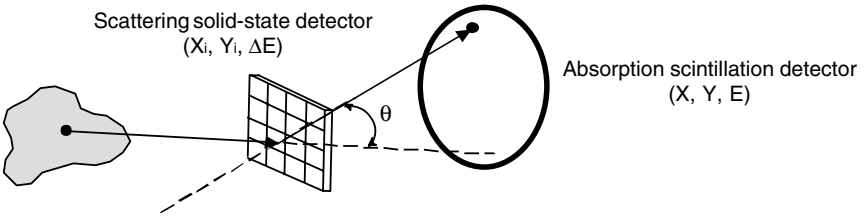


FIGURE 4. General principle of the Compton camera approach showing the scattering and absorption detectors.

having two detector modules where the gamma rays are first scattered in a solid-state detector and then absorbed in a second scintillation detector. After Compton scattering in the first detector, the scattered photon emerges from the interaction point with less energy and in a different direction than the incident photon. The tandem detectors in the Compton imager record the energies and interaction coordinates of the incident and scattered photons, respectively. This information can be used to calculate the scattering angle θ and the direction of the incident gamma ray. The precision in the measurement of the scattering angle θ depends mainly on the energy resolution of the first detector. The classical Compton equation expresses the energy of the scattered-photon as a function of the initial photon energy and the scattering angle θ, and assumes that the incident photon interacts with a free electron at rest.[33] There are corrections to that equation which take into account the fact that the electron actually is not at rest and is bound to an atom. The result is that the photons that scattered through a given angle actually have a distribution of energies sharply peaked about the value calculated by the classical Compton equation.[34] This effect, called Doppler broadening, constitutes an inherent limitation for Compton cameras. In Compton camera imaging, the measured scattering angle therefore is associated with an uncertainty which degrades the spatial resolution of reconstructed images. This uncertainty becomes bigger as the incident gamma ray energy decreases, thereby motivating the development of appropriate compensation methods for this effect.[35]

The first application of Compton imaging to nuclear medicine was proposed in 1974.[36] This was followed by series of seminal papers by Singh and co-workers describing analytical and experimental results of a Compton camera using pixelated germanium as the first detector and a standard Anger camera as second detector.[37] This work was continued at the University of Michigan in collaboration with CERN[38] leading to the development of the C-SPRINT: a prototype Compton camera system for low energy gamma ray imaging [39] and the design of a Compton camera for high energies with CZT detectors.[40] More recently, the potential of the Compton camera approach for scintimammography has also been reported.[41] It is also expected that working Compton cameras based on silicon microstrips and

segmented germanium detectors will be demonstrated in the near future.[39,42] In parallel, appropriate analytic[43,44] and iterative[45] image reconstruction algorithms were developed specifically for single-photon images acquired with electronic collimation. Significant development still is required before this technology can be considered to be practical and cost-effective in a clinical arena owing to the complexity of the detector technologies, data acquisition system, and image reconstruction techniques needed for Compton imaging. Nevertheless, significant progress has been made during the last decade in demonstrating the feasibility of medical imaging with a Compton camera.

4. Positron Emission Tomography

4.1 The Physical Principles of PET

When imaging positron-emitting radiopharmaceuticals, pairs of antiparallel 511 keV photons arising from electron-positron annihilations are recorded by block detectors surrounding the patient. A positron emission tomograph consists of a set of detectors usually arranged in adjacent rings that surround the field-of-view in order to image the spatial distribution of a positron-emitting radiopharmaceutical (Figure 5). In this case, the annihilation photons traverse a total tissue thickness that is equal to the body thickness intersected by the line between the two detector cells, also called the *line of response* (LOR).

After being sorted into parallel projections, the LORs defined by the coincidence channels are used to reconstruct the 3D distribution of the positron-emitter tracer within the patient. An event is registered if both crystals detect an annihilation photon within a fixed coincidence time window (between 6 and 12 ns) depending on the timing properties (decay time) of the scintillator. A pair of detectors is sensitive only to events occurring in the tube of response joining the two detectors, thereby registering direction information (electronic collimation). The finite positron range and the non-collinearity of the annihilation photons give rise to an inherent positional inaccuracy not present in conventional single-photon emission techniques. However, other characteristics of PET more than offset this disadvantage. There are many advantages associated to coincidence detection compared to single-photon detection devices; electronic collimation eliminates the need for physical collimation, thereby significantly increasing sensitivity and improving the spatial resolution. However, accurate corrections for the background and physical effects are essential to perform absolute measurements of tissue tracer concentration with PET.[46]

Coincidence events in PET fall into 4 categories: trues, scattered, randoms and multiples. A true coincidence is one in which both photons from a single annihilation event are registered by detectors in coincidence. In this case

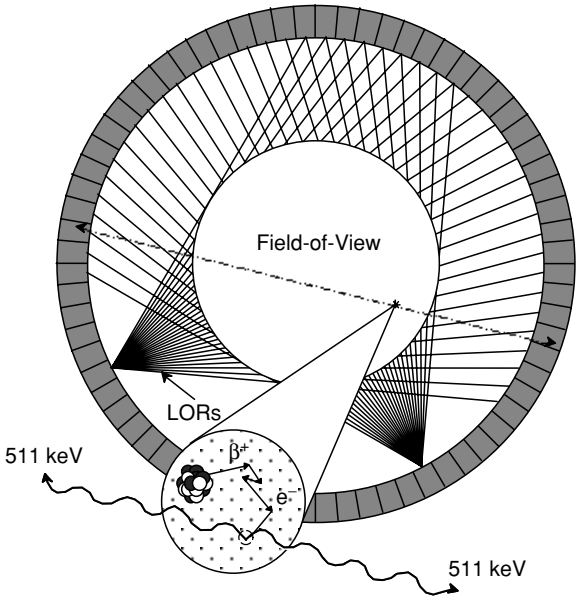


FIGURE 5. Originating from the decay of the radionuclide, a positron travels a few millimetres (depending on its energy and electronic tissue of the medium) before it annihilates with a nearby atomic electron, producing two 511 keV photons emitted in nearly opposite directions. A positron emission tomograph consists of a set of detectors usually arranged in adjacent rings surrounding the FOV. Pairs of annihilation photons are detected in coincidence. The size of the transaxial FOV is defined by the number of opposite detectors in coincidence.

neither photon undergoes any form of interaction prior to detection, and no other event is detected within the defined coincidence time window. A scattered coincidence occurs when at least one of the detected annihilation photons undergoes one or multiple Compton scatterings prior to detection. Since the direction of the Compton-scattered photon is changed, the resulting coincidence event will be assigned to the wrong LOR, thus decreasing the contrast and overestimating the activity concentration of the tracer. When two uncorrelated photons (i.e., photons not arising from the same annihilation event) are incident on block detectors within the system's coincidence time window, the coincidence is called random (or false). Finally, multiple coincidences occur when more than two photons are detected in different crystals within the same coincidence window. In this situation, it is not possible to determine the LOR to which the event should be assigned, and the event is rejected. Multiple events can also cause event mispositioning.[46]

4.2 Design Geometries for PET

There has been a remarkable progress in PET instrumentation design from a single ring of bismuth germanate (BGO) crystals with a spatial resolution of $\sim$ 15 mm, to multiple rings of detector blocks resulting in a spatial resolution of about 4-6 mm. Improvements in spatial resolution have been achieved by the use of smaller crystals and the efficient use of light sharing schemes to identify the active detector cell. Figure 6 illustrates possible geometric designs of PET imaging systems adopted by scanner manufacturers. These include large or pixelated detectors mounted on a rotating gantry, detectors arranged in a partial or full multi-ring geometry, and detectors assembled in a polygonal ring.[25]

Dedicated full-ring PET tomographs are still considered to provide state-of-the-art performance for whole-body imaging but have evolved through at least 4 generations since the design of the first PET scanner in the mid 1970s. The improved performance of full-ring systems compared to camera-based dual or triple-headed systems is due to higher overall system efficiency and

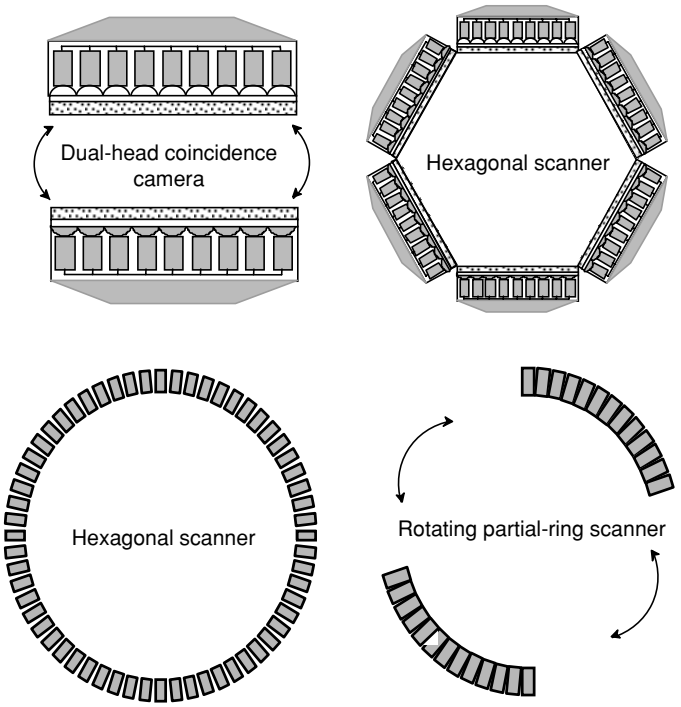


FIGURE 6. Illustration of the range of different geometries of PET imaging systems. The dual-head coincidence camera and partial ring tomographs require the rotation of the detectors to collect a full 180° set of projection data. The hexagonal design geometry could be designed using either continuous or pixelated crystals.

count rate capability which provides the statistical realization of the physical detector resolution and not a higher intrinsic physical detector resolution.[47] Obviously, this has some important design consequences since even if both scanner designs provide the same physical spatial resolution, the full-ring system will produce better quality images in patients as a result of its improved sensitivity.[48] The most important aspect related to the outcome of research performed in the field is the improvement of the cost/performance optimization of clinical PET systems.

A modern multi-ring tomograph with septa extended detects only $\sim 0.5\%$ of the annihilation photons emitted isotropically from the activity within the tomograph field-of-view. This increases to over 3% when the septa are retracted.[48] However, even if the detector system is 100% efficient for the detection of annihilation photons, the angular acceptance of modern scanners would record only 4-5% of the total coincidences. With the exception of dedicated high resolution PET tomographs (e.g. the high resolution research tomograph – HRRT[49]), the spatial resolution obtained with modern commercial tomographs is currently limited to 4-6 mm in all three directions.

4.3 Scintillation Crystals for PET

The critical component of PET tomographs is the scintillation detector.[50] The scintillation process involves the conversion of photons energy into visible light via interaction with a scintillating material. Photoelectric absorption and Compton scatter generate electrons of differing energy distributions. When scintillation detectors are exposed to a monoenergetic photon beam, the measured energy is not that of the electron generated by the initial interaction, but rather the total energy deposited by the photon in the detector through one or more cascaded interactions. In small crystals, photons may escape after depositing only part of their energy in the crystal. In practice, the energy distribution is also blurred by the finite energy resolution of the detection system as contributed by electronic noise and other effects. The energy resolution (in percent) of the system is defined as the ratio of the full-width at half-maximum (FWHM) of the photopeak and the mean (or peak) energy value of the photopeak. Increased light yield, faster rise and decay times, greater stopping power and improved energy resolution, in addition to low cost, availability, mechanical strength, moisture resistance, and machinability are the desired characteristics of scintillation crystals. Table 1 summarizes most of these properties for selected scintillators currently in use or under development for PET applications.[51,52] Improvements in these characteristics enable detectors to be divided into smaller elements, thus increasing resolution and minimising dead-time losses.

The choice of the scintillator is a fundamental element of a PET design. The selection is generally made after careful consideration of the physical characteristics including light yield, energy resolution, linearity of response with energy, light absorption coefficient at wavelength of emission,

TABLE 1. Characteristics of scintillation crystals under development and used in the design of current generation PET imaging systems.

Scintillator	BGO	LSO	GSO	LuAP	$LaBr_3$	LYSO
Formula	$Bi_4Ge_3O_{12}$	Lu_2SiO_5: Ce	Gd_2SiO_5: Ce	$LuAlO_3$: Ce	$LaBr_3$: Ce	$LuYSiO_5$: Ce
Density (g/cc)	7.13	7.4	6.71	8.34	5.3	7.1
Light yield (photons/keV)	9	25	8	10	61	32
Effective Z	75	66	60	65	46.9	64
Principal decay time (ns)	300	42	60	18	35	48
Peak wavelength (nm)	480	420	440	365	358	420
Index of refraction	2.15	1.82	1.95	1.95	1.88	1.8
Photofraction (%)*	41.5	32.5	25	30.6	15	34.4
Attenuation length (cm)*	1.04	1.15	1.42	1.05	2.13	1.12
Energy resolution (%)*	12	9.1	7.9	11.4	3.3	7.1
Hygroscopic	No	No	No	No	Yes	No

*@ 511 keV

mechanical properties and surface quality of crystals, availability and cost.[53] BGO has a very high physical density and effective atomic number, and is not hygroscopic. These properties rendered it the preferred scintillator for commercial PET units in the 1990s. Its major disadvantages are, however, the low light yield and only a moderately fast decay time that limits coincidence timing and count rate performance.

New detection technologies that are emerging include the use of new cerium doped crystals as alternatives to conventional BGO crystals, and the use of layered crystals and other schemes for depth-of-interaction (DOI) determination. It appears that cerium doped lutetium orthosilicate (LSO:Ce) produced by CTI Molecular Imaging (Knoxville, TN), lutetium yttrium orthosilicate (LYSO:Ce) produced by Photonic Materials Ltd. (Bellshill, UK) and cerium doped lanthanum bromide ($LaBr_3$:Ce), under development by Saint Gobain (France), are the most promising candidates. They combine high density and high atomic number necessary for an efficient photoelectric conversion of the annihilation photons, with a short decay time of the scintillation light, which is a key requirement for high counting rates. Phoswich detectors received considerable attention for the design of high resolution scanners dedicated for brain, female breast (positron emission mammography – PEM) and small animal imaging. This may be implemented with solid-state photodiode readouts, which also allows electronically collimated coincidence counting. Such a design has been implemented on the ECAT high resolution research tomograph (HRRT) with LSO crystals and PMT's-based readout.[49] Figure 7 illustrates the principle of the conventional detector block[54] and the phoswich approach[55] where two detectors are assembled in a sandwich-like design, the

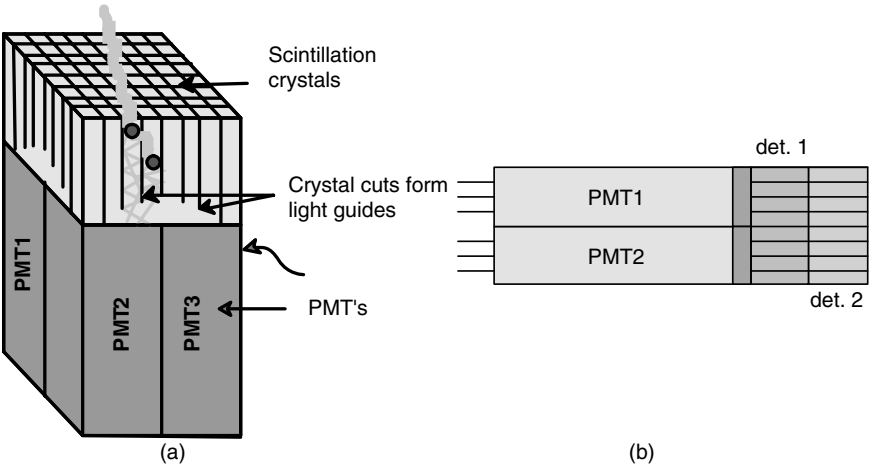


FIGURE 7. (a) Conventional block detector consisting of a set of crystals having cuts of different depths acting as light guides and segmenting the block into 64 (8×8) detection elements in this example. The block is optically coupled to four photomultiplier tubes at the back, and the crystal in which photon absorption occurs is identified by comparing the outputs of the four photomultiplier tubes (Anger logic). (b) Detector block consisting of a phoswich (detector 1 and 2) with depth-of-interaction measurement capability.

difference in decay time of the light is used to estimate depth in the crystal where the interaction occurred.

The design of high resolution imaging devices imposes some additional constraints with respect to the necessity for compact arrays of photodetectors; in turn, this has stimulated the development and use of multichannel PSPMT's and APDs. For the same application, some technologies introduced in the 1980s re-emerged, particularly the high density avalanche chamber (HIDAC), which uses lead converters contained within a multiwire proportional chamber.[56] Similarly detectors based on hybrid photodetectors and configurations using wavelength shifting fibres also have been developed.[57] Notwithstanding, the use of semiconductor detectors in PET is far from reaching acceptance, these devices are regarded as especially promising candidates for the design of PET cameras with the aim of overcoming the drawbacks of conventional PMT-based PET instrumentation.[25]

4.4 2D and 3D PET Data Acquisition

Most state-of-the-art full-ring PET scanners can be operated in both 2D and 3D data collection modes whereas newer tomographs operate essentially in fully 3D acquisition geometries. In the 2D mode, PET scanners have septa between detector planes to limit the detection of scattered photons, improve count rate performance and reduce random coincidence events. The use of

septa also simplifies the tomographic algorithm needed to reconstruct the data by requiring a technique that needs to account only for lines of response that are contained within one or a few slices in a 2D geometry, rather than in multiple slices as is needed for a 3D geometry. Figure 8 illustrates the conceptual difference between conventional 2D acquisition and volume data acquisition. In 2D PET, data acquisition is limited to coincidences detected within the same detector or adjacent detector rings: each 2D transverse section of the tracer distribution is reconstructed independently of adjacent sections. With the information obtained from detectors belonging to the same detector ring, images representing the tracer distribution in the planes of the detector rings (*direct planes*) are obtained. With the information obtained from detectors belonging to adjacent detector rings, we reconstruct

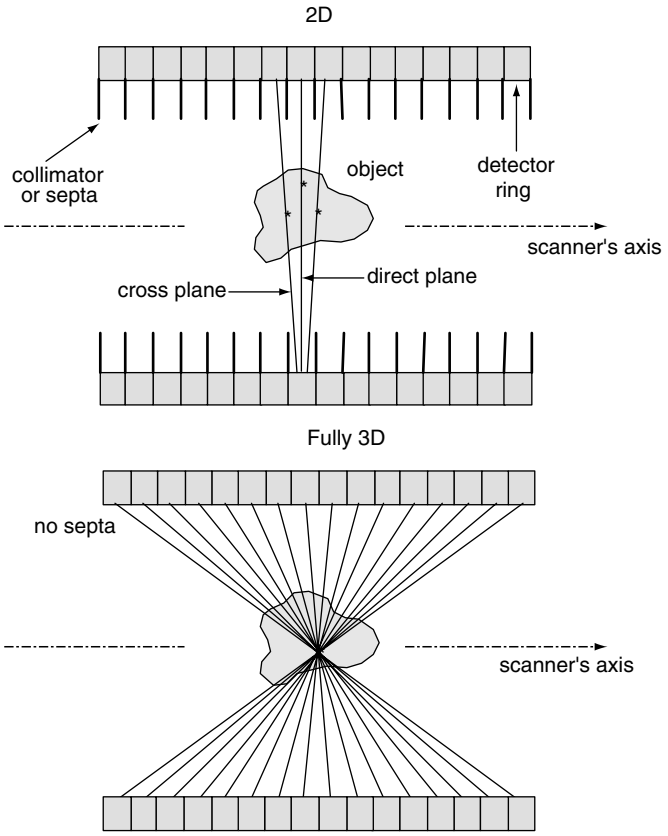


FIGURE 8. Diagram (not to scale) of the axial section of a 16-ring cylindrical PET tomograph operated in 2D mode (top) and 3D mode (bottom). In the 2D acquisition mode, the rings are separated by annular shielding (septa) and only in-plane and cross-plane LORs are allowed. In the 3D acquisition mode, the septa are retracted and coincidences allowed between any 2 rings.

images representing the tracer distribution in the planes between the detector rings (*cross planes*). Hence, a 3D PET scanner consisting of *N* detector rings gives (*2N-1*) images representing the tracer distribution in adjacent cross-sections of a patient. Transverse sections of the image volume obtained are then stacked together to allow the spatial distribution of the radiopharmaceutical to be viewed in 3D.

However, the use of septa also compromises the advantage of electronic collimation which otherwise records coincidence events efficiently within a cone-beam geometry and not just from the fan-beam geometry of the 2D mode. Thus, the 3D mode increases coincidence efficiency by about a factor of around five in comparison to a 2D acquisition. This can be accomplished by removing the interplane septa, at the expense of increasing scattered radiation. The septa themselves subtend an appreciable solid angle in the camera FOV, and as a result they have a significant shadowing effect on the detectors which can be as high as 50%. The geometric efficiency of the 3D mode requires that the detector be designed with a high count rate performance, which must be considered carefully in the selection of detector material (e.g. BGO have a relatively slow decay time). In fully 3D PET, the reconstruction algorithm must account for both oblique LORs formed from coincidences detected between different detector rings and for the increase in number of LORs which in turn depends on the number of crystal rings and degree of rebinning in comparison to 2D acquisitions. Different methods are used by the various scanner manufacturers to account for these characteristics. In addition, rebinning in 2D mode causes a variation in sensitivity along the axial FOV. In 3D mode, there is a much stronger variation in sensitivity (Figure 9) which peaks in the centre of the axial FOV.[58]

Likewise, different trues and randoms rates and different levels of scatter in 2D and 3D modes make it difficult to assess the real advantages of 3D vs 2D imaging. In order to more realistically assess these advantages, the concept of noise equivalent counts (NEC) has been proposed by Strother *et al.*[59] and was defined as:

$$NEC = \frac{[T(1 - SF)]^2}{T + 2fR} \tag{1}$$

where T and R are the total (i.e. trues+scatter) and randoms rates, SF is the scatter fraction and f is the fraction of the maximum transaxial FOV subtended by a centred 24 cm diameter field. The factor of two in the denominator results from randoms subtraction using data collected in a delayed coincidence window. The NEC quantifies the useful counts being acquired after applying perfect correction techniques for the physical effects and can be related to the global signal-to-noise ratio. As a result, the NEC has been extensively used to assess the real improvement obtained from 3D vs. 2D data collection strategies. The NEC does not, however, take into account all design or operational parameters that affect image quality.

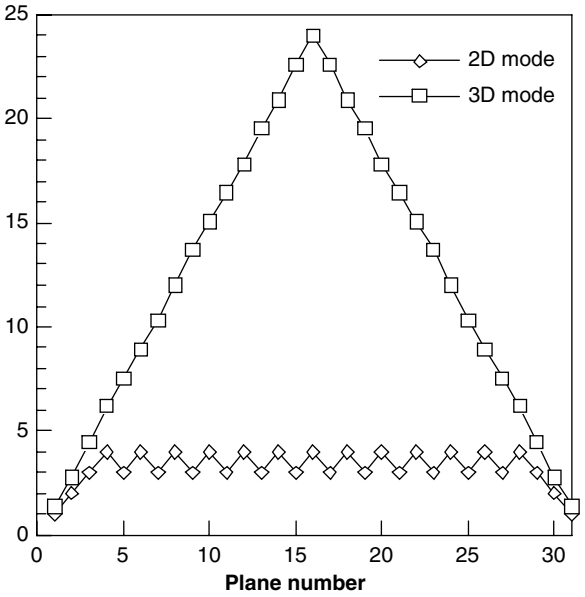


FIGURE 9. Relative sensitivity from the number of LORs in 2D and 3D acquisitions as a function of the plane number for a 16-ring PET tomograph. In this example, acquisition data are formed by LORs that arise from detector rings having a difference in location of ± 3 in 2D mode and ± 15 in 3D mode. As can be seen, at the edge of the axial FOV, 2D and 3D mode acquisitions have the same predicted sensitivity but in the centre of the FOV, the sensitivity in 3D mode is significantly higher.

The measured NEC using a 20 cm diameter uniform cylinder following the NEMA NU 2-2001 protocol[60] in different PET tomographs from the same manufacturer including the ECAT EXACT HR+ operating with septa extended and retracted is shown in Figure 10. It clearly shows that the maximum NEC is obtained at a lower activity concentration in 3D than in 2D as expected from the behaviour of the true coincidence rate.[58] As a result the signal-to-noise ratio will be higher in the low activity concentration range when operating in 3D mode than in 2D. This has important implications in clinical studies where reduction of the injected activities and hence the radiation dose to the patient is of concern.

5. Dedicated Small-Animal Imaging Devices

With the recent developments in radiochemistry and tracer production technology combined with progress made in molecular/cell biology, it has become possible to design specific tracers to image events non-invasively in small animals and humans to investigate disease processes *in vivo*. Murine models now have an essential role in formulating modern concepts of mammalian

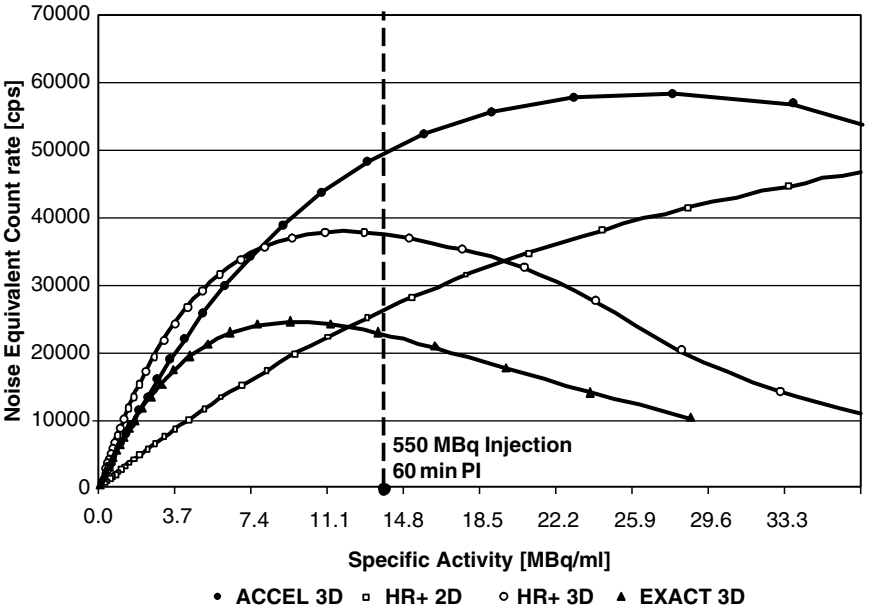


FIGURE 10. Noise equivalent count rate (NEC) for various multi-ring PET tomographs using LSO (ACCEL) and BGO (EXACT and EXACT HR+) detector technologies operating exclusively in 3D mode except the EXACT HR+ which is operated both in 2D and fully 3D modes. The measurements have been obtained with a 20 cm diameter uniform cylinder following the NEMA NU 2-2001 protocol. Typical activity concentration in the field of view 1h following injection of about 550 MBq is also shown (dashed line).

biology and human disease, and provide us with a realistic means of developing and evaluating new diagnostic and therapeutic techniques.[61] Moreover, transgenic and knock-out techniques now are available for manipulating the genome in a way that allows us to tailor animal models that accurately recapitulate biological and biochemical processes in the human. Mouse models now are available that allow investigators to study the development of tumors in mice in a way that represents virtually all major human cancers including those of the lung,[62,63] gastrointestinal system,[64,65] nervous system,[66] breast,[67,68] liver,[69] prostate,[70] pancreas,[71] and reproductive system.[72] Similarly, in cardiovascular research,[73-76] animal models are important for studying the hormonal pathways involved in regulation of hypertension and hypertension therapy,[77,78] cardiac electrophysiology,[79,80] mechanisms of apoptosis,[81] effects of exercise,[82] lipid metabolism and insulin sensitivity,[83,84] atherosclerosis,[85] and angiogenesis.[86] Likewise, the prospective impact of molecular imaging in many aspects of neuroscience research is well recognized.[87] The role of transgenic and knockout mice in biomedical research now has become profound and widespread, and transgenic animals (mice and rats)

at this time can be designed and created in a way that offers almost unlimited possibilities for addressing questions concerning the genetic, molecular, and cellular basis of biology and disease.[88-90]

With the ever increasing number and importance of human disease models, particularly in the smaller animals such as mice and rats,[88-90] the potential of high resolution nuclear medicine imaging technologies to contribute unique information is becoming apparent to many researchers. Radionuclide imaging can be performed using *in vitro* autoradiography. This technology offers exquisite spatial resolution, but is extremely time consuming and labour intensive, particularly when large volumes of tissue need to be sliced, exposed and digitized. However, unlike autoradiography, SPECT and PET can deliver functional and kinetic data from large volumes of tissue (in some cases the entire animal), with the results available within minutes of the end of the study. SPECT and PET also offer the critical advantage that they obtain functional information noninvasively, so each animal can be studied repeatedly. This allows each animal to serve as its own control in studies with a longitudinal design. Some animal models (particularly those involving pharmacological or surgical intervention) can exhibit high variability from one animal to another, thereby strongly supporting study designs in which disease progression or therapeutic response must be followed in an individual animal. The development of animal models often involves a large investment in time and expertise (particularly transgenic animals and study of gene therapy protocols). In these cases, researchers would welcome a tool that can noninvasively assess biological function to reduce study costs and to improve statistical power with limited numbers of animals because an individual animal can serve as its own control.

For these reasons, significant attention has focused on the use of radionuclide imaging for noninvasive investigations of cancer in small animals. Small-bore microPET systems now are available commercially with spatial resolutions nearing the fundamental limit of 1-2 mm for imaging tumour metabolism with ^{18}F-fluorodeoxyglucose, and can be used with other positron-emitting radionuclides when a radiochemistry laboratory and medical cyclotron are available. MicroSPECT also has important capabilities for cancer imaging in small animals using single-photon radiopharmaceuticals that are relatively inexpensive and can be obtained from commercial radiopharmacies without the need for an on-site cyclotron. Small animal imaging can be performed with widely available radiopharmaceuticals labelled with technetium-99m and with other radionuclides such as iodine-123 and indium-111 which also can be used to label disease-specific receptor-targeted agents for experimental studies with small animals. Hundreds of ^{125}I-labeled biochemicals are readily available from chemical supply companies (e.g., Fisher Scientific International, Inc.) for *in vitro* assays and autoradiography; these radiotracers also can be used for *in vivo* tumour localization studies in small animals with microSPECT especially in cases where the long half-life (60 days), relatively low specific activity, and higher-radiation dose of this

radionuclide are compatible with the goals of the study. SPECT and micro-SPECT imaging studies of small animals already are being used to investigate numerous aspects of cancer including multidrug resistance in tumours of the breast[91] and lung,[92] apoptosis as an early indicator of therapeutic outcome,[93] use of nucleoside analogs for imaging tissue proliferation,[94] delineating the role of anti-EGFR (epidermal growth factor receptor) in cell division, cancer progression, angiogenesis, and metastasis[95] and for development of pharmaceuticals for diagnosis, staging, and treatment of cancer.[93,96] Other investigators have developed noninvasive methods using high resolution SPECT and PET to assess cardiac volumes and ejection fraction,[97] progression of Parkinson's disease,[98] epilepsy,[99] development of pulmonary emboli and venous thrombi,[100] dopaminergic neurotransmission,[101-103] and development of myocardial infarction.[104] Finally, unlike microPET which can only detect 511 keV annihilation photons in coincidence, microSPECT can perform dual-isotope imaging that allows multiple processes to be monitored simultaneously during a single imaging study.

The use of radionuclide imaging for biological research with small animals presents several challenges beyond those faced in clinical nuclear medicine. First, small animal imaging typically is utilized in a research laboratory rather than a medical clinic. Therefore, the system should be reliable, easy to use, and relatively low in cost. Second, small animal radionuclide imaging must achieve millimeter, and ideally submillimeter spatial resolution to detect small lesions including micrometastases. Finally, a high level of detection efficiency allows procedures to be completed within 1 hour during which the animal can be safely anesthetized. Excellent detection efficiency also minimizes the amount of radiopharmaceutical needed for the study and limits the radiation dose which at higher levels can change tumor characteristics or cause lethality.[105]

To achieve these goals, several investigators have developed several methods to image single-photon radionuclides in small animals. High spatial resolution projection data suitable for imaging small animals can be obtained using pinhole collimation. The most direct approach simply adds a pinhole collimator to a conventional scintillation camera[6,9,106-113] and can provide excellent spatial resolution and with reasonable detection efficiency. The recent growth of pinhole SPECT-based cameras spurred the development of appropriate modelling strategies to enable the characterization and optimize the design of pinhole collimators.[110,114,115] Pinhole imaging also is performed with special-purpose compact scintillation cameras,[8,116-122] including the commercially available system from Gamma Medica, Inc. (Northridge, CA).[8,121,122] These compact systems have a small footprint and offer improved rotational stability for microSPECT which can be difficult to obtain with bulky clinical cameras. However, Funk *et al.*[105] have shown that these microSPECT systems require that the animal is administered with significant levels of radioactivity that contribute radiation doses that can change gene expression, and in some cases are near lethality for the animal. For these

reasons, other investigators[10,91,122-125] have developed small animal SPECT systems that offer both excellent spatial resolution and high detection efficiency with multiple pinhole detectors. Multipinhole configurations include those with multiple compact detectors each with an individual pinhole collimator, or those with a clinical scintillation camera or other large radionuclide imager with a multipinhole collimator. Several investigators have argued that the best approach for small animal imaging would combine multipinhole techniques with multiple compact detectors having a high level of intrinsic spatial resolution including the use of very small scintillator elements (YAP, CsI(Tl), or LSO) read out by position sensitive or multi-channel PMT's to achieve high spatial resolution.[10,122,123] While these advances in technology have improved performance and reduced cost, the size and low quantum efficiency of PMT's are often limiting factors in these designs. In view of this, solid-state detectors such as silicon p-i-n photodiodes (with unity gain) and silicon avalanche photodiodes are being actively considered as replacements for PMTs in nuclear medicine systems.

The demand for functional, metabolic, and molecular imaging of small animals also has stimulated the development of dedicated small-bore high-resolution PET systems for imaging mice, rats, small primates, and other mammalian species.[56,126-134] As in human imaging, both high detection sensitivity and excellent spatial resolution are priorities for PET imaging system design and are needed to achieve suitable levels of image quality and quantitative accuracy.[135] Thus, different PET designs have been suggested encompassing conventional small ring radius cylindrical block-detector based design with DOI capability and APDs readout, a renewed interest in the 3D HIDAC camera that achieves millimeter-scale spatial resolution[56] along with many other designs.[136] Several high-resolution small animal scanner designs have been or are being developed in both academic and corporate settings, with more than six such devices (both SPECT and PET) being offered commercially (Figure 11). More recently, advanced versions of these technologies have begun to be used across the breadth of modern biomedical research to study non-invasively small laboratory animals in a myriad of experimental settings. The commercially available microPET system,[137,138] developed originally at UCLA, consists of a cylindrical arrangement of $2 \times 2 \times 10\,\text{mm}^3$ LSO crystals read out by short optical fibres to multi-channel PMTs. The latest development of this design uses $\sim 1\,\text{mm}$ square crystals and achieves a spatial resolution of around 1 mm in the center of the field-of-view using a statistical reconstruction algorithm incorporating accurate system modelling.[130,132,133] With the introduction of commercial PET systems, small-animal imaging is becoming readily accessible and increasingly popular. The choice of a particular system being dictated in most cases by technical specifications, special attention has to be paid to methodologies followed when characterising system performance. Standardisation of the assessment of performance characteristics is thus highly desired.[139]

FIGURE 11. Photographs of dedicated small-animal PET scanners showing (a) the Hamamatsu SHR-7700 PET scanner based on BGO detector blocks designed for non-human primate imaging, (b) the multiwire proportional chamber technology-based HIDAC system, and (c) the microPET P4 scanner using LSO scintillation crystals. (Photographs courtesy of Hamamatsu Photonics KK, Japan, Oxford Positron Systems, UK and Concorde Microsystems, USA, respectively).

In parallel to these developments, dual-modality imaging, an approach where two different imaging techniques are integrated in a single unit that allows a small-animal to be imaged with both approaches during a single experimental procedure, such as SPECT/CT, PET/CT and PET/MRI are under development by different research groups and scanner manufacturers, and offer unique capabilities beyond those available from systems that perform radionuclide imaging alone (see chapter 2).

6. Future Directions for Nuclear Medicine Instrumentation

The capabilities of nuclear medicine instrumentation and techniques have undergone continual, and sometimes abrupt, improvements in performance and in their sophistication and complexity. A baseline of performance can be assessed by physicists who use objective measures of spatial resolution, energy resolution, count-rate response, and other parameters to assess the technical capabilities of commercial (and sometimes research) systems. The process of objective comparison has been facilitated by the wide acceptance and adoption of the NEMA standards, which provide an objective measure of performance parameters for nuclear imaging systems. While there may be only slight variations in image quality obtained in planar imaging with modern scintillation cameras, *in toto* the assessment of image quality is not as obvious given the wide variety of tasks encountered in the clinical environment. For example, the noise equivalent count rate can be measured to assess the physical performance of a PET scanner but does not give a comprehensive indication of image quality. Even though the NEMA

standards attempt to standardise the assessment of image quality using carefully designed phantoms, the measurement of physical parameters cannot account for special requirements of the clinical task being performed, and therefore requires additional assessments of clinical applicability and performance.[140]

The intrinsic physical performance of conventional radionuclide imaging systems is approaching their fundamental limits, given the performance of existing instrumentation and components. This has encouraged the development of innovative approaches capable of providing improved performance at a reduced or comparable cost to current technologies. For example, Braem *et al.*[30] have proposed a novel detector design which provides full 3D reconstruction free of parallax errors with excellent spatial resolution over the total detector volume. The key components are a matrix of long scintillator crystals coupled on both ends to hybrid photodetectors (HPDs) with matched segmentation and integrated readout electronics. Computer simulations and Monte Carlo modeling predict that the detector will achieve excellent spatial (x,y,z) and energy resolutions. The design also increases detection efficiency by reconstructing a significant fraction of events that undergo Compton scattering in the crystals. The 3D axial detector geometry (Figure 12) is configured from a matrix of 208 (13×16) long crystals each with a cross section of $3.2 \times 3.2\,\text{mm}^2$ and with an intercrystal spacing of 0.8 mm. Scintillation light produced after an interaction of an annihilation

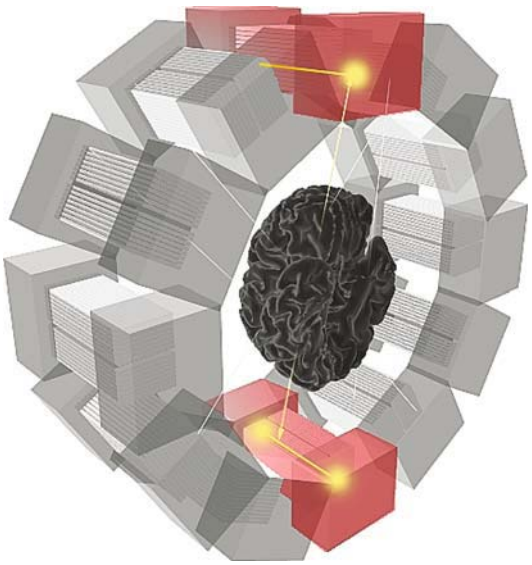


FIGURE 12. Illustration of principles of data acquisition for a brain PET camera using the proposed novel axial design based on 12 camera modules consisting of long scintillation crystals readout on both sides by HPDs. Reprinted with permission from ref.[30]

photon will propagate by total internal reflection to the ends of the crystal, where it will be detected by the HPD photodetectors. The transaxial resolution depends only on the crystal segmentation and not on its chemical composition whereas the axial resolution is closely related to the scintillator properties. The scintillator's optical bulk absorption length should be approximately equal to the crystal length to obtain both a high light yield and a significant light asymmetry required to decode the axial coordinate z of the photon interaction in the crystal matrix.

The design of whole-body PET imaging systems with larger axial field-of-view also is receiving considerable attention. In particular, the use of rotating flat panel LSO-based detector arrays is a promising design that, if successful, will enable a new generation of high-performance whole-body PET scanners.[141] The proposed design envisions several (3 to 5) panels mounted in a hexagonal configuration that can be rotated to acquire a full 3D dataset for tomographic reconstruction. Each panel contains 10,080 LSO crystals ($4 \times 4 \times 20\,mm^4$) coupled to an array of 88 PMT's which can identify individual crystal elements via a light sharing scheme. The detector panels will be coupled to electronics with fast detector readout and temporal resolution, and will be configured in a system capable of imaging the whole body significantly faster than is possible with current PET scanner designs.[50,142]

Finally, the field of nuclear medicine always has had a focus on objective evaluation and optimization of image quality, and on quantitative assessment of physiological parameters that can be extracted from radionuclide image data. This perspective primarily is the consequence of the fact that nuclear imaging has limited spatial resolution and signal-to-noise characteristics relative to other radiological techniques such as CT and MRI, which are excellent in depicting anatomical structure. Differences in performance between various scanners, as well as the degradation in performance of a single scanner, can be subtle and difficult to detect visually or by qualitative methods. The need to monitor and correct these changes have compelled both clinical users and research investigators to rely on objective measures of performance for system intercomparisons, to assess the impact of design changes, and for routine quality control and quality assurance tasks. The quantitative perspective of nuclear medicine also has been motivated strongly by its use in functional and metabolic assessments of biology and disease. By its inherent nature as an *in vivo* tracer technique, radionuclide imaging can be used to measure tissue perfusion, organ function, radiopharmaceutical biodistribution and kinetics, and other physiological parameters that require a level of quantification beyond anatomical metrics used with other radiological imaging techniques.

In this chapter, we have provided a brief overview of current state-of-the art developments in nuclear medicine instrumentation. We emphasize that many different design paths have been and continue to be pursued in both academic and corporate settings, that offer different trade-offs in terms of

their performance. It still is uncertain which designs will be incorporated into future clinical systems, but it is certain that technological advances will continue and will enable new quantitative capabilities in nuclear medicine imaging.

References

1. Jones T., The role of positron emission tomography within the spectrum of medical imaging. *Eur J Nucl Med* **23:** 207-211 (1996).
2. Anger H., Scintillation camera. *Rev Sci Instr* **29:** 27-33 (1958).
3. Moore S. C., Kouris, K. and Cullum, I., Collimator design for single photon emission tomography. *Eur J Nucl Med* **19:** 138-150 (1992).
4. Jaszczak R. J., Greer, K. L. and Coleman, R. E., SPECT using a specially designed cone beam collimator. *J Nucl Med* **29:** 1398-1405 (1988).
5. Tsui B. M. and Gullberg, G. T., The geometric transfer function for cone and fan beam collimators. *Phys Med Biol* **35:** 81-93 (1990).
6. Weber D. A. and Ivanovic, M., Ultra-high-resolution imaging of small animals: implications for preclinical and research studies. *J Nucl Cardiol* **6:** 332-344 (1999).
7. Green M. V., Seidel, J., Vaquero, J. J. *et al.*, High resolution PET, SPECT and projection imaging in small animals. *Comput Med Imaging Graph* **25:** 79-86 (2001).
8. MacDonald L. R., Patt, B. E., Iwanczyk, J. S. *et al.*, Pinhole SPECT of mice using the LumaGEM gamma camera. *IEEE Trans Nucl Sci* **48:** 830-836 (2001).
9. Wu M. C., Gao, D. W., Sievers, R. E. *et al.*, Pinhole single-photon emission computed tomography for myocardial perfusion imaging of mice. *J Am Coll Cardiol* **42:** 576-582 (2003).
10. Beekman F. J. and Vastenhouw, B., Design and simulation of a high-resolution stationary SPECT system for small animals. *Phys Med Biol* **49:** 4579-4592 (2004).
11. Kimiaei S. and Larsson, S. A., Optimal design of planar-concave collimators for SPECT-an analytical approach. *Phys Med Biol* **43:** 637-650 (1998).
12. Beekman F. J., Kamphuis, C., Hutton, B. F. *et al.*, Half-fanbeam collimators combined with scanning point sources for simultaneous emission-transmission imaging. *J Nucl Med* **39:** 1996-2003 (1998).
13. Formiconi A. R., Geometrical response of multihole collimators. *Phys Med Biol* **43:** 3359-3379 (1998).
14. Webb S., Binnie, D. M., Flower, M. A. *et al.*, Monte Carlo modelling of the performance of a rotating slit-collimator for improved planar gamma-camera imaging. *Phys Med Biol* **37:** 1095-1108 (1992).
15. Zeng G. L. and Gagnon, D., CdZnTe strip detector SPECT imaging with a slit collimator. *Phys Med Biol* **49:** 2257-2271 (2004).
16. Lodge M. A., Binnie, D. M., Flower, M. A. *et al.*, The experimental evaluation of a prototype rotating slat collimator for planar gamma camera imaging. *Phys Med Biol* **40:** 427-448 (1995).
17. Zeng G. L., Gagnon, D., Matthews, C. G. *et al.*, Image reconstruction algorithm for a rotating slat collimator. *Med Phys* **29:** 1406-1412 (2002).
18. Williams M. B., Goode, A. R., Galbis-Reig, V. *et al.*, Performance of a PSPMT based detector for scintimammography. *Phys Med Biol* **45:** 781-800 (2000).

19. Loudos G. K., Nikita, K. S., Giokaris, N. D. *et al*., A 3D high-resolution gamma camera for radiopharmaceutical studies with small animals. *Appl Radiat Isot* **58:** 501-508 (2003).
20. Singh M. and Horne, C., Use of a germanium detector to optimize scatter correction in SPECT. *J Nucl Med* **28:** 1853-1860 (1987).
21. Mauderli W. and Fitzgerald, L. T., Rotating laminar emission camera with Ge-detector: further developments. *Med Phys* **14:** 1027-1031 (1987).
22. Darambara D. G. and Todd-Pokropek, A., Solid state detectors in nuclear medicine. *Q J Nucl Med* **46:** 3-7 (2002).
23. Abe A., Takahashi, N., Lee, J. *et al*., Performance evaluation of a hand-held, semiconductor (CdZnTe)-based gamma camera. *Eur J Nucl Med Mol Imaging* **30:** 805-811 (2003).
24. Gagnon D., Zeng, G. L., Links, J. M. *et al*., ''Design considerations for a new solid-state gamma-camera: SOLSTICE'' *Proc*. IEEE Nuclear Science Symposium and Medical Imaging Conference, Oct. 4-10, San Diego, CA, Vol. 2; *pp* 1156-1160 (2001).
25. Humm J. L., Rosenfeld, A. and Del Guerra, A., From PET detectors to PET scanners. *Eur J Nucl Med Mol Imaging* **30:** 1574-1597 (2003).
26. Renker D., Properties of avalanche photodiodes for applications in high energy physics, astrophysics and medical imaging. *Nucl Instr Meth A* **486:** 164-169 (2002).
27. Joram C., Large area hybrid photodiodes. *Nucl Phys B* **78:** 407-415 (1999).
28. Weilhammer P., Silicon-based HPD development: sensors and front ends. *Nucl Instr Meth A* **446:** 289-298 (2000).
29. D'Ambrosio C. and Leutz, H., Hybrid photon detectors. *Nucl Instr Meth A* **501:** 463-498 (2003).
30. Braem A., Chamizo Llatas, M., Chesi, E. *et al*., Feasibility of a novel design of high-resolution parallax-free Compton enhanced PET scanner dedicated to brain research. *Phys Med Biol* **49:** 2547-2562 (2004).
31. Koral K. F., Zaidi, H. and Ljungberg, M., ''Medical imaging techniques for radiation dosimetry.'' in: *Therapeutic applications of Monte Carlo calculations in nuclear medicine*, edited by H Zaidi and G Sgouros Institute of Physics Publishing, Bristol, (2002), *pp* 55-83.
32. Singh M., An electronically collimated gamma camera for single photon emission computed tomography. Part I: Theoretical considerations and design criteria. *Med Phys* **10:** 421-427 (1983).
33. Evans R. D., *The atomic nucleus*, McGraw-Hill, New York, (1955).
34. Carlsson G. A., Carlsson, C. A., Berggren, K. F. *et al*., Calculation of scattering cross sections for increased accuracy in diagnostic radiology. I. Energy broadening of Compton-scattered photons. *Med Phys* **9:** 868-879 (1982).
35. Hirasawa M. and Tomitani, T., Effect of compensation for scattering angular uncertainty in analytical Compton camera reconstruction. *Phys Med Biol* **49:** 2083-2093 (2004).
36. Todd R. W., Nightingale, J. and Everett, D., A proposed g-camera. *Nature* **25:** 132 (1974).
37. Meier D., Czermak, A., Jalocha, P. *et al*., Silicon detector for a Compton camera in nuclear medical imaging. *IEEE Trans Nucl Sci* **49:** 812-816 (2002).
38. Martin J. B., Dogan, N., Gromley, J. *et al*., Imaging multi-energy gamma-ray fields with a Compton scatter camera. *IEEE Trans Nucl Sci* **41:** 1019-1025 (1994).

39. LeBlanc J. W., Clinthorne, N. H., Hua, C.-H. *et al*., C-SPRINT: a prototype Compton camera system for low energy gamma ray imaging. *IEEE Trans Nucl Sci* **45:** 943-949 (1998).
40. Du Y. F., He, Z., Knoll, G. F. *et al*., Evaluation of a Compton scattering camera using 3-D position sensitive CdZnTe detectors. *Nucl Instr Meth A* **457:** 203-211 (2001).
41. Zhang L., Rogers, W. and Clinthorne, N., Potential of a Compton camera for high performance scintimammography. *Phys Med Biol* **49:** 617-638 (2004).
42. Scannavini M., Speller, R., Royle, G. *et al*., A possible role for silicon microstrip detectors in nuclear medicine: Compton imaging of positron emitters. *Nucl Instr Meth A* **477:** 514-520 (2002).
43. Basko R., Zeng, G. L. and Gullberg, G. T., Application of spherical harmonics to image reconstruction for the Compton camera. *Phys Med Biol* **43:** 887-894 (1998).
44. Sauve A. C., Hero, A. O., III, Rogers, W. L. *et al*., 3D image reconstruction for a Compton SPECT camera model. *IEEE Trans Nucl Sci* **46:** 2075-2084 (1999).
45. Brechner R. R. and Singh, M., Iterative reconstruction of electronically collimated SPECT images. *IEEE Trans Nucl Sci* **37:** 1328-1332 (1990).
46. Meikle S. R. and Badawi, R. D., "Quantitative techniques in Positron Emission Tomography." in: *Positron Emission Tomography: Basic Science and Clinical Practice*, edited by P E Valk, D L Bailey, D W Townsend *et al*. Springer, London, (2003), *pp* 115-146.
47. Phelps M. E., PET: the merging of biology and imaging into molecular imaging. *J Nucl Med* **41:** 661-681 (2000).
48. Phelps M. E. and Cherry, S. R., The changing design of positron imaging systems. *Clin. Pos. Imag*. **1:** 31-45 (1998).
49. Wienhard K., Schmand, M., Casey, M. E. *et al*., The ECAT HRRT: performance and first clinical application of the new high resolution research tomograph. *IEEE Trans Nucl Sci* **49:** 104 -110 (2002).
50. Marsden P. K., Detector technology challenges for nuclear medicine and PET. *Nucl Instr Meth A* **513:** 1-7 (2003).
51. van Eijk C. W. E., Inorganic scintillators in medical imaging. *Phys Med Biol* **47:** R85-R106 (2002).
52. Moses W. W., Current trends in scintillator detectors and materials. *Nucl Instr Meth A* **487:** 123-128 (2002).
53. Derenzo S. E., Weber, M. J., Bourret-Courchesne, E. *et al*., The quest for the ideal inorganic scintillator. *Nucl Instr Meth A* **505:** 111-117 (2003).
54. Casey M. E. and Nutt, R., A Multicrystal two-dimensional BGO detector system for positron emission tomography. *IEEE Trans Nucl Sci* **33:** 460-463 (1986).
55. Dahlbom M., MacDonald, L. R., Schmand, M. *et al*., A YSO/LSO phoswich array detector for single and coincidence photon imaging. *IEEE Trans Nucl Sci* **45:** 1128-1132 (1998).
56. Jeavons A. P., Chandler, R. A. and Dettmar, C. A. R., A 3D HIDAC-PET camera with sub-millimetre resolution for imaging small animals. *IEEE Trans Nucl Sci* **46:** 468-473 (1999).
57. Worstell W., Johnson, O., Kudrolli, H. *et al*., First results with high-resolution PET detector modules using wavelength-shifting fibers. *IEEE Trans Nucl Sci* **45:** 2993-2999 (1998).

58. Bendriem B. and Townsend, D. W., *The theory and practice of 3D PET*, Kluwer Academic Publishers, The Netherlands, Dordrecht, (1998).
59. Strother S. C., Casey, M. E. and Hoffman, E. J., Measuring PET scanner sensitivity-relating count rates to image signal-to-noise ratios using noise equivalent counts. *IEEE Trans Nucl Sci* **37:** 783-788 (1990).
60. NEMA, Standards Publication NU 2-2001. Performance measurements of positron emission tomographs. National Electrical Manufacturers Association, 2001.
61. Hirst G. L. and Balmain, A., Forty years of cancer modelling in the mouse. *Eur J Cancer* **40:** 1974-1980 (2004).
62. Shmidt E. N. and Nitkin, A. Y., Pathology of mouse models of human lung cancer. *Comp Med* **54:** 23-26 (2004).
63. Kwak I., Tsai, S. Y. and DeMayo, F. J., Genetically engineered mouse models for lung cancer. *Annu Rev Physiol* **66:** 647-663 (2004).
64. Boivin G. P. and Groden, J., Mouse models of intestinal cancer. *Comp Med* **54:** 15-18 (2004).
65. Janssen K. P., Murine models of colorectal cancer: studying the role of oncogenic K-ras. *Cell Mol Life Sci* **60:** 495-506 (2003).
66. Dyer M. A., Mouse models of childhood cancer of the nervous system. *J Clin Pathol* **57:** 561-576 (2004).
67. Mant C. and Cason, J., A human murine mammary tumour virus-like agent is an unconvincing aetiological agent for human breast cancer. *Rev Med Virol* **14:** 169-177 (2004).
68. Gravekamp C., Sypniewska, R. and Hoflack, L., The usefulness of mouse breast tumor models for testing and optimization of breast cancer vaccines at old age. *Mech Ageing Dev* **125:** 125-127 (2004).
69. Bursch W., Grasl-Kraupp, B., Wastl, U. *et al*., Role of apoptosis for mouse liver growth regulation and tumor promotion: comparative analysis of mice with high (C3H/He) and low (C57Bl/6J) cancer susceptibility. *Toxicol Lett* **149:** 25-35 (2004).
70. Shappell S. B., Thomas, G. V., Roberts, R. L. *et al*., Prostate pathology of genetically engineered mice: definitions and classification. The consensus report from the Bar Harbor meeting of the Mouse Models of Human Cancer Consortium Prostate Pathology Committee. *Cancer Res* **64:** 2270-2305 (2004).
71. Leach S. D., Mouse models of pancreatic cancer: the fur is finally flying! *Cancer Cell* **5:** 7-11 (2004).
72. Arbeit J. M., Mouse models of cervical cancer. *Comp Med* **53:** 256-258 (2003).
73. Bader M., Bohnemeier, H., Zollmann, F. S. *et al*., Transgenic animals in cardiovascular disease research *Exp Physiol* **85:** 713-31 (2000).
74. Rao S. and Verkman, A. S., Analysis of organ physiology in transgenic mice. *Am J Physiol Cell Physiol* **279:** C1-C18 (2000).
75. Carmeliet P. and Collen, D., Transgenic mouse models in angiogenesis and cardiovascular disease. *J Pathol* **190:** 387-405 (2000).
76. James J. F., Hewett, T. E. and Robbins, J., Cardiac physiology in transgenic mice. *Circ Res* **82:** 407-415 (1998).
77. Lavoie J. L., Bianco, R. A., Sakai, K. *et al*., Transgenic mice for studies of the renin-angiotensin system in hypertension. *Acta Physiol Scand* **181:** 571-7 (2004).
78. Janssen B. J. and Smits, J. F., Autonomic control of blood pressure in mice: basic physiology and effects of genetic modification. *Am J Physiol Regul Integr Comp Physiol* **282:** R1545-64 (2002).

79. Gros D., Dupays, L., Alcolea, S. *et al*., Genetically modified mice: tools to decode the functions of connexins in the heart-new models for cardiovascular research. *Cardiovasc Res* **62:** 299-308 (2004).
80. Wessels A., Phelps, A., Trusk, T. C. *et al*., Mouse models for cardiac conduction system development. *Novartis Found Symp* **250:** 44-59; discussion 59-67, 276-279 (2003).
81. Russo G. L. and Russo, M., Ins and outs of apoptosis in cardiovascular diseases *Nutr Metab Cardiovasc Dis* **13:** 291-300 (2003).
82. Bernstein D., Exercise assessment of transgenic models of human cardiovascular disease. *Physiol Genomics* **13:** 217-26 (2003).
83. Kopecky J., Flachs, P., Bardova, K. *et al*., Modulation of lipid metabolism by energy status of adipocytes: implications for insulin sensitivity. *Ann N Y Acad Sci* **967:** 88-101 (2002).
84. Fruchart J. C. and Duriez, P., High density lipoproteins and coronary heart disease. Future prospects in gene therapy. *Biochimie* **80:** 167-72 (1998).
85. Daugherty A., Mouse models of atherosclerosis. *Am J Med Sci* **323:** 3-10 (2002).
86. Carmeliet P., Moons, L. and Collen, D., Mouse models of angiogenesis, arterial stenosis, atherosclerosis and hemostasis. *Cardiovasc Res* **39:** 8-33 (1998).
87. Jacobs A. H., Li, H., Winkeler, A. *et al*., PET-based molecular imaging in neuroscience. Eur J Nucl Med Mol Imaging **30:** 1051-1065 (2003).
88. Bernstein A. and Breitman, M., Genetic ablation in transgenic mice. *Molecular Biol Med* **6:** 523-530 (1989).
89. Doetschman T., Interpretation of phenotype in genetically engineered mice. *Lab Animal Sci* **49:** 137-143 (1999).
90. Hanahan D., Transgenic mice as probes into complex systems. *Science* **246:** 1265-1275 (1989).
91. Liu Z., Stevenson, G. D., Barrett, H. H. *et al*., Imaging recognition of multi-drug resistance in human breast tumors using 99mTc-labeled monocationic agents and a high-resolution stationary SPECT system. *Nucl Med Biol* **31:** 53-65 (2004).
92. Marsee D. K., Shen, D. H., MacDonald, L. R. *et al*., Imaging of metastatic pulmonary tumors following NIS gene transfer using single photon emission computed tomography. *Cancer Gene Ther* **11:** 121-127 (2004).
93. Blankenberg F. G., Mandl, S., Cao, Y.-A. *et al*., Tumor imaging using a standardized radiolabeled adapter protein docked to vascular endothelial growth factor. *J Nucl Med* **45:** 1373-1380 (2004).
94. Toyohara J., Hayashi, A., Sato, M. *et al*., Development of radioiodinated nucleoside analogs for imaging tissue proliferation: comparisons of six 5-iodonucleosides. *Nucl Med Biol* **30:** 687-696 (2003).
95. Schechter N., Yang, D., Azhdarinia, A. *et al*., Assessment of epidermal growth factor receptor with 99mTc-ethylenedicysteine-C225 monoclonal antibody. *Anticancer Drugs* **14:** 49-56 (2003).
96. Schottelius M., Wester, H., Reubi, J. *et al*., Improvement of pharmacokinetics of radioiodinated Tyr(3)-octreotide by conjugation with carbohydrates. *Bioconjug Chem* **13:** 1021-1030 (2002).
97. Vanhove C., Lahoutte, T., Defrise, M. *et al*., Reproducibility of left ventricular volume and ejection fraction measurements in rat using pinhole gated SPECT. *Eur J Nucl Med Mol Imaging* **32:** 211-220 (2005).

98. Popperl G., Tatsch, K., Ruzicka, E. *et al*., Comparison of alpha-dihydroergocryptine and levodopa monotherapy in Parkinson's disease: assessment of changes in DAT binding with [123I]IPT SPECT. *J Neural Transm* **111:** 1041-1052 (2004).
99. Hashizume K., Tsuda, H., Hodozuka, A. *et al*., Clinical and experimental studies of epilepsy associated with focal cortical dysplasia. *Psychiatry Clin Neurosci* **58:** S26-29 (2004).
100. Morris T. A., Marsh, J. J., Chiles, P. G. *et al*., Single photon emission computed tomography of pulmonary emboli and venous thrombi using anti-D-dimer. *Am J Respir Crit Care Med* **169:** 987-993 (2004).
101. Saji H., Iida, Y., Kawashima, H. *et al*., In vivo imaging of brain dopaminergic neurotransmission system in small animals with high-resolution single photon emission computed tomography. *Anal Sci* **19:** 67-71 (2003).
102. Acton P. D., Hou, C., Kung, M. P. *et al*., Occupancy of dopamine D2 receptors in the mouse brain measured using ultra-high-resolution single-photon emission tomography and [123]IBF. *Eur J Nucl Med Mol Imaging* **29:** 1507-1515 (2002).
103. Grunder G., Siessmeier, T., Piel, M. *et al*., Quantification of D2-like dopamine receptors in the human brain with 18F-desmethoxyfallypride. *J Nucl Med* **44:** 109-116 (2003).
104. Liu Z., Kastis, G. A., Stevenson, G. D. *et al*., Quantitative analysis of acute myocardial infarct in rat hearts with ischemia-reperfusion using a high-resolution stationary SPECT system. *J Nucl Med* **43:** 933-939 (2002).
105. Funk T., Sun, M. and Hasegawa, B. H., Radiation dose estimates in small animal SPECT and PET. *Med Phys* **31:** 2680-2686 (2004).
106. Jaszczak R. J., Li, J., Wang, H. *et al*., Pinhole collimation for ultra high-resolution, small field of view SPECT. *Phys Med Biol* **39:** 425-437 (1994).
107. Wu M. C., Tang, H. R., O'Connell, J. W. *et al*., An ultra-high resolution ECG-gated myocardial imaging system for small animals. *IEEE Tran Nucl Sci* **46:** 1199-1202 (1999).
108. Wu M. C., Hasegawa, B. H. and Dae, M. W., Performance evaluation of a pinhole SPECT system for myocardial perfusion imaging of mice. *Med Phys* **29:** 2830-2839 (2002).
109. Ogawa K., Kawade, T., Nakamura, K. *et al*., Ultra high resolution pinhole SPECT for small animal study. *IEEE Trans Nucl Sci* **45:** 3122-3126 (1998).
110. Weber D. A. and Ivanovic, M., Pinhole SPECT: ultra-high resolution imaging for small animal studies. *J Nucl Med* **36:** 2287-2289 (1995).
111. Ishizu K., Mukai, T., Yonekura, Y. *et al*., Ultra-high resolution SPECT system using four pinhole collimators for small animal studies. *J Nucl Med* **36:** 2282-2286 (1995).
112. Habraken J. B. A., de Bruin, K., Shehata, M. *et al*., Evaluation of high-resolution pinhole SPECT using a small rotating animal. *J Nucl Med* **42:** 1863-1869 (2001).
113. Schellingerhout D., Accorsi, R., Mahmood, U. *et al*., Coded aperture nuclear scintigraphy: a novel small animal imaging technique. *Mol Imaging* **1:** 344-353 (2002).
114. Metzler S. D., Bowsher, J. E., Smith, M. F. *et al*., Analytic determination of pinhole collimator sensitivity with penetration. *IEEE Trans Med Imaging* **20:** 730-741 (2001).

115. Accorsi R. and Metzler, S. D., Analytic determination of the resolution-equivalent effective diameter of a pinhole collimator. *IEEE Trans Med Imaging* **23:** 750-763 (2004).
116. Williams M. B., Zhang, G., More, M. J. *et al*., "Integrated CT-SPECT system for small animal imaging" Proc SPIE, Vol. 4142; *pp* 265-274 (2000).
117. Meikle S. R., Kench, P., Weisenberger, A. G. *et al*., A prototype coded aperture detector for small animal SPECT. *IEEE Trans Nucl Sci* **49:** 2167-2171 (2002).
118. Weisenberger A. G., Bradley, E. L., Majewski, S. *et al*., Development of a novel radiation imaging detector system for in vivo gene mapping in small animals. *IEEE Trans Nucl Sci* **45:** 1743-1749 (1998).
119. Weisenberger A. G., Wojcik, R., Bradley, E. L. *et al*., SPECT-CT system for small animal imaging. *IEEE Trans Nucl Sci* **50:** 74-79 (2003).
120. Welsh R. E., Brewer, P., Bradley, E. L. *et al*., "An economical dual-modality small animal imaging system with application to studies of diabetes" IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 3; *pp* 1845-1848 (2002).
121. MacDonald L. R., Iwanczyk, J. S., Patt, B. E. *et al*., "Development of new high resolution detectors for small animal SPECT imaging" IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 3; *pp* 21/75 (2002).
122. Beekman F. J., McElroy, D. P., Berger, F. *et al*., Towards in vivo nuclear microscopy: iodine-125 imaging in mice using micro-pinholes. *Eur J Nucl Med Mol Imaging* **29:** 933-938 (2002).
123. Furenlid L. R., Wilson, D. W., Chen, Y.-c. *et al*., FastSPECT II: a second-generation high-resolution dynamic SPECT imager. *IEEE Trans Nucl Sci* **51:** 631-635 (2004).
124. Acton P. D. and Kung, H. F., Small animal imaging with high resolution single photon emission tomography. *Nucl Med Biol* **30:** 889-895 (2003).
125. Schramm N. U., Ebel, G., Engeland, U. *et al*., High-resolution SPECT using multipinhole collimation. *IEEE Trans Nucl Sci* **50:** 315-320 (2003).
126. Cherry S. R., Shao, Y., Silverman, R. W. *et al*., MicroPET: a high resolution PET scanner for imaging small animals. *IEEE Trans Nucl Sci* **44:** 1161-1166 (1997).
127. Tornai M. P., Jaszczak, R. J., Turkington, T. G. *et al*., Small-animal PET: advent of a new era of PET research. *J Nucl Med* **40:** 1176-1179 (1999).
128. Chatziioannou A. F., Molecular imaging of small animals with dedicated PET tomographs. *Eur J Nucl Med Mol Imaging* **29:** 98-114 (2002).
129. Del Guerra A. and Belcari, N., Advances in animal PET scanners. *Q J Nucl Med* **46:** 35-47 (2002).
130. Shao Y., Cherry, S. R. and Chatziioannou, A. F., Design and development of 1 mm resolution PET detectors with position-sensitive PMTs. *Nucl Instr Meth A* **477:** 486-490 (2002).
131. Schelbert H. R., Inubushi, M. and Ross, R. S., PET imaging in small animals. *J Nucl Cardiol* **10:** 513-520 (2003).
132. Tai Y.-C., Chatziioannou, A., Yang, Y. *et al*., MicroPET II: design, development and initial performance of an improved microPET scanner for small-animal imaging. *Phys Med Biol* **48:** 1519-1537 (2003).
133. Yang Y., Tai, Y.-C., Siegel, S. *et al*., Optimization and performance evaluation of the microPET II scanner for in vivo small-animal imaging. *Phys Med Biol* **49:** 2527-2545 (2004).

134. Lee K., Kinahan, P. E., Miyaoka, R. S. *et al.*, Impact of system design parameters on image figures of merit for a mouse PET scanner. *IEEE Trans Nucl Sci* **51:** 27-33 (2004).
135. Miyaoka R. S., Dynamic high resolution positron emission imaging of rats. *Biomed Sci Instrum* **27:** 35-42 (1991).
136. Miyaoka R. S., Kohlmyer, S. G. and Lewellen, T. K., Performance characteristics of micro crystal element (MiCE) detectors. *IEEE Trans Nucl Sci* **48:** 1403-1407 (2001).
137. Tai Y., Chatziioannou, A., Siegel, S. *et al.*, Performance evaluation of the microPET P4: a PET system dedicated to animal imaging. *Phys Med Biol* **46:** 1845-1862 (2001).
138. Chatziioannou A., Tai, Y. C., Doshi, N. *et al.*, Detector development for microPET II: a 1 microl resolution PET scanner for small animal imaging. *Phys Med Biol* **46:** 2899-2910 (2001).
139. Weber S. and Bauer, A., Small animal PET: aspects of performance assessment. *Eur J Nucl Med Mol Imaging* **31:** 1545-1555 (2004).
140. Bergman S., The need for independent physics advice. *Eur J Nucl Med Mol Imaging* **30:** 491-493 (2003).
141. Nahmias C., Nutt, R., Hichwa, R. D. *et al.*, PET tomograph designed for five minute routine whole body studies. [abstract] *J Nucl Med* **43:** 11P (2002).
142. Townsend D. W., Carney, J. P. J., Yap, J. T. *et al.*, PET/CT today and tomorrow. *J Nucl Med* **45:** 4S-14 (2004).

2
Dual-Modality Imaging: More Than the Sum of its Components

B.H. Hasegawa* and H. Zaidi†

1. Introduction

The field of diagnostic radiology encompasses a wealth of imaging techniques that now are essential for evaluating and managing patients who need medical care. Traditional imaging methods such as plain film radiography and more recent techniques such as x-ray computed tomography (CT) and magnetic resonance imaging (MRI) can be used to evaluate a patient's anatomy with submillimeter spatial resolution to discern structural abnormalities and to evaluate the location and extent of disease. These methods also offer relatively fast scan times, precise statistical characteristics, and good tissue contrast especially when contrast media are administered to the patient. In addition, x-ray fluoroscopy and angiography can be used to evaluate the patency of blood vessels, the mechanical performance of the cardiovascular system, and structural abnormalities in the gastrointestinal or genitourinary systems. Similarly, CT and MRI can be performed with cardiac gating to the heart at different phases of the cardiac cycle. Computed tomography recently has experienced a significant increase in utilization with the advent of multislice helical scanning techniques that cover a large region of the patient's anatomy within a single breath-hold, with scan speeds that can capture both the arterial and venous phases of the contrast bolus. These increased scan speeds also enhance patient comfort, and contribute to patient throughput and cost effectiveness.

X-ray projection imaging, computed tomography, and magnetic resonance imaging differentiate disease from normal tissue by revealing structural differences or differences in regional perfusion of the administered contrast media. The interpretation of the images can be complicated when normal perfusion patterns are disrupted by prior surgery or radiation therapy, which can lead to tissue damage or necrosis where contrast patterns can mimic

*Prof. B.H. Hasegawa, Department of Radiology, University of California, San Francisco, CA, USA
†PD Dr H. Zaidi, Geneva University Hospital, Division of Nuclear Medicine, CH–1211 Geneva, Switzerland

those associated with neoplasia. This presents a significant challenge when the imaging techniques are used to define the anatomical extent of disease as is needed for planning highly conformal radiation treatment or for planning highly targeted therapeutic regimes.

In comparison to the anatomical imaging techniques described above, functional imaging methods including planar scintigraphy, single-photon emission computed tomography (SPECT), positron emission tomography (PET), and magnetic resonance spectroscopy (MRS), assess regional differences in the biochemical status of tissues.[1,2] In nuclear medicine, including SPECT and PET, this is done by administering the patient with a biologically active molecule or pharmaceutical which is radiolabeled and accumulated in response to its biochemical attributes. Radionuclide imaging and nuclear medicine rely on the tracer principle in which a minute amount of a radiopharmaceutical is administered to assess physiological function or the biomolecular status of a tissue, tumour, or organ within the patient. The amount of the radiopharmaceutical is sufficiently small so that its administration does not perturb the normal function of the patient. However, the radiopharmaceutical produces a radioactive signal that can be measured, and ideally imaged, using an external array of radiation detectors. By design, the radiopharmaceutical has a targeted action, allowing it to be imaged to evaluate specific physiological processes in the body. There now are many radiopharmaceuticals available for medical diagnosis, with additional radiotracers available for *in vivo* as well as *in vitro* biological experimentation.

Nuclear medicine relies on use of radionuclide tracers which emit radiation in amounts proportional to their regional concentration within the body. For this reason, radionuclide imaging often is called ''emission'' imaging. This is in contrast to x-ray or ''transmission'' imaging where an external radiation source transmits radiation through the body and onto a set of opposing detectors. The properties of emission imaging are best suited for imaging functional processes, while transmission imaging is best suited for visualizing anatomical structure. Because the amount of radiation that can be administered internally to the patient is limited by considerations of radiation dose and count rate, radionuclide images inherently have poor photon statistics, are produced with only modest spatial resolution, and require relatively long scan times. In addition, the visual quality of radionuclide images can be degraded by physical factors such as photon attenuation and scatter radiation.[3-10] In contrast, transmission images are produced with short scan times, excellent signal-to-noise characteristics, and submillimeter spatial resolution, but generally contain limited functional information. These considerations illustrate that x-ray (transmission) and radionuclide (emission) imaging provide quite different but complementary information about the patient.[11-14] This information can be interpreted by the diagnostician to detect, localize, and diagnose diseases. Similarly, biomedical research scientists can analyze these attributes to monitor and measure processes that are important for

biological research, drug discovery, and development of new diagnostic and therapeutic techniques in humans and animals.

Several investigators have developed methods which attempt to improve the correlation between anatomical and physiological information obtained using these x-ray transmission and radionuclide emission imaging studies. Software-based image registration[15-17] can fuse images from two or more different studies after they are acquired separately. Commonly, image registration techniques produce a single ''fused'' or ''combined'' image in which, for example, the radionuclide distribution is displayed in colour over a grey-scale CT image of the same anatomical region. The simplest form of image registration uses ''rigid-body'' translation and rotation to match the two image data sets. These techniques can be applied most successfully to neurological studies,[18,19] where the skull provides a rigid structure that maintains the geometrical relationship of structures within the brain. The situation is more complicated when image registration techniques are applied to other areas of the body, for example the thorax and abdomen, where the body can bend and flex, especially when the x-ray and radionuclide data are captured using different machines in separate procedures, often on different days (see chapter 9). Geometrical relationships between different anatomical regions can be affected by the shape of the patient table, the orientation of the body and limbs during the imaging procedure, and the respiratory state of the patient.[20] In these cases, image registration might match the patient anatomy in one region of the body, but not in all anatomical regions. Image warping can improve registration over a larger region of the patient's anatomy, but in most cases, software-based image registration can be challenging and, at most institutions, is not used routinely for clinical procedures.[16,17]

With the goal of improving the correlation of different types of image data, other investigators have developed instrumentation which integrate both x-ray and radionuclide imaging is in a single device.[1,21-35] This technique, often called dual-modality imaging, can combine a PET or SPECT system with a CT scanner, using a common patient table, computer, and gantry so that both the x-ray and radionuclide image data are acquired sequentially without removing the patient from the scanner. This technique thereby produces anatomical and functional images with the patient in the same position and during a single procedure, which simplifies the image registration and fusion processes.[1,25,28,32,34,36] In seeking to achieve accurate registration of the anatomical and functional data, dual-modality imaging offers several potential advantages over conventional imaging techniques.[37] First, the radionuclide and x-ray images are supplementary and complementary. Radionuclide data can identify areas of disease that are not apparent on the x-ray images alone.[14,38-41] X-ray images provide an anatomical context that interpreters use to differentiate normal radionuclide uptake from that indicating disease, and to help localize disease sites within the body. Second, the x-ray data can be used to generate a patient-specific map of attenuation coefficients and other *a priori* anatomical data which in turn is

used to correct the radionuclide data for errors due to photon attenuation, scatter radiation, and other physical effects.[1,21,22,25,27,42-44] In these ways, the x-ray data can be used to improve both the visual quality and the quantitative accuracy of the correlated radionuclide data.

2. Brief History of Dual-Modality Imaging

Whereas the advent of dedicated dual-modality imaging systems designed specifically for clinical use is relatively recent, the potential advantages of combining anatomical and functional imaging has been recognized for several decades by radiological scientists and physicians. Many of the pioneers of nuclear medicine, including Mayneord,[45-47] Anger,[48,49] Cameron and Sorenson,[50] and Kuhl[51] recognized that a radionuclide imaging system could be augmented by adding an external radioisotope source to acquire transmission data for anatomical correlation of the emission image. Furthermore, Kuhl *et al.*[52] added an external radionuclide source on his Mark IV brain scanner to produce anatomical images useful for both localizing regions of radionuclide uptake and to correct for soft tissue absorption in the radionuclide emission data. In a 1974 review of photon attenuation, Budinger and Gullberg[53] noted that a patient-specific attenuation map could be produced from transmission data acquired using an external radionuclide source or extracted from a spatially correlated CT scan of the patient. A specific implementation of a combined emission-transmission scanner was disclosed by Mirshanov[54] who produced an engineering concept diagram showing a semiconductor detector to record radionuclide emission data and a strip scintillator to record coregistered x-ray transmission data from a patient (Figure 1). In addition, Kaplan *et al.*[55] proposed a high performance scintillation camera to record both emission data from an internal radionuclide source and transmission data from an external x-ray source. However these concepts were never reduced to practice or implemented in either an experimental or a clinical setting.[56]

In late 1980's and early 1990's, Hasegawa *et al.*[23,26,57,58] at the University of California, San Francisco, pioneered the development of dedicated emission/transmission imaging systems which could record both radionuclide and x-ray data for correlated functional/structural imaging. The first prototype, developed by Lang *et al.*[26,58] used an array of HPGe detectors (Figure 2) with sufficient energy discrimination and count-rate performance to discriminate γ-rays emitted by an internally distributed radiopharmaceutical from x-rays transmitted through the body from an external x-ray source. Phantom experiments with this first prototype led to the development of a second prototype (Figure 3) having a 20-cm reconstruction diameter that could record emission and transmission data from stationary animal or object using a single HPGe detector array. Kalki *et al.*[59,60] used this second prototype in animal studies both to demonstrate the capability of the system to facilitate image correlation and to test the feasibility of correcting the

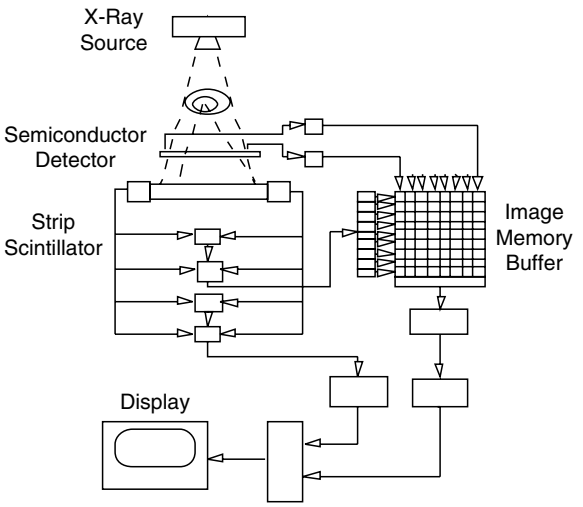


FIGURE 1. Concept drawing of transmission–emission tomography system proposed in the Soviet Union in 1987. System includes semiconductor detector for radionuclide imaging and strip scintillator for x-ray imaging, with electronics for combined recording and display of x-ray and radionuclide data sets. Reprinted from ref.[54]

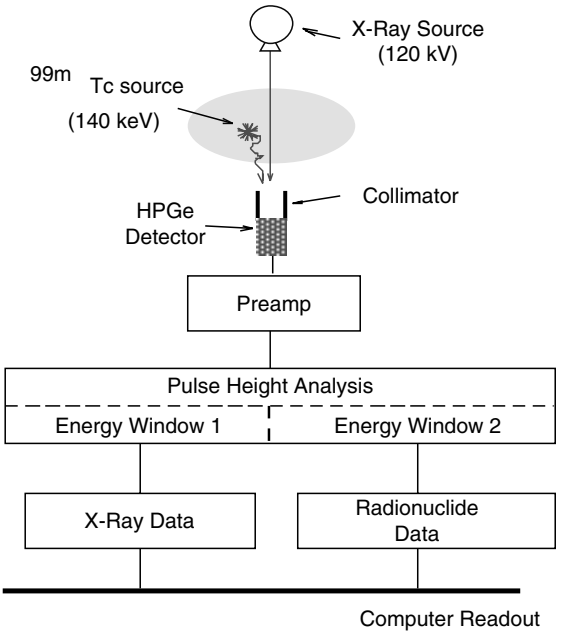


FIGURE 2. Schematic of data acquisition of combined emission-transmission imaging system developed at UCSF using single high-purity germanium detector array with fast pulse-counting electronics for simultaneous emission-transmission imaging.

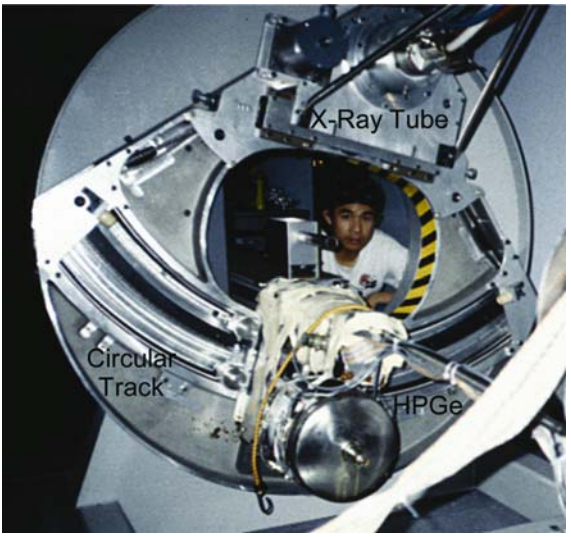


FIGURE 3. Early emission-transmission CT system at UCSF had an HPGe detector was translated across a circular arc to simulate an entire detector array and was rotated around an isocenter for tomographic imaging. Detector was read-out using fast pulse-counting electronics (Figure 2) using energy discrimination to separate the simultaneously acquired x-ray and radionuclide data. Reprinted with permission from ref.[1]

radionuclide data for photon attenuation using coregistered x-ray transmission data. Because the HPGe detector implemented in these first two prototypes was expensive and impractical for clinical use, the UCSF group next implemented a SPECT/CT scanner (Figure 4) for patient studies by sitting a GE 9800 Quick CT scanner in tandem with a GE 400 XR/T SPECT system.[21,22,42] This configuration allowed the patient to remain on a common patient table for radionuclide and x-ray imaging with separate detector technologies that already had been optimized for clinical use both in terms of technical performance and cost-effectiveness. The investigators used this system to demonstrate that CT data could produce a patient-specific attenuation map that could be incorporated into an iterative reconstruction algorithm for attenuation correction of the correlated radionuclide data. The system was used for imaging studies with phantoms, animals, and patients and demonstrated that the use of combined emission and transmission data could improve both the visual quality and the quantitative accuracy of radionuclide data in comparison to SPECT data alone.[21,22,42]

The first integrated PET/CT system (Figure 5) was developed by Townsend and co-workers at the University of Pittsburgh in 1998[28,33-35] by combining the imaging chains from a Somatom AR.SP (Siemens Medical Systems) CT system with an ECAT ART (CTI/Siemens) PET scanner. Both

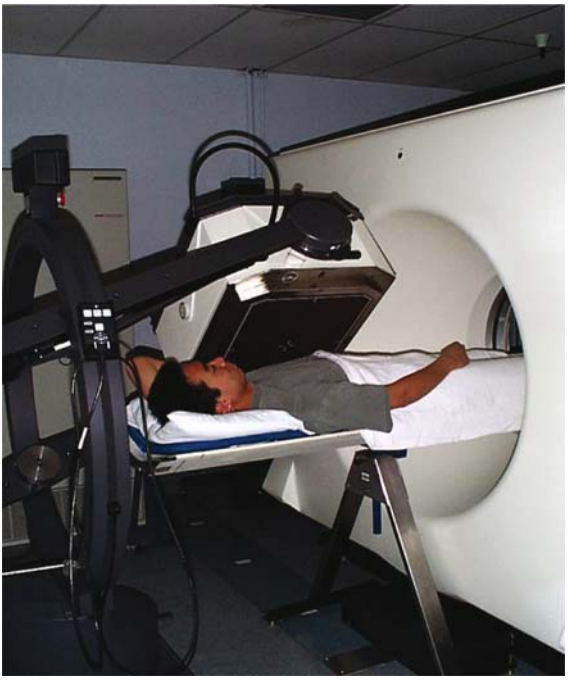


FIGURE 4. Combined SPECT/CT system configured at UCSF from a GE 9800 CT Quick CT scanner and a 400 XR/T SPECT system with a common patient table to translate the patient between the reconstruction volumes of the CT and SPECT systems. Reprinted with permission from ref.[1]

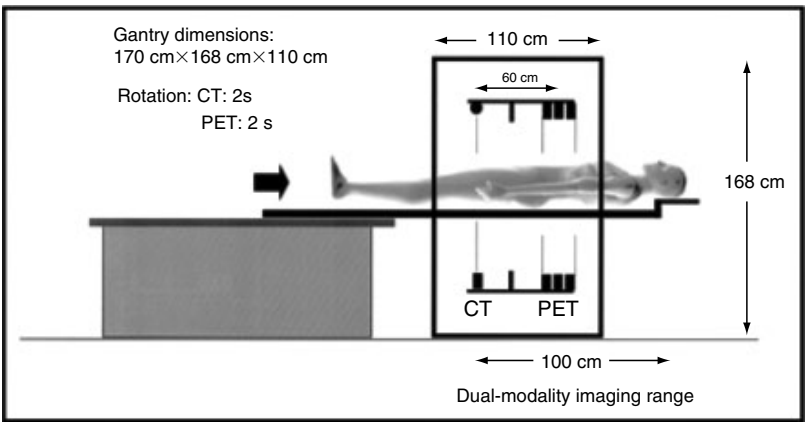


FIGURE 5. Schematic of combined PET/CT scanner developed at University of Pittsburgh by Townsend and co-workers. PET components were mounted on back of rotating CT scanner assembly. Centers of fields-of-view between PET and CT (vertical lines) were 60 cm apart. Combined PET/CT gantry was 110 cm deep, 170 cm high, and 168 cm wide. PET and CT data were acquired over 100 cm axial range of patient. Reprinted with permission from ref.[35]

the CT components and the PET detectors were mounted on opposite sides of the rotating stage of the CT system, and imaged a patient with a common patient table translated between the centers of the two tomographs which are offset axially by 60 cm.[33] The PET/CT prototype was operational at the University of Pittsburgh from May 1998 to August 2001, during which over 300 cancer patients were scanned. The success of these initial studies prompted significant interest from the major medical imaging equipment manufacturers who now all have introduced commercial PET/CT scanners for clinical use.

In 1999 and 2000, SPECT/CT and PET/CT dual-modality imaging systems were introduced by the major medical equipment manufacturers for clinical use, with approximately 400 systems of each type sold by midyear 2004. A significant success of PET/CT has been the improved image quality of FDG images for tumour localization.[38,39,41,61] The major use of SPECT/CT has been in reducing attenuation artefacts and improving the quality of myocardial perfusion imaging with ^{99m}Tc-sestamibi.[21,59,62-64] SPECT/CT also has demonstrated advantages for oncologic imaging with single-photon agents.[12,25,40,41,65-70] Both SPECT/CT and PET/CT have demonstrated their ability to facilitate attenuation correction using an x-ray based patient-specific attenuation map that can be produced faster and more accurately than attenuation maps generated with external radionuclide sources.[10,34] Clinical studies are underway to evaluate the applicability of x-ray based correction of photon attenuation,[71] and early results demonstrate improvement in sensitivity, specificity, and predictive accuracy in comparison to SPECT perfusion studies reconstructed without correction for photon attenuation. The anatomical information from PET/CT improves the differentiation of physiological (normal) uptake of FDG and other radiopharmaceuticals from that associated with disease, and thereby can reduce false positive errors in comparison to lesion characterization when radionuclide imaging is used alone. By providing high-resolution anatomical information from CT, dual-modality imaging also correlates functional and anatomical data to improve disease localization[14,38,39,41,61] and facilitates treatment planning for radiation oncology[12,72] or surgery.[68,73,74]

3. Capabilities of Dual-Modality Imaging

Dual-modality techniques offer a critical advantage over separate CT and radionuclide imaging systems in correlating functional and anatomical images without moving the patient (other than table translation). Dual-modality imaging also can account consistently for differences in reconstruction diameter, offsets in isocenter, image reconstruction coordinates, and image format (e.g., 512×512 vs. 128×128) between the CT and radionuclide image geometries to perform image coregistration and image fusion. Depending on the design of the system, image registration software also may be

needed to account for table sag or for misalignment when the patient moves between the CT and radionuclide image scans. Generally, the coordinate systems implicit in the radionuclide and CT image geometries are calibrated with respect to each other using fiducial markers that are scanned with both CT and radionuclide imaging. The image registration must be confirmed to avoid misregistration errors in the dual-modality images or in the radionuclide image reconstructed using CT-derived attenuation maps.

Dual-modality imaging also provides *a priori* patient-specific information that is needed to correct the radionuclide data for photon attenuation and other physical effects. Increasingly, x-ray sources are replacing external radionuclide sources for acquisition of the "transmission scan" to obtain projection data which can be reconstructed to produce a patient-specific map of linear attenuation coefficients so that the radionuclide (SPECT or PET) data can be reconstructed with a correction for photon attenuation.[5,75-80] Because the external radionuclide sources produce a limited fluence rate, the transmission scans often require several minutes and produce images that are noisy and photon-limited. Alternatively, transmission data can be acquired using an x-ray source having a small focus, a significantly higher photon fluence rate than a radionuclide source, and therefore produce tomographic reconstructions with 1 mm (or better spatial resolution) with excellent noise characteristics. The resulting CT images then can be calibrated to produce a patient-specific map of linear attenuation coefficients calculated for the energy of the radionuclide photons.[21-23,29,34,42-44] In this way, the CT data from a dual-modality imaging system facilitates, and in some ways simplifies, the process of correcting the radionuclide image for photon attenuation.

Several different CT-based attenuation correction techniques have been developed[21-23,29,34,42-44] and are reviewed in chapter 6. As an example, the technique developed by Blankespoor *et al.*[22,81] obtains CT calibration measurements from a phantom with cylindrical inserts containing water, fat-equivalent (ethanol), and bone equivalent material (K_2HPO_4). CT numbers extracted from each region are plotted against their known attenuation coefficients at the photon energy of the radionuclide to provide a piecewise linear calibration curve.[22,82,83] During the dual-modality imaging study, both CT and radionuclide data of the patient are acquired. The calibration curve described above is used to convert the CT image of the patient into an object-specific attenuation map. The resulting attenuation map then can be incorporated into the reconstruction of the radionuclide data[84] using ML-EM or other iterative algorithm, to correct the radionuclide data for perturbations due to photon attenuation. This or a similar process can be used to improve both the image quality and the quantitative accuracy of SPECT or PET images (see chapter 4). Important clinical applications include attenuation correction of myocardial perfusion images[80,85] to resolve false-positive defects caused by soft-tissue attenuation (Figure 6). In addition, the visual quality of oncologic images can be significantly improved when the ^{18}F-FDG data are reconstructed using attenuation correction.

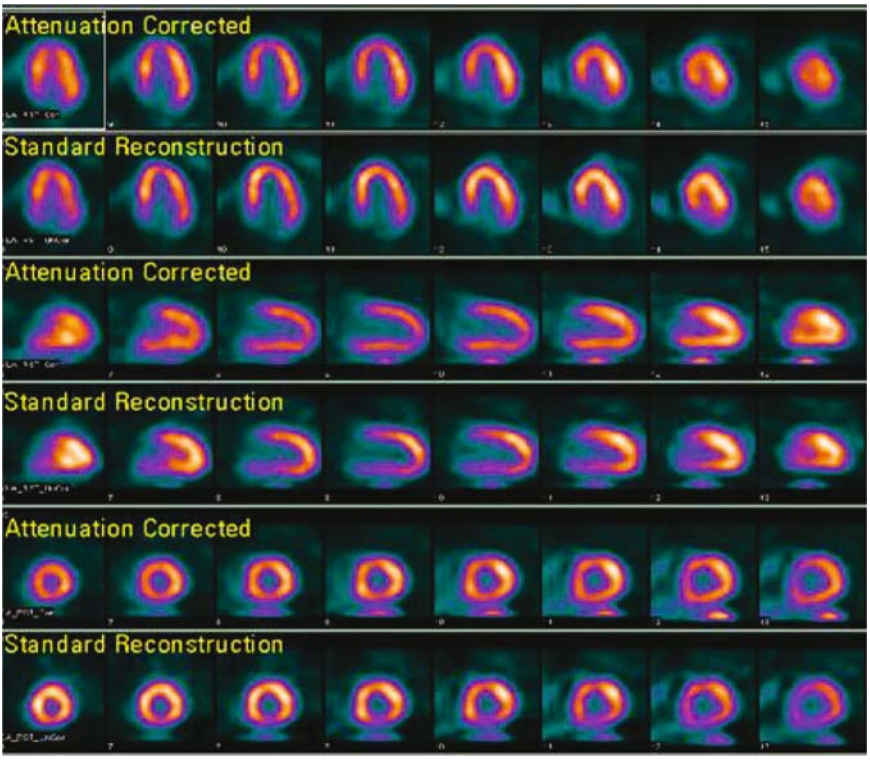


FIGURE 6. Myocardial perfusion scans acquired with GE Discovery VG using x-ray derived attenuation map. In each pair of rows, the top and bottom rows represent data reconstructed with and without attenuation correction respectively. In this study, the patient presented with chest pain with borderline normal ejection fraction of 47%. Resting ^{99m}Tc-sestamibi perfusion scan without attenuation correction shows mild inferior wall defect. However, myocardial perfusion images appear normal when reconstructed with x-ray based attenuation correction. Reprinted with permission from GE Healthcare Technologies, Waukesha, WI.

In addition to the energy conversion process described above, it is important to match the geometric characteristics of the CT-derived attenuation map and radionuclide tomograms. This typically requires the CT data to be resampled so that it can be presented in slices that have the same pixel format (e.g. 128×128, or 256×256) and same slice width as the radionuclide image. Accurate spatial registration of the CT and radionuclide data is important since slight differences in position of the attenuation map relative to the corresponding data can cause 'edge artefacts' which produce bright and dark 'rims' across edges of regions where the CT and radionuclide data are misaligned.

The process of generating attenuation maps as described above also assumes that the body is composed primarily of lean soft tissue (essentially

water-equivalent) and bone. However, CT studies of the head and body generally are obtained following the administration of intravenous and oral contrast media.[86] Since contrast media in the body can attenuate photons emitted by a radiopharmaceutical during a PET or SPECT study, it obviously is important to derive accurate attenuation maps that account for the presence of these materials in the body regardless of whether they are administered orally and/or intravenously.[42,69,86,87] It has been shown that SPECT and PET attenuation can be overestimated in the presence of positive contrast agents, which can generate significant artefacts (Figure 7). One can account for attenuation differences between iodine versus bone using a technique that generates a calibration curve for contrast media, using a method similar to that described above for soft tissue and bone alone. In this method, calibration data are obtained experimentally using CT to image a calibration phantom containing known concentrations of iodine contrast. The reconstructed CT values for each calibration region in the phantom are extracted from the CT scan, and are related to the known linear attenuation coefficients as a function

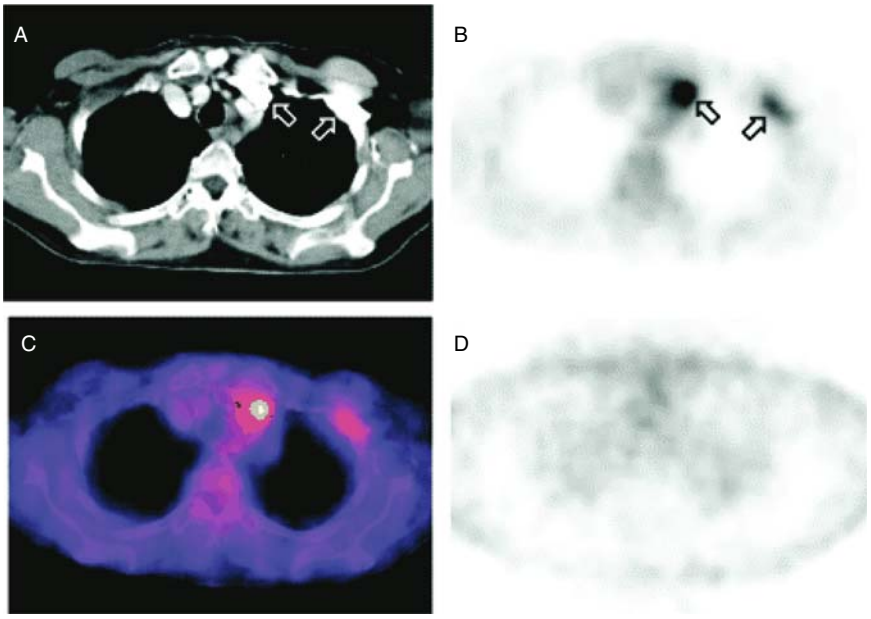


FIGURE 7. Contrast-enhanced related artefact in PET/CT imaging. Bolus passage of intravenous contrast agent in left subclavian and brachiocepalic veins on CT (**A**, arrows) led to areas of apparently increased glucose metabolism on CT-based attenuation corrected PET (**B**, arrows). On fused PET/CT images, this area of apparently increased glucose metabolism correlated with high-density contrast in venous system on CT (**C**). Reconstructed PET images without attenuation correction demonstrated homogeneous tracer distribution (**D**), demasking areas of apparently increased glucose metabolism as artefact. Reprinted with permission from ref.[86]

of CT number. Separate calibration curves are generated for each material (i.e., iodine vs. bone), for different x-ray potentials used to acquire CT scans, and for different photon energies (e.g. 140 keV, 364 keV, and 511 keV) encountered in nuclear medicine.[42,69] It is important to use image processing or other techniques to segment bony and iodine-containing regions in the patient's CT image so that bone and iodine can be scaled independently when forming the attenuation map. The attenuation maps with calibrated values for soft-tissue, fat, bone, and iodine, then are incorporated into an ML-EM or other iterative reconstruction algorithm for attenuation compensation of the radionuclide image. These methods have been tested in phantoms and in animals, and demonstrated in pilot studies on humans.

The x-ray data from a dual-modality imaging system also can be used to compensate the radionuclide data for contamination by scatter radiation (see chapter 7). One class of scatter compensation techniques characterizes the scatter distribution for different radionuclide and attenuation distributions with a spatially-dependent kernel, which can be used to convolve data from the photopeak window for estimating the scatter distribution from the acquired data. For example, the method proposed by Mukai *et al.*[88] assumes that the scatter radionuclide image ρ can be estimated from the "true" (i.e., "scatter-free") image η, as $\rho = P\eta$, where P is a matrix having elements $P(b,b')$ specifying the probability that a primary photon emitted from voxel $b\prime$ is scattered in voxel b with an energy suitable for detection within the primary energy window and which can be calculated by integrating the Klein–Nishina equation over the primary acquisition energy window taking into account the energy resolution of the emission system. Obviously, the "true" or "scatter-free" image is unknown, making it necessary to approximate this using the currently available image estimated by the tomographic reconstruction algorithm. The matrix P can be approximated from the coregistered CT-derived attenuation map of the object provided by the dual-modality imaging system to obtain the total attenuation ($\Delta(b,b\prime)$) along a line from the center of voxel b to the center of voxel $b\prime$. All the parameters needed can be determined accurately from x-ray based attenuation maps, illustrating how this process is facilitated by the availability of CT data from dual-modality imaging. Once the scatter distribution is calculated, it can be incorporated in the tomographic reconstruction algorithm to compensate the radionuclide data for the effect of this perturbation.

The visual quality and quantitative accuracy of the radionuclide data can be improved by using the methods described above by image fusion with CT, and use of x-ray data to correct the radionuclide data for perturbations arising from photon attenuation and Compton scatter. In addition, the radionuclide image can be improved by correcting the data for the geometric response of the radionuclide detector.[6,42,89,90] Traditionally, radionuclide quantification methods have focused on techniques that compensate the radionuclide image reconstruction process for physical errors such as photon attenuation, scatter radiation, and partial volume errors as discussed above. This may include the

use of recovery coefficients,[89] requiring *a priori* information about object size and shape, that generally are only used for simple target geometries (e.g., spheres). Other investigators have developed methods that incorporate physical models, not in image reconstruction, but rather into the quantification process itself. Several methods have been described in the literature combining SPECT and MRI[91] and PET and MRI[92,93] for brain imaging, and SPECT and CT for cardiac and oncologic imaging.[21,59,69] Several approaches were developed specifically for quantitation using planar imaging[65,94,95] and are described in chapter 13 of this book. One technique called "template projection"[42,90] can be used to quantify radioactivity in planar images. This process begins with dual-modality (i.e., radionuclide and CT) images of a target region surrounded by background. Regions-of-interest (ROIs) for the target (e.g., tumour) and background regions are defined on the high-resolution CT image, then are used to define "templates" which represent idealized radionuclide-containing objects (e.g., tumour vs. background) with unit radionuclide concentration. Planar imaging of these ideal radionuclide objects is modelled by mathematically projecting the templates onto the plane of the radionuclide detector using the known geometrical transformation between the CT coordinate system and the planar radionuclide image provided inherently by the dual-modality imaging system. This is performed using the projection machinery available from iterative reconstruction algorithms (e.g., ML-EM), including physical models of photon attenuation and non-ideal detector response, and which potentially can incorporate perturbations from detected scatter radiation, patient motion, and pharmaceutical kinetics. This process generates "projected templates" that are analogous to conventional ROIs, in that they delineate a target region over which events in the radionuclide image are integrated. Like conventional ROIs, the projected templates specify the geometry of the tumour and background regions on the projected planar radionuclide images; however the projected templates are defined on the high-resolution CT images rather than on the low-resolution radionuclide images. Furthermore, unlike traditional ROIs which are uniform, the projected templates are nonuniform and contain information about physical effects (photon attenuation, detector response, scatter radiation) included in the projector model.

Several methods using the projected templates can quantify activity in a planar radionuclide image. If we assume that the imaging process is linear, and consider the photons emitted from M different regions (e.g., tumours and background) where each region has a uniform activity concentration A_m, then the counts $p(d)$ measured in detector (i.e., pixel) d are estimated by $p^*(d)$ as

$$p^*(d) = \sum_{m=1}^{M} A_m \phi_m(d) \tag{1}$$

where $\phi_m(d)$ represents the value of the projected template at detector location d, which is physically equivalent to the relative number of photons

detected from radioactive region m in pixel location d. If we measure the value $p(d)$ from the planar emission data and given that we have calculated $\phi_m(d)$ from the template-projection technique, we can estimate the activity concentration A_m for region m (i.e., for the target lesion) by minimizing the weighted least squares difference

$$\chi^2 = \sum_{d=1}^{D} \left[\frac{p(d) - \sum_{m=1}^{M} A_m \phi_m(d)}{\sqrt{p(d)}} \right]^2 \tag{2}$$

where D is the number of detector elements (i.e., pixels) from which we estimate the activity concentrations (typically an area surrounding the target lesion). We note that a theoretical basis for this formulation has been described elegantly by Formiconi[96] and is similar to a method described by Liu.[97] An alternative formulation assuming Poisson statistics was suggested by Carson.[98] Specific details of the implementation described here can be found elsewhere.[42]

The template projection technique described above for planar imaging can be extended to quantify target regions in tomographic images. This technique, called ''template projection-reconstruction'',[42] begins with the ''template projection'' technique described above. This process is repeated for all angles sampled by the tomographic acquisition of the real radionuclide data. After the template projection data for the target and for the background structures are modelled, they are reconstructed with the same reconstruction algorithms (e.g., ML-EM or filtered-backprojection) that are used for reconstructing the emission data. The reconstructed templates contain information about physical effects (e.g., photon attenuation, scatter radiation, geometric response) included in the modelling process and can be used to quantify the emission tomographic data. Several methods are then available for estimating the object activity concentration. For example, the radionuclide content of the target region can be calculated using a technique analogous to that discussed above for planar imaging by assuming that both the imaging and reconstruction processes are linear. With this assumption, the reconstructed activity concentration $\rho(i)$ of each voxel of the radionuclide image is represented as a linear combination of the M radionuclide concentrations A_m (assumed to be uniform) in the different regions, weighted by their corresponding reconstructed template values $\gamma_m(i)$

$$\rho(i) = \sum_{m=1}^{M} A_m \gamma_m(i) \tag{3}$$

where $\gamma_m(i)$ represents the contribution of activity region m to measured counts in reconstructed voxel i, as estimated from the template projection-reconstruction process. The activity concentrations of each object may be

obtained by linear least squares fitting analogous to that given by Eq. 2. Alternatively, the corrected mean activity concentration ρ_t in the target region can be calculated through voxel-by-voxel rescaling after estimating or assuming the background activity concentration ρ_b.

$$\rho_t = \frac{1}{N} \sum_{i=1}^{N} \frac{\rho(i) - \rho_b \gamma_b(i)}{\gamma_t(i)} \tag{4}$$

where $\gamma_b(i)$ and $\gamma_t(i)$ represent the reconstructed template values for voxel location i and contributed by the background (b) and target (t) regions, respectively. The voxel-by-voxel correction is analogous to scaling the reconstructed data with object size- and shape-dependent recovery factors[42,89,99] (i.e., by division by $\gamma_t(i)$ and corrects for "spill-in" of background activity into the target lesion (by subtraction of the term $\rho_b \gamma_b(i)$). The radionuclide content of a given target region t, can be calculated as $\rho_t V_T$, where V_t is the actual target volume defined on the CT image.

At present, dual-modality imaging is most widely used to enable spatial registration of structural and functional information from CT and radionuclide imaging. Attenuation correction of radionuclide data using a CT-derived patient-specific map of attenuation coefficients also is available with all dual-modality imaging systems. However, very few clinical or biological studies have been conducted that use quantification methods such as template projection, template projection-reconstruction, or other model-based methods that extract *a priori* anatomical information for analysis of functional data from PET or SPECT. However, these techniques are being developed and are slowly being incorporated into research studies. For example, Koral *et al.*[99] have developed a technique that uses CT-SPECT image fusion with conjugate view imaging to quantify the uptake of single-photon radionuclides *in vivo*. In this analysis, a patient-specific attenuation map is derived from the correlated CT images to correct the radionuclide data for photon attenuation. In addition, volumes of interest delineating the extent of the tumour are defined anatomically on the CT scans, then are superposed on the SPECT data for radionuclide quantification[100,101]. Excellent results have been obtained with this method to assess radiation dosimetry of lymphoma patients undergoing radionuclide therapy with ^{131}I-tositumomab, also known as the anti-B1 monoclonal antibody.[65,94] An initial study showed that the percentage of infused dose in the tumour following therapeutic administration of the agent could be predicted within 8% by linear extrapolation of the uptake of tracer amounts infused before therapy.[94] A later study demonstrated that the probability of complete response using radionuclide therapy was correctly associated with a high tumour dose measured in the patient using the SPECT-CT analysis.[65] Furthermore, for individual tumours, the combined SPECT-CT method provided a statistically significant relationship between radiation dose and tumour volume reduction at 12 weeks whereas this relationship was not

significant for dosimetry estimated from a conventional pre-therapy conjugate view radionuclide imaging alone.[102,103] Koral *et al.* developed the quantification method using images obtained using separate SPECT and CT studies and that were combined using software-based image registration; however, the quantification technique certainly could be implemented using a dual-modality SPECT/CT system of the type described in this chapter. The reader is directed to a more extensive discussion of planar radionuclide quantification techniques in chapter 13. Radionuclide quantification also can be implementated using SPECT/CT, PET/CT, and other dual-modality imaging approaches of the type described in this chapter.[104]

4. General Design Features of Dual-Modality Imaging Systems

Modern dual-modality imaging systems incorporate subsystems for radionuclide imaging and for x-ray computed tomography that essentially use the same components as those in dedicated nuclear medicine and CT systems. In PET/CT, the radionuclide detector uses a scintillator (bismuth germinate, lutetium orthosilicate, gadolinium orthosilicate) coupled to an array of photomultiplier tubes for imaging the annihilation photons from the positron-emitting radiopharmaceuticals. Modern PET/CT scanners also include an x-ray source and detector identical to those used in modern multislice helical scanning CT scanners.[33-35] Similarly, SPECT/CT systems use conventional dual-headed scintillation cameras suitable for planar scintigraphy or tomographic imaging of single-photon radionuclides, or coincidence-imaging of PET radiopharmaceuticals. The first-generation clinical SPECT/CT scanners used a low-resolution CT detector[40,41,70,105] that offered relatively modest scan times (i.e., approximately 20 seconds per slice). However, newer SPECT/CT scanners now are becoming available that incorporate state-of-the-art multislice helical CT scanners identical to those used for diagnostic CT procedures.

The integration of the radionuclide and x-ray imaging chains in a dual-modality imaging system requires special considerations beyond those needed for scanners designed for single modality imaging alone. One challenge is offered by the presence of x-ray scatter from the patient that has the potential to reach and possibly damage the radionuclide detectors which are designed for the relatively low photon fluence rate encountered in radionuclide imaging.[34,35,64] To avoid this possibility, the radionuclide detector in a dual-modality system typically is offset in the axial direction from the plane of the x-ray source and detector. This distance can be relatively small when a modest x-ray tube is used such as the 140 kV, 1 mA tube used in the GE Millennium VG SPECT/CT system,[64] but can be 60 cm or more when a

diagnostic CT scanner is coupled to a modern PET scanner operated without septa.[28,30,31,34]

As noted above, all dual-modality systems rely on separate x-ray and radionuclide imaging chains that must be supported on a common mechanical gantry to maintain consistent spatial relationship between the two data sets, and allow the detectors to be rotated and positioned accurately for tomographic imaging. The requirements for translational and angular positioning accuracy are, of course, different for CT, SPECT, and PET. For example, CT requires approximately 1000 angular samples acquired with an angular position and center of rotation maintained with submillimeter accuracy. In comparison, SPECT and PET have spatial resolutions of a few millimetres, and therefore can be performed with an accuracy of slightly less than a millimetre for clinical imaging.

The mechanical gantry of the dual-modality imaging system obviously must be designed to satisfy the requirements for both the radionuclide image and for CT. This can be achieved in several ways. In first generation SPECT/CT systems,[40,41,70,105] the SPECT detectors and CT imaging chain were mounted on the same rotating platform and were used sequentially while rotated around the patient. This limited the rotational speed of the x-ray and radionuclide imaging chains to approximately 20 sec per rotation, but also had the advantage that it could be performed using a gantry similar to that used with a conventional scintillation camera. Second generation SPECT/CT systems that now are available include high-performance diagnostic CT subsystems. This requires the heavy SPECT detectors to be mounted on a separate rotating platform from the CT imaging chain which is rotated at speeds of 0.25 to 0.4 sec per revolution. While this design obviously increases the performance of the CT subsystem, it also increases the cost and complexity of the gantry.

In comparison to SPECT/CT in which the radionuclide detector is rotated around the patient during data acquisition, PET typically (but not always) is performed using a dedicated high-end stationary detector ring. A PET/CT system therefore can be configured by designing a gantry that mounts a stationary PET detector ring in tandem with a platform that rotates the CT imaging chain around the patient using a mechanical configuration similar to that used in a conventional diagnostic CT scanner. Alternatively, a partial ring of PET detectors can be rotated to acquire the PET data using the same rotating platform as the CT subsystem. This approach was taken by Townsend and his colleagues in their implementation of the first PET/CT system,[28,30,31] and is an alternative for a more economical dual-modality system in comparison to those that use a full-ring of PET detectors. All of these mechanical designs have been used in commercial dual-modality systems and obviously offer trade-offs in terms of their performance and cost.

The patient table is another seemingly simple, yet important element of a dual-modality scanner.[34,35] Most imaging systems use cantilevered patient tables to support the patient in the bore of the imaging system. Patient tables

are designed to support patients weighing up to 500 pounds, but obviously deflect to varying degrees when they are loaded and extended with normal adult patients. However, dual-modality systems use x-ray and radionuclide imaging chains in tandem and thereby require longer patient tables than conventional imaging systems. Moreover, the table extension and the degree of deflection can be different for the x-ray and radionuclide imaging chains which can introduce a patient-dependent inaccuracy in the registration of the x-ray and radionuclide images. This problem is overcome by several different methods. The first uses a patient table which is supported in front of the scanner, with a secondary support between or at the far end of the x-ray and radionuclide imaging chains to minimize table deflection. A second approach adopted by CTI Molecular Imaging and Siemens Medical Systems uses a patient table that can be fixed on a base that is translated across the floor to extend the patient into the scanner. Since the patient platform is stationary relative to its support structure (which acts as a fulcrum), the deflection of the patient table is identical when the patient is positioned in the radionuclide imager or the CT scanner.

Finally, in modern dual-modality scanners, the computer systems are well integrated in terms of system control, data acquisition, image reconstruction, image display, and data processing and analysis.[34,106] The dual-modality system must calibrate the CT data so that it can be used as an attenuation map to correct the radionuclide data for photon attenuation.[21,22,27,29,42,44] For physician review, the dual-modality system also typically registers the CT and radionuclide data and presents the radionuclide image as a colour overlay on the grey-scale CT image. Finally, software tools are provided that, for example, allow a cursor placed on the CT image by the operator with another cursor automatically placed in the identical position on the radionuclide image, and vice versa. These software functions allow the operator to utilize the dual-modality data in correcting, viewing, and interpreting and obviously are important design elements in modern dual-modality imaging systems.

5. PET/CT Imaging Systems

The first combined PET/CT system was developed by Townsend and co-workers at the University of Pittsburgh in 1998.[28,33,34] The system was configured by combining a Somatom AR.SP spiral CT scanner (Siemens Medical Systems) in tandem with the PET detectors from an ECAT ART PET system (CTI/Siemens). The PET subsystem consisted of two arrays of bismuth germanate (BGO) block detectors covering 16.2 cm in the axial direction with 24 partial detector rings operated without septa, allowing the PET data to be acquired in a fully 3-dimensional mode. The CT scanner was a third generation helical CT scanner that has an x-ray tube operated at 110-130 kVp with a 6.5 mm Al-equivalent filtration and having a xenon

x-ray detector with 512 elements. Both the CT components and the PET detectors are mounted on opposite surfaces of the rotating stage of the CT system, and during imaging are rotated continuously at a rate of 30 rpm. The system has a common patient table, with the patient translated between the centers of the CT and PET imaging planes which were offset axially by 60 cm. On the prototype system, an axial extent of 100 cm in the patient could be covered by simple table translation[33,34] with the PET and CT images acquired sequentially rather than simultaneously. The PET/CT prototype was operational at the University of Pittsburgh from May 1998 to August 2001, during which over 300 cancer patients were scanned. These pioneering studies by Townsend and his colleagues demonstrated both the feasibility and the clinical benefit of combined PET/CT scanning, and prompted significant interest from the major medical imaging equipment manufacturers who now all have introduced commercial PET/CT scanners for clinical use.

PET/CT scanners now are available from all of the major medical imaging equipment manufacturers (GE Medical Systems, CTI/Siemens Medical Systems, and Philips Medical Systems).[35,107] Current systems have up to 16 slice CT capability and have radionuclide detectors with either 2D or 3D PET imaging capability. The PET scanner can have either bismuth germinate (BGO), lutetium oxyorthosilicate (LSO), or gadolinium oxyorthosilicate (GSO) scintillators. The CT study typically is used for both localization of the FDG uptake[38,41,61] as well as for attenuation correction of the PET image. In addition, the use of CT in comparison to external transmission rod sources for producing the attenuation data increases patient throughput by approximately 30%.[107] As noted above, the PET/CT system also has a specially designed patient table that is designed to minimize deflection when it is extended into the patient port.

Figure 8 shows the 7 steps identified by Beyer *et al.*[108] that comprise a typical PET/CT scan, demonstrating the degree of integration available in a modern dual-modality imaging system. (1) The patient is prepared for imaging which commonly includes administration both with contrast media[86] and with the radiopharmaceutical, typically 370 to 555 MBq (10 to 15 mCi) of ^{18}F-fluro-deoxyglucose (FDG) in adults. (2) The patient then is asked to remove all metal objects that could introduce artefacts in the CT scan and then is positioned on the patient table of the dual/modality imaging system. (3) The patient then undergoes an ''overview'' or ''scout'' scan during which x-ray projection data are obtained from the patient to identify the axial extent of the CT and radionuclide study. (4) The patient undergoes a CT acquisition. (5) The patient then undergoes the nuclear medicine study approximately 1 hour after FDG administration. (6) The CT and PET data then are reconstructed and registered, with the CT data used for attenuation correction of the reconstructed radionuclide tomograms. (7) The images are reviewed by a physician who can view the CT scan, the radionuclide images, and the fused x-ray/radionuclide data, followed by preparation of the associated clinical report.

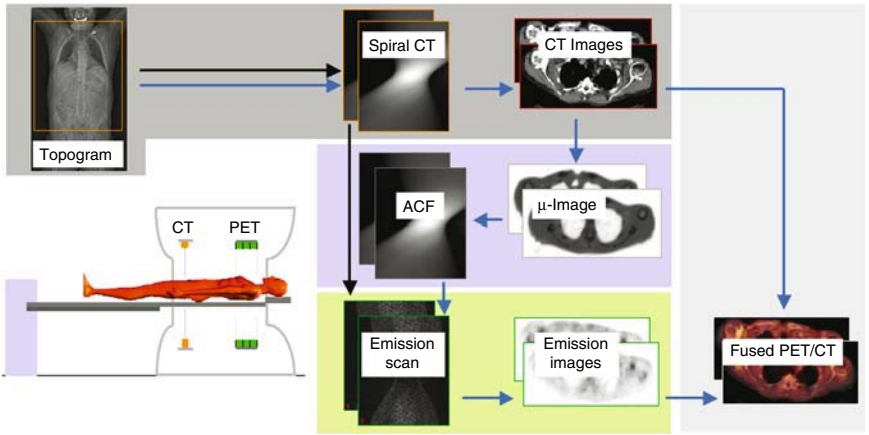


FIGURE 8. Illustration of standard PET/CT scanning protocol. The patient is positioned on a common specially designed patient table in front of the combined mechanical gantry. First, a topogram is used to define the co-axial imaging range (orange). The spiral CT scan (grey box) preceded the PET scan (green box). The CT data are reconstructed on-line and used for the purpose of attenuation correction of the corresponding emission data (blue box). Black and blue arrows indicate acquisition and data processing streams, respectively. Reprinted with permission from ref.[108]

There have been multiple studies which have demonstrated the role of PET/CT especially for oncologic applications.[110] For example, in a clinical study of PET/CT for the evaluation of cancer in 204 patients with 586 suspicious lesions, PET/CT provided additional information over separate interpretation of PET and CT in 99 patients (49%) with 178 sites (30%).[109] Furthermore, PET/CT improved characterization of equivocal lesions as definitely benign in 10% of sites and as definitely malignant in 5% of sites. It precisely defined the anatomical location of malignant FDG uptake in 6%, and led to the retrospective lesion detection on PET or CT in 8%. As a result, PET/CT had an impact on the management of 28 patients (14%), obviated the need for further evaluation in 5 patients, guided further diagnostic procedures in 7 patients, and assisted in planning therapy in 16 patients. Figure 9 shows an example of a patient study where the combined PET/CT images provided additional information, thus impacting the characterization of abnormal FDG uptake.

6. SPECT/CT Imaging Systems

As noted earlier, the first SPECT/CT imaging systems were developed at the University of California, San Francisco by Hasegawa and co-workers. The first prototype instruments (Figures 2 and 3) were configured as small-bore

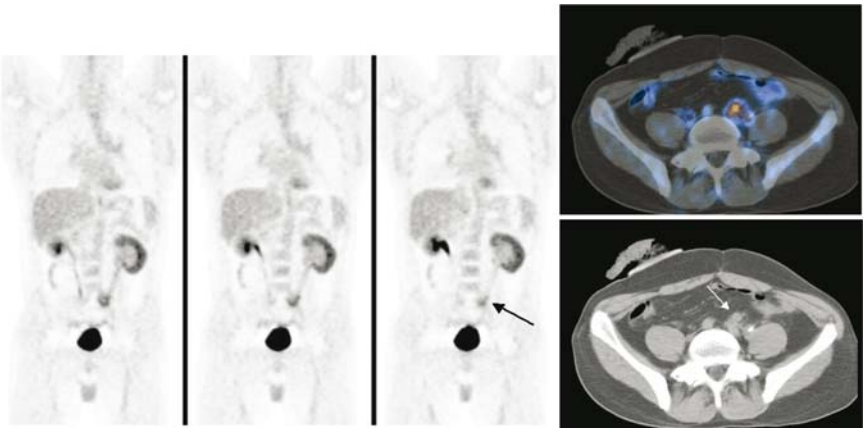


FIGURE 9. Precise characterization of increased ^{18}F-FDG uptake and retrospective lesion detection on CT, after PET/CT. The patient is a 35-y-old man, 22 months after treatment of colon cancer, with negative high-resolution contrast-enhanced CT and normal levels of serum tumour markers, referred for PET scan for assessment of pelvic pain. The coronal PET images (left) show area of increased ^{18}F-FDG uptake in left pelvic region (arrow), interpreted as equivocal for malignancy, possibly related to inflammatory changes associated with ureteral stent or to physiological bowel uptake. Hybrid PET/CT (right) transaxial image (top) precisely localized uptake to soft-tissue mass adjacent to left ureter, anterior to left iliac vessels. Mass (arrow) was detected only retrospectively on both diagnostic CT and CT component of hybrid imaging study (bottom). Patient received chemotherapy, resulting in pain relief and decrease in size of pelvic mass on follow-up CT. Reprinted with permission from ref.[109]

SPECT/CT systems with a small segmented high-purity germanium detector operated in pulse-counting mode for simultaneous x-ray and radionuclide imaging.[23,26,57,59] A clinical prototype SPECT/CT system (Figure 4) also was configured at UCSF by installing a GE X/RT SPECT camera and GE 9800 CT scanner in tandem with a common patient table,[21,22,42,69] and was used for studies on both large animals and patients.

The first dual-modality imaging system designed for routine clinical use was the Millennium VG (Figure 10) introduced by General Electric Healthcare Technologies (Waukesha, WI) in 1999.[40,41,70,105] This system has both x-ray and radionuclide imaging chains that were integrated on a common gantry for "functional-anatomical mapping". X-ray projection data are acquired by rotating the x-ray detector and low-power x-ray tube 220 degrees around the patient. Since the patient is not removed from the patient table, the x-ray and radionuclide images are acquired with a consistent patient position in a way that facilitates accurate image registration. The image acquisition of the CT data requires approximately five minutes, and the radionuclide data requires a scan time of approximately 30 to 60

FIGURE 10. GE Millennium VG SPECT system with x-ray anatomical mapping. Reprinted with permission from General Electric Healthcare Technologies, Waukesha, Wisconsin.

minutes. These SPECT/CT systems produce transmission data with significantly higher quality and better spatial resolution than those acquired with an external radionuclide source, and can be used for anatomical localization[14] and attenuation correction of the radionuclide data.

In June 2004, both Siemens Medical Systems and Philips Medical Systems introduced SPECT/CT systems offering the performance available on state-of-the-art diagnostic CT systems. For example, Siemens Medical Systems now offer their Symbia line of SPECT/CT systems that have a single-slice, 2-slice, or 6-slice CT options with 0.6 sec rotation speed coupled to a dual-headed variable-angle SPECT system. The 6-slice CT scanner will cover the abdomen in less than 6 seconds and allows both x-ray based attenuation correction and anatomical localization with the correlated diagnostic CT images. Philips Medical Systems has announced their Precedence SPECT/CT system which incorporates a dual-headed SPECT system with a 16 slice diagnostic CT scanner. The advanced CT capability has several potential benefits. First, they offer CT scan times that are compatible with the use of contrast media for improved lesion delineation with CT. Second, they offer improved image quality and spatial resolution typical of that offered by state-of-the-art diagnostic CT scanners. Finally, CT scanners offering 16 or more slices can be used for cardiac and coronary imaging. SPECT/CT scanners with multislice capabilities therefore offer improved performance for tumour imaging and have the potential to acquire and overlay radionuclide myocardial perfusion scans on a CT coronary angiogram showing the underlying arterial and cardiac anatomy in high spatial resolution.

7. Small-Animal Imaging Systems

The advantages of improved image quality and quantitative accuracy that are available for clinical applications of dual-modality imaging also are being investigated as a means of facilitating biological research, especially those that involve small animals such as mice and rats.[37,111] Advances in molecular biology and genetic engineering, and the modern growth of the pharmaceutical and biotechnology industries have increased the need for biological studies that involve mice. Approximately 90% of all research involving vertebrate biology now uses mice as the predominant mammalian model.[112] Approximately 25 million mice were raised in 2001 for experimental studies, and this number is expected to increase at a rate of 10% to 20% annually over the next decade. In addition, experimental methods now allow the genetic sequence of mice to be manipulated by knockout or transgenic techniques in which a genetic sequence from another individual or species can be deleted or inserted into that of an individual animal. Murine models now are available for a wide variety of biological conditions and transgenic animals[113-116] now account for a sizable and growing fraction of animal models used in biological research.

Traditional techniques of biological research including histology, organ sampling, and autoradiography, require that animals must be sacrificed for the measurement to be obtained. This precludes investigators from obtaining multiple measurements from an individual during the progression of disease or while an animal is monitored following diagnostic procedures or therapeutic interventions. This greatly increases the number of animals that must be used for an individual experiment, compromises statistical power, and requires that control and experimental measurements are made in different rather than in the same animal.

While noninvasive imaging techniques potentially could be used in these experiments to obtain the desired measurements, clinical medical systems do not have adequate spatial resolution and temporal resolution, and other specific characteristics needed to image small animals. For example, for cardiovascular studies, the left heart of a mouse in short axis view is approximately 5 mm in diameter with the left ventricular wall having a thickness of approximately 1 mm. In addition, the mouse heart rate is 600 to 900 beats per minute making it difficult to acquire image data with cardiac gating.[116,117] Similarly, small tumours can have a diameter of a few millimetres, within the spatial resolution limit of clinical CT and MRI systems but outside that of conventional radionuclide imaging such as SPECT or PET. For these reasons, investigators have developed high resolution imaging systems specifically designed for small animal imaging.[111,112,118-120] These include microCT systems[118,121-123] that incorporate a low power X-ray tube and a phosphor-coupled CCD camera or similar two-dimensional x-ray imaging detector to achieve spatial resolutions as high as 25 microns or

better. Similarly, high-resolution images can be obtained using microPET scanners[111,119] having high-resolution detectors operated in coincidence to obtain spatial resolutions in the range of 1 to 2 mm, whereas SPECT imaging of mice[124-129] can be performed using pinhole collimation to obtain spatial resolutions better than 1 millimetre. In general, these imaging systems provide characteristics similar to those used for clinical imaging of humans. In the specific case of radionuclide imaging, including microPET and micro-SPECT, the accumulation of radiotracer can be used to monitor physiological or biological interactions in the animal. Dual-modality imaging offers potential advantages in biological studies, similar to those offered in clinical protocols.[120,130-138] First, the availability of anatomical data improves localization and interpretation of studies involving radiopharmaceutical uptake. In addition, the visual quality and quantitative accuracy of small animal imaging can be improved using x-ray based techniques to correct the radionuclide data for physical errors contributed by photon attenuation, scatter radiation, and partial volume effects due to the limited spatial resolution of the radionuclide imaging system.

In response to these concerns, several investigators now are developing dual-modality imaging systems specifically designed for imaging small animals.[37] Dr. Cherry and his group, now at the University of California at Davis, have developed a microPET system for small-animal imaging. Commercial versions of these microPET systems suitable for imaging mice and rats now are available from CTI Concorde Microsystems (Knoxville, TN). The same group also has developed a microCT/microPET dual-modality imaging system.[131] The microPET detectors use LSO scintillator coupled through a fiber-optic taper to a position sensitive photomultiplier tube. These are placed on opposite sides of the animal with the annihilation photons from the positron emission detected in coincidence. The system also includes a microCT system having a microfocus x-ray tube and an amorphous selenium detector coupled to a flat panel thin film resistors readout array.[139] A schematic diagram of a prototype system and an image of a mouse obtained with the microPET/microCT system are shown in Figure 11.

SPECT/CT systems designed specifically for small animal imaging also are being developed by several investigators.[120,132-138,140,141] One of the first small animal SPECT/CT systems was developed by a consortium that includes The Thomas Jefferson National Accelerator Facility (Jefferson Laboratory), The University of Virginia, and researchers at the College of William and Mary.[132-134] These systems use a compact scintillation camera that operates with multiple Hamamatsu R3292 position-sensitive photomultiplier tubes (PSPMT's) coupled to a pixelated array of cesium iodide (CsI(Tl)) crystals using both pinhole and parallel-hole collimators. The x-ray data are acquired using a small fluoroscopic x-ray system (Lixi, Inc., Downers Grove, IL). Representative planar dual-modality images acquired on one of the systems mentioned above obtained from a 250-g rat injected with

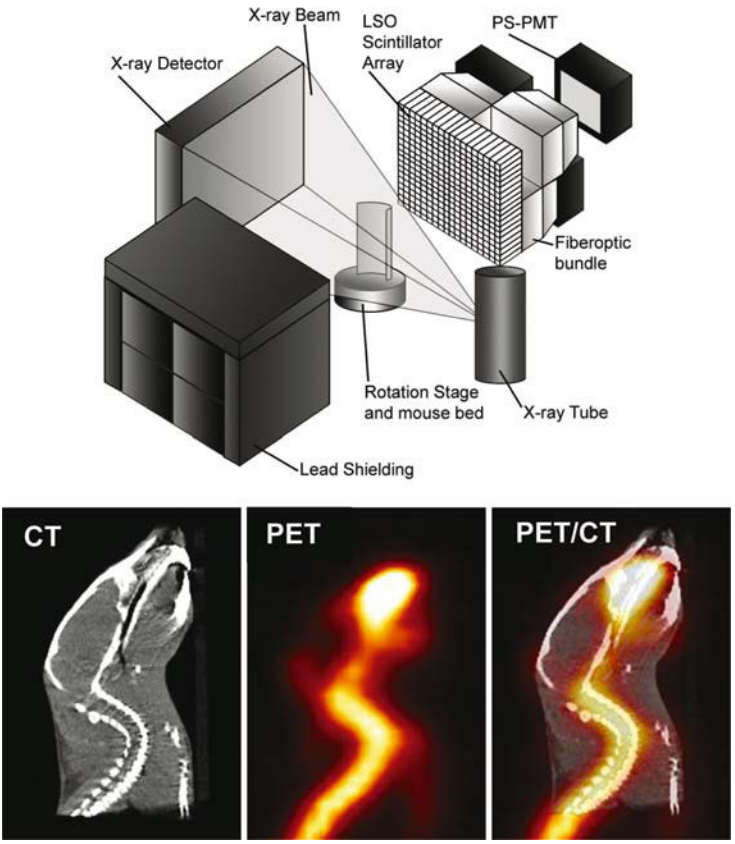


FIGURE 11. Top: schematic diagram of the experimental set-up. The CT and PET components are mounted in a coplanar geometry so that the same portion of the mouse is imaged by both systems at once. The back PET detector is shown without the lead shielding in place in order to show the design of the detector modules. The lead shielding on the entrance face is 1.5 mm thick. The detectors remain stationary while the object is rotated on the bed mounted on the rotation stage. Bottom: scan of a mouse 3 hours after injection of 26 MBq of ^{18}F bone seeking tracer. The images shown are from left to right: sagittal slices through the CT image, the ^{18}F PET image, and the fused PET/CT image. Note the overlap of the PET uptake with the location of the bones in the CT image. The scan was acquired using 400 views in a time of 38 min. Reprinted with permission from ref.[131]

^{99m}Tc-MDP are shown in Figure 12, demonstrating radiopharmaceutical uptake in the bone superimposed with an anatomical x-ray image of the rat.[133,142]

Gamma Medica™ Inc. (Northridge, CA) has developed and introduced a small animal SPECT/CT system[137,138] with two compact scintillation cameras[126,143,144] and a high-resolution CT subsystem[138] for dual-modality

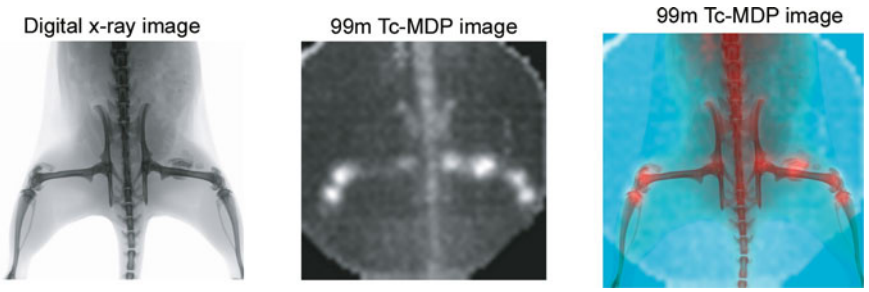


FIGURE 12. Superposition of digital x-ray radiography (left) and radionuclide ^{99m}Tc-MDP bone scan (middle) resulting in a fused image of the rear portion of a rat (right). (Courtesy of Dr Mark B. Williams, University of Virginia).

imaging of mice, rats, and other small animals (Figure 13). The scintillation camera can be operated with pinhole collimators that provide submillimeter spatial resolution in the reconstructed images or with parallel-hole collimators when higher detection sensitivity or whole body imaging is desired. The system includes a high-resolution microCT subsystem[138] configured with a CMOS x-ray detector coupled to a gadolinium oxysulfide scintillator and a low-power x-ray tube. The microCT system provides anatomical imaging with a spatial resolution of approximately 50 microns; the resulting x-ray data can be used both for attenuation correction and for anatomical localization of the radionuclide data (Figure 14). In addition, the radionuclide data can be acquired with ECG-gating for cardiovascular imaging applications where wall-motion abnormalities, ejection fraction calculations, or other assessments of ventricular function are necessary.

FIGURE 13. Gamma Medica X-SPECT system has dual-headed compact SPECT system and high-resolution microCT imaging capability. Reprinted with permission from Gamma Medica, Incorporated, Northridge, CA.

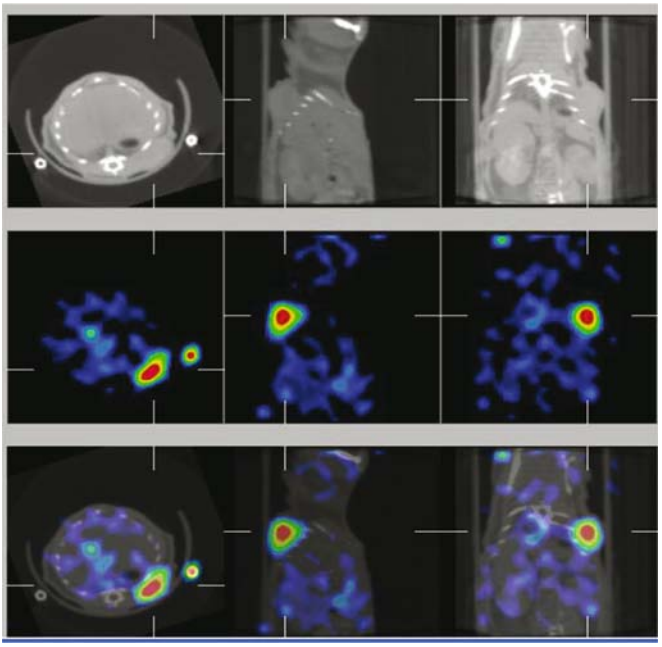


FIGURE 14. X-ray CT (top row), SPECT (middle row), and coregistered SPECT/CT (bottom row) obtained with small animal dual-modality system (X-SPECT, Gamma Medica, Northridge, CA) of colon carcinoma xenograft in right shoulder of mouse. SPECT images recorded 23 hours after administration of ^{123}I-labeled anti-CEA minibody[145] that localizes in colon carcinoma. Images shown in transverse (left column), sagital (middle column), and coronal (right column) tomograms. (Images from G Sundaresan, S Gambhir, AM Wu, with permission from AM Wu, UCLA, Los Angeles, CA, and reprinted with permission from ref.[146]

Finally, the Center for Gamma Ray Imaging at the University of Arizona, under the direction of Dr Barrett, has configured a high-resolution SPECT/CT system for small animal imaging.[135] High-resolution SPECT is performed with a modular semiconductor detector that consists of a $2.5 \times 2.5 \times 0.2\,cm^3$ slab of cadmium zinc telluride (CdZnTe) operated with a continuous gold electrode to apply bias on one side, and a 64×64 array of pixelated gold electrodes on the opposite side connected to an ASIC for readout of the individual pixel signals. The detector has a 380μm pixel pitch, and 330μm wide pixels, coupled to a 7 mm thick parallel-hole collimator for radionuclide imaging. The x-ray and radionuclide imaging subsystems are mounted with their image axes perpendicular to one another with the animal rotated vertically within the common field of view. The x-ray and radionuclide projection data are acquired sequentially, and corrected for distortions

and nonuniformities introduced by each of the detectors, then reconstructed with statistical iterative algorithms (OS-EM).

8. Need for Dual-Modality Imaging

Even though both PET/CT and SPECT/CT systems have been accepted commercially, the clinical role and need of these systems is still being debated. PET, for example, is a diagnostic technique that offers a sensitivity and specificity in excess of 90% for many malignant diseases, and some argue that an incremental improvement in specificity or sensitivity beyond that point probably cannot justify the cost of performing image fusion for all patients on a routine basis.[147] Furthermore, a patient receiving medical diagnosis typically undergoes both anatomical and functional imaging from commonly-available single-modality medical imaging systems. This means that the anatomical images can be viewed side-by-side with the functional images when necessary. In other cases, image registration techniques[15,16] can be used when a more direct fusion of anatomical and functional data following their acquisition without the use of the expensive new dual-modality imaging systems (see chapter 9).

Nonetheless, other practitioners now see important roles for dual-modality imaging in the sense that it offers important features for diagnostic studies and patient management.[37,73,109] First, the anatomical and functional information from dual-modality imaging are supplementary and not redundant. As noted earlier, anatomical imaging is performed with techniques such as CT or MRI that have excellent spatial resolution and signal-to-noise characteristics, but that may offer relatively low specificity for differentiating disease from normal structures. In contrast, radionuclide imaging generally targets a specific functional or metabolic signature in a way that can be highly specific, but generally lacks spatial resolution and anatomical cues which often are needed to localize or stage the disease, or for planning therapy.[36,39,41,148] Similarly, the availability of correlated functional and anatomical images improves the detection of disease by highlighting areas of increased radiopharmaceutical uptake on the anatomical CT or MRI scan, whereas regions that look abnormal in the anatomical image can draw attention to a potential area of disease where radiopharmaceutical uptake may be low. The information from CT supplements that from radionuclide imaging, and *vice versa*; therefore it generally is advantageous to view the CT and radionuclide images side-by-side during the diagnostic interpretation. In other cases, it can be valuable to view a fused dual-modality image in which the radionuclide data are presented as a colour overlay on the grey-scale CT image. The availability of a dual-modality imaging system facilitates this process in acquiring the functional and anatomical data with the patient in a consistent configuration and during a single study, in a way that is faster and more cost-efficient than attempting

to register the images by software after they are acquired on separate imaging systems. The dual-modality image data can be used to guide radiation treatment planning for example by providing anatomical and functional data that are important for defining the target volume as well as indicating normal regions that should avoid irradiation. Similar roles are played when the dual-modality data is used to guide surgical approach, biopsy, or other interventional procedures.[73]

In addition, the anatomical and functional information from a dual-modality imaging system are complementary in that together they provide information that cannot be easily discerned from one type of image alone. This is best illustrated in oncologic applications where anatomical imaging often is needed to differentiate whether the radiopharmaceutical has localized in response to disease (e.g. in the primary tumour, lymphatic system, or metastatic site)[38,39,41,68,73] or as part of a benign process (e.g. in the GI tract, urinary system, or in response to inflammation).[35] Moreover, the process of differentiating normal from abnormal uptake of ^{18}F-FDG can be complicated by the relative paucity of anatomical cues in the ^{18}F-FDG scan, making it necessary for the diagnostician to refer to anatomical images obtained from CT or MRI to correlate the structural and functional information needed to complete the analysis.

The potential advantages of dual-modality imaging are further illustrated by a meta-study of PET-CT fusion for radiotherapy treatment planning in non-small cell lung carcinoma (NSCLC). Paulino *et al.*[148] found that the use of FDG imaging with PET/CT may alter target volumes for radiation therapy treatment planning in 26 to 100% of patients with NSCLC compared with CT-based treatment planning alone. Approximately 15 to 64% had an increase in the planning target volume (PTV), whereas 21 to 36% had a decrease in PTV. Similarly, in a study at Duke University,[149] SPECT-CT was useful in detecting the 48% of patients with hypoperfused regions of the lung, and in 11% of patients this information was used to alter the radiation therapy fields to avoid highly functional lung tissue. In addition, PET/CT and SPECT/CT may have an important role in guiding radiation treatment planning in cancers of the brain, head and neck, cervix, and other areas including lymphoma and melanoma.[148] These examples are not exhaustive, but indicate the potentially important role of dual-modality imaging in guiding radiation treatment planning in comparison to techniques that use anatomical imaging with CT alone.

The complementary nature of dual-modality imaging also is demonstrated by the use of a CT-derived attenuation map and other *a priori* patient-specific information from CT that can be used to improve both the visual quality and the quantitative accuracy of the correlated radionuclide imaging. This occurs in myocardial perfusion imaging of ^{99m}Tc-sestamibi or thallium-201 where SPECT/CT can provide an attenuation map to compensate the radionuclide image for soft-tissue attenuation associated with false-positive errors. In both PET/CT and SPECT/CT, the use of CT-derived attenuation

map is also practical and cost-efficient in that with modern CT scanners it is generated in a few seconds, resulting in more efficient use of time than SPECT or PET scanners that use external radionuclide sources to record the transmission image. The resulting CT-derived attenuation map also is of significantly higher quality, and arguably is more accurate, than those acquired using scanning line sources available in conventional radionuclide imaging systems. As dual-modality imaging systems become available with faster and higher spatial-resolution CT systems, it may be possible to perform CT angiography of peripheral or coronary vessels that can be superimposed on a colour map of tissue perfusion or metabolism obtained with SPECT or PET.

9. Challenges of Dual-Modality Imaging

There are several challenges that face the use of dual-modality imaging, and that may represent inherent limitations in this technique. All currently-used dual-modality imaging systems record the emission and transmission data using separate detectors rather than a single detector. In addition, the x-ray and radionuclide imaging chains are separated by several tens of centimetres, to facilitate mechanical clearance and to avoid contamination of the radionuclide data by scatter radiation produced by the x-ray scan. One potential problem occurs when the patient moves either voluntarily or involuntarily between or during the anatomical and functional image acquisitions. This can occur, for example, when the patient adjusts his or her position while lying on the patient bed. Patient motion also occurs due to respiration, cardiac motion, peristalsis, and bladder filling, all of which can lead to motion blurring or misregistration errors between the radionuclide and CT image acquisitions.[110] Traditionally, CT data were acquired following a breath-hold, whereas PET data were acquired over several minutes with the patient breathing quietly. However, these breathing protocols can lead to misregistration artefacts due to anatomical displacements of the diaphragm and chest wall during a PET/CT or SPECT/CT scan. For example, if the position of the diaphragm is displaced in the CT scan, which then is used as an attenuation map for the radionuclide data, this displacement can lead to an underestimate or overestimate of radionuclide uptake in the reconstructed emission data.[63] The discrepancy in diaphragmatic position between PET and CT can result in the appearance of the so-called ''cold'' artefact at the lung base (Figure 15). A recent study[150] noted that in 300 patients with proven liver lesions; approximately 2% appeared to have the lesion localized in the lung due to respiratory motion. Care therefore must be taken when interpreting results from patients with disease in periphery of the lung where noticeable radiopharmaceutical uptake may be contributed by respiratory-induced motion artefacts rather than disease.

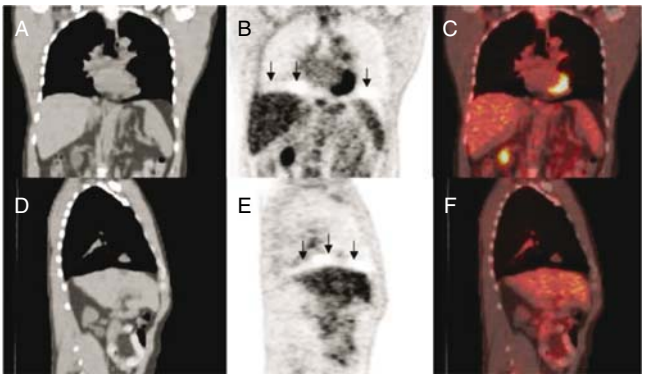


FIGURE 15. Demonstration of a cold artefact on PET images reconstructed with CT-based attenuation correction. (**A**) Coronal CT, (**B**) ^{18}F-FDG PET, and (**C**) fused, and (**D**) sagittal CT, (**E**) PET, and (**F**) fused images are displayed. A region of decreased metabolic activity is demonstrated in the diaphragmatic region (below vertical arrows), representing a "cold artefact". Reprinted with permission from ref.[110]

Modern dual-modality scanners now use CT technology that can acquire the anatomical data within a few seconds after the patient is positioned on the bed. For this reason, the CT acquisition rarely is the factor that limits the speed of the dual-modality image acquisition in comparison to SPECT or PET that can consume several minutes to complete. If additional increases in scan speed are needed, these must be implemented using faster radionuclide scans using newer detector technologies, faster scintillators, increased computing power, and more efficient scanner architectures or detector designs than are currently being used. In PET, this includes the possibility of replacing conventional PET block detectors with LSO panel detectors[35,151] which would cover a larger axial extent of the patient with the goal of achieving 5 min scan times and thereby would allow even faster scan times than are achievable with current systems. Regardless, faster scan speeds both improve patient comfort and limit the time during which patient motion can occur during the study. In addition, faster scan speeds can promote faster patient throughput and thereby increase system utilization and cost-effectiveness of the study.

Dual-modality imaging systems now commonly use the correlated CT scan as an attenuation map for the correction of the emission data acquired with PET or SPECT. The use of iodinated contrast media for the CT scan also presents a potential challenge when the CT scan is used as an attenuation map for correction of SPECT or CT data.[86] Since the x-ray energy of the CT data generally will be different than the photon energy of the radiopharmaceutical emission data, as described above, the CT data must be transformed to units of linear attenuation coefficient at the photon energy of the radiopharmaceutical emission data. This is performed by assuming

that image contrast in the CT data is contributed by mixtures of air, soft tissue, and bone. The presence of iodinated contrast complicates this process since two regions that have the same image contrast may indeed have different compositions, for example contributed by bone and soft tissue in one case and iodine contrast and soft-tissue in another situation. These artefacts are most severe in cases where the contrast media is concentrated, for example in abdominal imaging after the patient swallows a bolus of oral contrast. In this case, the higher densities contributed by the oral contrast media can lead to an overestimation of PET activity.[86] Some investigators have proposed using an image segmentation post-processing technique in which image regions corresponding to iodinated contrast are segmented using image processing from those contributed by bone.[87] In addition, other strategies including the acquisition of both pre-contrast and post-contrast CT scans, can be used to minimize possible artefacts contributed by the presence of contrast media when the CT scan is to be used as an attenuation map for correction of the PET data.[35]

Use of dual-modality has demanded that technologists be cross-trained in both modalities, and that medical specialists have expertise in interpreting both anatomical and functional image data. Also, optimal use of SPECT/CT and PET/CT requires a team approach involving physicians from radiology and nuclear medicine, as well as oncology, surgery, and other clinical disciplines.

10. Future of Dual-Modality Imaging

Over the past 10 to 15 years, dual-modality imaging has progressed steadily from experimental studies of technical feasibility to the development of clinical prototypes to its present role as a diagnostic tool that is experiencing widening commercial availability and clinical utilization. Dual-modality imaging has advanced rapidly, primarily by incorporating the latest technological advances of CT, SPECT, and PET. Currently, combined MRI/MRSI,[152,153] PET/CT,[34,35] and SPECT/CT[1,25] systems are available commercially and are used clinically. SPECT/CT systems now are available that include state-of-the-art dual-headed scintillation camera and multislice helical CT scanners for rapid patient throughput. Similarly, the PET/CT scanners now are widely available from all of the major equipment manufacturers that include high-speed helical CT scanners that offer up to 16-slice capability and submillimeter spatial resolution.[34,35] This has reduced the time needed to acquire the transmission data from several minutes with an external radionuclide source to a few seconds with x-ray imaging. In addition, the current generation of PET/CT scanners now incorporates state-of-the-art PET detector rings, some of which operate at high count rates and can perform volumetric imaging without septa. At the time when PET/CT was first introduced clinically in 2000, PET scan times required

approximately 1 hour to complete whereas a complete PET/CT study now can be completed in the range of 10 to 15 minutes. This has led to the potential of increasing patient throughput from 3 to 4 patients a day a few years ago to 10 to 12 patients a day with current-generation PET/CT scanners.[31] Moreover, the addition of advanced CT capability allows anatomical images to be acquired after the patient is administered with contrast media to improve lesion detection in oncologic imaging[86] or to visualize the cardiac chambers[154,155] as well as the coronary and peripheral vasculature.[156,157] With these capabilities, the next-generation PET/CT and SPECT/CT systems could produce high-resolution structural images of the cardiac chambers and of coronary and peripheral vasculature that can be correlated with myocardial perfusion and other functional assessments with radionuclide imaging. In addition, the use of contrast media could enable advanced radionuclide quantification techniques in clinical studies such as the template projection method discussed above.[21] These capabilities would have the potential of improving the quantitative assessment of cancer and cardiovascular disease in comparison to studies acquired with SPECT or CT alone.

Similarly, it is expected that the technology of small animal SPECT/CT will continue to advance. Current small animal radionuclide SPECT systems obtain submillimeter spatial resolution at the cost of reduced detection efficiency. Newer multipinhole SPECT systems[158-160] are under development and offer both improved geometric efficiency and spatial resolution in comparison to current radionuclide imaging approaches in a way that would improve image quality and reduce scan times for dynamic or ECG-gated cardiovascular imaging in small animals. Unfortunately, given the current state of high-resolution x-ray detectors and microfocus x-ray sources, microCT systems that allow cardiac gating and *in vivo* coronary imaging in small animals will be difficult to develop over the next few years. Nevertheless, these capabilities would be very useful for functional/structural imaging and quantitative radionuclide assessments of small animals, similar to those that we expect to develop for clinical dual-modality imaging. Finally, advances in computing power will enable the development and implementation of new anatomically-guided statistical reconstruction algorithms and data processing techniques that will offer advantages for both clinical and small animal imaging with dual-modality imaging.

While all clinical and commercial dual-modality systems have been configured in the form of PET/CT or SPECT/CT scanners, several investigators proposed and in some cases have implemented and tested prototype dual-modality systems that combine MRI with PET.[130,161-165] The development of combined PET/MRI imaging has several important motivations.[166] First, MRI produces high-resolution (1 mm or better) anatomical and structural images that offer better soft-tissue contrast sensitivity than CT, has excellent contrast between white and grey matter, and is capable of assessing flow, diffusion, and cardiac motion.[167-170] In addition, MRI can be combined with

magnetic resonance spectroscopy (MRS) to measure the regional biochemical content and to assess metabolic status or the presence of neoplasia and other diseases in specific tissue areas[171-173] including the prostate.[153,174-176] Finally, MRI does not use any ionizing radiation and therefore can be used in serial studies, for paediatric cases,[177] and in other situations where radiation dose should be limited. Radionuclide imaging (specifically PET) records the regional distribution of radiolabeled tracers, but unlike MRS cannot distinguish the specific molecular species to which the radionuclide is attached and unlike MRI provides only little anatomical information.

There are, however, several important challenges that must be overcome in implementing and operating a combined PET/MRI or SPECT/MRI imaging system.[32] In comparison to x-ray CT, MRI typically is more expensive, involves longer scan times, and produces anatomical images from which it is more difficult to derive attenuation maps for photon correction of the radionuclide data.[178] Furthermore, virtually all clinical radionuclide imaging detectors use photomultiplier tubes whose performance can be seriously affected in the presence of magnetic fields which are significantly smaller than those produced by modern MRI scanners. This is especially problematic in an MRI scanner which relies on rapidly switching gradient magnetic fields and radiofrequency (RF) signals to produce the magnetic resonance image. The presence of the magnetic field gradients and RF signals certainly could disrupt the performance of a photomultiplier tube and PET detector if they were located within or adjacent to the magnet of the MRI system. Similarly, the operation of the MRI system relies on a very uniform and stable magnetic field to produce the MRI image. The introduction of radiation detectors, electronics, and other bulk materials can perturb the magnetic field in a way that introduces artefacts in the MR image.

In spite of these challenges, several researchers are investigating methods to integrate a radionuclide imaging system directly in an MRI scanner. For example, it may be possible to design a radionuclide imaging system having detectors made from nonmagnetic materials that can be placed within the magnetic field of an MRI/MRS system.[32] For example, Shao *et al.*[130,179] developed a 3.8-cm ring of small scintillator crystals that was placed in the MR system for PET imaging. The crystals were optically coupled through 3 m long fiber optics to an external array of position-sensitive photomultiplier tubes, and which could be read-out through external processing electronics. By keeping the radiation-sensitive elements of the detector within the MR system, while operating the electronics away from the magnetic field, the combined system could perform simultaneous PET/MR imaging. Shao *et al.*[161] and Slates *et al.*[162] also performed simultaneous PET/MR imaging with a larger (5.6 cm-diameter) detector ring using the same design (Figure 16). Researchers at Kings College London placed the system inside of a 9.4-T NMR spectrometer to study metabolism in an isolated, perfused rat heart model. ^{32}P-NMR spectra were acquired simultaneously with PET images of ^{18}F-FDG uptake in the myocardium.[180,181] Slates *et al.*[162] are

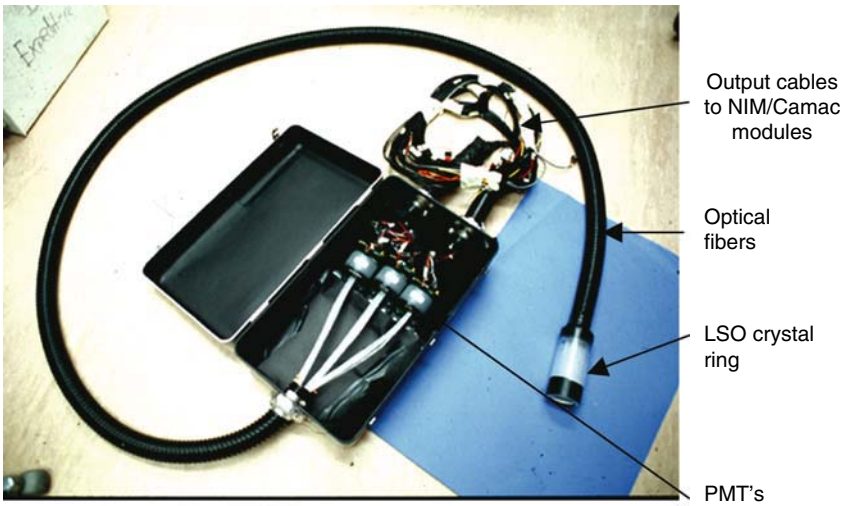


FIGURE 16. A photograph of the prototype MR compatible PET scanner (McPET) scanner developed by UCLA (CA) in collaboration with UMDS (London). Reprinted with permission from ref.[179]

extending this design concept to develop an MR-compatible PET scanner with one ring of 480 LSO crystals arranged in 3 layers (160 crystals per layer) with a diameter of 11.2 cm corresponding to a 5 cm diameter field of view, large enough to accommodate an animal within a stereotactic frame. The system is designed to offer adequate energy resolution and sensitivity for simultaneous PET/MRI imaging of small animals (Figure 17).[181,182]

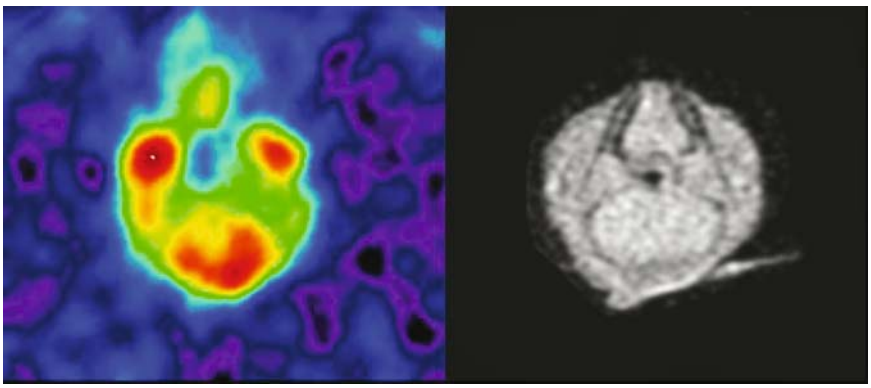


FIGURE 17. Small animal PET (left) and MRI (right) images obtained with MRI-compatible small animal PET detector (figure 16). Courtesy of Dr Simon R. Cherry (University of California, Davis) and reprinted with his permission.

Other investigators have proposed PET/MRI systems configured with suitable solid-state detectors that can be operated within a magnetic field for radionuclide imaging. For example, Pichler *et al.*[183] have tested APD's within a high-field (9.7 T) NMR spectrometer and have produced radionuclide data that appear to be free of distortion. However, it is still unknown whether the introduction of the APD's cause distortions in the magnetic field to an extent that would cause severe artefacts in the MRI image.[32]

11. Concluding Remarks

Dual-modality imaging is an approach that combines imaging modalities in a way that provides diagnostic information that is not available from a single imaging modality alone. Currently available dual-modality imaging systems include SPECT/CT or PET/CT, with which the radionuclide imaging technology (SPECT or PET) provides functional or metabolic information that is complementary to anatomical information provided by CT. In addition, the development of dual-modality systems that combine radionuclide imaging with magnetic resonance imaging is an area of active research. For spatial and temporal correlation, the dual-modality data can be presented as a fused image in which the radionuclide data are displayed in colour and are superposed on the gray-scale CT image. The resulting correlated data improves differentiation of foci of radionuclide uptake that can indicate disease from those representing normal physiological uptake that are benign.

Dual-modality imaging has been available clinically since the year 2000, and as such is a relatively recent development in the fields of diagnostic radiology and nuclear medicine. However, the commercial emergence of both PET/CT and SPECT/CT has been rapid and has benefited significantly from recent technological advances in the conventional SPECT, PET, and CT. The clinical growth of dual-modality imaging has been most dramatic in the area of PET/CT which now is available from all of the major medical imaging equipment manufacturers; some observers predict that all PET systems will be installed as PET/CT scanners in the near future. Over the past year, SPECT/CT has gained increased interest and has potential clinical applications in myocardial perfusion imaging and oncologic imaging. Newer high-resolution SPECT/CT and PET/CT systems also are becoming available for small animal imaging and are needed for molecular imaging, biological research, and pharmaceutical development in small animal models of human biology and disease.

At present dual-modality imaging is primarily used for structural-functional correlations. However, as this chapter has attempted to describe, dual-modality imaging also has important ramifications for radionuclide quantification, the major theme and focus of this volume. For example, dual-modality imaging provides x-ray CT data can be normalized to obtain a patient-specific map of attenuation coefficients that be used to compensate

the correlated radionuclide data for photon attenuation or for Compton scatter. In addition, regions of interest defined anatomically on the CT image can be used to quantify the correlated radionuclide data in a way that allows more precise target and background definition, and that can use model-based methods that correct the extracted quantitative data for partial-volume effects and other perturbations. The use and understanding of dual-modality imaging as an enabling concept for radionuclide quantification is just starting to emerge. The importance of this development will only be understood and manifest over the ensuing years as PET/CT, SPECT/CT, and other forms of dual-modality imaging become available and are utilized for clinical studies of humans as well as biological investigations involving animal models of biology and disease.

References

1. Hasegawa B. H., Iwata, K., Wong, K. H. *et al.*, Dual-modality imaging of function and physiology. *Acad Radiol* **9:** 1305-1321 (2002).
2. Schiepers C. and Hoh, C. K., Positron emission tomography as a diagnostic tool in oncology. *Eur Radiol* **8:** 1481-1494 (1998).
3. Zaidi H., Quantitative SPECT: Recent developments in detector response, attenuation and scatter correction techniques. *Phys Med* **12:** 101-117 (1996).
4. Larsson A., Johansson, L., Sundstrom, T. *et al.*, A method for attenuation and scatter correction of brain SPECT based on computed tomography images. *Nucl Med Comm* **24:** 411-420 (2003).
5. Huang S. C., Hoffman, E. J., Phelps, M. E. *et al.*, Quantitation in positron emission computed tomography. 1. Effects of inaccurate attenuation correction *J Comput Assist Tomogr* **3:** 804-814 (1979).
6. Hoffman E. J., Huang, S. C. and Phelps, M. E., Quantitation in positron emission computed tomography. I. Effect of object size. *J Comput Assist Tomogr* **3:** 299-308 (1979).
7. Tsui B., Frey, E., Zhao, X. *et al.*, The importance and implementation of accurate 3D compensation methods for quantitative SPECT. *Phys Med Biol* **39:** 509-530 (1994).
8. Tsui B. M. W., Zhao, X., Frey, E. *et al.*, Quantitative single-photon emission computed tomography: Basics and clinical considerations. *Semin Nucl Med* **24:** 38-65 (1994).
9. Rosenthal M. S., Cullom, J., Hawkins, W. *et al.*, Quantitative SPECT imaging. A review and recommendations by the focus committee of the Society of Nuclear Medicine computer and instrumentation council. *J Nucl Med* **36:** 1489-1513 (1995).
10. Zaidi H. and Hasegawa, B. H., Determination of the attenuation map in emission tomography. *J Nucl Med* **44:** 291-315 (2003).
11. Shreve P. D., Adding structure to function *J Nucl Med* **41:** 1380-1382 (2000).
12. Schillaci O. and Simonetti, G., Fusion imaging in nuclear medicine-Applications of dual-modality systems in oncology. *Cancer Biother Radiopharm* **19:** 1-10 (2004).

13. Aizer–Dannon A., Bar–Am, A., Ron, I. G. *et al*., Fused functional-anatomic images of metastatic cancer of cervix obtained by a combined gamma camera and an x-ray tube hybrid system with an illustrative case and review of the F-18-fluorodeoxyglucose literature. *Gynecol Oncol* **90:** 453-457 (2003).
14. Pfannenberg A. C., Eschmann, S. M., Horger, M. *et al*., Benefit of anatomical-functional image fusion in the diagnostic work-up of neuroendocrine neoplasms. *Eur J Nucl Med Mol Imaging* **30:** 835-843 (2003).
15. Maintz J. B. and Viergever, M. A., A survey of medical image registration. *Med Image Anal* **2:** 1-36 (1998).
16. Hutton B. F. and Braun, M., Software for image registration: algorithms, accuracy, efficacy. *Semin Nucl Med* **33:** 180-192 (2003).
17. Slomka P. J., Software approach to merging molecular with anatomic information. *J Nucl Med* **45 Suppl 1:** 36S-45S (2004).
18. Pellizzari C., Chen, G. T. Y., Spelbring, D. R. *et al*., Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. *J Comput Assist Tomogr* **13:** 20-26 (1989).
19. Pietrzyk U., Herholz, K., Fink, G. *et al*., An interactive technique for three-dimensional image registration: validation for PET, SPECT, MRI and CT brain studies. *J Nucl Med* **35:** 2011-2018 (1994).
20. Visvikis D. and Ell, P. J., Impact of techology on the utilisation of positron emission tomography in lymphoma: current and future perspectives. *Eur J Nucl Med Mol Imaging* **30:** S106-S116 (2003).
21. Da Silva A. J., Tang, H. R., Wong, K. H. *et al*., Absolute quantification of regional myocardial uptake of 99mTc-sestamibi with SPECT: experimental validation in a porcine model. *J Nucl Med* **42:** 772-779 (2001).
22. Blankespoor S. C., Wu, X., Kalki, K. *et al*., Attenuation correction of SPECT using x-ray CT on an emission-transmission CT system: Myocardial perfusion assessment. *IEEE Trans Nucl Sci* **43:** 2263-2274 (1996).
23. Hasegawa B. H., Gingold, E. L., Reilly, S. M. *et al*., Description of a simultaneous emission-transmission CT system. *Proc SPIE* Vol. 1231; *pp* 50-60 (1990).
24. Hasegawa B. H., Tang, H. R., Da Silva, A. J. *et al*., Dual-modality imaging. *Nucl Instr Meth Phys Res A* **471:** 140-144 (2000).
25. Hasegawa B. H., Wong, K. H., Iwata, K. *et al*., Dual-modality imaging of cancer with SPECT/CT. *Technol Cancer Res Treat* **1:** 449-458 (2002).
26. Lang T. F., Hasegawa, B. H., Liew, S. C. *et al*., Description of a prototype emission-transmission CT imaging system. *J Nucl Med* **33:** 1881-1887 (1992).
27. Beyer T., Kinahan, P. E., Townsend, D. W. *et al*., "The use of x-ray CT for attenuation correction of PET data" IEEE Nuclear Science Symposium 1577 (1994).
28. Beyer T., Townsend, D. W., Brun, T. *et al*., A combined PET/CT scanner for clinical oncology. *J Nucl Med* **41:** 1369-1379 (2000).
29. Townsend D., Kinahan, P. and Beyer, T., Attenuation correction for a combined 3D PET/CT scanner. *Physica Medica* **12:** (Suppl 1) 43-48 (1996).
30. Townsend D. W., Beyer, T., Kinahan, P. E. *et al*., "The SMART scanner: a combined PET/CT tomograph for clinical oncology" IEEE Nuclear Science Symposium and Medical Imaging Conference, Vol. 2; *pp* 1170-1174 (1998).
31. Townsend D. W., A combined PET/CT scanner: the choices. *J Nucl Med* **42:** 533-534 (2001).

32. Townsend D. W. and Cherry, S. R., Combining anatomy and function: the path to true image fusion. *Eur Radiol* **11:** 1968-1974 (2001).
33. Townsend D. W. and Beyer, T., A combined PET/CT scanner: the path to true image fusion. *Brit J Radiol* **75:** S24-S30 (2002).
34. Townsend D. W., Beyer, T. and Blodgett, T. M., PET/CT scanners: a hardware approach to image fusion. *Semin Nucl Med* **33:** 193-204 (2003).
35. Townsend D. W., Carney, J. P., Yap, J. T. *et al*., PET/CT today and tomorrow. *J Nucl Med* **45:** Suppl 1 4S-14S (2004).
36. Vogel W. V., Oyen, W. J. G., Barentsz, J. O. *et al*., PET/CT: Panacea, redundancy, or something in between? *J Nucl Med* **45:** 15S-24S (2004).
37. Phelps M. E., PET: The merging of biology and imaging into molecular imaging *J Nucl Med* **41:** 661-681 (2000).
38. Wahl R. L., Why nearly all PET of abdominal and pelvic cancers will be performed as PET/CT. *J Nucl Med* **45:** 82S-95S (2004).
39. Goerres G. W., von Schulthess, G. K. and Steinert, H. C., Why most PET of lung and head-and-neck cancer will be PET/CT. *J Nucl Med* **45:** 66S-71S (2004).
40. Even-Sapir E., Keidar, Z., Sachs, J. *et al*., The new technology of combined transmission and emission tomography in evaluation of endocrine neoplasms. *J Nucl Med* **42:** 998-1004 (2001).
41. Israel O., Keidar, Z., Iosilevsky, G. *et al*., The fusion of anatomic and physiologic imaging in the management of patients with cancer. *Semin Nucl Med* **31:** 191-205 (2001).
42. Tang H. R., "A Combined X-ray CT-Scintillation Camera System for Measuring Radionuclide Uptake in Tumors," Ph.D Thesis, University of California, 1998.
43. Kinahan P. E., Townsend, D. W., Beyer, T. *et al*., Attenuation correction for a combined 3D PET/CT scanner. *Med Phys* **25:** 2046-2053 (1998).
44. Kinahan P. E., Hasegawa, B. H. and Beyer, T., X-ray-based attenuation correction for position tomography/computed tomography scanners. *Semin Nucl Med* **33:** 166-179 (2003).
45. Mayneord W. V., The radiology of the human body with radioactive isotopes. *Brit J Radiol* **25:** 517-525 (1952).
46. Mayneord W. V., Radiological research. *Brit J Radiol* **27:** 309-317 (1954).
47. Mayneord W. V., Evans, H. D. and Newberry, S. P., An instrument for the formation of visual images of ionizing radiations. *J Scient Instruments* **32:** 45-50 (1955).
48. Anger H. O., Whole-body scanner Mark II. *J Nucl Med* **7:** 331-332 (1966).
49. Anger H. O. and McRae, J., Transmission scintiphotography. *J Nucl Med* **9:** 267-269 (1968).
50. Cameron J. R. and Sorenson, J., Measurement of bone mineral *in vivo*: an improved method. *Science* **142:** 230-232 (1963).
51. Kuhl D. E., Hale, J. and Eaton, W. L., Transmission scanning: A useful adjunct to conventional emission scanning for accurately keying isotope deposition to radiographic anatomy. *Radiology* **87:** 278-284 (1966).
52. Kuhl D. E., Edwards, R. O., Ricci, A. R. *et al*., The Mark IV system for radionuclide computed tomography of the brain. *Radiology* **121:** 405-413 (1976).
53. Budinger T. F. and Gullberg, G. T., Three-dimensional reconstruction in nuclear medicine emission imaging. *IEEE Trans Nucl Sci* **NS-21:** 2-20 (1974).
54. Mirshanov D. M., Transmission-Emission Computer Tomograph, USSR Patent 621.386:616-073 20.01.87-SU-181935 (January 20, 1987).

55. Kaplan C. H., Transmission/emission registered image (TERI) computed tomography scanners, International patent application No. PCT/US90/03722 (1989).
56. Hasegawa B. H., Tang, H. R., Iwata, K. *et al*., Correlation of structure and function with dual modality imaging. *Physica Medica* **17:** 241-248 (2001).
57. Hasegawa B. H., Reilly, S. M., Gingold, E. L. *et al*., Design considerations for a simultaneous emission-transmission CT scanner. [abstract] *Radiology* **173(P):** 414 (1989).
58. Lang T. F., Hasegawa, B. H., Liew, S. C. *et al*., ''A prototype emission-transmission imaging system'' IEEE Nuclear Science Symposium and Medical Imaging Conference, Santa Fe, NM, Vol. 3; *pp* 1902-1906 (1991).
59. Kalki K., Blankespoor, S. C., Brown, J. K. *et al*., Myocardial perfusion imaging with a combined x-ray CT and SPECT system. *J Nucl Med* **38:** 1535-1540 (1997).
60. Kalki K., Heanue, J. A., Blankespoor, S. C. *et al*., A combined SPECT and CT medical imaging system. *Proc SPIE* Vol. 2432; *pp* 367-375 (1995).
61. Cohade C., Osman, M. M., Leal, J. *et al*., Direct comparison of F-18-FDG PET and PET/CT in patients with colorectal carcinoma. *J Nucl Med* **44:** 1797-1803 (2003).
62. Fricke H., Fricke, E., Weise, R. *et al*., A method to remove artifacts in attenuation-corrected myocardial perfusion SPECT Introduced by misalignment between emission scan and CT-derived attenuation maps. *J Nucl Med* **45:** 1619-1625 (2004).
63. Takahashi Y., Murase, K., Higashino, H. *et al*., Attenuation correction of myocardial SPECT images with X-ray CT: effects of registration errors between X-ray CT and SPECT. *Ann Nucl Med* **16:** 431-435 (2002).
64. Blankespoor S. C., ''Attenuation correction of SPECT using x-ray CT for myocardial perfusion assessment,'' Master's thesis, University of California, 1996.
65. Koral K. F., Dewaraja, Y., Li, J. *et al*., Update on hybrid conjugate-view SPECT tumor dosimetry and response in ^{131}I-tositumomab therapy of previously untreated lymphoma patients. *J Nucl Med* **44:** 457-464 (2003).
66. Keidar Z., Israel, O. and Krausz, Y., SPECT/CT in tumor imaging? technical aspects and clinical applications. *Semin Nucl Med* **33:** 205-218 (2003).
67. Even-Sapir E., Lerman, H., Lievshitz, G. *et al*., Lymphoscintigraphy for sentinel node mapping using a hybrid SPECT/CT system. *J Nucl Med* **44:** 1413-1420 (2003).
68. Israel O., Mor, M., Gaitini, D. *et al*., Combined functional and structural evaluation of cancer patients with a hybrid camera-based PET/CT system using F-18 FDG. *J Nucl Med* **43:** 1129-1136 (2002).
69. Tang H. R., Da Silva, A. J., Matthay, K. K. *et al*., Neuroblastoma imaging using a combined CT scanner-scintillation camera and I-131 MIBG. *J Nucl Med* **42:** 237-247 (2001).
70. Bocher M., Balan, A., Krausz, Y. *et al*., Gamma camera-mounted anatomical x-ray tomography: technology, system characteristics, and first images. *Eur J Nucl Med* **27:** 619-627 (2000).
71. Liu Y., Wackers, F. J., Natale, D. *et al*., Validation of a hybrid SPECT/CT system with attenuation correction: A phantom study and multicenter trial. [abstract] *J Nucl Med* **44:** 209P (2003).
72. Rajasekar D., Datta, N. R., Gupta, R. K. *et al*., Multimodality image fusion in dose escalation studies in brain tumors. *J Appl Clin Med Phys* **4:** 8-16 (2003).

73. Schoder H., Larson, S. M. and Yeung, H. W. D., PET/CT in oncology: Integration into clinical management of lymphoma, melanoma, and gastrointestinal malignancies. *J Nucl Med* **45:** 72S-81S (2004).
74. Cook G. J. R. and Ott, R. J., Dual-modality imaging. *Eur Radiol* **11:** 1857-1858 (2001).
75. Chang L. T., A method for attenuation correction in radionuclide computed tomography. *IEEE Trans Nucl Sci* **NS-25:** 638-643 (1978).
76. Harris C. C., Greer, K. L., Jaszczak, R. J. *et al*., Tc-99m attenuation coefficients in water-filled phantoms determined with gamma cameras. *Med Phys* **11:** 681-685 (1985).
77. Malko J. A., Van Heertum, R. L., Gullberg, G. T. *et al*., SPECT liver imaging using an iterative attenuation correction algorithm and an external flood source. *J Nucl Med* **27:** 701-705 (1986).
78. Thompson C. J., Ranger, N. T. and Evans, A. C., Simultaneous transmission and emission scans in positron emission tomography. *IEEE Trans Nucl Sci* **36:** (1989).
79. Tsui B. M. W., Hu, H. B., Gilland, D. R. *et al*., Implementation of simultaneous attenuation and detector response correction in SPECT *IEEE Trans Nucl Sci* **NS-35:** 778-783 (1988).
80. Tsui B. M. W., Gullberg, G. T., Edgerton, E. R. *et al*., Correction of nonuniform attenuation in cardiac SPECT imaging *J Nucl Med* **30:** 497-507 (1989).
81. Blankespoor S., Wu, X., Kalki, K. *et al*., ''Attenuation correction of SPECT using x-ray CT on an emission-transmission CT system'' IEEE Nuclear Science Symposium and Medical Imaging Conference, Vol. 2; *pp* 1126-1130 (1995).
82. LaCroix K. J., ''An investigation of the use of single-energy x-ray CT images for attenuation compensation in SPECT,'' Master's thesis, University of North Carolina at Chapel Hill, 1994.
83. Perring S., Summers, Q., Fleming, J. *et al*., A new method of quantification of the pulmonary regional distribution of aerosols using combined CT and SPECT and its application to nedocromil sodium administered by metered dose inhaler. *Br J Radiol* **67:** 46-53 (1994).
84. Gullberg G. T., Huesman, R. H., Malko, J. A. *et al*., An attenuated projector-backprojector for iterative SPECT reconstruction. *Phys Med Biol* **30:** 799-816 (1985).
85. Ficaro E. P. and Wackers, F. J., Should SPET attenuation correction be more widely employed in routine clinical practice? *Eur J Nucl Med* **29:** 409-412 (2002).
86. Antoch G., Freudenberg, L. S., Beyer, T. *et al*., To enhance or not to enhance: [F-18]-FDG and CT contrast agents in dual-modality [F-18]-FDG PET/CT. *J Nucl Med* **45:** 56S-65S (2004).
87. Carney J. P., Beyer, T., Brasse, D. *et al*., Clinical PET/CT using oral CT contrast agents. [abstract] *J Nucl Med* **44:** 272P (2002).
88. Mukai T., Links, J. M., Douglass, K. H. *et al*., Scatter correction in SPECT using non-uniform attenuation data. *Phys Med Biol* **33:** 1129-1140 (1988).
89. Kessler R. M., Ellis, J. R. and Eden, M., Analysis of emission tomographic scan data: Limitations imposed by resolution and background. *J Comput Assist Tomogr* **8:** 514-522 (1984).
90. Tang H. R., Brown, J. K. and Hasegawa, B. H., Use of x-ray CT-defined regions of interest for the determination of SPECT recovery coefficients. *IEEE Trans Nucl Sci* **44:** 1594-1599 (1997).

91. Koole M., Laere, K. V., de Walle, R. V. *et al*., MRI guided segmentation and quantification of SPECT images of the basal ganglia: a phantom study. *Comput Med Imaging Graph* **25:** 165-172 (2001).
92. Rousset O. G., Ma, Y. and Evans, A. C., Correction for partial volume effects in PET: principle and validation. *J Nucl Med* **39:** 904-911 (1998).
93. Meltzer C. C., Kinahan, P. E., Greer, P. J. *et al*., Comparative evaluation of MR-based partial-volume correction schemes for PET. *J Nucl Med* **40:** 2053-2065 (1999).
94. Koral K. F., Li, J., Dewaraja, Y. *et al*., I-131 anti-B1 therapy/tracer uptake ratio using a new procedure for fusion of tracer images to computed tomography images. *Clin Cancer Res* **5:** 3004s-3009s (1999).
95. Koral K. F., Dewaraja, Y., Clarke, L. A. *et al*., Tumor-absorbed-dose estimates versus response in tositumomab therapy of previously untreated patients with follicular non-Hodgkin's lymphoma: preliminary report. *Cancer Biother Radiopharm* **15:** 347-355 (2000).
96. Formiconi A. R., Least squares algorithm for region-of-interest evaluation in emission tomography. *IEEE Trans Med Imag* **12:** 90-100 (1993).
97. Liu A., Williams, L. E. and Raubitschek, A. A., A CT assisted method for absolute quantitation of internal radioactivity. *Med Phys* **23:** 1919-1228 (1996).
98. Carson R. E., A maximum likelihood method for region-of-interest evaluation in emission tomography. *J Comput Assist Tomogr* **100:** 654-663 (1986).
99. Koral K. F., Zasadny, K. R., Kessler, M. L. *et al*., CT-SPECT fusion plus conjugate views for determining dosimetry in iodine-131-monoclonal antibody therapy of lymphoma patients. *J Nucl Med* **35:** 1714-1720 (1994).
100. Koral K. F., Lin, S., Fessler, J. A. *et al*., Preliminary results from intensity-based CT-SPECT fusion in I-131 anti-B1 monoclonal-antibody therapy of lymphoma. *Cancer* **80:** 2538-2544 (1997).
101. Koral K. F., Dewaraja, Y., Li, J. *et al*., Initial results for hybrid SPECT - conjugate-view tumor dosimetry in ^{131}I-anti-B1 antibody therapy of previously untreated patients with lymphoma. *J Nucl Med* **41:** 1579-1586 (2000).
102. Koral K. F., Francis, I. R., Kroll, S. *et al*., Volume reduction versus radiation dose for tumors in previously untreated lymphoma patients who received iodine-131 tositumomab therapy. Conjugate views compared with a hybrid method. *Cancer* **94:** 1258-1263 (2002).
103. Koral K. F., Kaminski, M. S. and Wahl, R. L., Correlation of tumor radiation-absorbed dose with response is easier to find in previously untreated patients. *J Nucl Med* **44:** 1541-1543 (2003).
104. Koral K. F., Zaidi, H. and Ljungberg, M., ''Medical imaging techniques for radiation dosimetry'' in: *Therapeutic applications of Monte Carlo calculations in nuclear medicine*, edited by H Zaidi and G Sgouros Institute of Physics Publishing, Bristol, (2002), *pp* 55-83.
105. Patton J. A., Delbeke, D. and Sandler, M. P., Image fusion using an integrated, dual-head coincidence camera with x-ray tube-based attenuation maps. *J Nucl Med* **41:** 1364-1368 (2000).
106. Ratib O., PET/CT image navigation and communication. *J Nucl Med* **45:** 46S-55S (2004).
107. Steinert H. C. and von Schulthess, G. K., Initial clinical experience using a new integrated in-line PET/CT system. *Br J Radiol* **73:** S36-S38 (2002).

108. Beyer T., Antoch, G., Muller, S. *et al*., Acquisition protocol considerations for combined PET/CT imaging. *J Nucl Med* **45:** 25S-35S (2004).
109. Bar-Shalom R., Yefemov, N., Guralnik, L. *et al*., Clinical performance of PET/CT in evaluation of cancer: Additional value for diagnostic imaging and patient management. *J Nucl Med* **44:** 1200-1209 (2003).
110. Cohade C. and Wahl, R. L., Applications of positron emission tomography/computed tomography image fusion in clinical positron emission tomography - clinical use, interpretation, methods, diagnostic improvements. *Semin Nucl Med* **33:** 228-237 (2003).
111. Cherry S. R., In vivo molecular and genomic imaging: new challenges for imaging physics. *Phys Med Biol* **49:** R13-R48 (2004).
112. Weissleder R. and Mahmood, U., Molecular imaging. *Radiology* **219:** 316-333 (2001).
113. Hanahan D., Transgenic mice as probes into complex systems. *Science* **246:** 1265-1275 (1989).
114. Sigmund C. D., Major approaches for generating and analyzing transgenic mice. *Hypertension* **22:** 599-607 (1993).
115. Wight D. C. and Wagner, T. E., Transgenic mice: a decade of progress in technology and research. *Mutation Res* **307:** 429-440 (1994).
116. James J. F., Hewett, T. E. and Robbins, J., Cardiac physiology in transgenic mice. *Circ Res* **82**: 407-415 (1998).
117. Wu M. C., Gao, D. W., Sievers, R. E. *et al*., Pinhole single-photon emission computed tomography for myocardial perfusion imaging of mice. *J Am Coll Cardiol* **42:** 576-582 (2003).
118. Paulus M. J., Sari-Sarraf, H., Bleason, S. S. *et al*., A new x-ray computed tomography system for laboratory mouse imaging. *IEEE Trans Nucl Sci* **46:** 558-564 (1999).
119. Cherry S. R., Shao, Y., Silverman, R. W. *et al*., MicroPET: a high resolution PET scanner for imaging small animals. *IEEE Trans Nucl Sci* **44:** 1161-1166 (1997).
120. Hasegawa B. H., Wu, M. C., Iwata, K. *et al*., Applications of penetrating radiation for small animal imaging. *Proc SPIE* Vol. 4786; *pp* 80-90 (2002).
121. Paulus M. J., Gleason, S. S., Kennel, S. J. *et al*., High resolution X-ray computed tomography: an emerging tool for small animal cancer research. *Neoplasia* **2:** 62-70 (2000).
122. Kennel S. J., Davis, I. A., Branning, J. *et al*., High resolution computed tomography and MRI for monitoring lung tumor growth in mice undergoing radioimmunotherapy: correlation with histology. *Med Phys* **27:** 1101-1107 (2000).
123. Paulus M. J., Gleason, S. S., Easterly, M. E. *et al*., A review of high-resolution X-ray computed tomography and other imaging modalities for small animal research. *Lab Anim (NY)* **30:** 36-45 (2001).
124. Ishizu K., Mukai, T., Yonekura, Y. *et al*., Ultra-high resolution SPECT system using four pinhole collimators for small animal studies. *J Nucl Med* **36:** 2282-2286 (1995).
125. Jaszczak R. J., Li, J., Wang, H. *et al*., Pinhole collimation for ultra high-resolution, small field of view SPECT. *Phys Med Biol* **39:** 425-437 (1994).
126. MacDonald L. R., Patt, B. E., Iwanczyk, J. S. *et al*., Pinhole SPECT of mice using the LumaGEM gamma camera. *IEEE Trans Nucl Sci* **48:** 830-836 (2001).

127. Ogawa K., Kawade, T., Nakamura, K. *et al*., Ultra high resolution pinhole SPECT for small animal study. *IEEE Trans Nucl Sci* **45:** 3122-3126 (1998).
128. Wu M. C., Hasegawa, B. H. and Dae, M. W., Performance evaluation of a pinhole SPECT system for myocardial perfusion imaging of mice. *Med Phys* **29:** 2830-2839 (2002).
129. Weber D. A. and Ivanovic, M., Pinhole SPECT: ultra-high resolution imaging for small animal studies. *J Nucl Med* **36:** 2287-2289 (1995).
130. Shao Y., Cherry, S. R., Farahani, K. *et al*., Simultaneous PET and MR imaging. *Phys Med Biol* **42:** 1965-1970 (1997).
131. Goertzen A. L., Meadors, A. K., Silverman, R. W. *et al*., Simultaneous molecular and anatomical imaging of the mouse *in vivo*. *Phys Med Biol* **21:** 4315-4328 (2002).
132. Weisenberger A. G., Wojcik, R., Bradley, E. L. *et al*., SPECT-CT system for small animal imaging. *IEEE Trans Nucl Sci* **50:** 74-79 (2003).
133. Williams M. B., Zhang, G., More, M. J. *et al*., Integrated CT-SPECT system for small animal imaging. *Proc SPIE* Vol. 4142; *pp* 265-274 (2000).
134. Welsh R. E., Brewer, P., Bradley, E. L. *et al*., "An economical dual-modality small animal imaging system with application to studies of diabetes" IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 3; *pp* 1845-1848 (2002).
135. Kastis G. A., Furenlid, L. R., Wilson, D. W. *et al*., Compact CT/SPECT small-animal imaging system. *IEEE Trans Nucl Sci* **51:** 63-67 (2004).
136. Iwata K., Wu, M. C. and Hasegawa, B. H., "Design of combined x-ray CT and SPECT systems for small animals" IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 3; *pp* 1608-1612 (1999).
137. Iwata K., Hwang, A. B., Wu, M. C. *et al*., "Design and utility of a small animal CT/SPECT system" IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 3; *pp* 1849-1852 (2002).
138. MacDonald L. R., Iwata, K., Patt, B. E. *et al*., Evaluation of x-ray detectors for dual-modality CT-SPECT animal imaging. *Proc SPIE* Vol. 4786; *pp* 91-102 (2002).
139. Andre M. P., Spivey, B. A., Martin, P. J. *et al*., Integrated CMOS-selenium x-ray detector for digital mammography. *Proc SPIE* Vol. 3336; *pp* 204-209 (1998).
140. Song X., Frey, E. C. and Tsui, B. M. W., "Development and evaluation of a microCT system for small animal imaging" IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 3; *pp* 1600-1604 (2002).
141. Hwang A. B., Iwata, K., Sakdinawat, A. E. *et al*., "Gantry specifications for a dual modality imaging system for small animals" IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 2; *pp* 1303-1307 (2003).
142. Varady P., Li, J. Z., Alden, T. D. *et al*., CT and radionuclide study of BMP-2 gene therapy-induced bone formation. *Acad Radiol* **9:** 632-637 (2002).
143. MacDonald L. R., Iwanczyk, J. S., Patt, B. E. *et al*., "Development of new high resolution detectors for small animal SPECT imaging" IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 3; *pp* 21/75 (2002).
144. McElroy D. P., MacDonald, L. R., Beekman, F. J. *et al*., Performance evaluation of A-SPECT: A high resolution desktop pinhole SPECT system for imaging small animals. *IEEE Trans Nucl Sci* **49:** 2139-2147 (2002).

145. Yazaki P. J., Shively, L., Clark, C. *et al*., Mammalian expression and hollow fiber bioreactor production of recombinant anti-CEA diabody and minibody for clinical applications. *J Immun Methods* **253:** 195-208 (2001).
146. Hasegawa B. H., Barber, W. C., Funk, T. *et al*., Implementation and applications of dual-modality imaging. *Nucl Instr Meth Phys Res A* **525:** 236-241 (2004).
147. Weber D. A. and Ivanovic, M., Ultra-high-resolution imaging of small animals: implications for preclinical and research studies. *J Nucl Cardiol* **6:** 332-344 (1999).
148. Paulino A. C., Thorstad, W. L. and Fox, T., Role of fusion in radiotherapy treatment planning *Semin Nucl Med* **33:** 238-243 (2003).
149. Munley M. T., Marks, L. B., Scarfone, C. *et al*., Multimodality nuclear medicine imaging in three-dimensional radiation treatment planning for lung cancer: challenges and prospects. *Lung Cancer* **23:** 105-14 (1999).
150. Osman M. M., Cohade, C., Nakamoto, Y. *et al*., Clinically significant inaccurate localization of lesions with PET/CT: frequency in 300 patients. *J Nucl Med* **44:** 240-243 (2003).
151. Nahmias C., Nutt, R., Hichwa, R. D. *et al*., PET tomograph designed for five minute routine whole body studies. [abstract] *J Nucl Med* **43:** 11P (2002).
152. Nelson S. J., Vigneron, D. B. and Dillon, W. P., Serial evaluation of patients with brain tumors using volume MRI and 3D ^{1}H MRSI. *NMR Biomed* **12:** 123-138 (1999).
153. Kurhanewicz J., Swanson, M. G., Nelson, S. J. *et al*., Combined magnetic resonance imaging and spectroscopic imaging approach to molecular imaging of prostate cancer. *J Magn Reson Imaging* **16:** 451-463 (2002).
154. Boxt L. M., Lipton, M. J., Kwong, R. Y. *et al*., Computed tomography for assessment of cardiac chambers, valves, myocardium and pericardium. *Cardiol Clin* **21:** 561-585 (2003).
155. Chan F. P., Cardiac multidetector-row computed tomography: principles and applications. *Semin Roentgenol* **38:** 294-302 (2003).
156. Nieman K., Rensing, B., Munne, A. *et al*., Three-dimensional coronary anatomy in contrast-enhanced multislice computed tomography. *Prev Cardiol* **5:** 79-83 (2002).
157. Foley W. D. and Karcaaltincaba, M., Computed tomography angiography: Principles and clinical applications. *J Comput Assist Tomogr* **27:** S23-S30 (2003).
158. Peterson T. F., Kim, H., Crawford, M. J. *et al*., "SemiSPECT: A small-animal imaging system based on eight CdZnTe pixel detectors" IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 3; *pp* 1844-1847 (2003).
159. Liu Z., Kastis, G. A., Stevenson, G. D. *et al*., Quantitative analysis of acute myocardial infarct in rat hearts with ischemia-reperfusion using a high-resolution stationary SPECT system. *J Nucl Med* **43:** 933-939 (2002).
160. Schramm N. U., Ebel, G., Engeland, U. *et al*., High-resolution SPECT using multipinhole collimation. *IEEE Trans Nucl Sci* **50:** 315-320 (2003).
161. Shao Y., Cherry, S. R., Farahani, K. *et al*., Development of a PET detector system compatible with MRI/NMR systems. *IEEE Trans Nucl Sci* **44:** 1167-1171 (1997).
162. Slates R., Cherry, S. R., Boutefnouchet, A. *et al*., Design of a small animal MR compatible PET scanner. *IEEE Trans Nucl Sci* **46:** 565-570 (1999).

163. Christensen N. L., Hammer, B. E., Heil, B. G. *et al*., Positron emission tomography within a magnetic field using photomultiplier tubes and light-guides. *Phys Med Biol* **40:** 691-697 (1995).
164. Hammer B. E., Christensen, N. L. and Heil, B. G., Use of a magnetic field to increase the spatial resolution of positron emission tomography. *Med Phys* **21:** 1917-1920 (1994).
165. Mackewn J. E., Strul, D., Hallett, W. A. *et al*., "Design and development of an MR-compatible PET scanner for imaging small animals" IEEE Nuclear Science Symposium and Medical Imaging Conference, Rome, Italy, (2004) *in press*
166. Gaa J., Rummeny, E. J. and Seemann, M. D., Whole-body imaging with PET/MRI. *Eur J Med Res* **30:** 309-312 (2004).
167. Maronpot R. R., Sills, R. C. and Johnson, G. A., Applications of magnetic resonance microscopy. *Toxicol Pathol* **32:** 42-48 (2004).
168. Johnson G. A., Cofer, G. P., Gewalt, S. L. *et al*., Morphologic phenotyping with MR microscopy: the visible mouse. *Radiology* **222:** 789-793 (2002).
169. Hu X. and Norris, D. G., Advances in high-field magnetic resonance imaging. *Annu Rev Biomed Eng* **6:** 157-184 (2004).
170. Lima J. A. and Desai, M. Y., Cardiovascular magnetic resonance imaging: current and emerging applications. *J Am Coll Cardiol* **15:** 1164-1171 (2004).
171. Hasumi M., Suzuki, K., Taketomi, A. *et al*., The combination of multi-voxel MR spectroscopy with MR imaging improve the diagnostic accuracy for localization of prostate cancer. *Anticancer Res* **23:** 4223-4227 (2003).
172. Yeung D. K., Yang, W. T. and Tse, G. M., Breast cancer: in vivo proton MR spectroscopy in the characterization of histopathologic subtypes and preliminary observations in axillary node metastases. *Radiology* **225:** 190-197 (2002).
173. Turner D. A., MR spectroscopy and imaging of breast cancer. A brief history. *Magn Reson Imaging Clin N Am* **2:** 505-510 (1994).
174. Coakley F. V., Teh, H. S., Qayyum, A. *et al*., Endorectal MR imaging and MR spectroscopic imaging for locally recurrent prostate cancer after external beam radiation therapy: preliminary experience. *Radiology* **233:** 441-448 (2004).
175. Coakley F. V., Qayyum, A. and Kurhanewicz, J., Magnetic resonance imaging and spectroscopic imaging of prostate cancer. *J Urol* **170:** S69-75; discussion S75-76 (2003).
176. Swanson M. G., Vigneron, D. B., Tran, T. K. *et al*., Magnetic resonance imaging and spectroscopic imaging of prostate cancer. *Cancer Invest* **19:** 510-523 (2001).
177. Vick G. W., Recent advances in pediatric cardiovascular MRI. *Curr Opin Pediatr* **15:** 454-462 (2003).
178. Zaidi H., Montandon, M.-L. and Slosman, D. O., Magnetic resonance imaging-guided attenuation and scatter corrections in three-dimensional brain positron emission tomography. *Med Phys* **30:** 937-948 (2003).
179. Slates R., Farahani, K., Shao, Y. *et al*., A study of artefacts in simultaneous PET and MR imaging using a prototype MR comptatible PET scanner. *Phys Med Biol* **44:** 2015-2027 (1999).
180. Garlick P. B., Marsden, P. K., Cave, A. C. *et al*., PET and NMR dual acquisition (PANDA): Applications to isolated perfused rat hearts. *NMR Biomed* **10:** 138-142 (1997).
181. Garlick P. B., Simultaneous PET and NMR-Initial results from isolated, perfused rat hearts. *Br J Radiol* 75 Spec No: S60-66 (2002).

182. Marsden P. K., Strul, D., Keevil, S. F. *et al*., Simultaneous PET and NMR. *Br J Radiol* 75 Spec No: S53-59 (2002).
183. Pichler B., Lorenz, E., Mirzoyan, R. *et al*., ''Performance tests of a LSO-APD PET module in a 9.4 Tesla magnet'' IEEE Nuclear Science Symposium and Medical Imaging Conference Record, Vol. 2; *pp* 1237-1239 (1997).

3
Analytic Image Reconstruction Methods in Emission Computed Tomography

B.M.W. Tsui* and E.C. Frey*

1. Introduction

Emission computed tomography (ECT) is a medical imaging modality that combines conventional nuclear imaging techniques with methods for image reconstruction from projections.[1] Depending on the radionuclide used, ECT can be divided into two major categories, positron emission tomography (PET) and single- photon emission computed tomography (SPECT). In PET, radiopharmaceuticals labeled with radionuclides that emit positrons are used. The annihilation of a positron and an electron results in two 511 keV photons traveling in directions that are 180° apart. Coincidence detection of the pair of 511 keV photons forms a line of response (LoR). The multiple detectors and coincidence electronic circuitry collect LoRs in many orientations. The collection of LoRs is then resorted and rebinned to form the projection dataset used in image reconstruction to compute the tomographic images.[2,3] In SPECT, radiopharmaceuticals labeled with radionuclides that emit gamma-ray photons are used. They include many of the agents that are routinely used in nuclear medicine clinics. Depending on the design geometry, a collimator that is placed in front of a position sensitive detector, typically a scintillation camera, accepts photons traveling in certain directions. The most commonly used parallel-hole collimator accepts photons incident from directions perpendicular to the detector face. By rotating the collimator-detector around the patient, projection data from different views are collected. The projection data from the different views are then used in image reconstruction.[4-6]

Emission computed tomography techniques were conceived early in the development of nuclear medicine imaging. A notable development was the MARK IV brain scanner by Kuhl and Edwards in the early 1960's.[7] The system consisted of four linear arrays of eight collimated detectors arranged

* Drs. B.M.W. Tsui and E.C. Frey, Division of Medical Imaging Physics, The Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, Maryland 21287, USA

in a square configuration. By rotating the arrays of detector around the patient's head, projection data were collected. At that time, the investigators were unable to solve the problem of image reconstruction from projections and the simple backprojection method was used to reconstruct the image data. The results were unimpressive and the full development of ECT was delayed until the development of computed tomography (CT) methods.

The advancement of ECT was boosted by the landmark development of both x-ray CT and image reconstruction techniques resulting in the award of the Nobel Prize in medicine in 1979.[8,9] At the same time, continuing improvements in both PET and SPECT systems and data acquisition methods were made by academic groups and commercial companies. These activities culminated in the introduction of the first commercial SPECT system by GE Medical Systems in 1981 and PET system designed at EG&G ORTEC in collaboration with Phelps and Hoffman in 1978. Since gaining FDA approval in 1999 for cancer staging and further fueled by the recent development of PET/CT, there has been a surge in interest in the clinical applications of PET and PET/CT. Today, SPECT and SPECT/CT systems are the major equipment in any clinical nuclear medicine clinic. The number of installed PET and PET/CT systems continues to increase at a healthy pace.

In this chapter, we review the fundamentals of image reconstruction from projections. Specifically, we describe the analytical image reconstruction methods which were the impetus for the development of ECT and have continued to play a major role in clinical ECT applications. The development of analytical image reconstruction methods can be traced back to Bracewell's work in radioastronomy in the 1950s. Finally, we discuss the limitations of analytical image reconstruction methods which have spurred the development of a new class of statistical image reconstruction methods and compensation techniques which are gaining much attention today (see chapter 4).

2. The Problem of Image Reconstruction in ECT

The methods of *image reconstruction from projections* are also called methods of *tomographic reconstruction*. *Tomography* has a Greek origin where *tomos* means slice and *graphy* means image or picture.

As shown in Figure 1, the object in ECT is a three-dimensional (3D) distribution of radioactivity. External detectors are used to collect radiation emitted from the radioactivity distribution to form projection data from different views around the object. In PET, coincidence detection and electronic collimation are used in the projection data acquisition. In SPECT, collimator-detectors are used to acquire the projection data. Depending on the system geometry, projection data from the object can be acquired into one-dimensional (1D) projection arrays, e.g., as in earlier generation of PET and SPECT systems, or two-dimensional (2D) arrays, e.g., modern PET and SPECT systems. In the following, we consider the 3D object as composed of

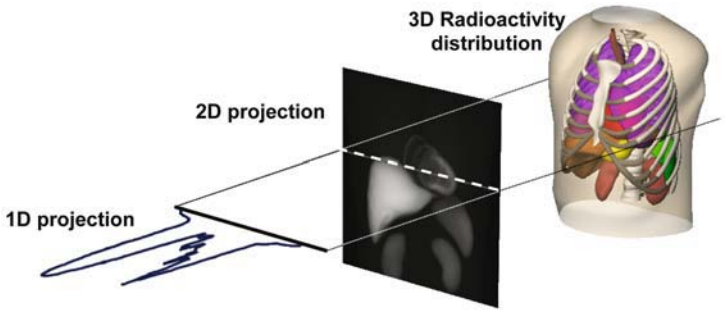


FIGURE 1. In emission computed tomography (ECT), the object is a 3D distribution of radioactivity. Projection data are obtained from different views around the object. Depending on the system geometry, 1D or 2D projection data are acquired. The image reconstruction problem in ECT is to seek estimate of the 3D radioactivity distribution from the multiple projections.

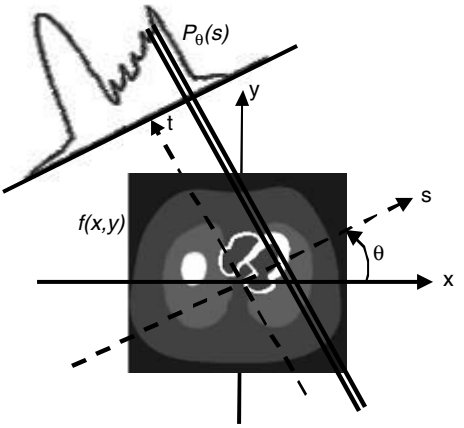


FIGURE 2. Geometry for the 2D image reconstruction problem. The 1D projection at angle θ,$P_\theta(s)$, is an array of integral of the object distribution $f(x,y)$ along the t-direction which is perpendicular to the detector face.

multiple slices of 2D distributions. The image reconstruction problem is thus reduced to recovering a 2D object distribution from a set of one-dimensional projections.

Figure 2 shows the geometrical configuration of the 2D image reconstruction problem. The 2D object distribution is represented by $f(x,y)$. It is mathematically convenient to formulate the problem in terms of the rotated coordinate system (s,t). In this system, the origin is coincident with the origin in the unrotated (x,y) system and the (s,t) axes are perpendicular and parallel, respectively, to the projection array. Suppose the 1D detector array lies along the rotated s-axis which makes an angle θ with the fixed x-axis. In PET, assume the detector response is perfect, i.e., the system spatial resolution is infinitely high and there is no photon scatter. Then the

acquired data at each point on the 1D projection array, $p'_\theta(s)$, is composed of sum of counts from LoRs that are aligned with and intersect the 1D detector array at the same point, or

$$p'_\theta(s) = \exp\left[-\int_{D(s)} \mu(s,t)dt\right] \cdot \int_{D(s)} f(s,t)dt \tag{1}$$

where $f(s,t)$ is the radioactivity distribution, $\mu(s,t)$ is the attenuation coefficient distribution and $D(s)$ is the thickness of the object in the direction parallel to the t-axis and intersect the detector array at point s. Note that in Equation (1), the exponential term represents the attenuation factor for the LoR and the second integral represents the sum of the radioactivity distribution along the LoR. Equation (1) can be written as

$$p_\theta(s) = \frac{p'_\theta(s)}{\exp\left[-\int_{D(s)} \mu(s,t)dt\right]} = \int_{D(s)} f(s,t)dt \tag{2}$$

where the modified 1D projection data $p_\theta(s)$ is integration of $f(s,t)$ along the t-axis.

In SPECT, assume a parallel-hole collimator with infinitely high spatial resolution and no scatter photon counts, the acquired 1D projection data array is given by

$$p'_\theta(s) = \int_{D(s)} f(s,t) \exp\left[-\int_{d(s)} \mu(s,t')dt'\right] dt \tag{3}$$

where $D(s)$ is the thickness of the object in the direction parallel to the t-axis and intersect the detector array at point s, $d(s)$ is the distance between the point (s,t) and the edge of $f(s,t)$ along the direction parallel to the t-axis and towards the detector. Different from PET, the exponential term represents the attenuation factor from the activity at position (s, t) to the detector at the projection position s. Equation (3) represents a major difficulty in SPECT image reconstruction due to photon attenuation which is discussed in detail in chapter 6. Here, we assume there is no attenuation effect, i.e., $\mu(s,t) = 0$, and Eqn. (3) reduces to

$$p_\theta(s) = \int_{D(s)} f(s,t)dt \tag{4}$$

where the $p_\theta(s)$ represents integration of $f(s,t)$ along the t-axis.

From Figure 2, it can be seen that

$$\begin{aligned} s &= x\cos\theta + y\sin\theta \\ t &= -x\sin\theta + y\cos\theta \end{aligned} \tag{5}$$

Equations (2) and (4) which have a similar form can be rewritten using δ function as

$$p_\theta(s) = \int_{D(s)} f(s,t)dt = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - s)dxdy \tag{6}$$

Equation (6) defines the 2D *Radon transform* of $f(x,y)$. It specifies the 1D projection of $f(x,y)$ at the projection angle θ. The problem of image reconstruction is to seek the solution of 2D *inverse Radon transform*, i.e., to find the 2D object distribution $f(x,y)$ given the set of 1D projection data, $p_\theta(s)$, obtained at different projection views, θ.

In the following section, we introduce the *Fourier slice theorem* which relates the 1D projections to the object distribution. The relationship is important in the derivation of the analytical image reconstruction methods described in the subsequent sections.

3. The Fourier Slice Theorem

From Eqns. (2) and (4), the 1D Fourier transform (FT) of $p_\theta(s)$ can be written as

$$\begin{aligned} P_\theta(\nu_s) &= \int_{-\infty}^{\infty} p_\theta(s) e^{-j\nu_s s} ds = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(s,t) dt \right] e^{-j\nu_s s} ds \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) e^{-j\nu_s (x\cos\theta + y\sin\theta)} dx dy \end{aligned} \tag{7}$$

Since $\nu_x = \nu_s \cos\theta$ and $\nu_y = \nu_s \sin\theta$, Eqn. (7) becomes

$$P_\theta(\nu_s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) e^{-j(\nu_x x + \nu_y y)} dx dy \tag{8}$$

which is the FT of $f(x,y)$ along the s-axis. Equation (8) can also be written as

$$P_\theta(\nu_s) = F(\nu) = F(\nu_s \cos\theta, \nu_s \sin\theta) \tag{9}$$

where $\nu = \nu_s(\cos\theta_{\sin}\theta)$. Equation (9) is the *Fourier (projection) Slice Theorem*, a major theorem in image reconstruction from projections.

The *Fourier Slice theorem* in Eqn. (9) states that the 1D FT of the projection at angle θ is equal to the cross-section of the 2D FT of the object distribution at the same angle γ. If one computes the 1D FT of all the projections $P_\theta(\nu)$, the 2D FT of the object distribution $F(\nu_x, \nu_y)$ along the same directions as the projection angles will be known. This important property provides insight into the problem of image reconstruction from projections. The collection of 1D FT of projections $P_\theta(\nu)$ form 'spoke' like cross-sections of $F(\nu_x, \nu_y)$ centered at $(\nu_x, \nu_y) = (0,0)$. In order to perform the inverse 2D FT of $F(\nu_x, \nu_y)$ to obtain the object distribution, the 'spoke' like samples of the $F(\nu_x, \nu_y)$ need to be interpolated into a square matrix. Since 2D interpolation is time-consuming, the direct image reconstruction method is not practical and is rarely used. Alternative methods have been devised that provide a more efficient approach to obtain the object distribution $F(\nu_x, \nu_y)$.

In the following, we describe the analytical reconstruction methods that are in common use. We start with the simple backprojection method for

historical reasons and demonstrate that the simple backprojection image is a blurred version of the object distribution. We then describe the filtered backprojection and the backprojection of the filtered projections. Finally, we will discuss equivalent image reconstruction methods that operate completely in the spatial domain.

4. Analytical Image Reconstruction Methods

4.1 Simple Backprojection Method

The simple backprojection method simply takes each of the measured projection data values and spreads it uniformly over the image pixels that lie along the projection ray passing through the projection bin. This is the method used by Kuhl and Edwards in reconstructing the first brain SPECT data acquired using the MARK IV system.7 To understand why the reconstructed images were less than desirable, we can write the simple backprojection image in polar coordinates as

$$\hat{f}_B(r,\phi) = Bpj\{p_\theta(s)\} = \int_0^\pi p_\theta(r\cos(\phi-\theta))d\theta$$
$$= \int_0^\pi \int_{-\infty}^{\infty} P_\theta(\nu_r)e^{j\nu_r t}d\nu_r d\theta = \int_0^\pi \int_{-\infty}^{\infty} P_\theta(\nu_r)e^{j\nu_r(x\cos\theta+y\sin\theta)}d\nu_r d\theta \quad (10)$$

where Bpj is the backprojection operation and $P_\theta(\nu_r)$ is the 1D FT of $p_\theta(r)$ and we have used the fact that $t = x\cos\theta + y\sin\theta$.

Since the 2D inverse FT, FT_{2D}^{-1}, of $P_\theta(\nu_r)$ in polar coordinates is given by

$$p(x,y) = \int_0^\pi \int_{-\infty}^{\infty} |\nu_r| P(\nu_r,\theta)e^{j\nu_r(x\cos\theta+y\sin\theta)}d\nu_r d\theta = FT_{2D}^{-1}\{P(\nu_r,\theta)\} \quad (11)$$

Equation (10) can be written as

$$\hat{f}_B(x,y) = \int_0^\pi \int_{-\infty}^{\infty} |\nu_r| \frac{P_\theta(\nu_r)}{|\nu_r|} e^{j\nu_r(x\cos\theta+y\sin\theta)}d\nu_r d\theta$$
$$= Bpj\left\{ FT_{2D}^{-1}\left[\frac{P_\theta(\nu_r)}{|\nu_r|}\right]\right\} \quad (12)$$

Using the relationship $\nu_x = \nu_s\cos\theta$ and $\nu_y = \nu_s\sin\theta$ Eqn. (9), we find

$$\hat{f}_B(x,y) = Bpj\{p_\theta(s)\} = Bpj\left\{ FT_{2D}^{-1}\left[\frac{P_\theta(\nu_s)}{|\nu_s|}\right]\right\} = FT_{2D}^{-1}\left[\frac{F(\nu_x,\nu_y)}{|\nu_s|}\right] \quad (13)$$

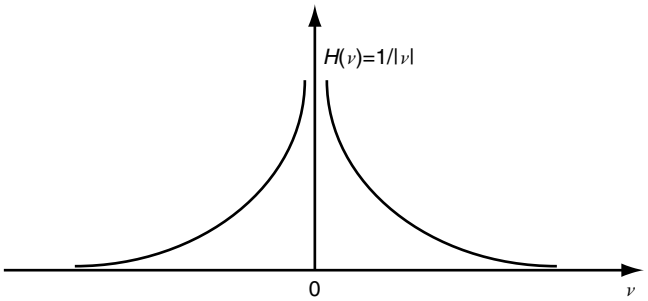


FIGURE 3. The blurring function that is applied to the object distribution in the frequency domain to give the reconstructed image obtained using the simple back-projection image reconstruction method.

Equation (13) shows that, the simple backprojection image, $\hat{f}_B(x,y)$, is the 2D inverse FT of $F(\nu_x, \nu_y)$, the FT of the object distribution, multiplied by the function $1/|\nu_s|$.

Figure 4a shows sample 2D slices through the computer generated 3D NCAT (NURB-based CArdiac Toro) phantom simulating the 3D Tc-99m labeled sestaMIBI distribution in a normal human. Figure 4b shows sample 2D projection data generated from the 3D radioactivity distribution from 128 equally spaced projection views over 180 degree around the phantom. The effects of collimator-detector response, photon attenuation and scatter were not included in the simulation.

The 2D projection images in Figure 4b can be considered as a collection of 1D projection arrays over all projection views with each corresponding to a

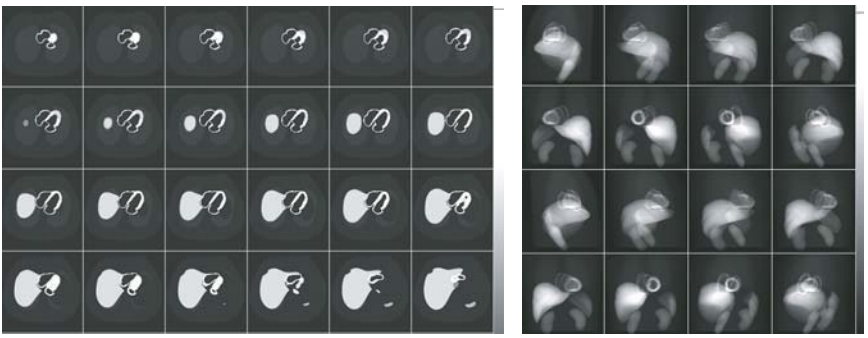


FIGURE 4. (a) Sample noise-free 2D slices through the 3D NCAT phantom simulating the 3D Tc-99m labeled SestaMIBI distribution in the normal human. (b) Sample 2D projection data simulated from the 3D radioactivity distribution shown in (a). The effects of system response and photon attenuation and scatter were not included in the simulation.

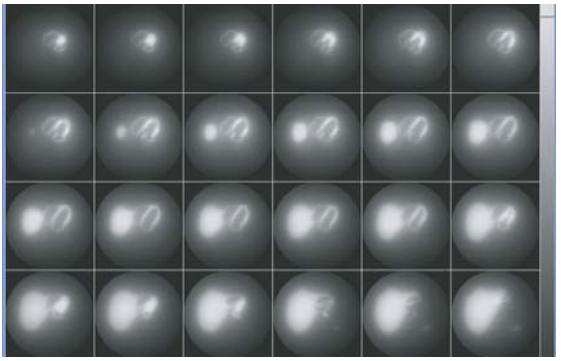


FIGURE 5. Reconstructed images obtained by applying the *simple backprojection* method to the 1D projection data sets shown in Figure 4b.

specific slice through the 3D phantom. In Figure 5, we show the 2D reconstructed images corresponding to the same slices through the 3D phantom shown in Figure 4a that are obtained by applying the *simple backprojection* method to the collections of 1D projection arrays. They demonstrate the blurring of the *simple backprojection* images as compared to the corresponding phantom slices shown in Figure 4b.

4.2 Filter of the Backprojection (FBpj)

Equation (13) shows that to obtain the 2D object distribution, we can simply multiply the 2D FT of $\hat{f}_B(x,y)$ by a *ramp function*, $|\nu|$, shown in Figure 6 and take the 2D inverse FT of the product, i.e.,

$$\begin{aligned}\hat{f}_{FB}(x,y) &= FT_{2D}^{-1}\{|\nu| \cdot FT_{2D}[\hat{f}_b(x,y)]\} = FT_{2D}^{-1}\left[|\nu| \cdot \frac{F(\nu_x,\nu_y)}{|\nu|}\right] \\ &= FT_{2D}^{-1}[F(\nu_x,\nu_y)] \end{aligned} \tag{14}$$

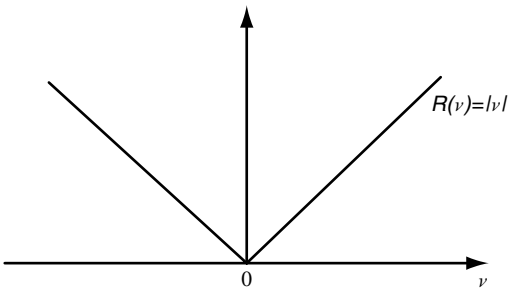


FIGURE 6. The ramp function that is applied to the FT of the 2D simple projection image to obtain the filter of the backprojection or FBpj image.

The multiplication of $\hat{f}_B(x,y)$ by a ramp function in the frequency domain is similar to applying a 'filter' function in a conventional image processing operation giving rise to the naming of the image reconstruction method as the *filter of the backprojection* (FBpj) method. However, it should be emphasized that the *ramp function* in Eqn. (14) is a result of the theory of image reconstruction and is needed to remove the blurring effect resulting from applying the simple backprojection operation. Its effect should not be regarded as a conventional high pass or edge enhancement filter that is applied directly to the projection data.

To demonstrate the effectiveness of the FBpj method, we use the same projection dataset generated from the radioactivity distribution of the 3D NCAT phantom shown in Figure 4a. Figure 7 shows the corresponding reconstructed images when the FBpj method is applied to the projection dataset. When compared to the phantom and the *simple backprojection* images shown in Figures 4 and 5, the effectiveness of the reconstruction method is clearly demonstrated.

The ramp function as shown in Figure 6 is unbounded, that is, its magnitude monotonically increases with ν. In practice, the object distribution in ECT is often assumed to be band-limited due to the band-limiting effects on the measured projection of the collimator-detector response and projection binning. When high frequency noise is present in the acquired projection data, the application of the ramp function amplifies the high frequency noise and gives rise to noisy reconstructed images. To reduce the reconstructed image noise, a noise smoothing filter is often applied to the projection data in addition to the ramp function as shown in Figure 8 and the 'average' smoothed FBpj image is given by

$$\begin{aligned}\hat{f}^F_{FB}(x,y) &= FT^{-1}_{2D}[W(\nu_x,\nu_y)\cdot F(\nu_x,\nu_y)] \\ &= FT^{-1}_{2D}\{W(\nu_x,\nu_y)\cdot|\nu|\cdot FT^{-1}_{2D}[\hat{f}_b(x,y)]\}\end{aligned} \tag{15}$$

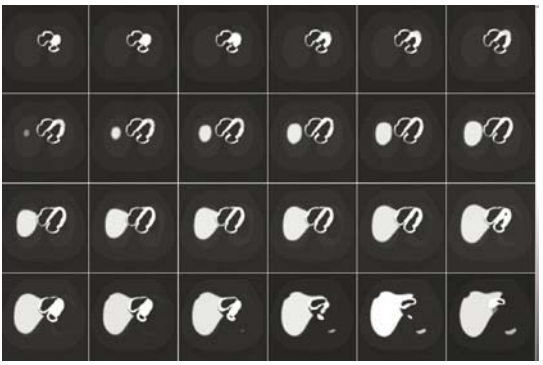


FIGURE 7. Reconstructed images obtained by applying the *filter of the backprojection* (FBpj) method to the 1D projection data sets shown in Figure 4b.

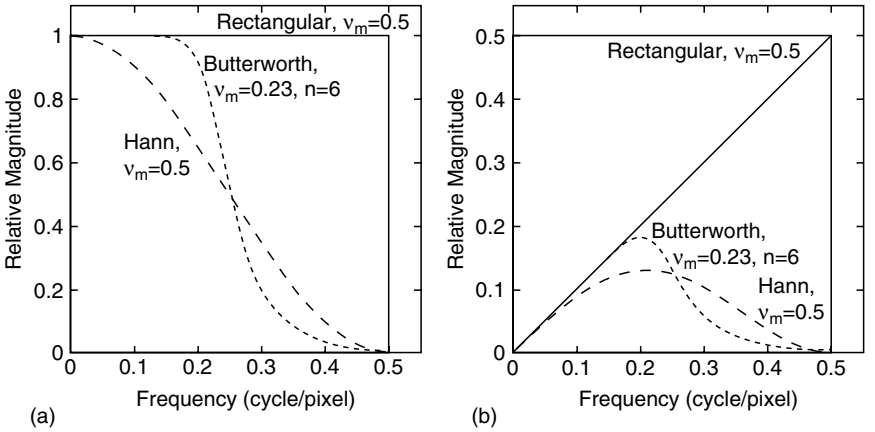


FIGURE 8. The typical filters used in ECT image reconstruction. (a) The rectangular, Han and Butterworth filters with cut-off frequency ν_m at 0.5, 0.5 and 0.23, respectively. The order n of the Butterworth filter is 6. (b) The products of the filters with the ramp function.

In Table 1 and Figure 8, several smoothing filters that are commonly used in ECT image reconstruction and their product with the ramp function are shown. The cut-off frequency ν_m of all the smoothing filters determines the amount of smoothing and loss of resolution in the reconstructed image, i.e., the lower the ν_m the more smoothing. The rectangular filter is the simplest and is the best in preserving high frequency information. However, it amplifies high frequency noise. The magnitude of the Hann filter starts to decrease from zero frequency earlier than the other filters and gives the most amount of smoothing and loss of resolution. The Butterworth filter has two parameters that determine its shape and has the most flexibility in tailoring the properties of the filter. In addition to the cut-off frequency ν_m, the order, n, determines how fast the filter rolls down from a value of unity at zero frequency towards zero at higher frequency. A smaller n value leads to a slower rolling down at higher frequency. At very large n values, the filter

TABLE 1. Commonly used smoothing filters in analytical image reconstruction methods in ECT.

	Filter Function Type		
	Rectangular	Hann	Butterworth
$W(\nu)$	$\begin{cases} 1 & \text{if } \lvert\nu\rvert \le \nu_m \\ 0 & \text{if } \lvert\nu\rvert > \nu_m \end{cases}$	$\begin{cases} 0.5 + 0.5\cos\left(\frac{\pi\nu}{\nu_m}\right) & \text{if } \lvert\nu\rvert \le \nu_m \\ 0 & \text{if } \lvert\nu\rvert > \nu_m \end{cases}$	$\frac{1}{\sqrt{1+(\frac{\nu}{\nu_m})^{2n}}}$
$W(\nu)\cdot\lvert\nu\rvert$	$\begin{cases} \lvert\nu\rvert & \text{if } \lvert\nu\rvert \le \nu_m \\ 0 & \text{if } \lvert\nu\rvert > \nu_m \end{cases}$	$\begin{cases} 0.5\lvert\nu\rvert + 0.5\lvert\nu\rvert\cos\left(\frac{\pi\nu}{\nu_m}\right) & \text{if } \lvert\nu\rvert \le \nu_m \\ 0 & \text{if } \lvert\nu\rvert > \nu_m \end{cases}$	$\frac{\lvert\nu\rvert}{\sqrt{1+(\frac{\nu}{\nu_m})^{2n}}}$

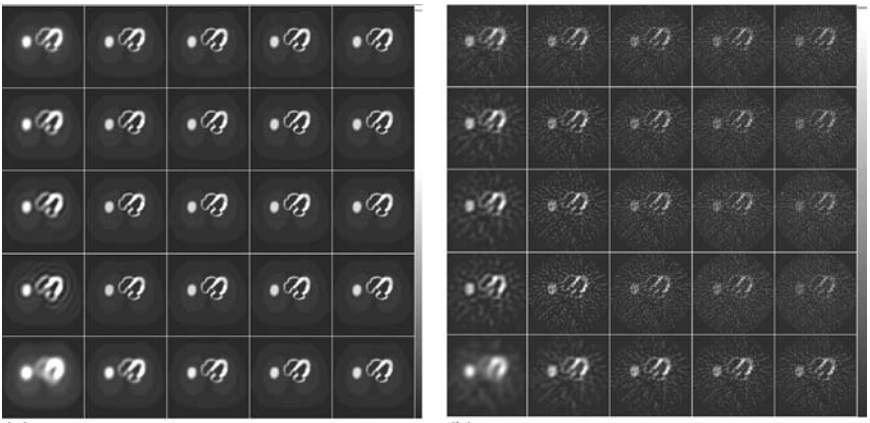


FIGURE 9. The backprojection of the filtered projections images obtained by applying different smoothing filter functions to images reconstructed from (a) noise-free and (b) noisy projection data. In both (a) and (b), the images in the top 4 rows were obtained using the Butherworth filter and, from top to bottom, with orders 2, 4, 8, and 32, respectively. The images in the bottom row were processed with the Hann filter. The images in the left to right column were obtained with cut-off frequencies, ν_m, of 0.1, 0.2, 0.3, 0.4 and 0.5 cycle/pixel, respectively.

shape has sharper corners which give rise to ringing artifacts in the reconstructed image.

Figure 9 shows the effectiveness of the Hann and the Butterworth filters in ECT image reconstruction as a function of the cut-off frequency and, for the Butterworth filter, the order n. Both noise-free and noisy projection data were used. For both smoothing filters, the reconstructed images are smoother as the cut-off frequency ν_m decreases. For the same ν_m, the Hann filter gives much smoother reconstructed images than the Butterworth filter. For the same ν_m, the Butterworth filter provides more preservation of low frequency features and high frequency noise smoothing when higher order n was used. However, ringing artifacts at edges of image structures can be found when the value of n was too high.

4.3 *Backprojection of Filtered Projections (BpjFP)*

Since the multiplication, the Fourier transform and the backprojection operations are linear and commutative, Eqn. (12) can be manipulated to provide an alternate means for image reconstruction. The procedure includes multiplying the FT of the 1D projections by the 1D ramp function in the frequency domain, and taking the 1D inverse transform of the products before backprojecting to form the 2D reconstructed image. The *backprojection of the filtered projections* (BpjFP) method can be represented by

$$\hat{f}_{BF}(x,y) = Bpj\{p_\theta^*(s)\} = Bpj\{FT^{-1}[|\nu_r| \cdot FT^1[p_\theta(s)]\}$$
$$= Bpj\{FT^{-1}[|\nu_s| \cdot P_\theta(\nu_s)]\} = FT^{-2}[\frac{|\nu_s| \cdot P_\theta(\nu_s)}{|\nu_s|}]$$
$$= FT^{-2}[F(\nu_x,\nu_y)] \tag{16}$$

where $p_\theta^*(s) = FT^{-1}[|\nu_r| \cdot FT^1[p_\theta(s)]$ are the *filtered projections*.

Although theoretically the BpjFP method is equivalent to the FBpj method, it offers several advantages. First, the multiplication of the ramp function and the FT operations are performed in 1D which is computationally more efficient as compared to performing similar operations in 2D. Also, since the operations can be performed on the projection data as they are acquired, the reconstructed image can be obtained almost as soon as the projection data are acquired. As a result, the BpjFP is the more popular analytical image reconstruction method and is often simply called the *filtered backprojection* (FBP) method.

For the BpjFB method, considerations of noise in the projection data and the application of an additional smoothing filter are similar to the FBpj method. Again, the difference is that the filtering operation is applied to the 1D projection data array instead of the 2D backprojection image.

4.4 Convolution Backprojection (CBpj)

In Eqn. (16), the BpjFP method involves the backprojection of the 1D inverse FT of the product of two functions, $|\nu_s|$ and $P_\theta(\nu_s)$. From the convolution theorem, this is equivalent to the backprojection of the convolution of two functions, the 1D inverse FT of $|\nu_s|$ and $p_\theta(s)$. Since the FT of $|\nu_s|$ is $-1/2\pi^2 r^2$, the filtered projection can be given by

$$p_\theta^*(s) = -\frac{1}{2\pi^2}\int_{-\infty}^{\infty}\frac{p_\theta(s')}{(s-s')^2}ds' \tag{17}$$

which contains a quadratic singularity. Using integration by part, Eqn. (17) becomes

$$p_\theta^*(s) = \frac{1}{2\pi^2}\int_{-\infty}^{\infty}\frac{\partial p_\theta(s')/\partial s'}{s-s'}ds' \tag{18}$$

Equation (18) shows that the *filtered projections* can be obtained by the convolution of the derivatives of $p_\theta(s)$ and the function *1/s*. The *convolution backprojection* (CBpj) method estimates the object distribution by the backprojection of $p_\theta^*(s)$, which is represented by the convolution given by Eqn. (17) or Eqn. (18).

The relationships between the object distribution and the projections found in the Fourier Slice theorem and Eqn. (18) were first derived by Radon in 1917. Equation (4) which is a transform relating the object distribution to the projections is also called the *Radon transform* and the solution of the image reconstruction problem is called the *inverse Radon transform* problem.

The implementation of the CBpj method involves handling the singularity at $s = s'$ in Eqn. (17) and approximating the derivatives of $p_\theta(s)$. The singularity of Eqn. (17) is caused by the FT of the ramp function $|\nu|$ which diverges at high frequencies. To avoid the divergence, we can replace $|\nu|$ by

$$C(\nu) = \begin{cases} |\nu|, & \text{for } |\nu| \leq \nu_m \\ 0, & \text{for } |\nu| \leq \nu_m \end{cases} \tag{19}$$

The use of $C(\nu)$ is justified since in practice $P_\theta(\nu)$ is usually band-limited. That is, $P_\theta(\nu)$ is non-zero at frequencies lower than some ν_m and zero at frequencies higher than ν_m. The inverse FT of $C(\nu)$ is given by

$$FT^{-1}[C(\nu)] = \frac{\nu_m}{\pi s} \sin(2\pi\nu_m s) - \frac{\sin^2(\pi\nu_m s)}{\pi^2 s^2} \tag{20}$$

which can be used in place of the inverse FT of $|\nu|$ as the convolution kernel in the CBpj method. Hence, the filtered projection can be written as

$$p^*_\theta(s) = \int_{-\infty}^{\infty} p_\theta(s') \left\{ \frac{\nu_m \sin[2\pi\nu_m(s-s')]}{\pi(s-s')} - \frac{\sin^2[\pi\nu_m(s-s')]}{\pi^2(s-s')^2} \right\} ds' \tag{21}$$

The first term on the right-hand side of Eqn. (21) is the convolution of $p_\theta(s)$ with a sinc function which has the effect of multiplying the FT of $p_\theta(s)$ with a rectangular function with a width ν_m. Assuming $P_\theta(\nu)$ is band-limited by the frequency ν_m, $p_\theta(s)$ is unchanged by the convolution. As a result, Eqn. (21) can be written as

$$p^*_\theta(s) = \nu_m p_\theta(s) - \int_{-\infty}^{\infty} p_\theta(s') \frac{\sin^2[\pi\nu_m(s-s')]}{\pi^2(s-s')^2} ds' \tag{22}$$

The digital implementation of the convolution kernel to form the filtered projection is given in ref.[10]

$$c(k) = \begin{cases} 1/4, & \text{if k} = 0, \\ -1/\pi^2 k^2, & \text{if k is odd}, \\ 0, & \text{if k is even}. \end{cases} \tag{23}$$

Other convolution kernels with some noise smoothing effect have also been derived[11] from the FT of the filter function $W(\nu) \cdot |\nu|$, where $W(\nu)$ is a smoothing filter such as the ones listed in Table 1.

In Table 2, we listed the major steps involved in the three major image reconstruction methods.

5. Analytical Image Reconstruction Methods for Special System Geometries

The analytical image reconstruction methods described in the previous sections assume the projection arrays are formed by the parallel beam geometry, i.e., the line integrals that form the projection are parallel to each other. In

TABLE 2. A comparison of the major steps involved in three analytical image reconstruction methods.

	Analytical Image Reconstruction Methods		
Steps	Filter of Backprojection (FBpj)	Backprojection of Filtered Projections (BpjFP)	Convolution Backprojection (CBpj)
1	Simple backprojection of projections	1D FT of projections	Convolve projections with convolution kernels
2	2D FT of backprojection image	Apply 1D ramp function and smoothing filter	Simple backprojection
3	Apply 2D ramp function and smoothing filter	Inverse 1D FT	
4	Inverse 2D FT	Simple backprojection	

conventional single-slice CT, the x-ray source and the 1D radiation detector form a 2D fan-beam geometry. In the recent multi-detector CT (MDCT) systems, the x-ray source and the 2D radiation detector form a 3D cone-beam geometry. In ECT, similar detection geometries can also be found. For example, in PET, LoRs from one detector element to multiple detector elements at the opposite side of the PET system form a 2D fan-beam geometry or 3D cone-beam geometry in single-slice or multi-slice systems, respectively. In SPECT, converging-hole collimators such as fan-beam[12,13] and cone-beam collimators[14-16] accept photons that travel in directions that form 2D fan-beam and 3D cone-beam geometries, respectively. They offer higher detection efficiency for the same spatial resolution as compared with parallel-hole collimator. The trade-off is reduced field-of-view for the same detector size.

Figure 10 shows the imaging configuration of x-ray CT where the focal point of the x-ray tube occupies the *focal point* of a fan that is completed by

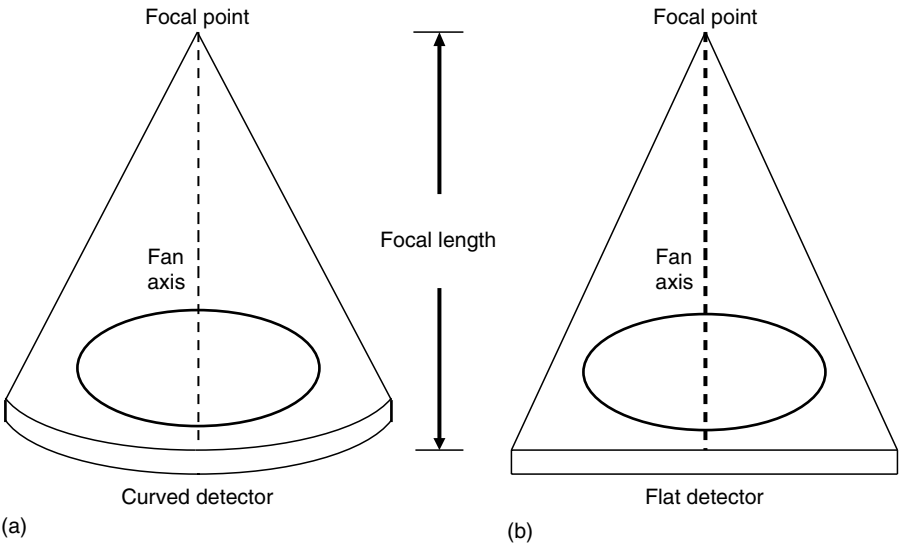


FIGURE 10. Fan-beam geometries with (a) curved detector and (b) flat detector.

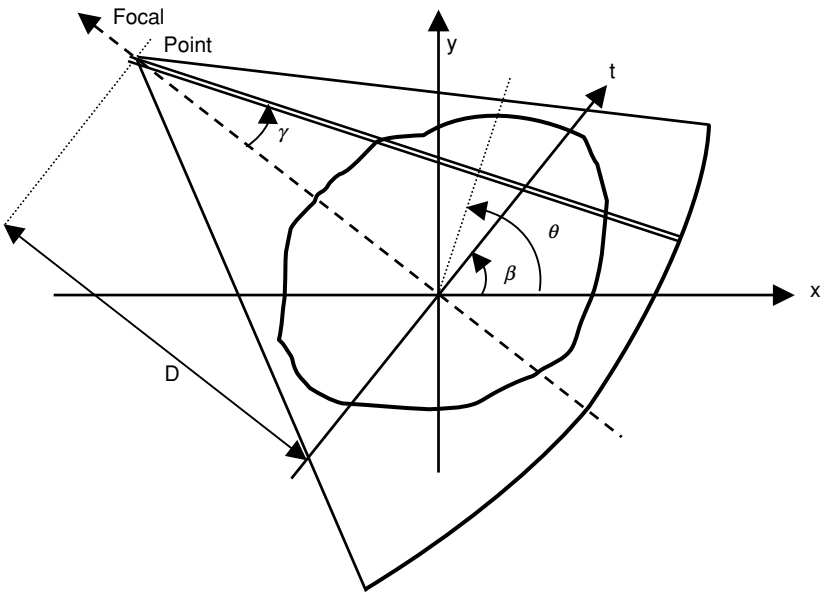


FIGURE 11. The fan-beam tomographic data acquisition geometry.

the detector array. Figure 11 shows the imaging configuration of SPECT using a fan-beam collimator. The axes of collimator holes that lie in a transverse plane perpendicular to the axis-of-rotation (AOR) of a fan-beam collimator converge to a focal point. The distance between the focal point and the detector is called *the focal length* of the collimator.

An equiangular fan-beam geometry with curved detector is used in conventional x-ray CT. A projection sampling geometry with projection bins having equal spatial widths and a flat detector is found in SPECT using fan-beam collimators.

5.1 Analytic Fan-Beam Image Reconstruction Methods

5.1.1 Rebinning Method

If $\Delta\beta = \Delta\gamma = \alpha$, i.e., the projection angle increments are the same as the sampling angle increments, then $\beta = m\alpha$, $\gamma = n\alpha$ where m and n are integers, then

$$r^{c}_{m\alpha}(n\alpha) = p_{(m+n)\alpha}(D \sin n\alpha) \tag{24}$$

where D is distance between the center-of-rotation (COR) and the focal point, β is the projection angle of the fan, γ is the angle between the axis of the fan and the designated projection ray, and α is the angular increment of the projection ray of interest from the fan axis.

The rebinning technique is straightforward and was used initially in fan-beam tomography. However, the method is very time-consuming and results in loss of resolution due to interpolation. For medical applications, it is important to develop fan-beam reconstruction techniques that are more efficient. Fan-beam filtered backprojection method was developed to allow direct image reconstruction of the fan-beam data without the time-consuming rebinning step.[17,18]

5.1.2 Direct Fan-Beam Filtered Backprojection Method

As described earlier, the projection ray indicated in the Figure is equivalent to the parallel projection ray in $p_\theta(t)$ with $t = D\sin\gamma$ and $\theta = \beta + \gamma$. To reconstruct, we note from parallel-beam reconstruction that:

$$f(x,y) = \frac{1}{2}\int_0^{2\pi}\int_{-S}^{S} p_\theta(s)h(r\cos(\theta-\phi)-s)ds d\theta \tag{25}$$

where S is the value of s for which $p_\theta(s) = 0$ for all $|s| > S$ in all projections, $h(r)$ is the convolution kernel. Note that the integration over θ is from 0 to 2π and the projection is zero outside the interval $[-S, S]$. The above equation can be rewritten in terms of γ and β:

$$f(r,\phi) = \frac{1}{2}\int_{-\gamma}^{2\pi-\gamma}\int_{-\sin^{-1}S/D}^{\sin^{-1}S/D} p_{\beta+\gamma}(D\sin\gamma)h(r\cos(\beta+\gamma-\phi)-D\sin\gamma)(D\cos\gamma)\gamma d\gamma d\beta \tag{26}$$

From Figure 12, let L be the distance from a point (x,y) to the focal point, then

$$L\cos\gamma' = D + r\sin(\beta-\phi) \text{ and } L\sin\gamma' = r\cos(\beta-\phi) \tag{27}$$

$$\text{where } \gamma' = \tan^{-1}\left(\frac{L\sin\gamma'}{L\cos\gamma'}\right) = \tan^{-1}\left(\frac{r\cos(\beta-\phi)}{D+r\sin(\beta-\phi)}\right) \tag{28}$$

$$\text{Also,we find } L(r,\phi,\beta) = \sqrt{(D+r\sin[\beta-\phi])^2 + r^2\cos^2(\beta-\phi)} \tag{29}$$

Using the above equations, Eqn. (28) can be written as

$$f(r,\phi) = \frac{1}{2}\int_0^{2\pi}\int_{-\gamma_M}^{\gamma_M} r_\beta^c(\gamma)h(L\sin[\gamma'-\gamma])D\cos\gamma d\gamma d\beta \tag{30}$$

Recognizing $h(t)$ is the inverse FT of $|\nu|$ in the frequency domain, Eqn. (30) can be rewritten as

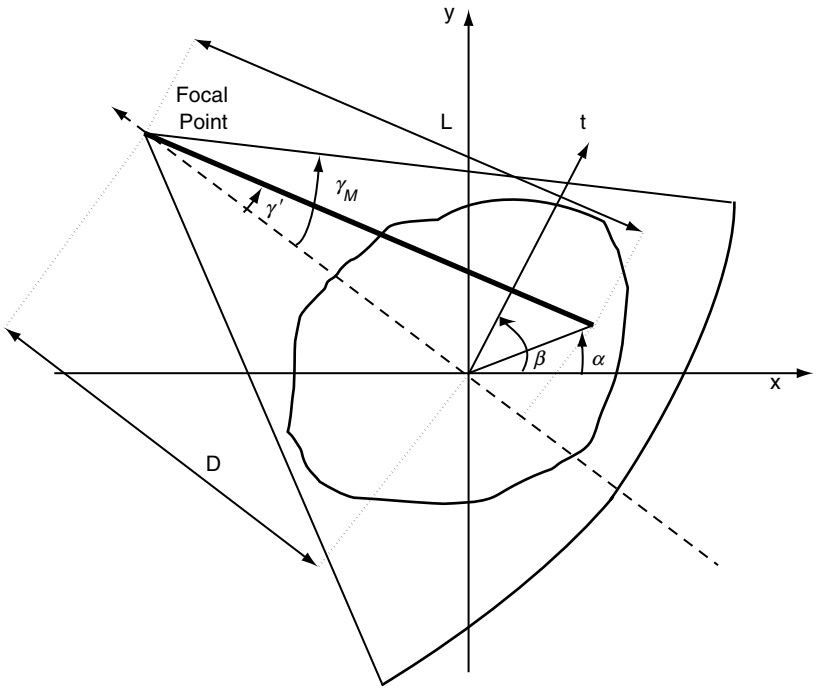


FIGURE 12. The parameters used in the derivation of the fan-beam image reconstruction methods.

$$f(r,\phi) = \int_0^{2\pi} \frac{1}{L^2} \int_{-\gamma_M}^{\gamma_M} r_\beta^c(\gamma) g^c(\gamma' - \gamma) D \cos\gamma d\gamma d\beta \tag{31}$$

where γ_M is the maximum of the angle γ, and

$$g^c(\gamma) = \frac{1}{2}\left(\frac{\gamma}{\sin\gamma}\right)^2 h(\gamma) \tag{32}$$

The fan-beam reconstruction algorithm for a curved detector can be written in the form of a weighted backprojection of the weighted and filtered projection algorithm, i.e.,

$$f(r,\phi) = \int_0^{2\pi} \frac{1}{L^2} q_\beta^c(r) d\beta \tag{33}$$

where the measured fan-beam projections $r_\beta^c(\gamma)$ are first multiplied by the weighting function $D\cos\gamma$ to obtain the weighted fan-beam projection $r_{\beta'}^c(\gamma)$, i.e.,

$$r_{\beta'}^c(\gamma) = r_\beta^c(\gamma) D \cos\gamma \tag{34}$$

and then convolved with the convolution kernel given by Eqn. (34) to give the filtered weighted fan-beam projections $q_\beta^c(\gamma)$, i.e.,

$$q_\beta^c(\gamma) = r_\beta^c{}'(\gamma) \otimes g^c(\gamma) \tag{35}$$

Using the similar approach, we can derive the weighted backprojection of the weighted filtered projection algorithm for a fan-beam geometry with flat detector as[17,18]

$$f(r,\phi) = \int_0^{2\pi} \frac{1}{U^2} q_\beta^f(s')d\beta$$

where

$$U(r,\phi,\beta) = \frac{D + r\sin(\beta - \phi)}{D}, \tag{36}$$

the weight function for the fan-beam geometry with flat detector is given by $D/\sqrt{D^2 + s^2}$ and the convolution function by $g^f(s) = h(s)/s$.

In summary, the weighted backprojection of the weighted filter projection algorithm for fan-beam image reconstruction with curve and flat detectors are given in Table 3.

5.2 *Analytic Image Reconstruction Methods in 3D PET*

In PET, the removal of the septa results in coincidences between detectors on different axial rings. The resulting LoRs make an oblique angle with the scanner axis. Since the three-dimensional (3D) activity distribution can be reconstructed using only the direct LoRs (i.e. those resulting from LoRs perpendicular to the scanner axis), the oblique projections are not required to reconstruct the activity distribution. However, using them has the potential to reduce the noise in the reconstructed image.

TABLE 3. Analytical image reconstruction methods for fan-beam geometries.

Steps	Curve Detector	Flat Detector
Multiply measured projection by weighting function	$r_\beta^c{}'(\gamma) = r_\beta^c(\gamma) D\cos\gamma$	$r_\beta^f{}'(\gamma) = r_\beta^f(\gamma)\frac{D}{\sqrt{D^2+s^2}}$
Filter weighted projection with convolution kernel	$q_\beta^c(\gamma) = r_\beta^c{}'(\gamma) \otimes \frac{1}{2}(\frac{\gamma}{\sin\gamma})^2 h(\gamma)$	$q_\beta^f(\gamma) = r_\beta^f{}'(\gamma) \otimes \frac{h(\gamma)}{\gamma}$
Perform weighted backprojection	$f(r,\phi) = \int_0^{2\pi} \frac{1}{L^2} q_\beta^c(r)d\beta$	$f(r,\phi) = \int_0^{2\pi} \frac{1}{U^2} q_\beta^f(s')d\beta$
	where $L(r,\phi,\beta)$ is given by $\sqrt{(D + r\sin[\beta - \phi])^2 + r^2\cos^2(\beta - \phi)}$	where $U(r,\phi,\beta)$ is given by $U(r,\phi,\beta) = \frac{D + r\sin(\beta - \phi)}{D}$

If the oblique projections are not truncated, they can be reconstructed by filtered backprojection using the Cosher filter.[19] This algorithm can be implemented either as filtered backprojection or backprojection of the filtered projections. In both cases, the backprojection operation is 3D. However, for real PET systems where the object imaged is longer than the axial length, the 3D reconstruction problem is complicated by the fact that all the oblique 2D projections are truncated, i.e. they are incompletely measured due to the finite axial extent of the tomograph. One way to handle this is by recognizing that the direct projections allow complete reconstruction of the activity distribution. The activity distribution reconstructed from these direct projections can then be reprojected to obtain the missing portions of the oblique projections. In practice, as much as 40% of the total reconstruction time is spent in estimating, and then backprojecting the unmeasured projection data. The algorithm combining this forward projection step with the Colsher filter is a 3D filtered backprojection algorithm referred to as 3DRP.[20]

An alternate exact approach to solve the problem of truncated 3D projections is the the FAst VOlume Reconstruction (FAVOR) algorithm.[21] However, FAVOR presents different noise propagation properties compared to other approaches, notably those using the Colsher filter. In particular, it exhibits a low-frequency 'hump' characteristic which includes most of the useful frequency range occurring in a typical image. This can be mathematically modified to introduce a specific bias by attributing a much greater weight to the direct projections than to the oblique ones. High-frequency statistical noise is, however, not affected since it occurs at frequencies higher than those corrected by the FAVOR filter.

While 3DRP works well and has served as a gold standard for 3D analytic reconstruction algorithms, it requires significant computational resources. As a result, a number of approximate methods have been proposed. Most of the well known approaches involve rebinning the data from the oblique sinograms into 2D data sets, thus allowing the use of 2D reconstruction methods and resulting in a significant decrease in reconstruction time. These algorithms rebin the 3D data into a stack of ordinary 2D direct sinograms and the image is reconstructed in 2D slice-by-slice, for instance using the standard 2D FBP or OSEM algorithms. The method used to rebin the data from the oblique sinograms into the smaller 2D data sets is the crucial step in any rebinning algorithm and has a significant influence on the resolution and noise in the reconstructed image.

The simplest rebinning algorithm is single-slice rebinning (SSRB).[22] It is based on the simple approximation that an oblique LoR can be projected to the direct plane halfway between the two endpoints of the LoR. In particular, SSRB assigns counts from an oblique LoR (where the detectors are on different axial detector rings) to the sinogram of the transaxial slice lying midway, axially between the two rings. This approximation is acceptable only near the axis of the scanner and for small apertures. For larger aper-

tures and farther from the axis of the scanner, SSRB results in position-dependent axial blurring.

Several alternative methods have been suggested to provide more accurate rebinning. Lewitt *et al.*[23] proposed a more refined rebinning approach, the multi-slice rebinning algorithm (MSRB), which is more accurate than SSRB, but suffers from instabilities in the presence of noisy data. The Fourier rebinning algorithm is a more sophisticated algorithm in which oblique rays are binned to a transaxial slice using the frequency-distance relationship of the data in Fourier space.24

6. Limitations of Analytical Image Reconstruction Methods

The analytical image reconstruction methods described above assume that no image degrading effects are present during the imaging process. That is, the imaging system has ideal imaging characteristics with perfect uniformity and infinitely high spatial resolution and there are no photon attenuation or scatter effects. Also, the ECT system has perfect mechanical alignment. In practice, however, many factors contribute to image degradation and the reconstructed images obtained from the analytical image reconstruction methods are less than ideal.

The detected counts in ECT are proportional to the number photons emitted from the radioactivity *in vivo*, that have traveled through the object, interacted with the body materials, passed through the collimator and been detected by the radiation detector. For the energies of photons emitted from the commonly used radionuclides in ECT, the predominate types of interactions are photo-electric and Compton scattering. The intrinsic resolution of the detector and the collimator used in SPECT, and the size of the detector elements, the positron range and non-colinearity of the dual 511 keV photon emissions in PET determine the collimator-detector response and resolution of the system.

In Eqns. (2) and (4), the projection data are written as line integral of the object or radioactivity distribution $f(x,y)$. In PET, the photon attenuation is factored out, and we assume that there is no effect from the system response and photon scatter in the object. In SPECT, we assume that there is no effect from the collimator-detector response and photon attenuation and scatter in the object.

Figure 13 shows the reconstructed images obtained by applying the analytical image reconstruction methods described above directly to SPECT projection data that include the effects of collimator-detector response and photon attenuation and scatter in the object. Simulated noise-free projection data from the 3D NCAT phantom were used. When the results are compared with Figures 4 and 7, they demonstrate the impact of the image degrading effects on the reconstructed images.

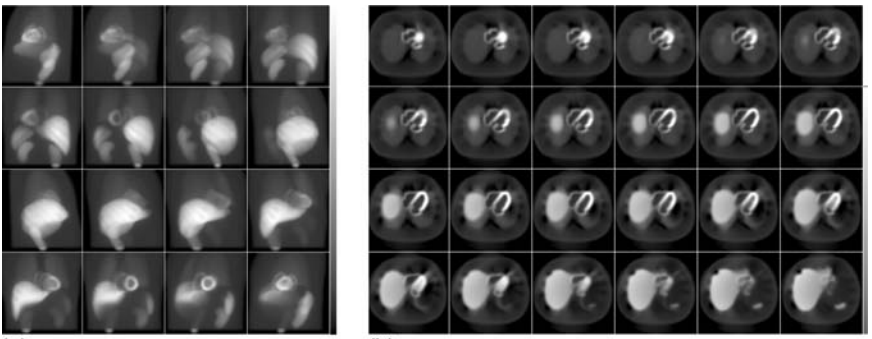


FIGURE 13. (a) Sample 2D noise-free projection data simulated from the 3D NCAT phantom with radioactivity distribution shown in Figure 4a. Here, the effects of system response and photon attenuation and scatter are included in the Monte Carlo simulation. (b) Sample 2D reconstructed images corresponding to the same slices shown in Figure 4a obtained using the backprojection of the filtered projections method. The images demonstrate the artifacts and distortions in the reconstructed image when image degrading effects are present in the acquired data.

The solution of the reconstruction problem with the image degrading effects including collimator-detector response and photon attenuation and scatter is difficult. In SPECT, suppose the object consists of a distribution of attenuation coefficient given by $\mu(s,t)$ and if the effect of attenuation is considered, instead of Eqn. (4), the projection data are given by Eqn. (3). The image reconstruction problem is then to seek the solution $f(s,t)$ of the inverse *attenuated Radon transform* expressed by Eqn. (3).

Consider a typical collimator-detector used in SPECT. The two-dimensional (2D) collimator-detector response function (CDRF), $cdrf(\vec{s};z)$, (see chapter 5 for a complete definition of the CDRF) is a function of the distance from the collimator face z and is shift-invariant across planes that are parallel to the face of the camera. Thus, the parameter $\vec{s}$ is the vector from the position of the projection of the source on the projection plane and the position of the detected photon. The projection data that include the effects of attenuation and collimator-detector response are given by

$$p_\theta(\vec{s}) = \int_{-\infty}^{+\infty} dt \int d\vec{s}' \times f(\vec{s}',t) \times cdrf(\vec{s}' - \vec{s};\, D+t) \times \exp\left[- \int_{t_d}^{t} \mu(\vec{s}',t')dt' \right] \quad (37)$$

where $\vec{s} = (s_u,\ s_v)$ is the intersection of the detector plane with the projection, D is the distance between the collimator and the center of rotation and $z = D + t$, and t_d is the t-coordinate of the detection plane just as for Eqn. (3). Inversion of Eqn. (37) is further complicated by the 3D collimator-detector response function and the 2D projection image. The image reconstruction problem is to find a solution of $f(s,t)$ from the projections described by Eqn. (37).

If scatter response is taken into account, the projection data will be given by

$$p_\theta(\vec{s}) = \int_{-\infty}^{+\infty} dt \int_{S_1}^{S_S} d\vec{s}' \times f(\vec{s}',t) \times srf(\vec{s}',t;\vec{s}) \times cdrf(\vec{s}' - \vec{s};D+t) \times \exp\left[-\int_{t}^{t_2} \mu(\vec{s}',t')dt'\right] \tag{38}$$

where $srf(\vec{s}',t;\vec{s})$ is the scatter response function and describes the probability that a photon emitted at position the position described by $(\vec{s}',t)$ is scattered and at position $\vec{s}$ on the detection plane. Equation (38) is a complex function especially since the scatter response function depends on the object being imaged and the source position of the emitting photon inside the body (i.e. it is not spatially variant). Analytical solutions to this problem have not been found yet.

Seeking accurate solutions to the image reconstruction problem in the presence of the image degrading factors has been the topic of much research in the past two decades. Significant progress has been made in the development of compensation methods for these image degrading factors. Since analytical solutions of the practical image reconstruction problem are difficult to find, statistical image reconstruction methods that incorporation accurate models of the noise statistics and components of the imaging process have been developed (see chapter 4). A major disadvantage of these iterative image reconstruction methods is the large amount of computations compared to analytic methods. However, with the advances in computer hardware and image reconstruction algorithms, these image reconstruction methods are gaining acceptance in clinical practice of ECT.

7. Summary

We have reviewed the 2D analytical image reconstruction methods used in ECT including PET and SPECT. The goal is to seek the 2D distribution of radioactivity *in vivo* from the measured 1D projection data from different views around the object. The basic theoretical formulation is similar to that used in single-slice x-ray CT. The projection data are 1D distribution of photons that are emitted from the radioactivity distribution and detected by radiation detectors. They are arranged in parallel-beam geometry. Assuming an ideal imaging situation where the effects of statistical noise, and photon attenuation and scatter are not present and perfect detector response, analytical image reconstruction methods provide reconstructed images that accurately represent the 2D radioactivity distribution *in vivo*.

Analytical image reconstruction methods consist of two major steps, simple backprojection and special processing of either the 1D projection data or the 2D backprojected images. This special processing can be equivalently applied in the spatial or spatial frequency domain. In the spatial frequency domain, the processing requires filtering with ramp function. With the use of fast Fourier transform algorithms, analytical image reconstruction methods that perform the filtering in the spatial frequency domain, i.e., the filter of the backprojection (FBpj) and the backprojection of the filtered projections (BpjFP) methods are often used. Among these two methods, the BpjFP method is preferred due to its simpler processing of the 1D projection data before backprojection and smaller data memory requirement. Additional filter function can be applied in conjunction with the ramp function to provide noise smoothing and edge enhancement to the reconstructed images.

A commonly used data acquisition geometry is the fan beam geometry. In a typical ring-type PET system, the LoRs from coincidence detection form fan-beam data arrays. In SPECT, fan- beam collimators are used to restrict the collection of photons that travel in a fan-beam geometry. The fan-beam projection data can be rebinned into parallel-beam projections before image reconstruction. However, the rebinning process is time-consuming and introduces data smoothing. Fan-beam reconstruction methods that apply directly to the fan-beam projection data are presented. They provide accurate reconstruction of images from fan-beam projections.

In practice, effects including those from the instrumentation, physical factors such as photon attenuation and scatter, statistical noise fluctuations and patient motions severely degrade the projection data and reconstructed image quality and quantitative accuracy in ECT. Analytical solutions of the image reconstruction problems that include these image degrading factors are complex and difficult. Iterative reconstruction methods that incorporate accurate models of the imaging process have been proposed. They have been shown to provide substantial improvement in both image quality and quantitative accuracy as compared with the analytical reconstruction methods. They are active areas of research and are discussed in more detail in the following chapters.

Acknowledgments. The authors acknowledge the support of the US Public Health Service grant RO1 EB00168.

References

1. Budinger T. F., Gullberg, G. T. and Huesman, R. H., "Emission Computed Tomography" in: *Image Reconstruction From Projections: Implementation and Applications*, edited by G. T. Herman Springer-Verlag, New York, (1979), *pp* 147-246.

2. Mandelkern K. and Phelps, M. E., "Methods and instrumentation for positron emission tomography" in: *Diagnostic Nuclear Medicine*, edited by A. Gottschalk, P.B. Hoffer, and E.J. Potchen Williams & Wilkins, Baltimore, (1988), *pp* 128-149.
3. Ter-Pogossian M. M., "Positron Emission Tomography (PET): General Principles" in: *Principles of Nuclear Medicine*, edited by H.N. Wagner, Z. Szabo, and J.W. Buchanan W.B. Saunders Company, Philadephia, (1995), *pp* 342-345.
4. Budinger T. F., "Single Photon Emission Computed Tomography" in: *Diagnostic Nuclear Medicine*, edited by A. Gottschalk, P.B. Hoffer, and E.J. Potchen Williams & Wilkins, Baltimore, (1988), *pp* 108-127.
5. Jaszczak R. J. and Tsui, B. M. W., "Single Photon Emission Computed Tomography (SPECT)" in: *Principles of Nuclear Medicine*, edited by H.N. Wagner, Z. Szabo, and J.W. Buchanan W.B. Saunders, Philadelphia, (1995), *pp* 317- 341.
6. Budinger T. F., "Single Photon Emission Computed Tomography" in: *Diagnostic Nuclear Medicine*, edited by M.P. Sandler, J.A. Patton, R.E. Coleman *et al*. Williams & Wilkins, Baltimore, (1996), *pp* 121-138.
7. Kuhl K. E. and Edwards, R. Q., Image separation radioisotope scanning. *Radiology* **80:** 653-661 (1963).
8. Cormack A. M., Representation of a function by its line integrals, with some radiological applications. *J Appl Phys* **34:** 2722-2727 (1963).
9. Hounsfield G. N., Computerized transverse axial scanning (tomography) Part 1: Description of system. *Br J Radiol* **46:** 1016-1022 (1973).
10. Ramachandran G. N. and Lakshminarayanan, A. V., 3- Dimensional reconstruction from radiographs and electron micrographs-Application of convolutions instead of Fourier transforms. *Proc Natl Acad Sci* **68:** 2236-2240 (1971).
11. Shepp L. A. and Logan, B. F., The Fourier reconstruction of a head phantom. *IEEE Trans Nucl Sci* **21:** 21-43 (1974).
12. Jaszczak R. J., Chang, L.-T., Stein, N. A. *et al*., Whole-body single-photon emission computed tomography using dual, large-field-of-view scintillation cameras. *Phys Med Biol* **24:** 1123-1143 (1979).
13. Tsui B. M., Gullberg, G. T., Edgerton, E. R. *et al*., Design and clinical utility of a fan beam collimator for SPECT imaging of the head. *J Nucl Med* **27:** 810-819 (1986).
14. Jaszczak R. J., Floyd, C. E., Jr., Manglos, S. H. *et al*., Cone beam collimation for single photon emission computed tomography: analysis, simulation, and image reconstruction using filtered backprojection. *Med Phys* **13:** 484-489 (1986).
15. Jaszczak R. J., Greer, K. L., Floyd Jr., C. E. *et al*., Imaging characteristics of a high resolution cone beam collimator. *IEEE Trans Nucl Sci* **33:** 644-648 (1988).
16. Jaszczak R. J., Greer, K. L. and Coleman, R. E., SPECT using a specially-designed cone-beam collimator. *J Nucl Med* **29:** 1398-1405 (1988).
17. Gullberg G. T., Crawford, C. R. and Tsui, B. M. W., Reconstruction algorithm for fan-beam with a displaced conter-of-rotation. *IEEE Trans Med Imaging* **5:** 23-34 (1986).
18. Crawford C. R., Gullberg, G. T. and Tsui, B. M., Reconstruction for fan beam with an angular-dependent displaced center-of-rotation. *Med Phys* **15:** 67-71 (1988).
19. Colsher J. G., Fully three-dimensional positron emission tomography. *Phys Med Biol* **25:** 103-115 (1980).
20. Kinahan P. E. and Rogers, J. G., Analytic 3D image reconstruction using all detected events. *IEEE Trans Nucl Sci* **36:** 964-968 (1989).

21. Defrise M., Townsend, D. W. and Clack , R., ''Favor: a fast reconstruction algorithm for volume imaging in PET.'' *Conf. Rec.* 1991 IEEE Medical Imaging Conference, Santa Fe, USA, *pp* 1919-1923 (1992).
22. Daube-Witherspoon M. E. and Muehllehner, G., Treatment of axial data in three-dimensional PET. *J Nucl Med* **28:** 1717-1724 (1987).
23. Lewitt R. M., Muehllehner, G. and Karp, J. S., Three-dimensional image reconstruction for PET by multi-slice rebinning and axial image filtering. *Phys Med Biol* **39:** 321-339 (1994).
24. Defrise M., Kinahan, P. E., Townsend, D. W. *et al*., Exact and approximate rebinning algorithms for 3-D PET data. *IEEE Trans Med Imaging* **16:** 145-58 (1997).

4
Iterative Reconstruction Methods

B.F. Hutton*, J. Nuyts§ and H. Zaidi†

1. Introduction

The direct reconstruction in emission tomography (ET) using analytical methods was described in chapter 3. This included filtered back projection (FBP), until recently the most commonly used method of reconstruction in clinical practice. Analytical reconstruction usually assumes a relatively simple model of the emission and detection processes and would become quite complex if rigorous models were applied. FBP has further limitations due to the presence of streak artefacts that are particularly prominent near hot structures and the noise enhancement that is inherent in the reconstruction. An alternative to analytical reconstruction is the use of iterative reconstruction techniques, which can more readily incorporate more complex models of the underlying physics and also can better accommodate assumptions regarding the statistical variability of acquired data. Unlike analytical reconstruction where FBP dominates, there are many approaches to iterative reconstruction. Difficulties in understanding the classification of these algorithms have led to considerable confusion in the choice of appropriate algorithms, particularly in routine clinical application. This chapter is intended to provide a general overview of iterative reconstruction techniques with emphasis on practical issues that may assist readers in making an informed choice. Iterative reconstruction can be applied equally well to single-photon emission computed tomography (SPECT) or positron emission tomography (PET) or indeed any tomographic data. The coverage in this chapter will be general, although some specific issues relating to either PET or SPECT will be highlighted. Reconstruction of transmission data will also be briefly addressed. For additional coverage readers are referred to other recent texts that specifically address iterative reconstruction.[1-3]

*Prof. B.F. Hutton, Institute of Nuclear Medicine, University College London, London WIE 6BT, UK; also Centre for Medical Radiation Physics, University of Wollongong, NSW 2522, Austalia
§Dr J. Nuyts, Katholieke Universiteit Leuven, Nuclear Medicine Department, U.Z. Gasthuisberg, Herestraat 49, B30000 Leuven, Belgium
†PD Dr H. Zaidi, Division of Nuclear Medicine, Geneva University Hospital, 1211

2. What is Iterative Reconstruction?

2.1 General Iterative Techniques

Iterative techniques are common in problems that involve optimization. The reconstruction problem can be considered a particular case where one is trying to determine the 'best' estimate of the distribution of activity based on the measured projections. An assumption underlying analytical reconstruction techniques is that there is a unique solution. However in practice, due to presence of noise, there are normally a number of possible solutions to the reconstruction problem. The 'best' reconstruction is determined by defining some criterion that measures goodness of fit between the reconstruction estimate and measurements and by adopting an algorithm that finds the optimal solution. Iterative techniques are well suited to solving this sort of problem.

A flow-chart that illustrates the general iterative technique used in reconstruction is presented in Figure 1. The iterative algorithm involves a feedback process that permits sequential adjustment of the estimated recon-

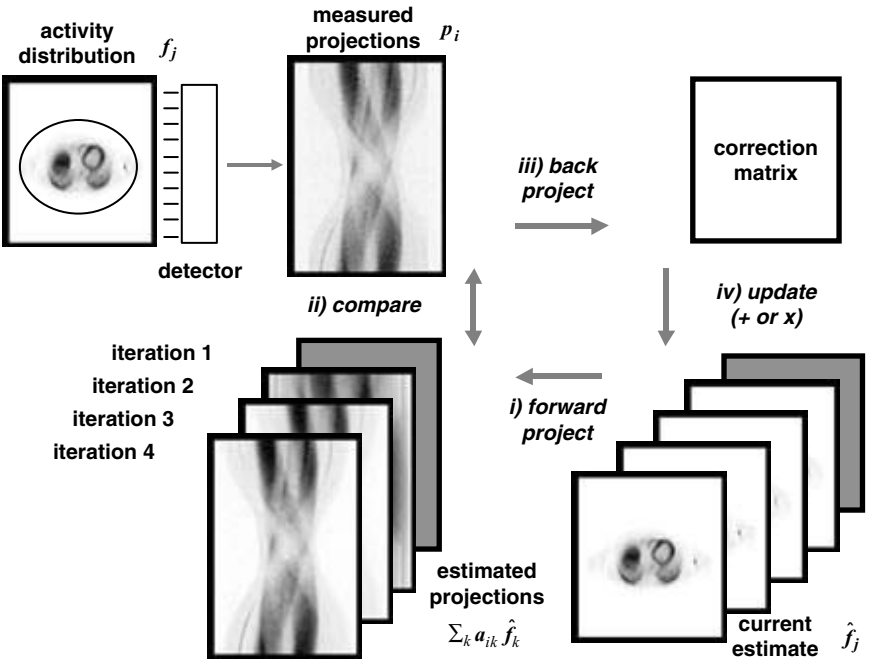


FIGURE 1. Schematic of general iterative reconstruction algorithm. Starting from a uniform grey image, estimated projections are constructed by forward projection. These are compared with the measured projections and a correction matrix is constructed by back projection. The reconstruction estimate is update either by summation or multiplication and this becomes the starting point for the next iteration.

struction so as to improve its correspondence with the measured projections. The iterative process begins with an initial estimate of the object count distribution, which may either be simply a uniform arbitrary count level or a previous reconstruction (e.g. using FBP). Provided a suitable model of the emission and detection physics is incorporated, the projections that would arise from this initial object can be estimated (by forward projection): this is effectively an estimate of what the detectors would measure given the initial object. Clearly this estimate of projections will differ from the actual measured projections, unless the initial guess happens to coincide exactly with the actual object. The difference between estimated and measured projections can be used to modify the original estimate of the object by use of suitable additive or multiplicative corrections at each point (usually via back projection). The adjusted object then becomes the starting point for a second iteration. This proceeds as the first with the forward projection so as to re-estimate the projections that would derive from this updated object. The process continues for multiple iterations using the continuous feedback loop until a final solution is reached (where usually a predetermined number of iterations have been completed, resulting in very little object change between iterations). This iterative process is central to all iterative reconstruction algorithms.

2.2 Attractions of Iterative Methods

Iterative reconstruction has a number of potential advantages that make it attractive in comparison with analytical methods. Foremost is the limiting assumption in analytical techniques that the measured data are perfectly consistent with the source object, a requirement that is never true in practice given the presence of noise and other physical factors (e.g. attenuation). The lack of consistency generally results in problems with noise and possible problems in finding a unique analytical solution. In contrast, iterative techniques function with either consistent or inconsistent data and can attempt to model noise directly, with improved noise properties in the final reconstruction. In addition iterative algorithms are well suited to handling complex physical models of the emission and detection processes, including position variant attenuation or distance dependent resolution. This ability to directly model the system, including some consideration of the noise characteristics, provides considerable flexibility in the type of data that can be reconstructed.

Iterative reconstruction methods have various additional properties that make them attractive, although some properties (e.g. non-negativity) are specific to the particular type of reconstruction and will be dealt with in the appropriate section. In general though, iterative methods provide reconstructions with improved noise characteristics. They greatly reduce the streaking artefacts that are common in analytical reconstruction and are better able to handle missing data (such as may occur due to truncation or inter-detector gaps). They generally provide accurate reconstructions that

can be used for quantification, provided some caution is exerted in choice of reconstruction parameters.

The main limitation in using iterative algorithms is the execution speed, which is significantly slower than FBP. However improvements in computer speed in combination with acceleration techniques (described in section 7) permits reconstruction in a clinically acceptable time.

2.3 Classification of Iterative Reconstruction Methods

It is convenient at the outset to introduce a general classification for the various iterative reconstruction methods. These broadly form three logical groups, defined mainly by the underlying assumptions regarding the nature of the data, whether statistical or non-statistical. The earliest developments involved no assumptions regarding the origins of noise in emission data and led to the various classical algebraic reconstruction techniques, which are non-statistical. These are described in section 3. The statistical techniques can be broadly divided into two main groups, those that assume Gaussian noise (involving least squares solutions: section 4) and those that assume Poisson noise (involving maximum likelihood solutions: section 5). These two fairly basic assumptions have led to two distinct classes of reconstruction that involve different algorithms and different properties. Variations on these algorithms specifically address issues such as noise control and acceleration (described in sections 6 and 7 respectively).

2.4 Nomenclature for Discrete Iterative Reconstruction

It is useful to establish nomenclature with reference to the general iterative algorithm in Figure 1. Consider the measurement of a set of projections $p(s,\phi)$, where s represents distance along a projection and ϕ represents the angle of acquisition, originating from a distribution of activity in the corresponding single slice in an object given by $f(x,y)$. It is convenient to consider the equivalent discrete set of projection pixel values p_i for counts originating from the object voxel activity concentration f_j. The process of estimating projections from a known (or estimated) activity distribution requires a description of the physical model that adequately describes the emission and detection processes. This essentially defines the probability of detecting an emitted photon, originating from location j, at any particular position, i, on the detector, which will depend on many factors including detector geometry, attenuation and resolution. A single matrix, a_{ij}, referred to as the transition matrix or system matrix, can conveniently describe this relationship. The process of forward projection can then simply be expressed as

$$p_i = \sum_j a_{ij} f_j \tag{1}$$

Similarly the back projection operation can be expressed as:

$$f_j = \sum_i a_{ij} p_i \tag{2}$$

3. Algebraic Reconstruction Algorithms

Algebraic reconstruction techniques are based on relaxation methods for solving systems of linear equalities or inequalities. Recent developments suggest that with proper tuning, the convergence of these algorithms can be very fast. However, the quality of the reconstructions has not been investigated thoroughly. In this section, a summary overview of algebraic reconstruction techniques is first set out, followed by a description of the algorithmic implementation of the classical algebraic reconstruction technique (ART) and its variants.

3.1 ART

The ART algorithm is based on the Kaczmarz method and uses a successive over-relaxation method well-known in numerical linear algebra.[4] It was first proposed as a reconstruction method in 1970 by Bender *et al.*[5] and was applied mainly to X-ray photography and electron microscopy. The early developments addressed mainly reconstruction of X-ray CT data but more recently, ART-type algorithms have been proposed specifically for PET.[6] The principle of the basic ART algorithm consists of describing every iteration point by point (row-action method), and in correcting all voxels in the image which are found on a projection ray, so as to minimise the difference between the values of the calculated and measured projections at the point under consideration. The process comes to a stop when a certain criterion becomes relatively small. For example the sum of squared differences between the calculated and measured projections can be used as such a criterion.

In fact, ART consists of guessing at a value for all the voxels f_j, and then modifying each element along each ray by a factor which compensates for the discrepancy between the measured and the calculated ray sum:

$$f_j^{new} = f_j^{old} \frac{p_i}{\sum_k a_{ik} f_k^{old}} \tag{3}$$

Here f_j^{new} and f_j^{old} refer to the current and previous estimates of the reconstructed object respectively. If the calculated ray sum is the same as the measured value, it implies that the guessed value is correct for a particular projection; however, for another projection there might be a large discrepancy. Thus the pixels of the last views (while lying in the ray for the new view) will be modified according to the discrepancy between the new ray and the measured value. Thus, each ray from each projection is examined and values of f falling within that ray are changed iteratively for all the

projections for a certain number of iterations. It is evident that the computational effort of the ART algorithm is relatively low, but data matrices can be several hundred of Megabytes in size.

It has been shown that by careful adjustment of the relaxation parameters and the order in which the collected data are accessed during the reconstruction procedure, ART can produce high-quality reconstructions with excellent computational efficiency.[4] In particular, the choice of the projection data access scheme proved to be crucial for improvement of low-contrast object detection.[7] An important modification of ART consists of setting to zero those values in the array that are clearly zero because they correspond to a ray sum that was observed as zero. This is an important boundary condition for any of the iterative techniques.

3.2 Other Variants

A large number of ART variants have subsequently been proposed in literature. Eq. (3) above is called multiplicative ART (MART). Another method of correcting the discrepancy between the measured and calculated projections consists of adding the difference between them. This is called the additive form of ART (AART). The diverse variants of ART correspond to row-action techniques, as each iteration only calls for the use of one equation at a time. Other algorithms like the simultaneous iterative reconstruction technique (SIRT) consist of correcting simultaneously each voxel for all rays passing through it. So, these corrections are incorporated by using data from all of the projections simultaneously. The simultaneous version of MART (SMART) is similar to the ML-EM approach (see section 5.1).

The block-iterative version of ART (BI-ART) is obtained by partitioning the projection set and applying the ART equation to each subset (a block represents a subset of the projection data). The correction is computed for each block, and may be interpreted as a weighted back-projection of the difference between the computed image projection and the acquired projection. This correction is performed only after an entire projection image is computed. Other variants of ART include the block-iterative version of SMART (BI-SMART) and the rescaled block-iterative SMART (RBI-SMART).[8] With BI-SMART in PET, an issue is the way to define a block to ensure the convergence of the algorithm, at least for the consistent case.

4. Statistical Algorithms-Gaussian Assumption

4.1 Least Squares Solutions

In nuclear medicine, the count rates are usually fairly low, and as a result, the data tend to be rather noisy. As mentioned above, one cannot hope to recover the true solution from such data, and some criterion to define the

'best solution' is needed. In statistical reconstruction, the best solution is defined as the one that is most likely, given the data. Thus, the reconstruction is obtained by finding the image $\hat{f}$ that maximises the conditional probability prob$[f|P]$, where P are the measured projections. Using Bayes' rule, this objective function can be rewritten in a more convenient form:

$$\text{prob}[f|P] = \frac{\text{prob}[P|f]\text{prob}[f]}{\text{prob}[P]} \tag{4}$$

Because P is constant, maximising prob$[f|P]$ is equivalent to maximising prob$[P|f]$prob$[f]$. The factor prob$[P|f]$ is called the "likelihood" and tells how well the data agree with the image. The factor prob$[f]$ is the "prior", and tells what is known already about the image f, prior to the measurement. Finally, prob$[f|P]$ is called the "posterior", and represents what is known by combining the a-priori knowledge with the information obtained from the measurement. If it is assumed that nothing is known in advance, then the prior is constant and maximum-a-posteriori (MAP) reconstruction is equivalent to maximum-likelihood (ML) reconstruction. The discussion of the posterior is deferred to a later section.

In emission and transmission tomography, the noise on the counts measured at two different detector locations is uncorrelated (the noise is "white"). This allows for factorisation of the likelihood:

$$\text{prob}[P|f] = \prod_i \text{prob}[p_i|f] \tag{5}$$

Maximising a function is equivalent to maximizing its logarithm. Therefore, the maximum-likelihood reconstruction $\hat{f}$ is obtained by maximising

$$\text{log-likelihood} = \sum_i \ln \text{prob}[p_i|f] \tag{6}$$

Assuming that the noise can be well approximated as a Gaussian distribution with known standard deviation makes the maximum-likelihood solution identical to the least squares solution. The mean of the Gaussian distribution for detector i is computed as the projection of the image $\hat{f}$. Consequently, the least squares solution is obtained by minimizing (over F):

$$L_G(P,F) = \sum_i \frac{\left(p_i - \sum_j a_{ij} f_j\right)^2}{2\sigma_i^2} = \frac{1}{2}(P-AF)'C^{-1}(P-AF) \tag{7}$$

where P is a column matrix with elements p_i, A is the system matrix with elements a_{ij} and F is a column matrix with elements f_j, and prime denotes transpose. C is the covariance matrix of the data, which is assumed to be diagonal here, with elements $c_{ii} = \sigma_i^2$. A straightforward solution is obtained by setting the first derivative with respect to f_j to zero for all j. This solution is most easily written in matrix form:

$$\hat{F} = \left[A' C^{-1} A\right]^{-1} A' C^{-1} P \tag{8}$$

The problem with this expression is the large size of the matrices. The inverse of the so-called Fisher information matrix $A' C^{-1} A$ is required.[9] It has $J \times J$ elements, where J is the number of voxels in the image to be reconstructed. Its matrix elements equal

$$\text{FIM(j,k)} = \sum_i \frac{a_{ij} a_{ik}}{\sigma_i^2} \tag{9}$$

which is mostly non-zero: although the diagonal dominates, this matrix is not sparse at all and direct matrix inversion is prohibitive.

As an exercise, it can be assumed (although totally unacceptable) that all standard deviations σ_i are identical and equal to 1. Moreover, consider an idealized parallel-hole projection, with perfect resolution and no attenuation or other degrading effects. The image minimizing L_G then becomes

$$\hat{F} = [A' A]^{-1} A' P \tag{10}$$

The operator $A' A$ computes the backprojection of the projection of a point. In this idealized case, the operator $A' A$ is shift invariant, and it can be shown that its inverse is (a digital version of) the inverse Fourier transform of the ramp filter. Consequently, this analysis reveals that FBP computes the solution of the unweighted least squares problem.

In emission tomography, it is not realistic to assume that all the standard deviations σ_i are identical. As a result, the Fisher information and its inverse become more complex, position dependent operators, that cannot be implemented with shift invariant filters (such as a ramp filter). Consequently, one has to turn to numerical, iterative optimisation techniques.

Another problem is the estimation of the standard deviations σ_i. The Poisson distribution is well approximated as a Gaussian with variance equal to the mean. But the mean (noise-free) projection count is unknown. There are two approaches to deal with this problem. The first one is to estimate σ_i from the noisy counts p_i. One could simply set $\sigma_i = {p_i}^{0.5}$. However, this leads to a negative bias in the reconstruction: all counts that happen to be smaller than the mean will be assigned a smaller standard deviation, and hence a higher weight in the weighted least squares computations. These noisy weights may even cause streak artefacts in the reconstruction. So the noise on the weights has to be reduced, e.g. by estimating σ_i from a smoothed version of the data. The second approach is to estimate σ_i from the current reconstruction, rather than from the data. This is closer to the physical reality, because the measurement is assumed to be a noise realization of the noise-free projection of the tracer distribution. However, it is also more complex, because now the weights depend on the unknown solution of the problem. Possibly because of this, the data-based estimate of the weights has received more attention in the literature.

4.2 Approaches to Minimization

Many optimisation algorithms can be applied to this problem (e.g. conjugate gradient or steepest descent), but it is beyond the scope of this text to present these in detail. One thing that most of these algorithms have in common is that they will modify each voxel such as to decrease the value of L_G in every iteration. To do that, the value added to voxel j must have the opposite sign to the derivative of L_G with respect to f_j. This can be written as

$$\hat{f}_j^{new} = \hat{f}_j^{old} - \frac{1}{\alpha_j} \frac{\partial L_G}{\partial f_j}\bigg|_{\hat{f}_j^{\text{old}}} \tag{11}$$

where α_j is some positive value, that may change during iteration, and that has to be tuned in order to guarantee convergence. To show the typical form of the optimisation algorithms, the derivatives of the two following cost functions are computed:

$$L_{G1}(P,F) = \sum_i \frac{(p_i - \hat{p}_i)^2}{2\sigma_i^2} \text{ and } L_{G2}(P,F) = \sum_i \frac{(p_i - \hat{p}_i)^2}{2\hat{p}_i} \text{ with } \hat{p}_i = \sum_j a_{ij}\hat{f}_j \tag{12}$$

L_{G1} is the cost function with pre-computed (data based) standard deviation. L_{G2} is the version where the variance is estimated from the current reconstruction. The derivatives with respect to f_j equal:

$$\frac{\partial L_{G1}}{\partial f_j} = -\sum_i a_{ij} \frac{p_i - \hat{p}_i}{\sigma_i} \text{and} \frac{\partial L_{G2}}{\partial f_j} = -\sum_i a_{ij} \frac{(p_i - \hat{p}_i)(p_i + \hat{p}_i)}{2\hat{p}_i^2} \tag{13}$$

The difference between algorithms is mainly in how the step size α_j is tuned. Note that (weighted) least squares algorithms do not "naturally" produce non-negative solutions. If non-negativity is desired, it must be imposed during the iterations, which can adversely affect convergence. Sauer and Bouman[10] and Fessler[11] proposed an effective method, which has been used successfully by others.[12,13] It updates the reconstruction voxels sequentially (as in ART), sets the step size α_j equal to the element *FIM(j,j)* of the Fisher information matrix and produces a non-negative solution. Other solutions have been proposed, e.g. scaled steepest descent by Kaufman[14] and a preconditioning algorithm by Chinn and Huang.[15]

5. Statistical Algorithms-Poisson Assumption

5.1 Maximum Likelihood Solution: the ML-EM Algorithm

The random nature of radioactive decay suggests that a Poisson model is more appropriate for emission data (although this is well approximated by a Gaussian provided measured counts are reasonably high). An appealing

consequence of using the Poisson model is that non-negativity is assured even at low count levels.

The basic Poisson model provides the probability of measuring a particular count, c, given an expected measurement, r:

$$\text{prob}[c|r] = \frac{e^{-r}r^c}{c!} \tag{14}$$

Using this Poisson model the probability of acquiring the projection count distribution that was measured, P, given an estimated distribution of activity in the emission object, f, can be represented by the product of probabilities for individual projection pixels. This conditional probability is referred to as likelihood, L

$$L(P|f) = \text{prob}[P|f] = \prod_i \exp\left[-\sum_j a_{ij} f_j\right] \left(\sum_j a_{ij} f_j\right)^{p_i} (p_i!)^{-1} \tag{15}$$

As argued above (section 4.1, eq (4)-(6)), maximising (the logarithm) of prob$[P|f]$ is equivalent to maximising prob$[f|P]$, provided the *a priori* distribution can be assumed to be constant. This provides the most likely distribution of emissions that represents the original activity distribution, given the measured projections.

There are various approaches to determine the maximum likelihood (ML) solution but the most commonly used is the expectation maximization (EM) algorithm, which unifies various previous statistical approaches.[16] This involves an iterative process with the attractive property that convergence is guaranteed. ML-EM was originally applied to emission tomography in the early eighties[17,18] but continues to be widely used. Other groups also deserve credit for much of the early application of ML-EM to emission tomography (e.g. Miller *et al.*[19]). The EM algorithm involves two distinct steps. First the expected projections are calculated by forward projection using the appropriate system / transition matrix, based on the estimate of the activity distribution from the previous iteration (an initial guess in the case of the first iteration). Second the current estimate is updated so as to maximise the likelihood, achieved by multiplication of the previous estimate by the back projection of the ratio of measured over estimated projections. The resultant ML-EM equation is derived elsewhere, as cited above, and is given by

$$f_j^{new} = \frac{f_j^{old}}{\sum_l a_{lj}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old}} \tag{16}$$

5.2 Properties of ML-EM Reconstruction

The EM algorithm results in an iterative process for estimation consistent with the general iterative flow-chart given in Figure 1. In this case the update is multiplicative, not unlike SMART, with the update simply being the ratio

of measured over estimated projections. As the number of iterations increases the likelihood increases, providing an estimate that theoretically is more likely to be close to the true object distribution. In practice, however, the image reaches an optimal visual quality at typically around 16 iterations and, in the absence of any noise constraint, appears progressively more noisy at higher number of iterations (Figure 2). The noise characteristics are appealing, with the variance remaining proportional to number of counts rather than being approximately position-independent as in FBP.[9,20] This tends to favour lesion detection in low count areas where the signal to noise ratio can be markedly improved. The ML solution after a large number of iterations is not the most 'desirable' solution as it reflects the actual noisy distribution of emitted counts rather than the underlying activity, whose distribution is likely to be much less variable. In clinical practice it is common to stop at a small number of iterations in order to limit noise. However it should be recognised that the reconstruction does not converge at the same rate for all points. Halting the reconstruction early runs a risk of reducing reconstruction accuracy, which can be avoided by using a larger number of iterations with post-reconstruction smoothing.[21] Alternative approaches to controlling noise are discussed in section 6.

There are several attractive theoretical properties of ML-EM although in practice these rarely offer real advantage due to approximations in the system model and the relatively high level of noise in most emission tomography studies. The use of a multiplicative update guarantees positive values and also means that areas outside the object, where zero counts are expected,

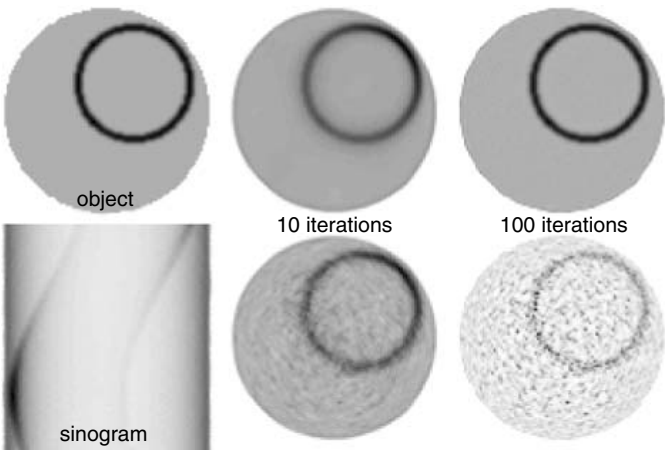


FIGURE 2. PET simulation of circular object in a uniform background, with uniform attenuation. ML-EM reconstructions at 10 and 100 iterations are shown, obtained from a sinogram without noise (top) and with Poisson noise (bottom). At 10 iterations, convergence is clearly incomplete. At 100 iterations, the noise level is becoming unacceptable.

are constrained so as to remain zero. This can result in problems near object edges when the projection counts are low as any zero pixel on a projection defines a projected line where counts are set to zero and remain unchanged, even when intersected by non-zero projections. The conservation of counts that occurs, results in the possibility of misplaced counts (usually hot spots) near the object edge.

A definite limitation with ML-EM reconstruction is the time taken for reconstruction. Each of the forward and back projection steps in all iterations takes approximately the same time as FBP; twenty iterations ML-EM would therefore take forty times as long as FBP. Fortunately steps can be taken to significantly accelerate the reconstruction as outlined in section 7.

6. Approaches to Controlling Noise

As mentioned above, the ML-criterion does not prevent noise propagation, and in many clinical cases the noise becomes unacceptably high at high iteration numbers. Stopping iterations early is dangerous: because the convergence of ML-EM depends on position and even orientation, the spatial resolution can be very position dependent at low iteration numbers. This is illustrated by the PET simulation in Figure 2. The true object is a circular ring of activity, embedded in a large disc with background activity. Attenuation within the disc is uniform. Because the attenuation along central projection lines is higher, the convergence of ML-EM is slower in the centre. At 10 iterations, the spatial resolution in the centre is clearly lower than near the edge. At 100 iterations, the resolution has become uniform. However, when Poisson noise is present, the noise level at 100 iterations is very high.

6.1 MAP with Gibbs Prior

As discussed above, maximising the likelihood is equivalent to maximising the posterior, if one can assume that nothing is known about the image, prior to the measurement. But obviously, some a-priori knowledge is available: the image should not be too noisy. Somehow, this prior knowledge must be translated into a mathematical expression, which can be inserted in equation (4).

A convenient way to define a prior distribution favouring smooth solutions is via a Markov random field (or equivalently: a Gibbs random field). In a Markov random field, the probability of a particular voxel depends on the voxel values in a neighbourhood of that voxel. The dependence is defined with a Gibbs distribution of the following form (e.g. Geman and McLure[22]):

$$\text{prob}\left[f_j\right] = \frac{1}{Z} e^{-\beta U(f)} \text{ with } U(f) = \sum_{k \in N_j} V\left(f_j, f_k\right) \tag{17}$$

where prob[f_j] denotes the conditional probability for the value in voxel j given the rest of the image, Z is a normalization constant, β a factor that determines the strength of the prior (1/β is the "temperature"), and N_j contains the voxels neighbouring voxel j. U is the so-called energy function, and higher energies are less likely. In fact, more complex forms of U are allowed, but this one is nearly always used. To favour smooth images, noisy ones should be penalized with a higher energy (hence the alternative name "penalized likelihood"). An obvious choice for V is the "quadratic penalty":

$$V_Q(f_j, f_k) = \frac{(f_j - f_k)^2}{2\sigma^2} \tag{18}$$

which makes *prob[f]* maximum for a perfectly uniform image. A possible disadvantage of V_Q is that it heavily penalizes large differences, causing strong smoothing over edges. To avoid that, the Huber function (e.g. Mumcuoğlu *et al.*[23]) quadratically penalizes "small" differences, but penalizes "large" differences only linearly:

$$\begin{aligned} V_H(f_j, f_k) &= \frac{(f_j - f_k)^2}{2\sigma^2} && \text{if} |f_j - f_k| < \sigma \\ &= \frac{|f_j - f_k| - \sigma/2}{\sigma} && \text{if } |f_j - f_k| \geq \sigma \end{aligned} \tag{19}$$

A function with a similar edge-preserving characteristic (but with a gradual transition to the linear behaviour) is the logcosh function (see e.g. de Pierro[24]): $V_L = \ln(\cosh((f_j - f_k)/\sigma))$. Functions have been devised that preserve edges better than V_H or V_L, e.g. by applying a constant instead of a linear penalty for large differences. However, in contrast to the functions mentioned above, functions applying a constant penalty are not concave; they introduce multiple local maxima, making the final MAP-reconstruction dependent upon the initial image and the particular optimisation algorithm. For that reason, these functions are rarely used in emission tomography.

The functions V_Q, V_H and V_L penalize in proportion to the (square of) absolute differences between neighbouring voxels. But because the absolute tracer uptake values vary with scan time, injected dose, attenuation, patient weight and metabolism, it can be difficult to select a good value for the parameter σ. As one can see from the equations, ML-EM is insensitive to the absolute counts in the sinogram: if the sinogram is multiplied with a factor, the corresponding ML-EM reconstruction will simply scale with the same factor. This is no longer true for the MAP-reconstruction based on the quadratic or Huber prior. To avoid this problem, a concave function penalizing relative differences rather than absolute ones has been proposed:[25]

$$V_R(f_j, f_k) = \frac{(f_j - f_k)^2}{f_j + f_k + \gamma|f_j - f_k|} \tag{20}$$

Similar to the Huber prior, V_R has some edge-preserving features: the penalty decreases slower when the relative difference $|f_j - f_k|/(f_j + f_k)$ is large compared to $1/\gamma$.

Having defined the prior, an algorithm is needed to maximise the posterior, or equivalently, its logarithm: $\ln \text{prob}[P|f] + \ln \text{prob}[f]$. Direct application of the EM strategy yields the following expression:

$$f_j^{new} = \frac{f_j^{old}}{\sum_l a_{lj}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old} + \beta \frac{\partial U(f)}{\partial f_j}} \tag{21}$$

where $U(f)$ is the energy function of the Gibbs prior. The problem is that the derivative must be evaluated in f_j^{new}, which complicates the solution dramatically. Green[26] proposed the one step late (OSL) approach, which is to evaluate the derivative in the available reconstruction image f_j^{old}. This approach is effective if β is not too large, but convergence is not guaranteed, and with large β the algorithm can actually diverge. de Pierro[24] found a new (non-statistical) derivation for the classical ML-EM algorithm, which is based on the construction of a surrogate function at each iteration. This is a function that coincides with the likelihood function up to the first derivate and which is easier to maximise than the likelihood. Each iteration, the surrogate is maximised, and one can show that this guarantees an increase of the likelihood. The same approach can be applied to most Gibbs priors, and hence, to the posterior. This has led to the development of new and more robust algorithms for MAP-reconstruction (see Ahn and Fessler[27] and the references therein).

6.2 *Post-smoothing Unconstrained ML-EM*

As shown in Figure 2, possible convergence problems of ML-EM are eliminated by iterating sufficiently long. This leads to unacceptable noise levels, but the noise can be reduced considerably with moderate smoothing,[21,28] as illustrated in Figure 3. Since this is high frequency noise, smoothing is even more effective than with uncorrelated noise.[29]

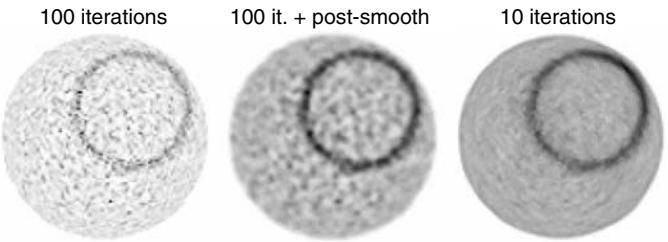


FIGURE 3. Same PET simulation as in Figure 2, comparing post-smoothing ML-EM after 100 iterations with the reconstruction without post-smoothing (left) and the ML-EM reconstruction at 10 iterations.

An advantage of post-smoothed ML-EM is its wide availability. It is currently available on most SPECT and PET systems. As will be discussed below, an additional advantage (for some applications) is that it can impose nearly uniform spatial resolution. A disadvantage is that many iterations are required to achieve this resolution uniformity.

6.3 Expectation-Maximisation-Smooth (EMS)

Some groups have investigated the application of a smoothing filter after every ML-EM iteration, called EMS (expectation-maximisation-smooth) or inter-iteration filtering.[28,30,31] Although there is no proof of convergence, this algorithm always appears to converge in practice. The image characteristics are somewhat similar to those of penalized likelihood reconstruction.

6.4 Uniform Resolution

For some applications in emission tomography it is desirable that the spatial resolution in the images is predictable, independent of the object and usually also independent of position. Examples are quantitative analysis of PET studies, where one has to assume that changes in numbers are due to changes in metabolism, and not to changes in characteristics of the imaging system.

A problem with most MAP-algorithms (and also with EMS) is that they tend to produce object and position dependent resolution. The likelihood term seeks ideal resolution, while the prior (penalty) wants no resolution. The final resolution depends on the relative strength of both terms. The strength of the prior is usually chosen to be constant, but the strength of the likelihood depends on the object, on the position within the object and even on the orientation. To obtain the same balance between the two everywhere, the strength of the prior has to follow that of the likelihood. Such algorithms have been designed, and compared to post-smoothed ML-EM.[32-34] It is found that at matched resolution, the noise characteristics of these penalized- likelihood algorithms are similar to those of post-smoothed ML-EM. Post-smoothed ML-EM is much easier to program, and is already widely available. However, the penalty improves the conditioning of the inverse problem, which can be exploited to obtain faster convergence.[34] So for applications requiring uniform resolution, penalized-likelihood may still have its role as an acceleration technique.

6.5 Anatomical Priors

Many attempts have been undertaken to exploit anatomical knowledge, available from registered MRI or CT images, during the reconstruction of PET or SPECT data. The aim is to avoid the resolution loss due to the regularization, or even to recover the resolution of the emission reconstruction, by making use of the superior resolution of the anatomical images.

These attempts are only meaningful if one can assume that there is a fair overlap between anatomical and functional (tracer uptake) boundaries. Fortunately, this seems to be a valid assumption in many applications.

The anatomical information has been used to tune the noise suppressing prior in a MAP-algorithm, by limiting smoothing to within organ boundaries revealed by the anatomical data.[35-40] In addition, the segmented MR image can be used for attenuation and scatter correction purposes.[41] If the finite resolution of the emission tomograph is modelled, then these algorithms can produce a strong resolution recovery near anatomical boundaries. To further push resolution recovery, Sastry and Carson[42] introduced a tissue composition model, which considers each (coarse) PET-voxel as composed of one or more tissue classes, which are obtained from a segmented MRI image of the same patient. The PET-activity in every voxel is then a weighted sum of tissue class activities. Figure 4 shows an example obtained with a similar approach.

6.6 *Median Root Prior*

Alenius *et al.*[45,46] have proposed a penalized-likelihood method based on a median filter. The algorithm can be written in the following form:

$$f_j^{new} = \frac{f_j^{old}}{\sum_i a_{ij}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old} + \beta \frac{f_j^{old} - M_j}{M_j}} \tag{22}$$

where M_j is the value of voxel j, obtained by median filtering the image f_j^{old}. Following the OSL interpretation discussed above, the difference between a voxel and the median of its neighbourhood is used as the gradient of some energy function. The derivation is empirical, and in fact, the corresponding energy function does not exist. Hsiao[47] recently proposed a very similar

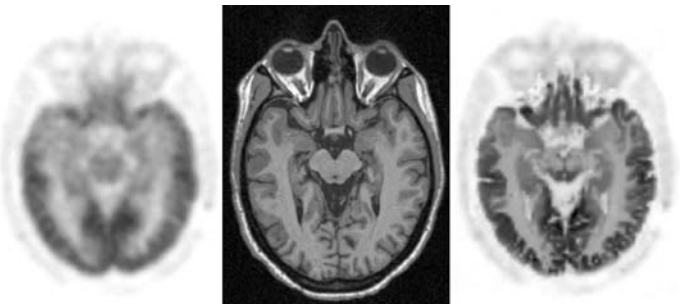


FIGURE 4. PET image obtained with MAP-reconstruction without (left) and with (right) the use of anatomical information from a registered MRI image (center), using a tissue composition model.[43] The 3D sinogram was first rebinned with FORE,[44] then the MAP-reconstructions were computed.[12]

algorithm, which does minimize a well defined objective function. Albeit empirical, the algorithm is very effective and has some interesting features. In contrast to most priors, it does not strive towards a completely flat image (which is not a very realistic prior assumption). Instead, it only requires that the image be locally monotonic. Locally monotonic images do not have small hot spots, but they can have sharp and/or smooth edges. In addition, similar to the relative difference prior, it penalizes relative differences, making β a "unit-less", easily tuned parameter. Stated intuitively, the MRP algorithm essentially suppresses all hot or cold spots that are small compared to the size of its median filter. It follows that, when applied in hot spot detection applications (such as PET whole body imaging in oncology), the mask should be chosen sufficiently small and β not too large.

Figure 5 illustrates the behaviour of ML-EM with moderate post-smoothing, MRP and the relative difference prior. Large homogenous regions (such as the liver and the mediastinum) are much better visualized in the MAP-reconstruction images, in particular with MRP. Smaller structures, such as the blood vessels, are somewhat attenuated by both priors. A very hot and small spot, such as the lesion in the neck, is strongly suppressed by the MRP-penalty. It is better preserved by the relative difference prior, because of its tolerance for large voxel differences. But it is best recovered by the noisier ML-EM image,

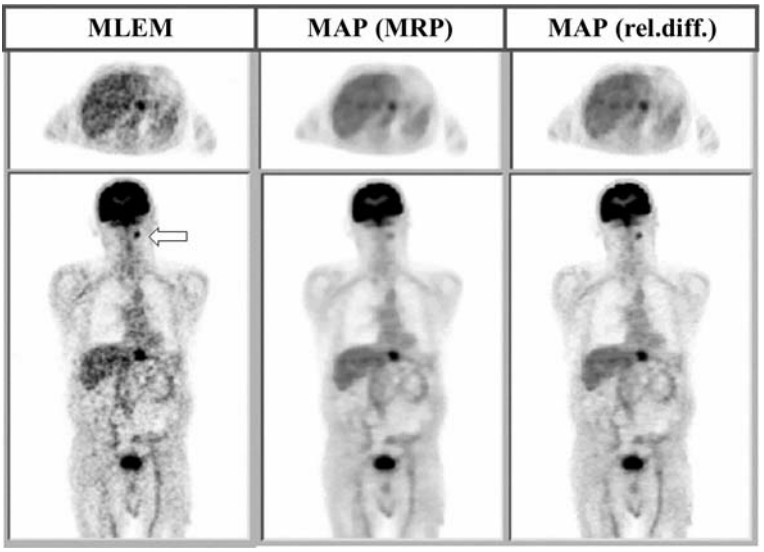


FIGURE 5. Three reconstructions of the same patient's PET scan. Left: post-smoothed ML-EM; center: MAP with median root prior; and right: MAP with relative difference prior (V_R). The median root prior has superior noise suppression and edge preservation for larger structures. However, hot spots that are small compared to the median filter size are suppressed (see the lesion in the neck, indicated by the arrow).

because only moderate smoothing was applied. The penalties should be applied with caution and must be optimized for each particular task.

7. Approaches to Acceleration of ML-EM

7.1 Ordered Subsets EM: OS-EM

The computation time for iterative reconstruction algorithms is typically orders of magnitude longer than for FBP. Fortunately steps can be taken to significantly accelerate these algorithms. The most widely used acceleration technique, perhaps due to its simplicity, is the ordered subsets EM or OS-EM algorithm.[48] Ordered subsets or block-iterative methods have been applied to a number of iterative algorithms with OS-EM being similar in some respects to MART. The essential difference between OS-EM and ML-EM is the use of only a subset of projections for updating rather than comparison of all estimated and measured projections. For OS-EM one iteration is normally considered the use of all data once; consequently use of only part of the data during the update process is termed a sub-iteration. In OS-EM the reconstruction proceeds by utilising subsets of the projections, chosen in a specific order that attempts to maximise the new information being added in sub-iterations. Iteration proceeds by using different projections in each subsequent subset until all projections are used. The resulting equation is very similar to the standard ML-EM equation, the only difference being the use of subsets, S_n, where $n \in N$, the total number of projections divided by the number of projections per subset or subset size:

$$f_j^{new} = \frac{f_j^{old}}{\sum\limits_{i \in S_n} a_{ij}} \sum_{i \in S_n} a_{ij} \frac{p_i}{\sum\limits_k a_{ik} f_k^{old}} \tag{23}$$

Note that sub-iterations recycle through S_n for subsequent iterations. Provided subset size is not too small and subset order is chosen carefully, the reconstruction estimate of OS-EM at each sub-iteration is almost indistinguishable from the reconstruction for a corresponding number of full iterations of ML-EM (Figure 3). The computation time for each sub-iteration is roughly equal to the time for a conventional ML-EM iteration divided by N, due to the reduced number of forward and back projection operations in the subset. Consequently for 128 projections and subset size of 4, the acceleration factor is approximately 32.

There are some theoretical concerns regarding OS-EM. For example, there is no proof that OS-EM in its standard form converges to the same solution as ML-EM. In fact OS-EM reaches a limit cycle, where the result depends to a small extent on the point in the subset sequence at which iteration is stopped. However, with the noisy data typical of clinical practice, this has not been found to be of particular concern, since the possible

solutions tend to be very similar, provided subset size is not too small. Choosing subset order is important as a necessary condition is that there be subset balance, which requires that $\sum_{i \in S_n} a_{ij}$ is independent of n. In practice the use of four well spaced projections is a reasonable compromise between speed and reconstruction quality.[49]

7.2 Variants of OS-EM

There have recently been a number of variants of OS-EM that address some of its theoretical limitations. The rescaled block-iterative (RBI) algorithm[50] does converge to a single solution in the absence of noise, independent of subset order. The equation for RBI is

$$f_j^{new} = f_j^{old} + \frac{f_j^{old}}{\max \sum_{i \in S_n} a_{ij}} \sum_{i \in S_n} a_{ij} \left(\frac{p_i}{\sum_k a_{ik} f_k^{old}} - 1 \right) \tag{24}$$

which can be seen to reduce to exactly OS-EM when $\sum_{i \in S_n} a_{ij}$ is constant for all n. Alternatively the row-action maximum likelihood algorithm (RAMLA) is designed to avoid the limit cycle by using under-relaxation (deliberately decreasing the influence of the update as iteration advances).[51] The equation again has a very similar form:

$$f_j^{new} = f_j^{old} + \lambda_k f_j^{old} \sum_{i \in S_n} a_{ij} \left(\frac{p_i}{\sum_k a_{ik} f_k^{old}} - 1 \right), \quad 0 < \lambda_k \leq \sum_{i \in S_n} a_{ij}^{-1} \tag{25}$$

where λ_k decreases with iteration k, but remains constant throughout the sub-iterations within each main iteration. More recent work suggests that λ could be varied with sub-iteration.[52] An alternative approach to controlling the limit cycle is to increase the subset size with iteration,[48] providing results similar to over-relaxation.[53] The similarity in these algorithms leads to some confusion in their applicability, particularly as some of the theoretical concerns are of little clinical relevance e.g. with noisy data the ML solution is undesirable. Nevertheless, as seen from the above equations, some level of standardisation in the approaches is emerging.

7.3 Acceleration for Regularized Reconstruction Algorithms

The same approaches to acceleration can be applied to regularised statistical algorithms as to ML-EM since their formulations are generally similar. The OS-GP was suggested as the regularised OS-EM equivalent for MAP-OSL reconstruction based on the Gibbs prior.[48] Similarly block-iterative approaches have been used for the basis of developing the RBI-MAP

algorithm[54] and the block sequential regularized EM algorithm (BSREM) as a regularized form of RAMLA.[55] There is some debate regarding the choice of regularized acceleration algorithm, with suggestion that BSREM type algorithms have theoretical advantages.[27,55] Still further work is necessary to confirm which algorithm is optimal for clinical application.[56]

8. Using ML with Non-Poisson Data

8.1 Transmission

The ML-EM algorithm described above is based on the assumption that the data can be considered as linear combinations of the unknown tracer distribution, and that they are Poisson distributed. In transmission tomography, the counts t_i are Poisson distributed, but they are a non-linear function of the image μ_j. For that purpose, dedicated ML and MAP algorithms have been developed for transmission tomography. The likelihood function to be optimised is now:

$$L_t = \sum_i t_i \ln \hat{t}_i - \hat{t}_i, \text{ with } \hat{t}_i = b_i \exp\left(-\sum_j l_{ij}\mu_j\right) + s_i \tag{26}$$

where t_i is the measured transmission count rate, b_i is the blank scan, l_{ij} is the intersection length of projection line i with the image voxel j, and s_i contains additive contributions, such as Compton scatter, random coincidences or emission activity in post- injection transmission scanning in PET.

One can apply the expectation-maximisation (EM) strategy again.[57,58] However, in this case, it leads to a somewhat cumbersome expression. To obtain more elegant algorithms, other optimisation methods have been used. Mumcuoğlu *et al.*[59] proposed a conjugate gradient method. Using de Pierro's surrogate function approach, Fessler[60] derived a suitable class of algorithms. The unconstrained and non-parallel version can be written as:

$$\mu_j^{new} = \mu_j^{old} + \frac{\sum_i l_{ij}(1 - t_i/\hat{t}_i)(\hat{t}_i - s_i)}{d_j} \text{ with } d_j = \sum_i \frac{l_{ij}^2 (t_i - s_i)^2}{\alpha_{ij} t_i} \tag{27}$$

where α_{ij} can be chosen, as long as $\alpha_{ij} \geq 0$ and $\sum_j \alpha_{ij} > 0$.

Setting $\alpha_{ij} = l_{ij}\mu_j^{old} / \sum_k l_{ik}\mu_k^{old}$ yields the convex algorithm of Lange and Fessler,[18] choosing $\alpha_{ij} = l_{ij} / \sum_k l_{ik}$ produces the algorithm used in Nuyts.[61] These algorithms behave similarly to the ML-EM algorithm in emission tomography: they have similar noise propagation; they need a comparable amount of computation time and show similar convergence speed. They can be accelerated with ordered subsets as well.[62]

As in emission tomography, the noise can be suppressed by combining the likelihood with a prior. However, in transmission tomography more prior

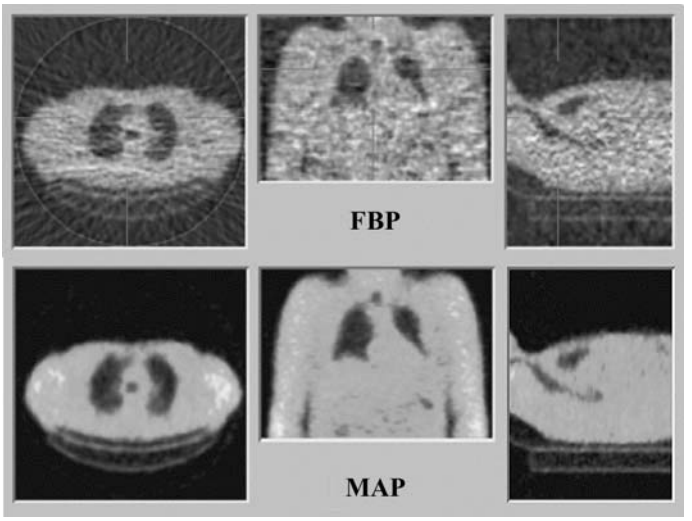


FIGURE 6. Comparison of FBP (above) and MAP (below) reconstruction of very short PET transmission scan.

information is available: the attenuation of tissue and bone is known. In addition, attenuation images are much smoother than emission images. For this reason, more aggressive priors can be applied.[59-61] An example is shown in Figure 6. The effect of these strong priors can be considered as a segmentation, thus producing transmission maps resembling those of the so-called "segmented attenuation correction" algorithms.[63-66] The main difference with the MAP-approach is that the segmentation is done during, not after the reconstruction.

8.2 Corrected Emission

To reduce processing time and data storage, it is often convenient to work with pre-corrected data. For example, careful reconstruction on a PET system requires separate storage of at least three sinograms: one containing the raw, Poisson distributed data, one containing multiplicative effects (including attenuation and detector efficiency) and another one containing additive effects (Compton scatter and randoms). In particular in 3D PET, this represents an impressive amount of data, requiring considerable processing time for reconstruction. An attractive alternative is to produce a single smaller sinogram by pre-correcting the data and applying Fourier rebinning. But then the data are no longer Poisson distributed. It has been shown that ML-EM yields suboptimal images from such data.[12]

An obvious solution is to abandon ML-EM and use a weighted least squares algorithm instead (e.g. Fessler[11]). However, because the ML-EM algorithm is more widely available, an alternative solution has been devised.

It is possible to manipulate the data in order to restore the Poisson distribution approximately, and then undo the manipulation during ML-EM reconstruction. Consider a random number x, and assume that estimates for its mean and variance are available. The distribution can be shifted, such as to make the variance equal to the mean, as required by a Poisson distribution: mean $(x + s) = (x + s)$. It follows that the shift amount to $s = (x) - \text{mean}(x)$. This approach, for example, is used to deal with randoms correction in PET.[67] Alternatively, one can scale x, such that mean$(a\ x) =$ var$(a\ x)$, which yields $a = \text{mean}(x)\ /\ \text{var}(x)$. This approach can be used for reconstruction of transmission scans, if no dedicated transmission algorithm is available.[68]

9. Application to Specific Problems

9.1 More Complex Emission Models

The attraction of iterative reconstruction is that it can be relatively easily modified so as to be applicable to specific emission models. The limitation tends to be that the computational speed is directly related to the complexity of the underlying model. Not only does the computation per iteration increase but also the number of iterations required. Fortunately the combination of faster computers and effective acceleration models means that tackling more complex problems can now be contemplated.

The incorporation of information on distance-dependent resolution in the transition matrix[69,70] results in the iterative reconstruction effectively operating in 3D, allowing for the probability that photons travelling obliquely to the detector still can be detected. Incorporation of resolution loss in the system model in theory suggests that the final reconstruction should have resolution recovery, although in practice this is limited by slow convergence. This will be discussed in more detail elsewhere. The correction of scatter is a particularly demanding problem due to the complex dependence of scatter on the non-homogeneous attenuation in tissue and the broad 3D distribution of scatter. Modelling scatter for incorporation in the transition matrix is difficult since it is object dependent, although appropriate models have been developed[71,72] and efficient Monte Carlo solutions have been proposed.[73] It is worthwhile noting that there are advantages to incorporating measured scatter estimates (such as might be obtained from multiple energy windows) directly in the projector.[74] The incorporation of estimated scatter ($\hat{s}$) in the ML-EM equation is given by:

$$f_j^{new} = \frac{f_j^{old}}{\sum_i a_{ij}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old} + \hat{s}} \tag{28}$$

Since the scatter estimate is simply used to provide consistency between estimated and measured projections the resultant reconstruction is much more resistant to noise than if the scatter was directly subtracted from the measured projections. Indeed the direct subtraction introduces concern that the projections are no longer Poisson distributed (as discussed further below). A similar approach can be taken to account for measured randoms in PET or indeed any source of measured counts that is not accounted for by the transition matrix.

Extensions to the system modelling have been developed to accommodate the more complex situation with either dual radionuclides[75] or multiple-energy radionuclides.[76] In dual radionuclide studies two distributions are simultaneously determined allowing for cross-talk between the radionuclides being imaged. In the case of a multiple-energy radionuclide a single estimate can be determined taking into account the different attenuation that occurs at different emission energies.

A further complexity to iterative reconstruction can be introduced if a fourth dimension is included in the model to account for tracer redistribution with time or motion occurring during gated acquisition. Provided a model of the time variation is well understood this can be incorporated in the transition matrix so as to account for inconsistencies that may otherwise occur between estimated and measured projections. The interested reader is referred to relevant publications.[77,78]

9.2 3D Reconstruction

Some of the complexities mentioned above already suggest the need for a 3D reconstruction (or even in some cases 4D) in order to cater for the possibility of non-planar detection. Clearly this becomes even more evident with detectors that specifically are designed to acquire data in 3D (e.g. cone-beam SPECT, pinhole SPECT or septa-less PET). The attraction of iterative reconstruction is that the basic algorithms do not require any specific change except for definition of the transition matrix, unlike FBP, which usually requires the derivation of specific non-standard filters. A specific example where iterative reconstruction has been widely applied is in 3D PET where usually some form of rebinning is first performed (e.g. Fourier rebinning or FORE) in combination with iterative reconstruction.[44] Alternatives based on direct 3D iterative reconstruction have also been implemented.[79,80]

9.3 Motion Correction

An area where iterative reconstruction is proving particularly useful is the correction of motion, determined either from independent measurement[81] or via estimation directly from the measured projections.[82] Since patient movement can be considered as equivalent to detector movement, the reconstruction necessitates the incorporation of data at specific angular positions

relative to the reconstructed object space. Provided most projections can be considered to correspond to a stationary patient at various locations, reconstruction can be accomplished rather easily using iterative reconstruction with OS-EM selecting subsets of projections so as to correspond to specific patient/detector positions. More exact correction necessitates either multiple short-duration acquisitions to minimise the possibility of movement in an individual acquisition frame or use of list-mode[83] where each individual event can be corrected for movement so as to locate the exact ray-path. The EM list mode equation[84,85] can be formulated to be very similar to the conventional EM equation:

$$f_j^{new} = \frac{f_j^{old}}{\sum_i a_{ij}} \sum_{i \in L} a_{ij} \frac{1}{\sum_k a_{ik} f_k^{old}} \tag{29}$$

Note, however, that the main summation is restricted to the elements in the list-mode data rather than all projection bins; this effectively ignores zero projection bins and consequently the numerator reduces to unity. The OS-EM version simply divides the list-mode data into sub-lists.

10. Clinical Evaluation

Evaluation and clinical validation of image reconstruction algorithms is inherently difficult and sometimes unconvincing. There is a clear need for guidelines to evaluate reconstruction techniques and other image processing issues in emission tomography. A particular concern in clinical studies is the tendency to compare not only different algorithms but different approaches to processing, without effort to isolate the effects due to the reconstruction algorithm itself. Examples are the comparison of iterative algorithms including attenuation correction with analytic approaches without attenuation correction or the comparison of different iterative algorithms where the fundamental implementation differs (e.g. use of "blobs"[86] rather than pixels in the projector). This simply adds to the confusion in interpreting results. A further common problem is the comparison of clinical images where the reconstruction algorithm results in different signal to noise properties, typically dependent on the number of iterations utilised. Evaluation for a range of parameters tends to provide more objective results where the trade-off in noise and signal (e.g. recovery coefficient) can be more meaningfully compared.

Most of the algorithms developed so far have been evaluated using either simulated or experimentally measured phantom studies, in addition to qualitative evaluation of clinical data. This has been extended more recently to objective assessment of image quality using Receiver Operating Characteristics (ROC) analysis based on human or computer observers,[87] evaluation of the influence of reconstruction techniques on tracer kinetic

parameter estimation[88] and voxel-based analysis in functional brain imaging using statistical parametric mapping.[89,90]

10.1 Validation Using Simulation and Phantom Studies

A simulator is an efficient tool that can be used to generate data sets in a controllable manner in order to assess the performance of different reconstruction algorithms. Medical imaging simulation tools have been shown to be very useful for validation and comparative evaluation of image reconstruction techniques since it is possible to obtain a reference image to which reconstructed images should be compared. The ability to theoretically model the propagation of photon noise through emission tomography reconstruction algorithms is crucial in evaluating the reconstructed image quality as a function of parameters of the algorithm. Two broad categories of simulation packages have emerged: simplified analytical modelling tools and sophisticated Monte Carlo simulations. In the first class of simulators, several simplifying approximations are adopted to improve ease of use and speed of operation (e.g. Gaussian noise distribution and scatter model based on an analytical point spread function).[91] On the other hand, the Monte Carlo method is widely used for solving problems involving statistical processes and is very useful in medical imaging due to the stochastic nature of radiation emission, transport and detection processes. Many general purpose and dedicated Monte Carlo programs have been in use in the field of nuclear imaging with many of them available in the public domain.[92,93] Although variance reduction techniques have been developed to reduce computation time, the main drawback of the Monte Carlo method is that it is extremely time-consuming.

Software and physical phantoms used in medical imaging were historically limited to simple point, rod, and slab shapes of sources and attenuating media. Such simple geometries are useful in studying fundamental issues of image reconstruction, but clinically realistic distributions cannot be evaluated by such simple geometries. A precise modeling of the human body requires appropriate information on the location, shape, density, and elemental composition of the organs or tissues. Several software phantoms modeling different parts of the human body have been developed over the years to assess the accuracy of reconstruction procedures and are described in detail in Chapter 11 of this book. Interested readers are also refereed to textbooks referenced above discussing extensively issues related to Monte Carlo modeling in nuclear medicine.

10.2 Subjective and Objective Assessment of Image Quality

Different approaches have been suggested to judge image quality when evaluating image reconstruction algorithms. As the 'best' reconstruction algorithm can only be selected with respect to a certain task, different

'basic' performance measures can be used. Basically, there is no single figure of merit that summarises algorithm performance, since performance ultimately depends on the diagnostic task being performed. Well-established figures of merit known to have a large influence on many types of task performance are generally used to assess image quality.[94]

In a clinical environment, the evaluation is further hampered by the multiplicity of the medical purposes for which the corrections may be studied. In the simplest approach, trained nuclear medicine physicians carry out observer performance studies and are asked to rank images by their degree of quality. A common method to assess image quality with respect to a detection task is the use of observer studies where the performance of any observer (human or computer algorithm) is characterized by ROC analysis. In such studies, observers rate images based on their confidence that a defect/lesion exists in a large set of images. Curve-fitting methods are then used to fit the rating data to receiver operating characteristic (ROC) curves, which plot the covariation in ''true positive'' and ''false positive'' conditional probabilities across changes in the decision-variable's criterion for a ''positive'' binary decision.[87,95,96] The estimated area under the fitted ROC curve is often used as a general index of image quality or performance accuracy for any alternative classification task. ROC and localization ROC (LROC) techniques have been extensively used to evaluate lesion detectability.[97] For example, using simulated MCAT phantom data and randomly located 1 cm-diameter lesions, Lartizien *et al.*[98] demonstrated that the FORE+AWOSEM (attenuation weighted OS-EM) algorithm results in the best overall detection and localization performance for 1-cm-diameter lesions compared with the FORE+OSEM and FORE+FBP algorithms in PET imaging. The major drawback of this approach is that it is costly and complex, since a reasonable number of experienced observers should be used to analyse many images under carefully controlled conditions. In addition, such techniques rely on experimental phantom measurements or simulated data since the ground truth needs to be known. Furthermore, for optimisation of reconstruction algorithms in which possible parameter settings suffer a combinatorial explosion, human psychophysical studies are simply not viable. Therefore, most qualitative assessment studies are restricted to subjective observer rating scores by experienced physicians. The diversity of available algorithms also makes comparison difficult as results can be inconclusive when limited cross-comparison is involved.

There are an increasing number of published articles that verify the benefit of utilising iterative algorithms in clinical situations. These verify many of the theoretical advantages outlined in section 2.2, including ease of incorporating attenuation correction,[99] reduction in streak artefacts,[100] tolerance to missing data[101] and most importantly, noise reduction in low count areas of the image. It should be noted that, since noise generated in iterative reconstruction is signal-dependent, the signal to noise gains are restricted

to low count areas, whereas high count areas may at best be unaffected.[102,103] In many nuclear medicine studies there is interest in derived parameters rather than individual images (e.g. extracted functional parameters or significance maps reflecting changes in activity distribution). It therefore is of some interest to identify which reconstruction algorithms are optimal for these applications and to verify that the reconstruction algorithms provide quantitative results.[104,105] An interesting approach in comparative evaluation studies for functional brain imaging is to carry out voxel-based statistical analysis using statistical parametric mapping (SPM).[106] A recent study on the impact of model-based scatter correction and iterative reconstruction on spatial distribution of ^{18}F-[FDG] in reconstructed brain PET images of healthy subjects using this kind of analysis demonstrated that OS-EM reconstruction does not result in significant changes when compared to FBP reconstruction procedures, while significant differences in ^{18}F-[FDG] distribution arise when images are reconstructed with and without explicit scatter correction for some cerebral areas.[89] Other related studies showed that iterative reconstruction has the potential to increase the statistical power and to give the best trade-off between signal detection and noise reduction in PET activation studies as compared with FBP reconstruction.[90]

11. Future Prospects

The progress in iterative reconstruction has been immense in the past ten years, the main opportunities arising from the availability of both improved processing speed and faster algorithms. This has permitted much more ambitious algorithms that tackle not just conventional 2D reconstruction but a range of applications in 3D (e.g. Liu *et al.*[80]) and even 4D (including the time domain).[78] The appeal of iterative reconstruction in adapting to different acquisition geometry or physical situation has already been well demonstrated. The practicality of performing such reconstruction is revitalising the consideration of alternative approaches to imaging such as use of multiple pinhole collimators or so-called 'Compton' imaging (e.g. Braem *et al.*[107]). Although currently focussed on high resolution imaging for small animals there is potential for future developments to address specific clinical applications.

Anticipating even greater gains in computer power, even more complex and ambitious reconstruction becomes clinically realistic. More extensive use of fully 3D reconstruction algorithms can be expected, with more exact correction likely to provide optimal quality of reconstruction. The feasibility of using Monte Carlo-based approaches has been demonstrated[108] and previous work has demonstrated that symmetries can be used to greatly accelerate the Monte Carlo calculations.[73] Whether this degree of sophistication in accurately modelling the emission and detection processes is warranted in clinical practice remains to be seen.

The demands placed on reconstruction algorithms continue to expand. There is a strong evolution towards more complex acquisition hardware, in particular in PET (more and smaller detector crystals, larger axial extent of the detectors, better energy resolution allowing multi-energy acquisitions). In some of these complex geometries, the assumptions of Fourier rebinning are no longer valid and more accurate rebinning algorithms are needed.[44] The huge size of the data sets compensates for the growing computer speed and, since acceleration of 3D reconstruction has probably reached its limits, the computational demand continues to push technology to its limit.

The availability of dual-modality instruments is likely to influence the approaches to reconstruction, with algorithms that combine anatomical information likely to increase in popularity. These approaches have previously relied on accurate registration and ready access to anatomical data. The advent of dual modality instruments makes this much more practical.

Accompanying these developments is the increasingly difficulty task of understanding the intricacies of the ever-increasing range of algorithms. In the clinical environment the uncertainty in choice of algorithm has acted as a deterrent to their use. To some extent there is 'convergence' in the various algorithms and it is to be anticipated that there may be some future consensus as to the best algorithm to choose for specific applications. The availability of open platforms and more widely available software, specifically designed for clinical application, should at some stage accelerate the acceptance of algorithms. However, at the time of writing OS-EM remains the most widely used iterative algorithm in clinical practice, with variants specifically tailored for use with PET or transmission reconstruction. It is clear, however, that there is great need for development of standard approaches for evaluation of algorithms e.g. the availability of standard validated data for inter-lab comparison.

Acknowledgments. One of the co-authors (HZ) acknowledges the support of Grant SNSF 3152A0-102143 awarded by the Swiss National Science Foundation.

References

1. Lalush D. S., "Iterative image reconstruction" in: *Emission Tomography: The Fundamentals of PET and SPECT*. edited by M Wernick and J Aarsvold Academic Press, San Diego, (2004) *in press*
2. Fessler J. A., "Statistical image reconstruction methods for transmission tomography" in: *Handbook of Medical Imaging*, edited by M. Sonka and J. M. Fitzpatrick SPIE, Bellingham, (2000) *pp* 11-70.
3. Lewitt R. M. and Matej, S., Overview of methods for image reconstruction from projections in emission computed tomography. *Proceedings of the IEEE* **91:** 1588-1611 (2003).

4. Herman G. T. and Meyer, L. B., Algebraic reconstruction techniques can be made computationally efficient. *IEEE Trans Med Imaging* **12:** 600-609 (1993).
5. Bender R., Gordon, R. and Herman, G. T., Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography. *J Theor Biol* **29:** 471-482 (1970).
6. Lu H. H., Chen, C. M. and Yang, I. H., Cross-reference weighted least square estimates for positron emission tomography. *IEEE Trans Med Imaging* **17:** 11-8 (1998).
7. Guan H., Gordon, R. and Zhu, Y., Combining various projection access schemes with the algebraic reconstruction technique for low-contrast detection in computed tomography. *Phys Med Biol* **43:** 2413-2421 (1998).
8. Byrne C. L., Block-iterative methods for image reconstruction from projections. *IEEE Trans Image Proc* **5:** 792-794 (1996).
9. Barrett H. H., Wilson, D. W. and Tsui, B. M. W., Noise properties of the EM algorithm. I. Theory. *Phys Med Biol* **39:** 833-846 (1994).
10. Sauer K. and Bouman, C., A local update strategy for iterative reconstruction from projections. *IEEE Trans Sign Proc* **41:** 534-548 (1993).
11. Fessler J. A., Penalized weighted least-squares image reconstruction for positron emission tomography. *IEEE Trans Med Imaging* **13:** 290290-300 (1994).
12. Comtat C., Kinahan, P. E., Defrise, M. *et al*., Fast reconstruction of 3D PET data with accurate statistical modeling. *IEEE Trans Nucl Sci* **45:** 1083-1089 (1998).
13. Meikle S. R., Fulton, R. R., Eberl, S. *et al*., An investigation of coded aperture imaging for small animal SPECT. *IEEE Trans Nucl Sci* **48:** 816-821 (2001).
14. Kaufman L., Maximum likelihood, least squares, and penalized least squares for PET. *IEEE Trans Med Imaging* **12:** 200-214 (1993).
15. Chinn G. and Huang, S. C., A general class of preconditioners for statistical iterative reconstruction of emission computed tomography. *IEEE Trans Med Imaging* **16:** 11-10 (1997).
16. Dempster A. P., Laird, N. M. and Rubin, D. B., Maximum likelihood from incomplete data via the EM algorithm. *J Royal Stat Soc B* **38:** 11-38 (1977).
17. Shepp L. A. and Vardi, Y., Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imag* **1:** 113-122 (1982).
18. Lange K. and Fessler, J. A., Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Trans Imag Proc* **4:** 1430-1438 (1995).
19. Miller M. I., Snyder, D. L. and Miller, T. R., Maximum likelihood reconstruction for single-photon emission computed tomography. *IEEE Trans Nucl Sci* **32:** 769769- 778 (1985).
20. Wilson D. W., Tsui, B. M. W. and Barrett, H. H., Noise properties of the EM algorithm. II. Monte Carlo simulations *Phys Med Biol* **39:** 847-871 (1994).
21. Beekman F. J., Slijpen, E. T. and Niessen, W. J., Selection of task-dependent diffusion filters for the post–processing of SPECT images. *Phys Med Biol* **43:** 1713-1730 (1998).
22. Geman S. and McClure, D. E., Statistical methods for tomographic image reconstruction. *Bull Int Stat Inst* **52-4:** 55-21 (1987).
23. Mumcuoglu E. U., Leahy, R. M. and Cherry, S. R., Bayesian reconstruction of PET images: methodology and performance analysis. *Phys Med Biol* **41:** 1777-1807 (1996).

24. de Pierro A. R., On the convergence of an EM-type algorithm for penalized likelihood estimation in emission tomography. *IEEE Trans Med Imaging* **14:** 762-765 (1995).
25. Nuyts J., Beque, D., Dupont, P. *et al.*, A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography. *IEEE Trans Nucl Sci* **49:** 56-60 (2002).
26. Green P. J., Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Trans Med Imaging* **9:** 84-93 (1990).
27. Ahn S. and Fessler, J. A., Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms. *IEEE Trans Med Imaging* **22:** 613-626 (2003).
28. Slijpen E. T. P. and Beekman, F. J., Comparison of post-filtering and filtering between iterations for SPECT reconstruction. *IEEE Trans Nucl Sci* **46:** 2233-2238 (1999).
29. Nuyts J., On estimating the variance of smoothed MLEM images *IEEE Trans Nucl Sci* **49:** 714-721 (2002).
30. Silverman B. W., Jones, M. C., Wilson, J. D. *et al.*, A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography. *J R Statist Soc B* **52:** 271-324 (1990).
31. Mustafovic S., Thielemans, K., Hogg, D. *et al.*, ''Object dependency of resolution and convergence rate in OSEM with filtering'' Nuclear Science Symposium Conference Record, 2001 IEEE, Vol. 3; *pp* 1786-1790 (2001).
32. Fessler J. A. and Rogers, W. L., Spatial resolution properties of penalized-likelihood image reconstruction: space-invariant tomographs. *IEEE Trans Imag Proc* **5:** 1346-1358 (1996).
33. Nuyts J. and Fessler, J. A., A penalized-likelihood image reconstruction method for emission tomography, compared to postsmoothed maximum-likelihood with matched spatial resolution. *IEEE Trans Med Imaging* **22:** 1042-1052 (2003).
34. Stayman J. W. and Fessler, J. A., Compensation for nonuniform resolution using penalized-likelihood reconstruction in space-variant imaging systems. *IEEE Trans Med Imaging* **23:** 269-284 (2004).
35. Johnson V. E., Wong, W. H., Hu, X. *et al.*, Image restoration using Gibbs priors: boundary modeling, treatment of blurring, and selection of hyperparameter. *IEEE Trans Pattern Anal Machine Intell* **13:** 413-425 (1991).
36. Gindi G., Lee, M., Rangarajan, A. *et al.*, Bayesian reconstruction of functional images using anatomical information as priors. *IEEE Trans Med Imaging* **12:** 670-680 (1993).
37. Ardekani B. A., Braun, M., Hutton, B. F. *et al.*, Minimum cross-entropy reconstruction of PET images using prior anatomical information. *Phys Med Biol* **41:** 2497-2517 (1996).
38. Som S., Hutton, B. F. and Braun, M., Properties of minimum cross-entropy reconstruction of emission tomography with anatomically based prior. *IEEE Trans Nucl Sci* **45:** 3014-3021 (1998).
39. Rangarajan A., Hsiao, I.–T. and Gindi, G., A bayesian joint mixture framework for the integration of anatomical information in functional image reconstruction. *J Mat Imag Vision* **12:** 119-217 (2000).
40. Comtat C., Kinahan, P. E., Fessler, J. A. *et al.*, Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels. *Phys Med Biol* **47:** 11-20 (2002).

41. Zaidi H., Montandon, M.-L. and Slosman, D. O., Magnetic resonance imaging-guided attenuation correction in 3D brain positron emission tomography. *Med Phys* **30:** 937-948 (2003).
42. Sastry S. and Carson, R. E., Multimodality Bayesian algorithm for image reconstruction in positron emission tomography: a tissue composition model. *IEEE Trans Med Imaging* **16:** 750-761 (1997).
43. Baete K., Nuyts, J., Van Paesschen, W. *et al*., "Anatomical based FDG-PET reconstruction for the detection of hypometabolic regions in epilepsy" *Nuclear Science Symposium Conference Record*, 2002 IEEE, Vol. 3; *pp* 1481-1485 (2002).
44. Defrise M., Kinahan, P. E., Townsend, D. W. *et al*., Exact and approximate rebinning algorithms for 3-D PET data. *IEEE Trans Med Imaging* **16:** 145-158 (1997).
45. Alenius S. and Ruotsalainen, U., Bayesian image reconstruction for emission tomography based on median root prior. *Eur J Nucl Med* **24:** 258-265 (1997).
46. Alenius S. and Ruotsalainen, U., Generalization of median root prior reconstruction. *IEEE Trans Med Imaging* **21:** 1413-1420 (2002).
47. Hsiao I. T., Rangarajan, A. and Gindi, G., A new convex edge-preserving median prior with applications to tomography. *IEEE Trans Med Imaging* **22:** 580-585 (2003).
48. Hudson H. M. and Larkin, R. S., Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* **13:** 601-609 (1994).
49. Hutton B. F., Hudson, H. M. and Beekman, F. J., A clinical perspective of accelerated statistical reconstruction. *Eur J Nucl Med* **24:** 797-808 (1997).
50. Byrne C. L., Accelerating the EMML algorithm and related iterative algorithms by rescaled block-iterative methods. *IEEE Trans Imag Proc* **7:** 100-109 (1998).
51. Browne J. and de Pierro, A. B., A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *IEEE Trans Med Imaging* **15:** 687-699 (1996).
52. Tanaka E. and Kudo, H., Subset-dependent relaxation in block-iterative algorithms for image reconstruction in emission tomography. *Phys Med Biol* **48:** 1405-1422 (2003).
53. Schmidlin P., Bellemann, M. E. and Brix, G., Subsets and overrelaxation in iterative image reconstruction. *Phys Med Biol* **44:** 1385-1396 (1999).
54. Lalush D. S., Frey, E. C. and Tsui, B. M., Fast maximum entropy approximation in SPECT using the RBI-MAP algorithm. *IEEE Trans Med Imaging* **19:** 286-294 (2000).
55. de Pierro A. R. and Yamagishi, M. E. B., Fast EM-like methods for maximum "a posteriori" estimates in emission tomography. *IEEE Trans Med Imaging* **20:** 280-288 (2001).
56. Leahy R. and Byrne, C. L., Recent developments in iterative image reconstruction for PET and SPECT. *IEEE Trans Med Imaging* **19:** 257-260 (2000).
57. Lange K. and Carson, R., EM reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr* **8:** 306-316 (1984).
58. Manglos S. H., Gagne, G. M., Krol, A. *et al*., Transmission maximum-likelihood reconstruction with ordered subsets for cone beam CT. *Phys Med Biol* **40:** 1225-1241 (1995).
59. Mumcuoglu E. U., Leahy, R., Cherry, S. R. *et al*., Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images. *IEEE Trans Med Imaging* **13:** 687-701 (1994).

60. Fessler J. A., Ficaro, E. P., Clinthorne, N. H. *et al.*, Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction. *IEEE Trans Med Imaging* **16:** 166-175 (1997).
61. Nuyts J., Dupont, P., Stroobants, S. *et al.*, Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. *IEEE Trans Med Imaging* **18:** 393-403 (1999).
62. Lalush D. S. and Tsui, B. M., Performance of ordered-subset reconstruction algorithms under conditions of extreme attenuation and truncation in myocardial SPECT. *J Nucl Med* **41:** 737-744 (2000).
63. Meikle S. R., Dahlbom, M. and Cherry, S. R., Attenuation correction using count-limited transmission data in positron emission tomography. *J Nucl Med* **34:** 143-150 (1993).
64. Tai Y.-C., Lin, K.-P., Dahlbom, M. *et al.*, A hybrid attenuation correction technique to compensate for lung density in 3-D total body PET. *IEEE Trans Nucl Sci* **43:** 323-330 (1996).
65. Xu M., Cutler, P. and Luk, W., An adaptive local threshold segmented attenuation correction method for whole-body PET imaging. *IEEE Trans Nucl Sci* **43:** 331-336 (1996).
66. Zaidi H., Diaz-Gomez, M., Boudraa, A. E. *et al.*, Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging. *Phys Med Biol* **47:** 1143-1160 (2002).
67. Yavuz M. and Fessler, J. A., Statistical image reconstruction methods for randoms-precorrected PET scans. *Med Image Anal* **2:** 369-378 (1998).
68. Nuyts J., Michel, C. and Dupont, P., Maximum-likelihood expectation-maximization reconstruction of sinograms with arbitrary noise distribution using NEC-transformations. *IEEE Trans Med Imaging* **20:** 365-375 (2001).
69. Tsui B. M., Frey, E. C., Zhao, X. *et al.*, The importance and implementation of accurate 3D compensation methods for quantitative SPECT. *Phys Med Biol* **39:** 509-530 (1994).
70. Hutton B. F. and Lau, Y. H., Application of distance-dependent resolution compensation and post-reconstruction filtering for myocardial SPECT. *Phys Med Biol* **43:** 1679-1693 (1998).
71. Beekman F. J., Kamphuis, C. and Viergever, M. A., Improved SPECT quantitation using fully three-dimensional iterative spatially variant scatter response compensation. *IEEE Trans Med Imaging* **15:** 491-499 (1996).
72. Frey E. C. and Tsui, B. M. W., A practical method for incorporating scatter in a projector-backprojector for accurate scatter compensation in SPECT. *IEEE Trans Nucl Sci* **40:** 1107-1116 (1993).
73. Beekman F. J., de Jong, H. W. and van Geloven, S., Efficient fully 3-D iterative SPECT reconstruction with Monte Carlo-based scatter compensation. *IEEE Trans Med Imaging* **21:** 867-877 (2002).
74. Bowsher J. E. and Floyd, C. E., Jr., Treatment of Compton scattering in maximum-likelihood, expectation-maximization reconstructions of SPECT images. *J Nucl Med* **32:** 1285-1291 (1991).
75. de Jong H. W., Beekman, F. J., Viergever, M. A. *et al.*, Simultaneous (99m)Tc/(201)Tl dual-isotope SPET with Monte Carlo-based down-scatter correction. *Eur J Nucl Med Mol Imaging* **29:** 1063-1071 (2002).

76. Hudson H. M., Hutton, B. F., Larkin, R. S. *et al*., Investigation of multiple energy reconstruction in SPECT using MLEM. *J Nucl Med* **37:** 171P (abstract) (1996).
77. Carson R. E. and Lange, K., The EM parametric image reconstruction algorithm. *J Am Statist Assoc* **80:** 20-22 (1985).
78. Lalush D. S. and Tsui, B. M., Block-iterative techniques for fast 4D reconstruction using a priori motion models in gated cardiac SPECT. *Phys Med Biol* **43:** 875-886 (1998).
79. Matej S. and Browne, J. A., ''Performance of a fast maximum likelihood algorithm for fully 3D PET reconstruction'' in: *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, edited by P. Grangeat and J-L. Amans Kluwer Academic Publishers, Dordrecht, (1996) *pp* 297-315.
80. Liu X., Comtat, C., Michel, C. *et al*., Comparison of 3-D reconstruction with 3D-OSEM and with FORE+OSEM for PET. *IEEE Trans Med Imaging* **20:** 804-814 (2001).
81. Fulton R. R., Hutton, B. F., Braun, M. *et al*., Use of 3D reconstruction to correct for patient motion in SPECT. *Phys Med Biol* **39:** 563-574 (1994).
82. Hutton B. F., Kyme, A. Z., Lau, Y. H. *et al*., A hybrid 3-D reconstruction/registration algorithm for correction of head motion in emission tomography. *IEEE Trans Nucl Sci* **49:** 188-194 (2002).
83. Bloomfield P. M., Spinks, T. J., Reed, J. *et al*., The design and implementation of a motion correction scheme for neurological PET. *Phys Med Biol* **48:** 959-978 (2003).
84. Barrett H. H., White, T. and Parra, L. C., List-mode likelihood. *J Opt Soc Am A* **14:** 2914-2923 (1997).
85. Reader A. J., Erlandsson, K., Flower, M. A. *et al*., Fast accurate iterative reconstruction for low-statistics positron volume imaging. *Phys Med Biol* **43:** 835-846 (1998).
86. Matej S. and Lewitt, R. M., Practical considerations for 3-D image reconstruction using spherically symmetric volume elements. *IEEE Trans Med Imaging* **15:** 68-78 (1996).
87. Swensson R. G., Unified measurement of observer performance in detecting and localizing target objects on images. *Med Phys* **23:** 1709-1725 (1996).
88. Boellaard R., van_Lingen, A. and Lammertsma, A. A., Experimental and clinical evaluation of iterative reconstruction (OSEM) in dynamic PET: quantitative characteristics and effects on kinetic modeling. *J Nucl Med* **42:** 808-817 (2001).
89. Montandon M.-L., Slosman, D. O. and Zaidi, H., Assessment of the impact of model-based scatter correction on ^{18}F-[FDG] 3D brain PET in healthy subjects using statistical parametric mapping. *Neuroimage* **20:** 1848-1856 (2003).
90. Mesina C. T., Boellaard, R., Jongbloed, G. *et al*., Experimental evaluation of iterative reconstruction versus filtered backprojection for 3D [15O]water PET activation studies using statistical parametric mapping analysis. *Neuroimage* **1:** 1170-1179 (2003).
91. Beekman F. J. and Viergever, M. A., Fast SPECT simulation including object shape dependent scatter. *IEEE Trans Med Imag* **14:** 271-282 (1995).
92. Ljungberg M., Strand, S.-E. and King, M. A., ''Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging'' (Institute of Physics Publishing, Bristol, 1998).

93. Zaidi H. and Sgouros, G., "Therapeutic applications of Monte Carlo calculations in nuclear medicine" (Institute of Physics Publishing, Bristol, 2002).
94. Barrett H. H., Denny, J. L., Wagner, R. F. *et al.*, Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance. *J Opt Soc Am A* **12:** 834-852 (1995).
95. Metz C. E., ROC methodology in radiologic imaging. *Invest Radiol* **21:** 720-733 (1986).
96. Barrett H. H., Abbey, C. K. and Clarkson, E., Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions. *J Opt Soc Am A Opt Image Sci Vis* **15:** 1520-1535 (1998).
97. Gifford H. C., King, M. A., de_Vries, D. J. *et al.*, Channelized hotelling and human observer correlation for lesion detection in hepatic SPECT imaging. *J Nucl Med* **41:** 514-521 (2000).
98. Lartizien C., Kinahan, P. E., Swensson, R. *et al.*, Evaluating image reconstruction methods for tumor detection in 3-dimensional whole-body PET oncology imaging. *J Nucl Med* **44:** 276-290 (2003).
99. Lonneux M., Borbath, I., Bol, A. *et al.*, Attenuation correction in whole-body FDG oncological studies: the role of statistical reconstruction. *Eur J Nucl Med* **26:** 591591-598 (1999).
100. Blocklet D., Seret, A., Popa, N. *et al.*, Maximum-likelihood reconstruction with ordered subsets in bone SPECT. *J Nucl Med* **40:** 19781978-1984 (1999).
101. Hatton R., Hutton, B. F. and Angelides, S., Improved tolerance to missing data in myocardial perfusion SPECT using OSEM reconstruction. *Eur J Nucl Med Mol Imaging* **31:** 857-861 (2004).
102. Riddell C., Carson, R. E., Carrasquillo, J. A. *et al.*, Noise reduction in oncology FDG PET images by iterative reconstruction: a quantitative assessment. *J Nucl Med* **42:** 1316-1323 (2001).
103. Fulton R. R., Meikle, S. R., Eberl, S. *et al.*, Signal to noise ratio in neurological FDG PET: Comparison of FBP and OS-EM reconstruction algorithms. *J Nucl Med* **37:** 220P (abstract) (1996).
104. Visvikis D., Cheze_LeRest, C., Costa, D. C. *et al.*, Influence of OSEM and segmented attenuation correction in the calculation of standardised uptake values for [18F]FDG PET. *Eur J Nucl Med* **28:** 1326-1335 (2001).
105. Iida H. and Eberl, S., Quantitative assessment of regional myocardial blood flow with thallium-201 and SPECT. *J Nucl Cardiol* **5:** 313-331 (1998).
106. Friston K., Holmes, A., Worsley, K. *et al.*, Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping* **2:** 189-210 (1995).
107. Braem A., Chamizo Llatas, M., Chesi, E. *et al.*, Feasibility of a novel design of high-resolution parallax-free Compton enhanced PET scanner dedicated to brain research. *Phys Med Biol* **49:** 2547-2562 (2004).
108. Buvat I., Lazaro, D. and Breton, V., "Fully 3D Monte Carlo reconstruction in SPECT: proof of concept and is that worthwhile?" *Conf. Proc. of the VIIth International Meeting on Fully Three-dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 29 June-4 July 2003, Saint-Malo, France. Available on CD-ROM.

5
Collimator-Detector Response Compensation in SPECT

E.C. Frey* and B.M.W. Tsui*

1. Introduction

The collimator-detector response (CDR) of a single-photon nuclear medicine imaging system refers to the image generated from a point source of activity. The shape of the CDR is the primary factor determining the image resolution in SPECT. As a result, a great deal of work has gone into developing methods to compensate for the CDR. The goal of this chapter is to first discuss the CDR, its components, and their effect on SPECT images. Next, compensation methods, both analytical and iterative, are discussed followed by a discussion of the efficacy of compensation methods.

2. The Collimator Detector Response

The CDR of a single-photon nuclear medicine imaging system has 4 components: the intrinsic response, due to uncertainty in position estimation in the detector system, and the geometric, septal penetration and septal scatter components that depend on the collimator characteristics. Suppose that we have a point at position $\vec{r}$ and we wish to know the probability that photons emitted at this point will be detected at some point in the detection plane, $\vec{x}$. This is described by the collimator-detector response function (CDRF), $d(\vec{x}, \vec{r})$. We can decompose the CDRF into several components:

$$d(\vec{x}, \vec{r}) = \int\int i(\vec{x},\vec{x}')(g(\vec{x}', \vec{r}) + p(\vec{x}', \vec{r}) + s(\vec{x}', \vec{r}))d\vec{x}' \tag{1}$$

In Eq. 1, $i(\vec{x}, \vec{x}')$ is the intrinsic point-source response function of the camera (IRF) and describes the probability that a photon incident at position $\vec{x}'$ is detected at $\vec{x}$,$g(\vec{x}',\vec{r})$, $p(\vec{x}', \vec{r})$, and $s(\vec{x}', \vec{r})$ are the geometric (GRF), septal

*Drs. E.C. Frey and B.M.W. Tsui, Division of Medical Imaging Physics, The Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, MD, 21287–0859, USA

penetration (SPRF), and septal scatter (SSRF) response functions describing the probability that a photon emitted at $\vec{r}$ will pass through the collimator holes (GRF), pass through the septa without interacting (SPRF), or scatter in the septa (SSRF), and result in a photon incident on the detection plane at position $\vec{x}'$. For the purposes of this chapter, we adopt the convention (which is not generally adopted in the literature) that the response functions are normalized for a source with unit intensity and we can thus think of the response functions as probability densities. In this convention the sensitivity of the collimator-detector is included in the CDRF and, in fact, the integral of the CDRF is related to the sensitivity by a factor that depends on the units used for activity and time.

It is often assumed that the response functions are spatially invariant in planes parallel to the face of the collimator and that the intrinsic response is spatially invariant in the detection plane. In this case, we can write Eq. 1 as:

$$d(\vec{x};D) = i(\vec{x}) \otimes (g(\vec{x};D) + p(\vec{x};D) + s(\vec{x};D)) \tag{2}$$

where D is the distance from the detection plane to the plane containing the source. In the following we discuss the individual components of the CDR and the validity of the assumption of spatial invariance in planes parallel to the detector.

2.1 Intrinsic Response Function

The intrinsic response function (IRF) is the response of the scintillation camera, excluding the collimator, to a pencil beam of radiation. It is determined by two factors: the uncertainty of position estimation in the camera-detector system and the effects of scattering in the crystal. For low energy incident photons the scattering in the crystal is a relatively small effect. However, at higher energies, e.g. for medium energy photons emitted from isotopes such as In-111, scattering in the crystal becomes important. The uncertainty of light estimation is determined by the noise in the signals from the photomultiplier tubes (PMTs), resulting from the statistical variation in the production and collection of scintillation photons produced in the gamma camera, and the position estimation method used. For higher energy photons, a significant fraction of photons undergo Compton scattering in the crystal and the resulting spread in energy deposition also contributes to the intrinsic resolution. The integral of the IRF represents the efficiency of the crystal in detecting photons and is a function of the energy of the incident photon, the energy window used, and the thickness and composition of the crystal.

Typically the full-width-at-half-maximum (FWHM) of the IRF, often referred to as the intrinsic resolution, of modern gamma cameras is better than 4 mm. There is typically significant spatial variation due to the varying efficiency of light collection. For example, in acceptance testing recently

performed on a camera delivered to our institution, the intrinsic resolution varied from 3.1 to 4.6 mm over the face of the camera. As described below, the IRF is much narrower than the other (distant-dependent) components of the collimator-detector response for the distances from the collimator to the source typical of SPECT. As a result, the assumption of spatial invariance of the intrinsic response is often adequate.

2.2 Geometric Response Function

The geometric response function (GRF) is the portion of the total collimator response function that represents the distribution of detected photons that travel through the collimator holes without interacting with or passing through the collimator septa. The integral of the geometric response function, when normalized according to the convention described above, gives the geometric efficiency of the collimator.

Of the three components of the collimator response, the geometric response is the easiest to treat theoretically. In general, the GRF is determined by two factors: the GRF of a given hole and the hole pattern of the collimator. In collimators designed for low energy photons, the septal thickness is generally small compared to the intrinsic resolution and the aspect ratio of the collimator, i.e., the ratio of hole size to hole length, is usually large. In these cases, the average GRF, obtained by averaging the GRFs from all possible source positions relative to the center of a collimator hole, is often used. This averaging can be carried out during the derivation of the GRF and is equivalent to moving the collimator during acquisition. Formulations for the GRF have been derived using this moving collimator approximation for arbitrary hole shapes including parallel-[1], cone-, and fan-beam collimators.[2,3] Recently, a formulation has been developed to compute the collimator response function without the averaging.[4,5]

For parallel-hole collimators, the Fourier transform of the GRF resulting from the moving collimator approximation is related to the product of the Fourier transform of the aperture function which describes the collimator holes:

$$G(\vec{\nu};D) = \varepsilon \left| \frac{A\left(\frac{D+L+B}{L}\vec{\nu}\right)}{A(0)} \right|^2 \tag{3}$$

where $G(\vec{\nu};D)$ is the 2D Fourier Transform of the GRF for a source at distance D from the face of the collimator, $A(\vec{\nu})$ is the 2D Fourier Transform of the aperture function, L is the collimator thickness, B is the distance from the back face of the collimator to the detection plane, ε is the efficiency of the collimator, and $\vec{\nu}$ is the spatial frequency.

From Eq. 3, we see that the shape of the GRF remains essentially the same while the size of the GRF is linearly related to the distance, D. Also, the GRF will have the same symmetry as the aperture holes. Thus, except for

round holes, the GRF is not radially symmetric. However, it has been shown that the GRF for hexagonal holes can be approximated using the GRF for round holes having the same hole area.[1] The use of round holes has the advantage that the spatial version of the response function can be computed analytically:

$$g(r;D) = \frac{\varepsilon}{\pi}\left(2\cos^{-1}\frac{|r_T|}{2R} - \frac{|r_T|}{R}\sqrt{1-\frac{|r_T|^2}{4R^2}}\right) \tag{4}$$

where R is the radius of the collimator hole, r is the distance in the detection plane from the intersection of the line perpendicular to the detection plane containing the source, and r_T describes the displacement of the centers of the projected aperture functions from the front face and back face of the collimator and is given by:

$$r_T = r\frac{L}{D+L+B} \tag{5}$$

Figure 1 shows an example of GRF for a LEHR collimator with round holes ($L = 4.1$ cm, $R = 0.19$ cm, $B = 0.64$ cm) at a distance of $D = 10$ cm from the face of the collimator and the same GRF convolved with a 4 mm FWHM Gaussian IRF. Figure 2 shows a profile through the center of these response functions in addition to a Gaussian fit to each. Note that the GRF alone is not fit very well by a Gaussian but when convolved with the IRF, the combination is reasonably Gaussian in shape. Figure 3 shows the FWHM of the GRF and the GRF convolved with a Gaussian IRF as a function of distance. Note that the latter function can be well fit by a function having the form:

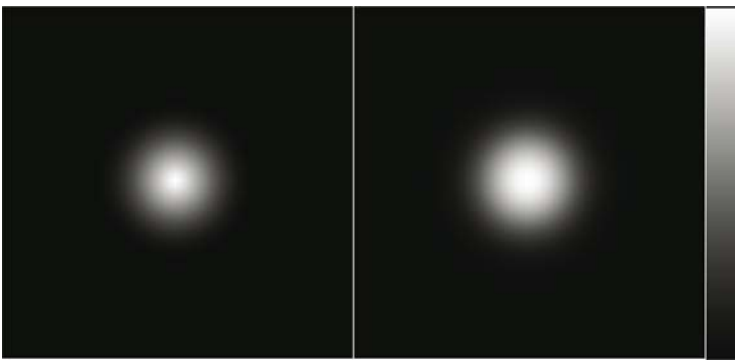


FIGURE 1. Image showing the GRF (left) and the GRF convolved with the IRF (right) for a LEHR collimator with circular holes at a distance of 10 cm from the face of the collimator. The side length of the image is 3 cm. The FWHMs of the response functions shown in these images are (left to right) 5.7 and 7.5 mm.

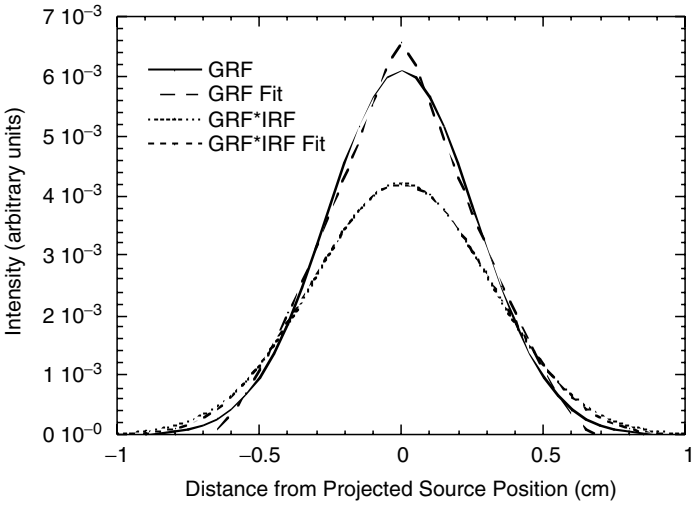


FIGURE 2. Profiles and Gaussian fits to profiles through center of the GRF and the GRF convolved with the IRF (GRF*IRF) shown in Fig. 1.

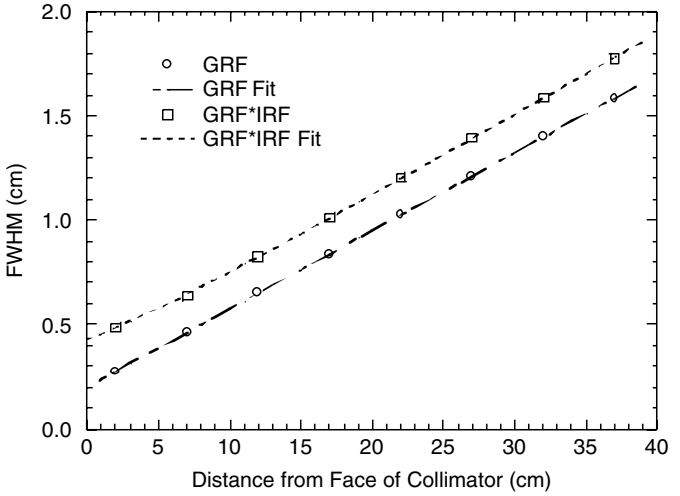


FIGURE 3. Plot of FWHM of GRF and GRF convolved with the IRF for the same LEHR collimator as shown in Fig. 1.

$$FWHM_{GRF^*IRF}(D) = \sqrt{(aD+b)^2 + c^2} \tag{6}$$

where a, b, and c are fitting parameters. Note that if the GRF was modelled by a Gaussian with a width linearly related to D, then fitting parameters a, b, and c can be interpreted as the resolution gradient, the FWHM of the GRF

at the face of the collimator, and the FWHM of the intrinsic resolution, respectively.

In understanding the properties of SPECT reconstructed images, it is often useful to know the behaviour of the GRF in the frequency domain. The geometric transfer function (GTF) is the Fourier Transform of the GRF divided by the geometric efficiency. For a round collimator hole, from Eq. 3 we find that the GTF is given by:

$$\begin{aligned} \mathrm{GTF}(\nu) &= 4\frac{\mathrm{J}_1^2(\beta\nu)}{\beta^2\nu^2}, \text{ and} \\ \beta &= 2\pi R\left(1+\frac{(D+B)}{L}\right), \end{aligned} \tag{7}$$

where $J_1(x)$ is the Bessel function of order 0, ν is spatial frequency and the other symbols are described above. Note that the GTF in Eq. 7 has the value 1 at $\nu = 0$. A sample profile of the GTF for round collimator holes for a LEHR and LEGP collimator at a distance of 10 cm is shown in Figure 4. From this, we find the frequency at which the GTF has its first zero, ν_0, which is given by:

$$\nu_0 = \frac{3.8312}{2\pi R\left(1+\frac{(D+B)}{L}\right)} \tag{8}$$

To a good approximation, there is very little information for frequencies above this retained in the projection of the point source at a distance D from the face of the collimator.

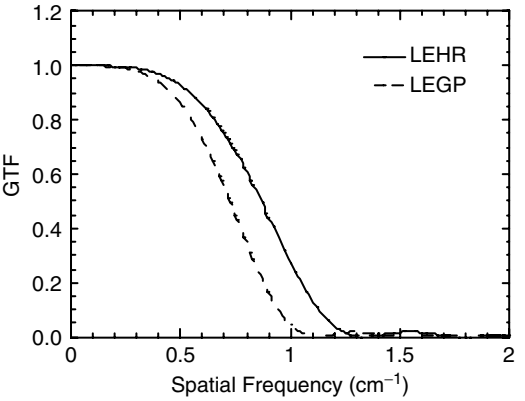


FIGURE 4. Plot of profile through the center of the GTF for LEHR and LEGP collimators with round holes for a point source at 10 cm from the face of the collimator.

2.3 *Septal Penetration and Septal Scatter Response Functions*

The septal penetration response function (SPRF) describes the contribution to the total collimator-detector response by photons that penetrate through the collimator septa. To date there has been no analytic treatment of the septal penetration response for multi-hole collimators. However, the SPRF can be studied using ray tracing or Monte Carlo (MC) simulation techniques.[66-9] The septal scatter response function (SSRF) describes photons that scatter in the collimator septa and are subsequently detected in the energy window of interest. The SSRF is even more difficult to treat theoretically, but can also be analyzed using MC simulation techniques. Both of these effects are most important for isotopes emitting medium and high energy photons such as In-111, Ga-67, or I-131.

Figure 5 shows images representing the total and various components of the response function for a high-energy general-purpose (HEGP) and medium-energy general-purpose collimator (MEGP) for a 364 keV source at 10 cm from the face of the collimator. The collimators had hexagonal holes. These response functions were obtained using MC simulation methods and include the effects of scatter in the detector, but not the portion of the IRF due to imprecision in estimating the interaction position. Also note that these images are shown on a logarithmic grey scale in order to emphasize the features in the tails of the response functions. Note that, for

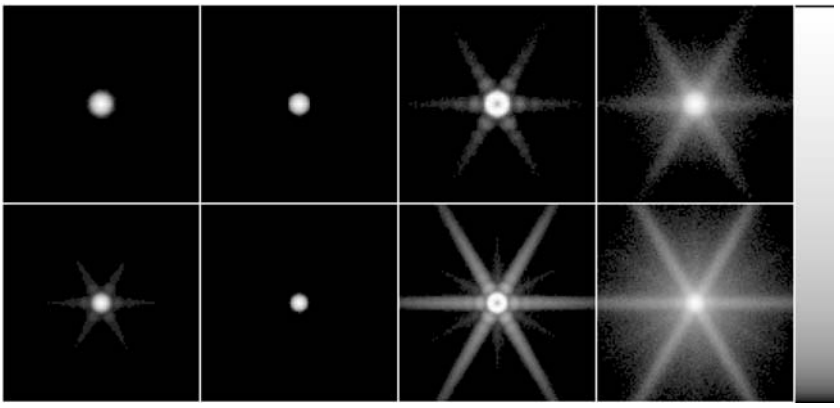


FIGURE 5. Images showing (from left to right) TRF, GRF, SPRF, SSRF in a HEGP (top row) and MEGP (bottom row) collimator for imaging a 364 keV source at 10 cm from the collimator face. These response functions were obtained using a moving collimator and MC simulation techniques. They do not include the effects of the IRF, though do include scattering within the scintillation crystal and in the area of the camera behind the crystal. The images are individually normalized and shown on a logarithmic greyscale to emphasize the features in the tails of the response functions. The side length of the image is 20 cm.

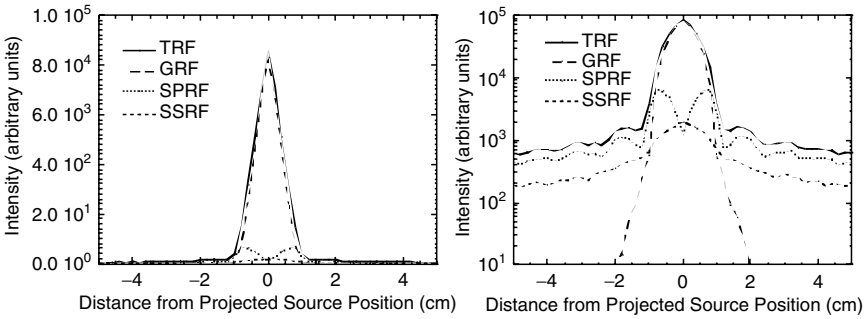


FIGURE 6. Horizontal profile through the response function images of the MEGP collimator shown in Fig. 5 plotted in linear (left) and logarithmic (right) horizontal scales.

the HEGP collimator, the major effect of penetration and scatter is to widen the tail of the response function: the shape of the central peak of the response function is largely unchanged and is similar to that for the geometric response. For the MEGP collimator, it is possible to see the SPRF and SSRF in the total response function (TRF), but the effect is relatively minor. For both collimators, the SPRF has similar features, though is much wider and larger in magnitude for the MEGP than the HEGP collimator for the same 364 keV photons.

In Figure 5, one also notes the significant effects of the hole shape and hole lattice in the orientation of the spokes in both the SPRF and SSRF. To illustrate these effects more quantitatively, horizontal profiles through the SPRF image are plotted in Figure 6. Note that scatter and penetration increase the width of the TRF compared to the GRF. While the height of the SPRF and SSRF tails may seem small, the geometric-to-total ratios, i.e., the ratios of the integrals of SPRF and SSRF divided by the integral of the GRF; at 10 mm are .695 and .446 for the HEGP and MEGP collimators, respectively. This means that a very large fraction of the detected photons have passed through or scattered in the collimator septa. However, the photons that penetrate or scatter and are detected near the GRF will have less impact on image quality than those far off in the tails.

3. Effect of Collimator-Detector Response on SPECT Images

As demonstrated in section 2, the collimator-detector response depends on the source-to-collimator distance. As a result, the effect on SPECT images is relatively complex. Consider first low energy photons where the penetration and collimator scatter components can be neglected. Figure 7 shows the

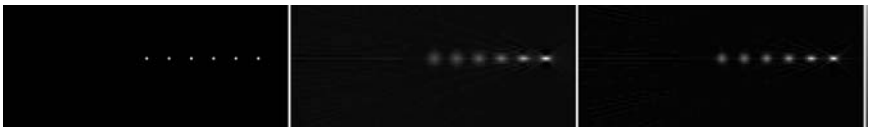


FIGURE 7. Effect of the GRF on SPECT image resolution. The image on the left is a point-source phantom. The middle and right images were reconstructed using FBP from projections obtained using LEGP and LEHR collimators, respectively. The image has been truncated at the top and bottom.

effects of GRF on image resolution. Projections were simulated of a phantom containing 6 point sources 5 cm apart using a radius-of-rotation of 27 cm and LEHR and LEGP collimators. The images were reconstructed using filtered backprojection (FBP) without any compensation of the CDRF. For the purposes of this chapter, FBP implies no compensation. Tangential and radial profiles, which, for these sources, correspond to the vertical and horizontal directions in Figure 7, respectively, are shown in Figure 8. The images and profiles demonstrate the spatial variance of the reconstructed resolution. In particular, the resolutions in the tangential and axial directions improve with distance from the center of rotation while the resolution in the radial direction remains relatively constant.

4. Analytic Compensation Methods

Several analytic methods (i.e., those having a closed form solution and not requiring iteration) have been proposed to compensate for the effects of the CDRF on SPECT images. To date, all of these methods have involved approximations about the CDRF. In general they can be divided into methods that assume that the CDRF is spatially invariant and those that explicitly model the distance dependence.

Methods based on the spatially invariant approximation are also known as restoration filters. The basic idea is to write the measured image (either projection or reconstructed image), $m(\vec{x})$, as the convolution of the CDRF, $\mathrm{cdrf}(\vec{x})$, and the image that would be obtained with a perfect collimator (i.e., where the CDRF is a delta function), $o(\vec{x})$:

$$m(\vec{x}) = \mathrm{cdrf}(\vec{x}) \otimes o(\vec{x}) \tag{9}$$

Taking the Fourier transform (FT) of Eq. 9 and dividing by the FT of the CDRF, we obtain:

$$O(\vec{\nu}) = M(\vec{\nu})\mathrm{CDRF}^{-1}(\vec{\nu}) \tag{10}$$

where the functions with capital letters indicate the FTs of the corresponding functions with lower case letters. There are two problems with this formulation. First, the FT of the CDRF has zeros in it. In addition, the measured

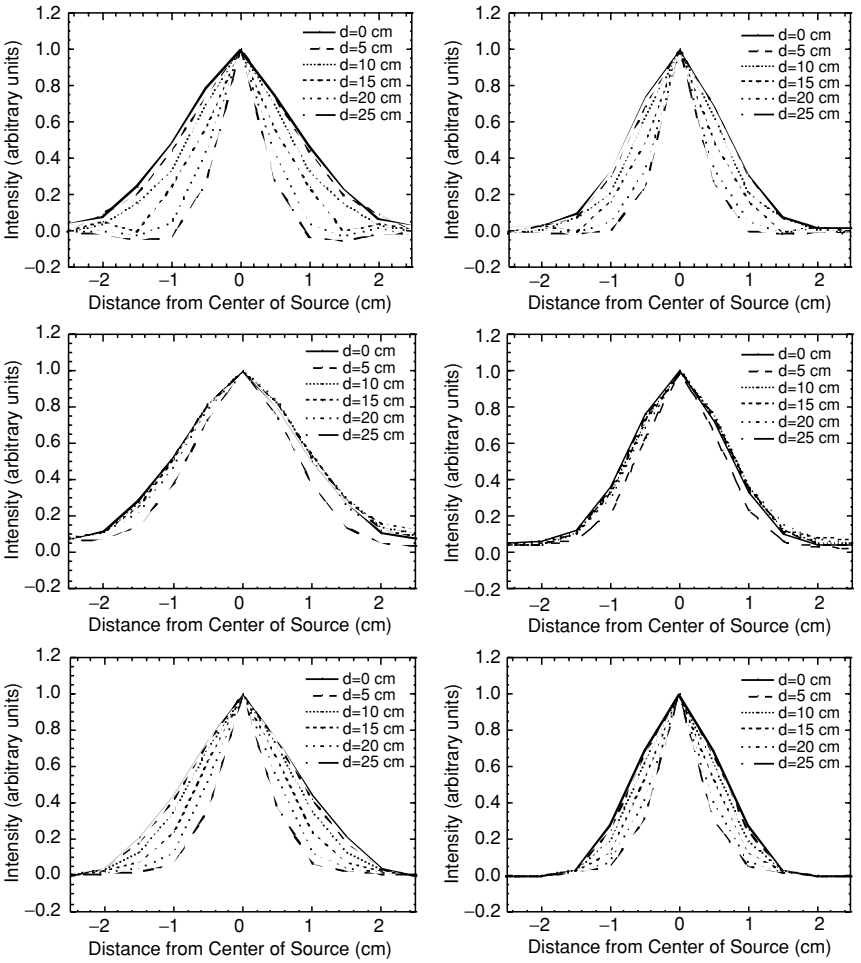


FIGURE 8. Tangential (top row), radial (middle row) and axial (bottom row) profiles through reconstructed point sources of Fig. 5 for LEGP (left) and LEHR (right) collimators. The distance from the center of rotation to the point, d, is indicated in the legend. The profile for each point is normalized to unit intensity.

data are noisy and the high frequencies of the image, whether projections or reconstructed image, will be dominated by noise. One solution to both of these problems is to roll off the high frequencies using a low pass filter. Two forms of low pass filters have been applied in SPECT, the Metz and Wiener Filters. The Metz filter, Metz($\vec{\nu}$), has the form:

$$\text{Metz}(\vec{\nu}) = \text{CDRF}^{-1}(\vec{\nu})\left[1 - \left(1 - \text{CDRF}(\vec{\nu})^2\right)^n\right] \tag{11}$$

where $n > 1$ is the order of the filter. The term in the square brackets acts as a low-pass filter. For $n = 1$, the filter becomes CDRF($\vec{\nu}$) and is often called a

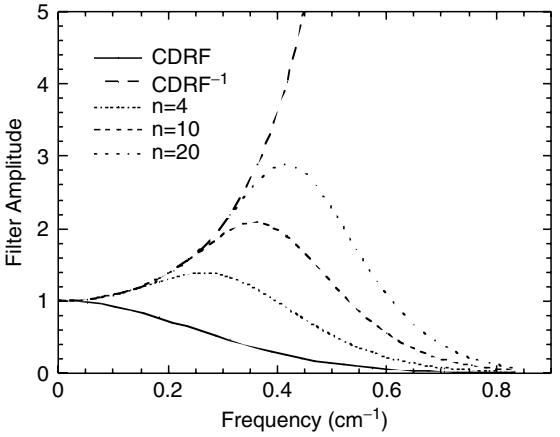


FIGURE 9. Plot of filter amplitude versus frequency for Metz filters with orders 1 (equal to the CDRF), infinity (equal to the CDRF^{-1}), 4, 10, and 20. In all cases a Gaussian CDRF with a FWHM of 1.5 cm was used.

matched filter; as $n \to \infty$, the Metz filter becomes the inverse filter. Thus, for low n, the low pass term in square brackets dominates, while for higher orders the filter behaves more like an inverse filter. This is illustrated in Figure 9. King *et al.*[10,11] have applied the Metz filter to nuclear medicine images with a count-dependent, non-integral order.

An alternative filter that has been applied in the context of SPECT imaging is the Wiener filter.[12,13] The Wiener filter is the filter that minimizes the mean squared error between the restored and true projections. It has the form:

$$\text{Wiener}(\vec{\nu}) = \text{CDRF}^{-1}(\vec{\nu})\left[\frac{\text{CDRF}(\vec{\nu})^2}{\text{CDRF}(\vec{\nu})^2 + N(\vec{\nu})/|O(\vec{\nu})|}\right] \tag{12}$$

where $N(\vec{\nu})$ is the noise power spectrum and $|O(\vec{\nu})|$ is the object power spectrum. Again, the term in square brackets is a low pass filter, in this case depending on both the noise power spectrum and the object power spectrum (which, paradoxically, is what we are trying to recover). Note that for frequencies where the noise power spectrum is much smaller than the object power spectrum, Eq. 12 becomes the inverse filter. However, at frequencies where the noise dominates, the Wiener filter approaches zero. One of the difficulties with the Wiener filter is that it requires estimates of the noise and object power spectra. One solution to this problem is to assume that noise is stationary. In this case, the noise power spectrum for the projections, assuming Poisson noise statistics, is flat with a value equal to the mean number of counts. For reconstructed images, approximate expressions for the noise power spectrum can be used.[14,15] The object power spectrum can be approximated in the region where the noise power is high

by extrapolating a fit of the image power spectrum minus the approximate noise power spectrum.[12]

One important issue in applying either a Metz or Wiener filter is the selection of the CDRF. Since the true CDRF is spatially varying, an approximation must be made. Among several approaches are using the CDRF for a distance from the face of the collimator equal to the center of rotation or half the distance from the center of rotation. King *et al.*[10,11] have fit a parameterized CDRF to minimize the mean square error of the measured images. Finally, the spatial variance of the CDRF can be improved by taking the geometric mean of opposing views.

One limitation of these restoration filters is that, since they are spatially invariant, they will tend to overcompensate in some areas and undercompensate in others. Also, they do not improve the asymmetry of the reconstructed point response (except perhaps by degrading it).

The second class of analytic methods explicitly includes the distance dependence of the CDRF. The basic idea is to write expressions for the projections of a function that include the CDRF and attempt to invert them, i.e. to solve for the object in terms of the projections. Since attenuation has an important effect on the measured projections, the attenuation distribution is assumed to be zero or uniform in the development of the expressions for the projections; analytic CDRF compensation methods for nonuniform attenuators have not yet been developed. To date, a closed form exact solution has not been found. However, several approximate methods have been derived.

An important basis for a number of these methods is the frequency-distance relation (FDR).[16] This approximation relates the position of a point source to the FT of the sinogram of the source. Note that since the angular coordinate in the sinogram is periodic, a Fourier series transform can be used in this direction. The FDR predicts that the energy in this Fourier transform will be concentrated along a line in Fourier space through the origin and with a slope linearly related to the distance from the point source to the center of rotation (COR). For a distributed source, this implies that the activity from all sources at a specific distance from the COR lies along a single line in Fourier space. For a circular orbit, the activity at a fixed distance from the COR is also at a fixed distance from the collimator face. Thus, one could perform inverse filtering along these lines using the appropriate CDRF. However, it should be noted that the derivation of the FDR requires invoking an approximation, the stationary phase approximation. Thus, methods based on the FDR will necessarily be approximate.

Lewitt *et al.*[17,18] were the first to use the FDR to compensate for the spatially varying CDRF. Glick *et al.*[19] exploited the FDR in combination with restoration filtering using the Wiener filter. This method has the advantage that it explicitly includes noise control. It was shown to provide improved symmetry in the reconstructed point spread function. However, one limitation of all these methods is that they tend to result in correlated

noise with a texture that appears to have unfavourable properties compared to the noise correlations introduced by iterative methods.[20]

A second method for deriving analytical filters is to assume a form for both the shape and distance dependence of the CDRF that allows analytic inversion. Appledorn[21] has analytically inverted the attenuated Radon transform equation assuming a CDRF shaped like a Cauchy function:

$$\text{Cauchy}(x) = \left(1 + \left(\frac{x}{w}\right)^2\right)^{-1.5} \tag{13}$$

where w is width parameter. It should be noted that the Cauchy function having the same FWHM as a Gaussian will have longer tails and be somewhat more sharply peaked in the center of the response function. Soares *et al.*[22] have implemented this method and evaluated it with and without combined analytic attenuation compensation. While the method did reduce the spatial variance of the response, it seemed to do so largely by degrading the resolution in the tangential direction.

A second class of methods approximates the CDRF as a Gaussian, not an unreasonable approximation as demonstrated above. However, in order to invert the equations, an approximate form for the distance dependence is assumed. van Elmbt *et al.*[23] used the approximation that the square of the width of the CDRF is proportional to the distance from the center-of-rotation. The result is a modification of Bellini's method[24] for uniform attenuators that includes approximate compensation for the spatially variant CDRF.

Pan *et al.*[25] have developed a method that improves on this approximation, assuming that the change in width of the CDRF over the object was less than the width at the center of the object. This assumption leads to an expression for the CDRF of the form:

$$\sigma^2(D) = \sigma_0^2 + 2\sigma_0\sigma_1(D - r) \tag{14}$$

where σ is the width parameter for the Gaussian CDRF, D, is the distance to the face of the collimator, r is the distance from the collimator face to the center-of-rotation, σ_0 is the width of the CDRF at the center of rotation, and σ_1 is the slope of the assumed linear relationship between D and the width of the Gaussian CDRF. This method has been evaluated by Wang *et al.*[26,27] It was found that, even for noise-free data, the assumptions about the shape and distance-dependence of the Gaussian FWHM resulted in high-frequency artefacts unless a low-pass filter was applied. In the presence of noise, low pass filtering was even more essential. Pan's method[25] resulted in the greatest improvement in resolution near the center of rotation, where the approximation in Eq. 14 is most accurate. The radial resolution improved but the tangential resolution degraded with distance away from the center of rotation. These effects are likely due to the approximate nature of the model for the detector response.

5. Iterative Compensation Methods

Iterative reconstruction-based compensation for the CDRF is performed by including the effects of CDRF in the model of the image formation process. For simplicity, consider the maximum likelihood-expectation maximization (MLEM) algorithm[28] (see chapter 4):

$$f_j^{(n+1)} = \frac{f_j^{(n)}}{\sum_l a_{ij}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{(n)}} \tag{15}$$

In this expression, $f_j^{(n+1)}$ is the reconstructed image in the j-th voxel after $n+1$ iterations, p_i is i-th projection bin from the measured projection data, and a_{ij} is the probability that a photon emitted in the j-th voxel is detected in the i-th projection bin. The key, then, to iterative-based compensation is modelling the CDRF in the matrix A whose elements are a_{ij}.

The matrix A is generally large and, when the CDRF is modelled, its sparseness is greatly reduced compared to the case when no effects or attenuation alone are modelled. Consider the reconstruction of a $128 \times 128 \times 128$ image from 128×128 projections acquired at 128 views. Assuming that the average size of the PSF is 7×7, the matrix A has $128^4 \times 7^2 \approx 1.3 \times 10^{10}$ elements. This matrix is currently too large to store on current general purpose computers. As a result, implementation of CDRF compensation has typically involved the use of projection and back-projection operators that model the CDRF. These codes implement the projection and backprojection operations described in Eqs. 1 and 2 of chapter 4 during each step of the iterative reconstruction process.

One of the first methods to model the distance-dependent CDRF in a projector-backprojector pair involved the use of ray-tracing techniques.[29] Projector-backprojectors based on these techniques are often referred to as ray-driven. Multiple equiangularly spaced projection rays were propagated from each projection bin back through the reconstruction matrix, as illustrated in Figure 10. The amount summed into the projection bin for each voxel crossed is equal to the product of the value in the voxel, the line length through the voxel, and the value of CDRF at that distance and offset from the central ray. This method was initially implemented in a two-dimensional geometry, assuming that objects have infinite axial extent (e.g., modelling voxels as lines instead of points). One advantage of this ray-tracing approach is that it is straightforward to model nonuniform attenuation. However, the spacing of the projection rays must be such that voxels are not missed in order to avoid artefacts in the projections. In addition, as the pixel size decreases, and if the modelling is performed in three-dimensions, the number of rays per bin and the time to compute the projections will become large.

An alternative to the approach of tracing rays from the projection bins through the reconstruction volume is to trace rays from the center of each

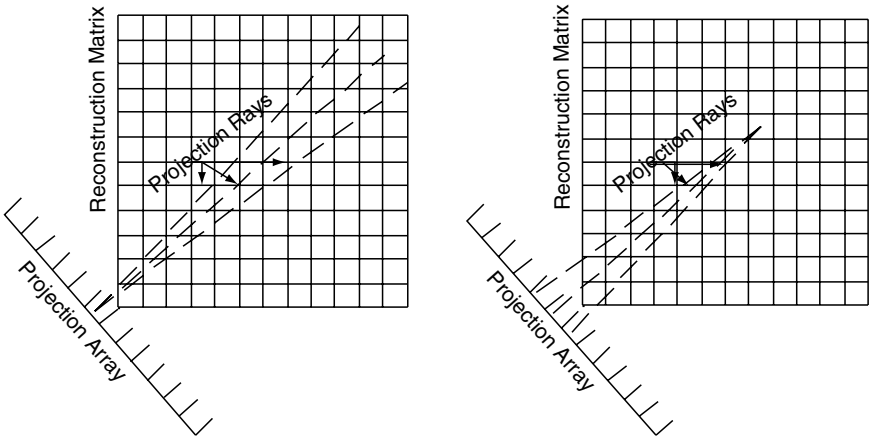


FIGURE 10. Illustration of ray-driven (left) and pixel-driven (right) projector implementing CDRF models.

voxel to the center of projection bins where the CDRF is non-zero. A projector based on this method is often known as pixel- or voxel-driven. In this case, ray tracing is used to calculate the attenuation factor for the rays as they traverse the object and the amount summed into each projection bin is the product of the voxel value and the CDRF appropriate for the source-detector distance and offset from the central ray. This method avoids the problem of missing voxels, but still is very computationally intensive. However, it can be speeded up by noticing that the attenuation factors for all the rays from a given voxel with respect to a given projection angle will be almost the same. Thus, the attenuation factor needs to be calculated only for the central ray (and this value could be calculated once and stored for use in subsequent iterations). It should be noted that this will result in some modelling errors, especially when the CDRF is wide. Even with this change, it still requires a great deal of time to compute the projection of the activity distribution.

A significant reduction of the time required to model the CDRF was realized with the development of rotation-based projector-backprojector pairs.[30,31] These algorithms exploit the fact that, for parallel-beam geometries, the CDRF is spatially invariant in planes parallel to the collimator face. By rotating the projection matrix so the sampling grid is parallel to the detector, the distance dependence of the CDRF can be modelled by convolving each such plane with the appropriate CDRF. This convolution can be done in the frequency[31] or spatial[32] domains. Generally, spatial domain convolution is faster when the CDRF has a relatively limited spatial extent in comparison to the size of the projections. The rotation-based projection process is illustrated in Figure 11. After convolution with the CDRF, the columns perpendicular to the detector are summed to form the projection

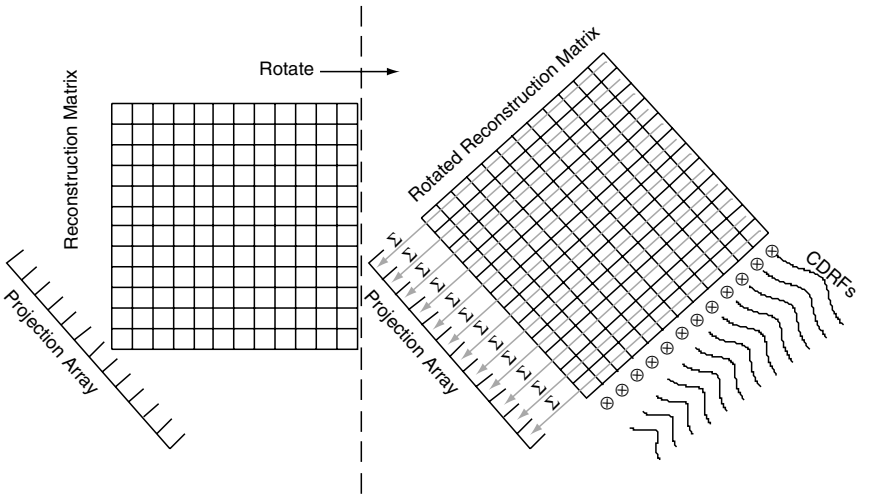


FIGURE 11. Illustration of rotation-based projector that models the CDRF.

image. If attenuation is to be modelled in the projector, the attenuation map can be rotated and the sums modified to include the attenuation factors for each voxel. Selection of the image rotation algorithm is important. Bilinear interpolation can be used, but other interpolation methods may prove better.[33] Image rotation based on shears gives a high quality reconstructed image with good computational efficiency.[34]

While rotation-based CDRF modelling provides a major increase in computational efficiency, further improvement can be obtained if the CDRF is approximated as a Gaussian. In this case the convolution can be performed by separating the convolution into two orthogonal one-dimensional convolutions[35] or via the use of Gaussian diffusion.[36,37] In the incremental blurring or diffusion method, the blurring of the rotated reconstruction matrix starts in the plane farthest from the collimator, plane N. The current estimate in this plane is treated as a 2D image and convolved with a small kernel representing the change in the CDRF from plane N to plane $N-1$. The resulting blurred 2D image is added to the 2D image in plane $N-1$. The 2D image in plane $N-1$ (which now includes the blurred version of plane N) is convolved with another small kernel, representing the change in the CDRF from plane $N-1$ to $N-2$, and added to plane $N-2$. This process of convolution with a small kernel and addition to the next closer plane is repeated until the plane closest to the collimator is reached. This last plane is then convolved with the CDRF corresponding to the distance from it to the collimator face. The net result is that the planes further from the detector are blurred by many small kernels that result in blurring equivalent to that from the appropriate CDRF; planes closer to the detector are blurred by a smaller number of the same kernels and are thus

blurred less, but with an amount that is equivalent to direct blurring with the CDRF. For a CDRF with a Gaussian shape, the incremental kernels can be calculated analytically.[30]

The backprojection can also be performed with a rotation-based algorithm. This is accomplished by performing the steps in the reverse order as described for the projector. First, the values in the projection bin are spread along the corresponding columns in a rotated frame like the one on the right in Figure 11. The pixel values are then attenuated and each plane is convolved with the appropriate CDRF. The resulting image is then rotated back to the unrotated-frame (at the left in Figure 11), and summed into the reconstructed image.

With the combination of all these acceleration methods, compensation for GRF can be achieved on modern personal computers for a $64 \times 64 \times 64$ reconstruction in times under 10 seconds per iteration. As a result, though clearly requiring more computational resources than analytic methods, iterative reconstruction-based CDRF compensation can be performed in times that are clinically realistic, especially when used with rapidly converging algorithms such as the ordered-subsets expectation-maximization (OSEM) algorithm.

6. Efficacy of CDRF Compensation

There have been a number of studies evaluating the efficacy of CDRF compensation. These include qualitative evaluations,[38] evaluations in terms of quantitative of the effect of CDRF compensation on performance in estimation[4242-45] and detection tasks.[4646-50]

Tsui demonstrated qualitative improvements in image quality in terms of noise and image resolution obtained using iterative 3D reconstructions with CDRF compensation compared to 2D reconstructions with or without compensation.[38] Iterative reconstruction with CDRF compensation was demonstrated to quantitatively improve the image resolution as measured by the FWHM.[40] The minimum FWHM achieved was limited by the collimator design and pixel size, with higher resolution collimators requiring a smaller pixel size to achieve maximal improvements in image resolution. This work also demonstrated that, even after compensation, the reconstructed spatial resolution varied spatially. In the context of myocardial-perfusion imaging, a comparison of reconstructed image quality in terms of the FWHM and standard deviation showed that iterative-based compensation provided improved resolution with less noise compared to analytic FDR-based compensation.[36]

The images shown in Figures 12 and 13 demonstrate the improved resolution that can be obtained with iterative CDRF compensation as compared to FBP. These images are reconstructions of projection data based on the phantom shown in Figure 7 with the display zoomed to the region

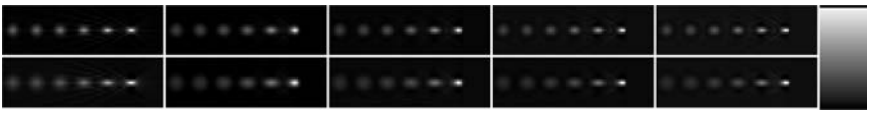


FIGURE 12. Efficacy of iterative CDRF compensation for LEHR (top row) and LEGP (bottom row) collimators. The images shown are transaxial images from reconstructions of projections of the phantom shown in Fig. 7. They have been truncated to the region of the point sources. The leftmost source is located at the center of rotation. The columns were reconstructed using (left to right) FBP and OSEM with 64 subsets per iteration at 5, 10, 20, and 30 iterations, respectively.

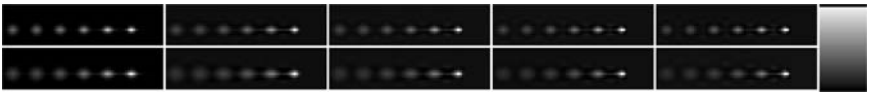


FIGURE 13. Same as Fig. 12, but showing slices parallel to the axis of rotation, demonstrating the resolution in the axial direction.

containing the point sources. Only the GRF was modelled in the projection and reconstruction process. Note that, for the iterative reconstruction, the point sources are placed on top of a flat background with 100 times the amplitude, reconstructed, and then the background is subtracted. This procedure is used to avoid problems with nonlinearities in the reconstruction giving an inaccurate representation of the true resolution characteristics. Note that the resolution is improved. Also note that the resolution improves with iterations. In general, more iterations are required for convergence when CDRF compensation is performed as compared to iterative reconstruction with no compensation.

The images in Figure 14 show reconstructions of noisy and noise-free projections of a uniform cylinder of activity obtained with LEHR and LEGP collimators and reconstructed using OS-EM with and without GRF compensation. They illustrate some additional properties of iterative CDRF compensation. The images demonstrate improved sharpness at the edge of the disk, but also the presence of Gibbs-like ringing artefacts. The nature of the artefacts depends on the collimator used. In the noisy images, notice the significant difference in noise texture in the images obtained with CDRF compensation; the noise texture is different for the two collimators. Also note that with CDRF compensation the noise texture is spatially varying.

Since iterative reconstruction is nonlinear and spatially variant, traditional measures of image quality such as the modulation transfer function (MTF), noise power spectrum (NPS), and noise equivalent quanta (NEQ), which is the ratio of the square of the MTF and the noise power spectrum, do not apply. However, Wilson[39] has proposed the use of local versions of these quantities. The local MTF is computed from the local point spread function (PSF). Since iterative reconstruction using MLEM-like algorithms

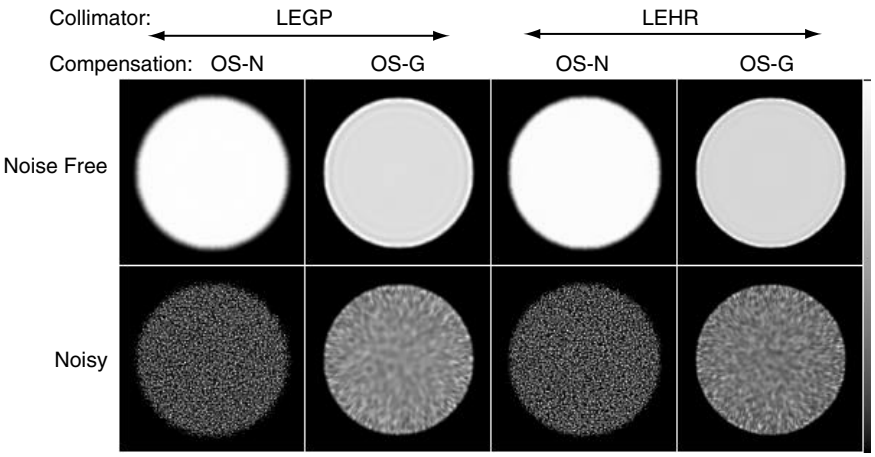


FIGURE 14. Reconstructed images from analytically-simulated projections of a cylinder obtained using LEGP and LEHR collimators. Only the GRF was modelled in the projections. Images were reconstructed from projection data with and without the addition of Poisson noise. Five iterations of OSEM were used with 64 subsets per iteration (4 angles per subset) with no compensation (OS-N) and compensation for the GRF (OS-G).

are nonlinear, the PSF cannot be estimated by simply reconstructing the projections of a point source. Instead, the point source projections are treated as a perturbation to a background image. After reconstruction of the perturbed image, the reconstruction of the background image is subtracted providing a local PSF; the magnitude of the FT of this local PSF provides the local MTF. The local NPS is computed from the covariance matrix for the point of interest. It was found that iterative reconstruction with CDRF compensation produces local MTFs with improved response at mid-frequencies. Subsequent unpublished work has demonstrated that, with large numbers of iterations, all frequencies up to approximately the first zero in the MTF obtained by filtered backprojection can be restored. This is illustrated by Figures 15 and 16, which show the local MTF for a point source reconstructed with FBP and iterative reconstruction with CDRF compensation. Note that for the iterative method, there is some overcompensation at mid-frequencies (MTF > 1) and that there is some artifactual increase at frequencies above the point where the MTF for FBP goes to zero.

The noise properties are also an important factor in evaluating CDRF compensation. It has been found that the local NPS increases at the frequencies that are restored, similar to what is found with restoration filtering. This gives rise to the lumpy noise texture seen in the images in Figure 14.[39] In fact, the boosting of the noise is such that the NEQ is not greatly improved by CDRF compensation. In recent work using these measures, Wilson

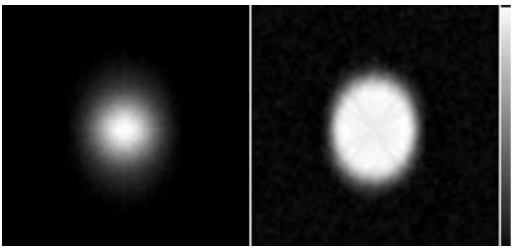


FIGURE 15. Images of local MTF in transaxial plane for a point source located 15 cm from the center of rotation. The projections were generated analytically and simulated a LEGP collimator. The image on the left is from FBP with no low pass filter and the one on the right is from 20 iterations of OSEM with 64 subsets per iteration (4 angles per subset) with GRF compensation.

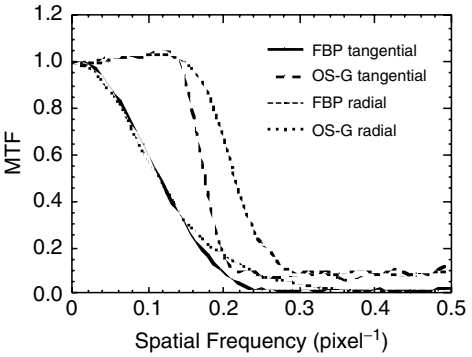


FIGURE 16. Profiles through the MTF images in Fig. 15. The tangential direction is along the vertical axis and the radial direction along the horizontal one in Fig. 15.

et al.[41] have pointed out that improper modelling of the CDRF can result in poorer MTFs and noise spectra, while the NEQ was relatively insensitive to these modelling errors.

Several studies have demonstrated an improvement in performance on estimation tasks with CDRF compensation. In the context of myocardial perfusion imaging, both analytic and iterative CDRF compensation were found to provide improved uniformity in the myocardial wall and improved estimates of wall thickness compared to no compensation. Iterative reconstruction-based compensation provided better results than analytic compensation based on the FDR.[43] Pretorius *et al.*[42] evaluated iterative and analytic FDR-based CDRF compensation in terms of recovery of the maximum and total counts in a volume. They found that iterative methods provided better recovery and reduced variance in estimates compared to the analytic methods, but that recovery coefficients were more spatially variant for the iterative methods.

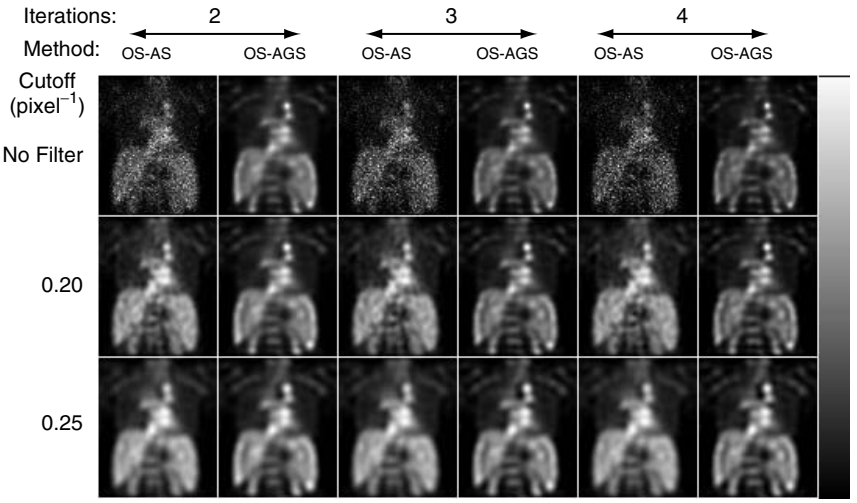


FIGURE 17. Coronal slice of SPECT image of distribution of In-111 ibritumomab tiuxetan, a radiolabeled antibody. The images were reconstructed using OSEM with 32 subsets per iteration (4 angles per subset) with attenuation and scatter (AS) or attenuation, geometric response and scatter (AGS) compensation. The number of iterations and the cut-off frequency of an order 8 3-D Butterworth post-reconstruction filter are also indicated.

Figure 17 illustrates why there might be improved quantitation with CDRF compensation. The task to be performed with this In-111 ibritumomab tiuxetan image is to estimate organ uptakes for use in dosimetry calculations. Note the difference in noise properties and the sharpness of the image obtained with and without CDRF compensation. The images with CDRF compensation have improved spatial resolution, even after low pass filtering, compared to the images without. This improved resolution allows more accurate definition of organ regions and reduced partial volume effects.

Finally, a number of studies have evaluated the efficacy of CDRF compensation in the context of defect detection tasks. Gifford *et al.*[49] showed improved performance for a tumour detection and localization task analyzed using localization receiver operating characteristics (LROC) analysis. The study used simulated Ga-67 images. It was found that iterative reconstruction-based CDRF compensation gave improved performance compared to FBP or FDR-based analytic compensation followed by iterative reconstruction. For myocardial defect detection, it has been shown using both simulated[48,50] and clinical[47] data with mathematical and human observers that iterative reconstruction-based CDRF compensation results in improved observer performance. Note that in all these studies, regularization by a post-reconstruction filter was required for optimal performance.

Figure 18 illustrates the difference in image quality obtained with attenuation compensation alone and attenuation combined with CDRF compensation. Note that the images that include CDRF compensation have better resolution and noise properties, even after low-pass filtering. Figure 19

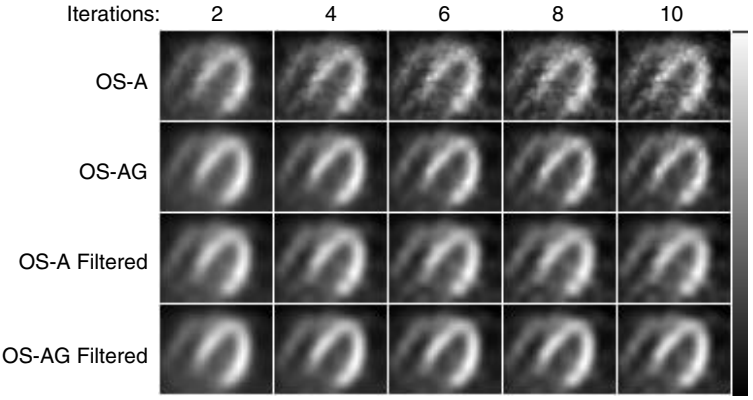


FIGURE 18. Reconstructed images from a clinical Tc-99m sestamibi scan. The images were reconstructed with OSEM using 8 subsets (4 angles per subset) with compensation for attenuation alone (OS-A) or with attenuation and geometric response (OS-AG). The number of iterations is indicated above. Images obtained using a 3-D Butterworth post-reconstruction filter with order 8 and cut-off 0.24 pixel^{-1} are also shown.

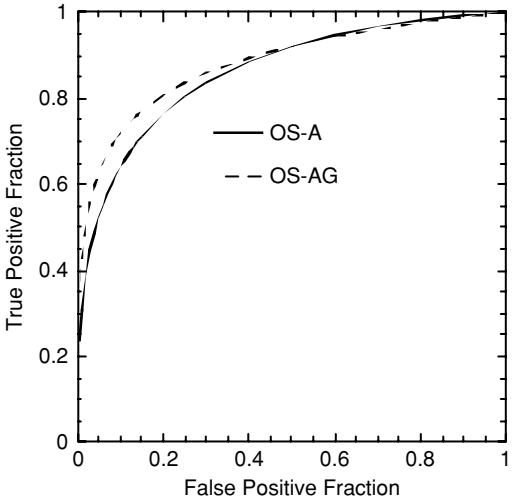


FIGURE 19. ROC curve for a human observer study comparing OSEM reconstruction with attenuation compensation (OS-A) and attenuation plus geometric response compensation (OS-AG) in a myocardial perfusion defect detection task. This study used simulation images and is described in ref.[48]

illustrates the effect of CDRF compensation on myocardial perfusion detection. The areas under the ROC curve without and with CDRF compensation were 0.863 and 0.882, respectively.

7. Summary

The CDRF in SPECT has several components: the geometric, septal penetration, septal scatter and intrinsic response function components. Analytical formulations for the geometric component have been developed, while Monte Carlo simulation has allowed estimation of the septal penetration and septal scatter components. The geometric response dominates the resolution characteristics for low energy photons and medium and higher energy photons when used with well-designed collimators. The width of the geometric response is proportional to source-to-collimator distance. In SPECT, the intrinsic component plays a relatively small role compared to the geometric response due to the relatively large distances to the collimator from many of the organs of interest.

The distance-dependence of the CDRF results in a spatially varying reconstructed resolution in SPECT. In general, the tangential and axial resolutions improve with distance from the center of rotation while the radial resolution is relatively constant.

Methods of compensating for the CDRF can be divided into two general classes: analytical and iterative. Analytic methods are fast, but ones developed to date involve approximations about the shape or spatial variation of the CDRF. Iterative methods require more computational resources but have the potential to allow for compensation of CDRFs with arbitrary shapes as well as image degrading phenomenon such as attenuation and scatter.

A number of studies have been performed evaluating the various CDRF compensation methods. Both iterative and analytical methods improve the resolution in the resulting images, but at the cost of introducing noise correlations. Task-based evaluations of the CDRF compensation methods have demonstrated an improvement in performance for quantitative tasks such as activity estimation as well as defect detection tasks. Generally these evaluations have shown that iterative methods provide better performance than the analytical methods.

Acknowledgements. The authors are grateful to Dr. Yong Du for his careful proofreading of the manuscript.

References

1. Metz C. E., The geometric transfer function component for scintillation camera collimators with straight parallel holes. *Phys Med Biol* **25:** 10591059-1070 (1980).
2. Tsui B. M. W. and Gullberg, G. T., The geometric transfer-function for cone and fan beam collimators. *Phys Med Biol* **35:** 81-93 (1990).
3. Frey E. C., Tsui, B. M. W. and Gullberg, G. T., Improved estimation of the detector response function for converging beam collimators. *Phys Med Biol* **43:** 941-950 (1998).
4. Formiconi A. R., Geometrical response of multihole collimators. *Phys Med Biol* **43:** 3359-3379 (1998).
5. Formiconi A. R., Passeri, A. and Calvini, P., Theoretical determination of the collimator geometrical transfer function for the reconstruction of SPECT data. *IEEE Trans Nucl Sci* **46:** 1075-1080 (1999).
6. De Vries D. J., Moore, S. C., Zimmerman, R. E. *et al.*, Development and validation of a Monte Carlo simulation of photon transport in an Anger camera. *EEE Trans Med Imaging* **9:** 430-438 (1990).
7. Du Y., Frey, E. C., Wang, W. T. *et al.*, Combination of MCNP and SimSET for Monte Carlo simulation of SPECT with medium- and high-energy photons. *IEEE Trans Nucl Sci* **49:** 668-674 (2002).
8. Wilderman S. J., Dewaraja, Y. and Koral, K. F., Accurate modeling of nuclear-medicine collimators in Monte Carlo simulation of high-energy photons. *Nucl Instr Meth A* **422:** 745-750 (1999).
9. Wang W. T., Frey, E. C., Tsui, B. M. W. *et al.*, Parameterization of Pb X-ray contamination in simultaneous Tl-201 and Tc-99m dual-isotope imaging. *IEEE Trans Nucl Sci* **49:** 680-692 (2002).
10. King M. A., Doherty, P. W., Schwinger, R. B. *et al.*, Fast count-dependent digital filtering of nuclear medicine images: concise communication. *J Nucl Med* **24:** 1039-1045 (1983).
11. King M. A., Penney, B. C. and Glick, S. J., An image-dependent Metz filter for nuclear medicine images. *J Nucl Med* **29:** 19801980-1989 (1988).
12. King M. A., Doherty, P. W., Schwinger, R. B. *et al.*, A Wiener filter for nuclear medicine images. *Med Phys* **10:** 876-880 (1983).
13. King M. A., Schwinger, R. B., Doherty, P. W. *et al.*, Two-dimensional filtering of SPECT images using the Metz and Wiener filters. *J Nucl Med* **25:** 1234-1240 (1984).
14. Riederer S. J., Pelc, N. J. and Chesler, D. A., Noise power spectrum in computed x-ray tomography. *Phys Med Biol* **23:** 446-454 (1978).
15. Hanson K. M. and Boyd, D. P., Characteristics of computed tomographic reconstruction noise and their effect on detectability. *IEEE Trans Nucl Sci* **25:** 160-163 (1978).
16. Edholm P. R., Lewitt, R. M. and Lindholm, B., Novel properties of the Fourier decomposition of the sinogram. *SPIE Proceedings* **671:** 88-18 (1986).
17. Xia W. S., Lewitt, R. M. and Edholm, P. R., Fourier correction for spatially variant collimator blurring in SPECT. *IEEE Trans Med Imaging* **14:** 100-115 (1995).
18. Lewitt R. M., Edholm, P. R. and Xia, W., Fourier method for correction of depth dependent collimator blurring. *SPIE Proceedings* **1092:** 232-243 (1989).
19. Glick S. J., Penney, B. C. and King, M. A., Non-iterative compensation for the distance-dependent detector response and photon attenuation in SPECT imaging. *IEEE Trans Med Imaging* **13:** 363-374 (1994).

20. Soares E. J., Glick, S. J. and King, M. A., Noise characterization of combined bellini-type attenuation correction and frequency-distance principle restoration filtering. *IEEE Trans Nucl Sci* **43:** 3278-3290 (1996).
21. Appledorn C. R., "An Analytical solution to the nonstationary reconstruction problem in single photon emission computed tomography. *Prog Clin Biol Res* **363**: 69-79 (1991).
22. Soares E. J., Byrne, C. L., Glick, S. J. *et al*., Implementation and evaluation of an analytical solution to the photon attenuation and nonstationary resolution reconstruction problem in SPECT. *IEEE Trans Nucl Sci* **40:** 1231-1237 (1993).
23. van Elmbt L. and Walrand, S., Simultaneous correction of attenuation and distance-dependent resolution in SPECT - an analytical approach. *Phys Med Biol* **38:** 1207-1217 (1993).
24. Bellini S., Piacentini, M., Cafforia, C. *et al*., Compensation of tissue absorption in emission tomography. *IEEE Trans ASSP* **27:** 213-318 (1979).
25. Pan X. and Metz, C. E., A Class of analytical methods that compensate for attenuation and spatially-variant resolution in 2D SPECT. *IEEE Trans Nucl Sci* **43:** 2244-2254 (1996).
26. Wang W. T., Tsui, B. M. W., Frey, E. C. *et al*., "Comparison of an analytical and an iterative based collimator-detector response compensation method in SPECT", *Conf Rec* of IEEE Nuclear Science Symposium, Vol. 2; *pp* 1382-1386 (1998).
27. Wang W. T., "An Evaluation of an Analytical Collimator-Detector Response Compensation Method in SPECT," Masters Thesis, The University of North Carolina, 1999.
28. Shepp L. A. and Vardi, Y., Maximum likelihood estimation for emission tomography. *IEEE Trans Med Imaging* **1:** 113-121 (1982).
29. Tsui B. M. W., Hu, H. B., Gilland, D. R. *et al*., Implementation of simultaneous attenuation and detector response correction in SPECT. *IEEE Trans Nucl Sci* **35:** 778-783 (1988).
30. McCarthy A. W. and Miller, M. I., Maximum-Likelihood SPECT in clinical computation times using mesh-connected parallel computers. *IEEE Trans Med Imaging* **10:** 426-436 (1991).
31. Zeng G. L. and Gullberg, G. T., Frequency domain implementation of the three-dimensional geometric point source correction in SPECT imaging. *IEEE Trans Nucl Sci* **39:** 1444-1453 (1992).
32. Frey E. C., Ju, Z.-W. and Tsui, B. M. W., A fast projector-backprojector pair modeling the asymmetric, spatially varying scatter response function for scatter compensation in SPECT imaging. *IEEE Trans Nucl Sci* **40:** 1192-1197 (1993).
33. Wallis J. W. and Miller, T. R., An optimal rotator for iterative reconstruction. *IEEE Trans Med Imaging* **16:** 118-122 (1997).
34. DiBella E. V. R., Barclay, A. B., Eisner, R. L. *et al*., A comparison of rotation-based methods for iterative reconstruction algorithms. *IEEE Trans Nucl Sci* **43:** 3370-3376 (1996).
35. Beekman F. J., Eijkman, E., Viergever, M. A. *et al*., Object shape dependent PSF model for SPECT imaging. *IEEE Trans Nucl Sci* **40:** 31-39 (1993).
36. Kohli V., King, M. A., Glick, S. J. *et al*., Comparison of frequency-distance relationship and Gaussian-diffusion-based methods of compensation for distance-dependent spatial resolution in SPECT imaging. *Phys Med Biol* **43:** 1025-1037 (1998).

37. King M. A., Pan, T. S. and Luo, D. S., An investigation of aliasing with Gaussian-diffusion modeling of SPECT system spatial resolution. *IEEE Trans Nucl Sci* **44:** 1375-1380 (1997).
38. Tsui B. M. W., Frey, E. C., Zhao, X. D. *et al.*, The importance and implementation of accurate three-dimensional compensation methods for quantitative SPECT. *Phys Med Biol* **39:** 509-530 (1994).
39. Wilson D. W., "Noise and Resolution Properties of FB and ML-EM Reconstructed SPECT Images," Ph.D Dissertation, University of North Carolina at Chapel Hill, 1994.
40. Tsui B. M. W., Zhao, X. D., Frey, E. C. *et al.*, "Characteristics of reconstructed point response in three-dimensional spatially variant detector response compensation in SPECT" in: *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, edited by P. Grangeat and J-L Amans Kluwer Academic Publishers, (1996), *pp* 509-530.
41. Wilson D. W. and Barrett, H. H., The effects of incorrect modeling on noise and resolution properties of ML-EM images. *IEEE Trans Nucl Sci* **49:** 768-773 (2002).
42. Pretorius P. H., King, M. A., Pan, T. S. *et al.*, Reducing the influence of the partial volume effect on SPECT activity quantitation with 3D modelling of spatial resolution in iterative reconstruction. *Phys Med Biol* **43:** 407-420 (1998).
43. Kohli V., King, M. A., Pan, T.-S. *et al.*, Compensation for distance-dependent resolution in cardiac-perfusion SPECT: impact on uniformity of wall counts and wall thickness. *IEEE Trans Nucl Sci* **45:** 1104-1110 (1998).
44. Ljungberg M., Sjogreen, K., Liu, X. W. *et al.*, A 3-dimensional absorbed dose calculation method based on quantitative SPECT for radionuclide therapy: Evaluation for I-131 using Monte Carlo simulation. *J Nucl Med* **43:** 1101-1109 (2002).
45. Ljungberg M., Frey, E., Sjogreen, K. *et al.*, 3D absorbed dose calculations based on SPECT: Evaluation for 111-In/90-Y therapy using Monte Carlo simulations. *Cancer Biother Radiopharm* **18:** 99-107 (2003).
46. Pretorius P. H., Gifford, H. C., Narayanan, M. V. *et al.*, Comparison of detection accuracy of perfusion defects in SPECT for different reconstruction strategies using polar-map quantitation. *IEEE Trans Nucl Sci* **50:** 1569-1574 (2003).
47. Narayanan M. V., King, M. A., Pretorius, P. H. *et al.*, Human-observer receiver-operating-characteristic evaluation of attenuation, scatter, and resolution compensation strategies for Tc-99m myocardial perfusion imaging. *J Nucl Med* **44:** 1725-1734 (2003).
48. Sankaran S., Frey, E. C., Gilland, K. L. *et al.*, Optimum compensation method and filter cutoff frequency in myocardial SPECT: A human observer study. *J Nucl Med* **43:** 432-438 (2002).
49. Gifford H. C., King, M. A., Wells, R. G. *et al.*, LROC analysis of detector-response compensation in SPECT. *IEEE Trans Med Imaging* **19:** 463-473 (2000).
50. Frey E. C., Gilland, K. L. and Tsui, B. M. W., Application of task-based measures of image quality to optimization and evaluation of three-dimensional reconstruction-based compensation methods in myocardial perfusion SPECT. *IEEE Trans Med Imaging* **21:** 1040-1050 (2002).

6
Attenuation Correction Strategies in Emission Tomography

H. Zaidi* and B.H. Hasegawa†

1. The Problem of Photon Attenuation in Emission Tomography

The physical basis of the attenuation phenomenon lies in the natural property that photons emitted by the radiopharmaceutical will interact with tissue and other materials as they pass through the body. For photon energies representative of those encountered in nuclear medicine (i.e., 68 to 80 keV for ^{201}Tl to 511 keV for positron emitters), photons emitted by radiopharmaceuticals can undergo photoelectric interactions where the incident photon is completely absorbed. In other cases, the primary radionuclide photon interacts with loosely bound electrons in the surrounding material and are scattered. The trajectory of the scattered photon generally carries it in a different direction than that of the primary photon. However, the energy of the scattered photon can be lower than (in the case of incoherent scattering) or be the same as (in the case of coherent scattering) that of the incident photon. It is worth emphasizing that for soft tissue (the most important constituent of the body), a moderately low-Z material, we note two distinct regions of single interaction dominance: photoelectric below and incoherent above 20 keV. Moreover, the percentage of scattered events which undergo Compton interactions in the object is more than 99.7% at 511 keV for water, in which the number of interactions by photoelectric absorption or coherent scattering is negligible.

Mathematically, the magnitude of photon transmission through an attenuating object can be expressed by the exponential equation:

$$\Phi = \Phi_o \exp\left[-\int_s \mu(x,y)\, dr\right] \tag{1}$$

*PD Dr H. Zaidi, Geneva University Hospital, Division of Nuclear Medicine, CH-1211 Geneva, Switzerland
†Prof. B.H. Hasegawa, Department of Radiology, University of California San Francisco, CA, USA

where Φ_o and Φ are the incident and transmitted photon fluences (in units of photons per unit area) and dr is a differential of the thickness of tissue encountered as the beam of photons passes through the body along path S. The parameter μ is the linear attenuation coefficient, which represents the probability that the photon will undergo an interaction while passing through a unit thickness of tissue. Therefore, the linear attenuation coefficient is a measure of the fraction of primary photons which interact while traversing an absorber and is expressed in units of inverse centimetres (cm^{-1}).

Linear attenuation coefficients often are referred to as narrow-beam (collimated) or broad-beam (uncollimated) depending on whether or not the transmitted photon fluence includes scattered photons that escape and are transmitted through the object. The "build-up factor" caused by the broad-beam conditions of nuclear medicine imaging is defined as the ratio of the transmitted photons divided by the value predicted from the ideal narrow-beam measurement in which scatter is excluded from the transmitted beam. Therefore, the build-up factor is equal to 1 for narrow-beam geometry but it will increase with depth for broad beam geometries until a plateau is reached. Narrow-beam transmission measurements are ideally required for accurate attenuation correction in emission tomography. This is of course, determined by the geometry of the transmission data acquisition system.

Furthermore, the situation for photon attenuation is different for PET than in SPECT. When a radionuclide distribution is measured in planar scintigraphy or in SPECT, the amount of attenuation depends on the tissue pathlength and the type of tissue (e.g., soft tissue, bone, vs. lung) that the photon encounters as it travels between the point of emission and the point of detection. When positron-emitting radiopharmaceuticals are used for the imaging study, the imaging system records two antiparallel 511 keV photons that are emitted after electron-positron annihilation. In this case, the annihilation photons traverse a total tissue thickness that is equal to the body thickness intersected by the line between the two detectors, also called the line of response.

Figure 1 shows the narrow-beam attenuation as a function of the source depth in water for representative photon energies for radionuclides encountered in nuclear medicine and for different attenuating media calculated using data from the XCOM photon cross section library[1] and ICRU Report 44.[2] The data in Figure 1 show that attenuation of emission photons is severe for both gamma-emitting and positron-emitting radionuclides (singles detection). These values also emphasize that photon attenuation is an unavoidable process, which can affect the quality of the diagnostic information that we gather from radionuclide imaging in a direct and profound way. In a clinical setting, since the thickness of tissue varies for different regions of the patient's anatomy, the magnitude of the error introduced by photon attenuation can also vary regionally in the radionuclide image. Therefore, a lesion located deep within the body will produce a signal that is attenuated to a greater degree than that for a superficial lesion. Similarly, a tissue region with uniform

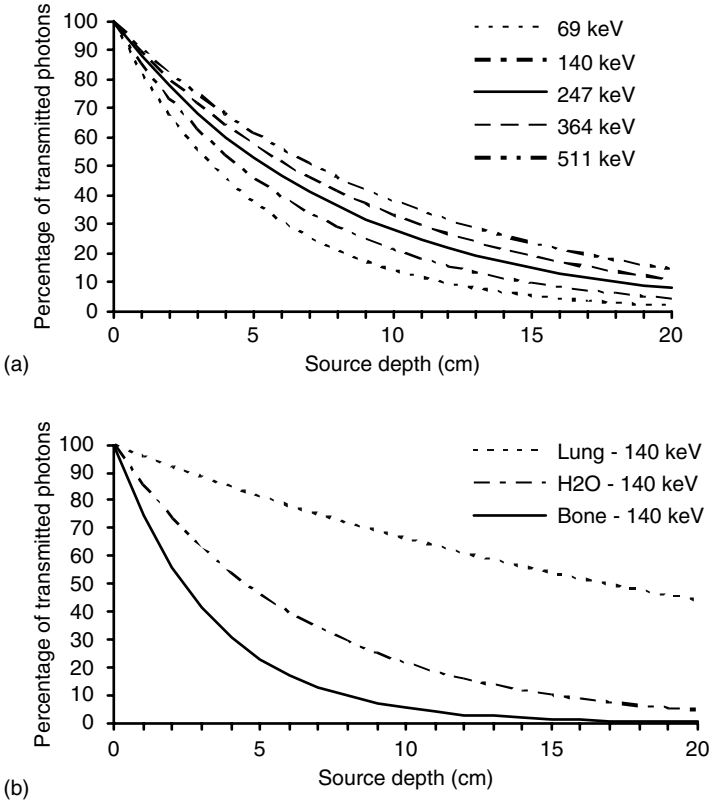


FIGURE 1. (*a*) The narrow-beam attenuation as a function of the source depth in a cylindrical water phantom for photons of different energies of interest in nuclear medicine imaging. (*b*) Same as above for 140 keV and different attenuating media.

radionuclide content that lies below tissue having a variable thickness will generate an image with variable count density. This can occur in myocardial perfusion imaging when soft-tissue attenuation due to the diaphragm or breast tissue can cause false-positive defects. Reconstruction of tomographic images without attenuation correction can cause erroneously high-count densities and reduced image contrast in low-attenuation regions such as the lung. All of these effects can introduce artefacts into radionuclide images that can complicate visual interpretation and can cause profound accuracy errors when radionuclide images are evaluated quantitatively. For this reason, it is important to understand both the physical processes that underlie photon attenuation and the methods that can be used to correct radionuclide images for these physical factors. Attenuation correction in emission tomography is now widely accepted by the nuclear medicine community as vital for achieving the goal of producing artefact-free, quantitatively accurate data. While, this is no longer the subject of debate in cardiac SPECT,[3-5] there are still some

controversies regarding its usefulness in routine clinical PET oncology studies.[6-8]

2. Determination of the Attenuation Map in Emission Tomography

Reliable attenuation correction methods for emission tomography require determination of an "attenuation map", which represents the spatial distribution of linear attenuation coefficients and accurately delineates the contours of structures in the body for the region of the patient's anatomy that is included in the radionuclide imaging study. After the attenuation map is generated, it can then be incorporated into the radionuclide reconstruction algorithm to correct the emission data for errors contributed by photon attenuation, scatter radiation, or other physical perturbations.

The methods for generating the attenuation map generally can be described as falling within two main classes. The first class includes "transmissionless" correction methods based on assumed distribution and boundary of attenuation coefficients (calculated methods), statistical modelling for simultaneous estimation of attenuation and emission distributions or consistency conditions criteria. The second class include correction methods based on transmission scanning including an external radionuclide source, X-ray CT, or segmented magnetic resonance images (MRI). These methods vary in complexity, accuracy, and computation time required. To date, the most accurate attenuation correction techniques are based on measured transmission data acquired before (pre-injection), during (simultaneous), or after (post-injection) the emission scan.

2.1 Transmissionless Approaches

In some cases, attenuation maps can be generated without adding a separate transmission scan to the emission acquisition. Algorithms in this class of methods either assume a known body contour in which a (uniform) distribution of attenuation coefficients is assigned or try to derive the attenuation map directly from the measured emission data. Only methods widely used in clinical routine and implemented on commercial systems will be described in this section. Other sophisticated and computer intensive approaches exist for generating an attenuation map without a separate transmission measurement. These include methods that apply consistency conditions and statistical modelling for simultaneous estimation of emission and transmission distributions that will be addressed in section 6 on future developments.

2.1.1 Calculated Methods

The assumption of uniform attenuation is straightforward in imaging the brain and abdominal areas where soft-tissues are the dominant constituent,

as opposed to regions like the thorax which is more heterogeneous. In these regions, if the body contour can be determined from the emission data, then the region within the contour can be assigned a uniform linear attenuation coefficient value corresponding to that of water or soft tissue to generate the corresponding attenuation map. The body contour can be determined either manually or with automatic edge detection methods.

2.1.1.a Manual contour delineation

The simplest manual method consists of approximating the object outline by an ellipse drawn around the edges of the object. Uniform attenuation is then assigned within the contour to generate the attenuation map. An irregular contour can also be drawn manually by an experienced technologist. The method is generally only appropriate for brain studies and is implemented on approximately all commercial SPECT and PET systems.[9] Although empirical, the method has some attractive properties: it is relatively quick, easy to use, and increases patient throughput, which is a relevant issue in a busy clinical department. However, the manually-defined contour generally will not fit the actual patient anatomy exactly and does not account for variations in the distribution of the attenuation coefficient within the patient.

2.1.1.b Automatic edge detection methods

A variation of the calculated attenuation correction is an automated technique that traces the edge of the object in projection space using an appropriate edge-detection algorithm. This allows the attenuation map to form any convex shape with the advantage that automated edge detection reduces the burden on the operator. In addition, lung regions can sometimes be delineated from the emission data in which case a more accurate attenuation map can be defined. Algorithms proposed for estimating the patient contour include those that define the contour: (i) based on the acquisition of additional data in the Compton scatter window,[10-12]; (ii) directly from the photopeak data only,[13-16] or (iii) by segmentation of the body and lung regions either by an external wrap soaked in ^{99m}Tc,[17] or using both scatter and photopeak window emission images.[18] Other methods use a set of standard outline images[19] to define the shape of the attenuation map. Assigning known attenuation coefficients to the soft tissue and lung regions then forms the attenuation map. Because it is generally difficult to define the patient contour from emission data alone without the use of transmission data, transmissionless techniques have had limited clinical application using these methods.

In the case of brain imaging, automated methods also allow for a certain thickness of higher attenuation material to be added to the calculation to account for the skull.[13] More recently, an automated method was proposed to compute a 3-component attenuation map for brain PET imaging.[20] The

technique generates an estimated skull image by filtered backprojection of the reciprocal of an emission sinogram. The thickness and radius of the skull is then estimated from profiles extracted from the image. The resulting thickness and radius values are then used to generate a model of the brain, skull, and scalp. Appropriate linear attenuation coefficients are then assigned to estimate the attenuation map for the head. More refined methods make use of an optical tracking system to derive 3D patient-specific head contour.[21] A previously acquired reference attenuation map is then transformed to match the contour of the reference head with the target head using the thin plate spline technique. A practical advantage of the optical tracking system is that it can also be utilized for motion correction.

It is generally well accepted that transmission-based nonuniform attenuation correction can supply more accurate attenuation maps than transmissionless techniques. However, it is not entirely clear whether nonuniform attenuation maps provide specific benefits in the routine clinical practice of tomographic brain imaging. Comparisons made by independent observers have shown no significant differences in subjective quality assessment between images reconstructed with uniform and non-uniform attenuation maps.[22] Hooper *et al.*[23] have shown using clinical PET data that calculated attenuation correction[13] gave rise to appreciable bias in structures near thick bone or sinuses when compared to the clinical 'gold standard' (transmission-based attenuation correction). Some authors reported that uniform attenuation-corrected studies provided unreliable regional estimates of tracer activity.[24] The same study concluded that estimation of the attenuation map from a segmented reconstruction of a lower-energy Compton scatter window image was reported as the next most accurate clinical method and can be reliably used when transmission scanning cannot be used.[24] In contrast, semi- quantitative analysis of images reconstructed using transmissionless attenuation maps produces results that are very similar in ^{99m}Tc-ECD uptake values for healthy volunteers in comparison to those obtained with a transmission-based method.[25] However, special attention should be paid to the choice of the optimal effective broad-beam attenuation coefficient (μ_{eff}) to use when combining attenuation and scatter corrections[26] for reconstruction of emission data that may have been perturbed by scatter and attenuation in the human skull. The attenuation of the skull has been evaluated by many investigators,[25,27,28] all suggesting the use of a lower value of μ_{eff} than for a uniform soft tissue medium. The choice of the 'optimal value' of the linear attenuation coefficient was studied in an elegant paper by Kemp *et al.*[27] where the use of an effective bone and tissue attenuation coefficients to compensate ^{99m}Tc-HMPAO brain SPECT resulted in images of improved uniformity and increased count density. In another study using an anthropomorphic phantom, the 'best' choice of the effective linear attenuation coefficient was found to be slice-dependent and reliant on the skull thickness and the methods used for attenuation and scatter corrections.[29] Van Laere *et al.*[30] used an attenuation coefficient of

0.105 cm^{-1} determined from experimental studies using the 3D Hoffman brain phantom and 0.09 cm^{-1} for clinical studies[25] indicating that results obtained on phantom studies cannot be extrapolated directly for application on human data. Similarly, using the same experimental set-up, Montandon *et al.*[31] determined that the optimal effective μ value in 3D PET brain imaging is much lower compared to the theoretical value for scatter corrected data ($\mu = 0.096 cm^{-1}$) and was found to be equal to 0.06 cm^{-1}. The deviation from the theoretical value may, in all cases, be explained by non-optimal scatter corrections.

2.1.2 Other Methods

Another approach, which is receiving considerable attention, is to compute the attenuation map directly from the emission data, eliminating the transmission scan from the acquisition protocol. The problem of "transmissionless" image reconstruction in ECT has a long history, starting from a pioneering work by Censor *et al.*[32], where alternating iterations of the reconstruction algorithm were used to reconstruct emission tomograms and attenuation maps from a set of emission projections alone. Many researchers, who applied various optimization techniques, also have used similar philosophies in generating emission tomograms and attenuation maps. For instance, Nuyts *et al.*[33] formulated the problem as an optimization task where the objective function is a combination of the likelihood and an *a priori* probability. The latter uses a Gibbs prior distribution to encourage local smoothness and a multimodal distribution for the attenuation coefficients. Other methods included the use of the EM algorithm, as was done in ref.[34], or penalty functions.[35,36] The techniques have had limited success, but often produce artefacts in the form of cross-talk between the emission image and the attenuation map.

Other transmissionless reconstruction methods attempt to avoid cross-talk between the emission image and attenuation map by reconstructing the emission image and the attenuation map independently. A more general technique applies the consistency conditions for the range of the attenuated Radon transform to obtain the attenuation map from SPECT data. Implementations based on previously used continuous conditions have shown that reconstructions did not converge to an acceptable solution.[37] Bronnikov[38] recently suggested an original approach that strengthens the paradigm of the consistency conditions by setting them in the framework of a discrete representation of the problem. It should be pointed out that even if much worthwhile research has been performed in this area, there is no clear evidence from the published literature regarding the applicability of these techniques in a clinical environment. Further research is clearly needed to convince the nuclear medicine community and gain confidence in this approach. Potential future developments are addressed in section 6 below.

2.2 Transmission-based Approaches

In clinical and research applications, where the attenuation coefficient distribution is not known *a priori*, and for areas of inhomogeneous attenuation such as the chest, more direct methods must be used to generate the attenuation map. This includes transmission scanning,[39] segmented MRI data[40,41] or appropriately scaled X-ray CT scans acquired either independently on separate[42-45] or simultaneously on multimodality imaging systems[46-49]. These methods are described in more detail in the following sections.

2.2.1 Radionuclide Transmission Scanning

As reported by Bailey,[39] the use of *transmission scanning* using an external radionuclide source dates back to the pioneering work of Mayneord in the 1950's.[50] A more refined approach for acquiring transmission data for use in conjunction with conventional emission scanning was implemented by Kuhl in 1966.[51] Transmission scanning now is commonly available on commercial SPECT and PET systems, allowing it to be performed in clinical departments on a routine basis, especially when it is combined with simultaneous emission scanning. In a clinical environment, the most widely used attenuation correction techniques use transmission data acquired either before (pre-injection),[52,53] during (simultaneous),[54,55] or after (post-injection)[23,56] the emission scan. Interleaving emission and transmission scanning has proved to be very practical in oncology studies where multiple bed positions are needed. Sequential emission-transmission scanning is technically easier to perform than simultaneous scanning, but it increases the imaging time and suffers from image registration problems caused by patient misalignment or motion. Simultaneous acquisition requires no additional time for the emission and transmission measurements, which is important for routine clinical studies. However, errors may be introduced by cross-talk between the transmission and emission data. It has been shown that the attenuation coefficients and activity concentrations are not significantly different when estimated with sequential and simultaneous emission transmission imaging.[54] Since the reconstructed attenuation coefficients are energy-dependent, the reconstructed attenuation coefficients are transformed to the coefficients of the appropriate isotope energy using suitable techniques. The accuracy of the transmission and emission maps produced using different transmission-emission source combinations has been the subject of a long debate.[30,57] In addition, various approaches have been proposed to eliminate contamination of emission data by transmission photons and to reduce spillover of emission data into the transmission energy window.[57-59] Several transmission scanning geometries have emerged for clinical implementation for SPECT,[39,60] hybrid SPECT/PET[61] and dedicated PET,[39,62] as illustrated in Figures 2-4.[63] The following sections describe the different transmission sources and data acquisition geometries that have been proposed so far.

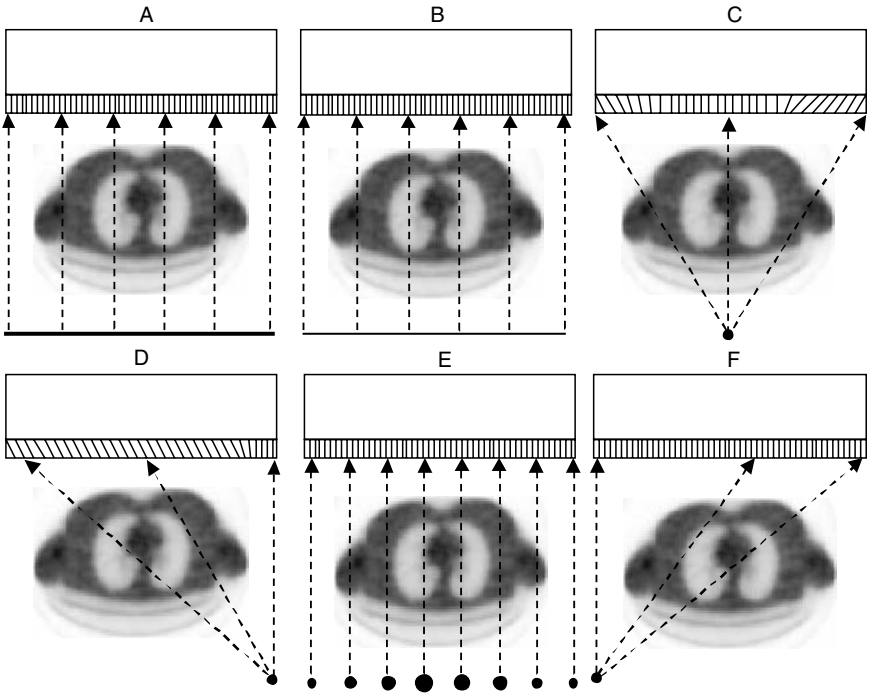


FIGURE 2. Different configurations of Htransmission scanning geometries for SPECT are shown in this diagram. **A.** Sheet source; **B.** scanning line source; **C.** fixed line source and converging collimation; **D.** point or line source and asymmetric fan-beam collimation; **E.** multiple line sources where the source size is proportional to the relative activity of the source, and finally **F.** point source and septal penetration of parallel collimation.

2.2.1.a SPECT

Radionuclide transmission-based methods in SPECT include both sequential and simultaneous scanning using external ^{57}Co, ^{99m}Tc, ^{133}Ba, ^{139}Ce, ^{153}Gd, ^{201}Tl, or ^{241}Am sources. Early designs of transmission systems for research with SPECT cameras used uncollimated flood or sheet sources. The main advantage of this configuration is that the source fully irradiates the opposite head, and therefore requires no motion of the source other than that provided by the rotation of the camera gantry. These geometries also have drawbacks associated with the high proportion of scattered photons in the transmission data due to the broad-beam imaging conditions. As a result, the attenuation map estimates an "effective" linear attenuation coefficient rather than the value that would be calculated from narrow-beam geometry. This difficulty can be overcome in part by collimating the transmission source[64] to produce a geometry that more accurately represents a narrow beam transmission geometry.

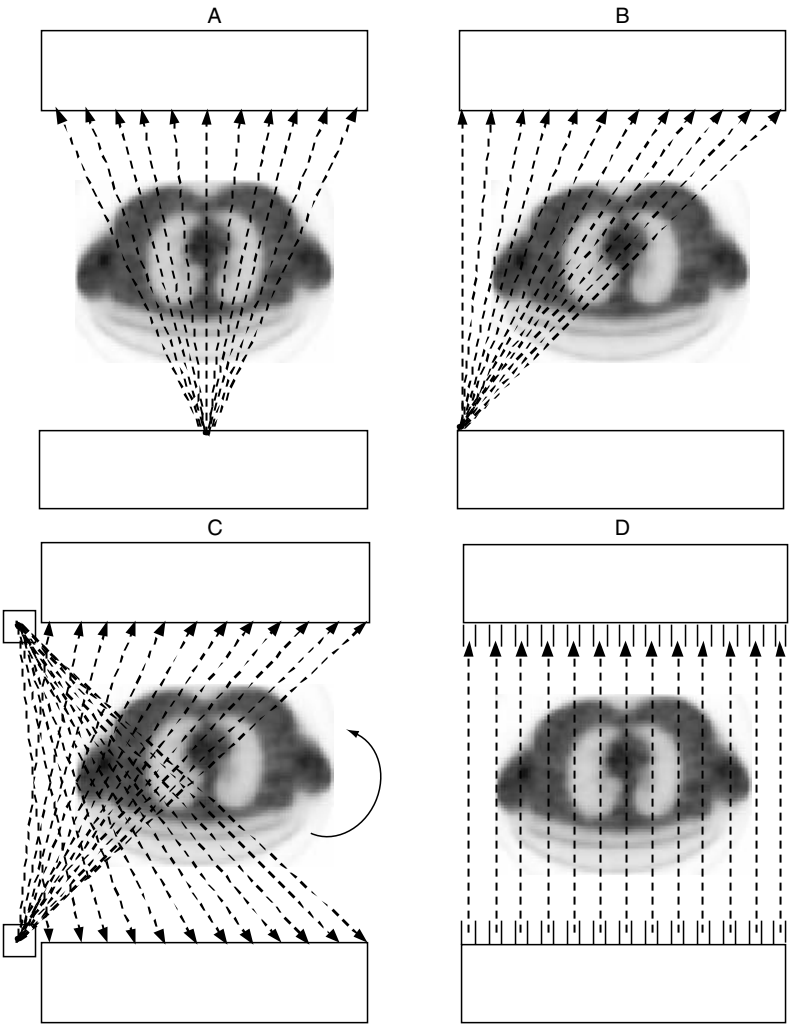


FIGURE 3. Different configurations of transmission scanning geometries for hybrid SPECT/PET cameras are shown. **A.** Symmetric point source and fan beam geometry which truncates the imaged object; **B.** Offset point source covering the entire field-of-view where the flux of radiation includes the center of rotation for sampling to be complete; **C.** 2 scanning point sources translating axially and located far enough to the side so that the flux of radiation does not include the center of rotation. Truncation is avoided by lateral movement of the bed and 360° rotation; **D.** Multiple point sources inserted between existing septa and placed along a line parallel to the axis of rotation near the edge of one camera. The septa provide axial collimation for the sources so that the transmission system operates in a two-dimensional offset fan-beam geometry.

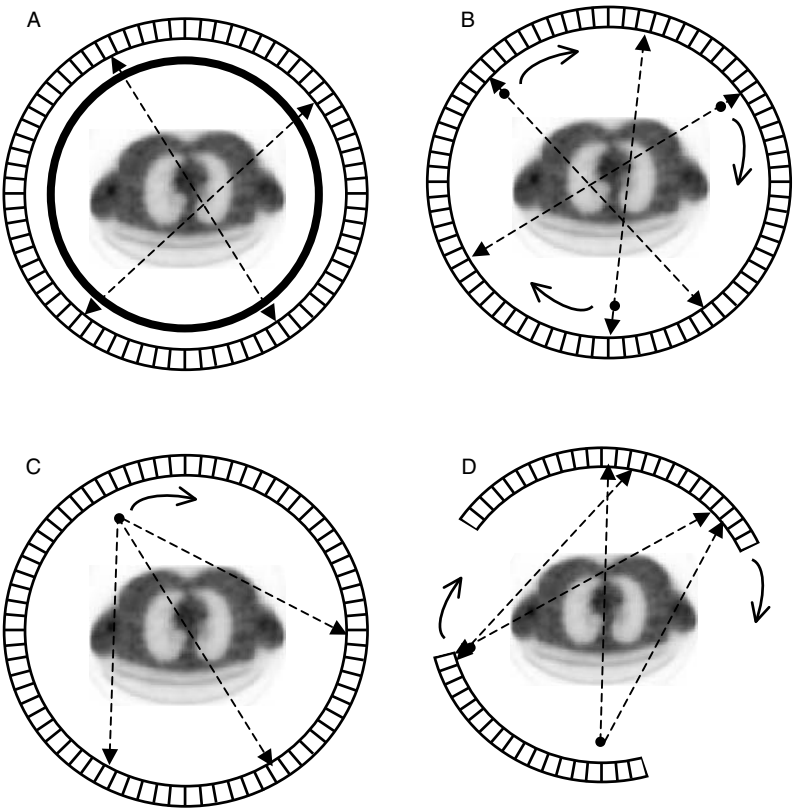


FIGURE 4. Different configurations of transmission scanning geometries for PET are shown in the diagram. **A.** Ring sources of a positron-emitting radionuclide measuring transmission in coincidence mode, **B.** rotating positron-emitting rods measuring transmission in coincidence mode, **C.** single-photon source producing coincidence events between the known source position and photons detected on the opposing side of the detector ring, and finally **D.** fixed positron-emitting rod (left) or single-photon (right) sources on a rotating, partial-ring scanner.

Traditionally, SPECT systems used either ^{99m}Tc or ^{201}Tl transmission sources that produce accurate attenuation maps for these respective emission radionuclides.[30,65] More recently, transmission data acquired with ^{153}Gd or ^{133}Ba external sources or with a rotating X-ray tube have been used to compute the attenuation map. The main radionuclide-based configurations include: (i) stationary line source fixed at the collimator's focus with convergent collimation on a triple-detector system,[54] (ii) scanning line sources with parallel-hole collimation,[66] (iii) a multiple line source array with parallel-hole collimation,[67] (iv) scanning point sources using either fan-beam and offset fan-beam geometry[68] or (v) asymmetric fan-beam geometry acquired by using a high-energy source that emits transmission photons capable of

penetrating the septa of a parallel-hole collimator.[69] Previously published reviews can be consulted for a more detailed survey of these geometries.[5,39,60,63] The most widely implemented configuration for commercial transmission acquisition is the scanning line source geometry. However, each configuration has its unique advantages and drawbacks and camera manufacturers are still optimizing the apparatus used to acquire the transmission data.

2.2.1.b Hybrid SPECT/PET

Several schemes have been proposed to perform transmission scanning on coincidence gamma camera systems.[61] Laymon *et al.*[70] developed a ^{137}Cs-based transmission system for a dual-headed coincidence camera that can be used for post-injection transmission scanning. Commercial SPECT/PET systems use single-photon emitting ^{133}Ba ($T_{1/2} = 10.5$ yrs, $E_\gamma = 356$ keV) point sources[69] or ^{153}Gd ($T_{1/2} = 240$ days, $E_\gamma = 97$-101 keV) line sources.[67] As noted above, ^{153}Gd is commonly used as a transmission source for SPECT. In addition, transmission data obtained using ^{153}Gd can be scaled to provide an attenuation map for coincidence imaging of positron emitters. In comparison, ^{133}Ba has a long half-life and does not have to be replaced as do shorter-lived radionuclide transmission sources such as ^{153}Gd and ^{57}Co, ^{99m}Tc, or ^{201}Tl. Furthermore, ^{133}Ba has the potential advantage that it is a radionuclide source with a photon energy (356 keV) that is between those encountered for emission imaging in PET (511 keV) and in SPECT (e.g., ^{201}Tl, ^{99m}Tc, ^{111}In, ^{123}I, ^{131}I). Therefore, it may be more suitable for obtaining transmission data for SPECT imaging than annihilation photons from a positron source used with PET, but may be more difficult to shield than other single-photon transmission sources commonly used with SPECT. It is important to note that any potential advantages or disadvantages of ^{133}Ba as a transmission source for hybrid SPECT/PET cameras have not been demonstrated. Finally, it also is possible to use an X-ray tube as a transmission source in hybrid SPECT/PET systems, such as that implemented with the General Electric Discovery VH dual-headed camera. The use of an X-ray tube offers the advantages of higher photon fluence rates and faster transmission scans, with anatomical imaging and localization capability that cannot be obtained using radionuclide transmission sources. However, the use of the X-ray tube also requires a separate X-ray detector since its photon fluence rate far exceeds the count rate capabilities of current scintillation camera technology, and as a point source its use is not compatible with parallel-hole collimators required for SPECT imaging.

2.2.1.c PET

The early PET scanners used transmission ring sources of the positron-emitting radionuclides ^{68}Ga/^{68}Ge ($T_{1/2} = 68$ min and 270.8 days, respectively), which co-exist in secular equilibrium. In this case, annihilation

photons are acquired in coincidence between the detector adjacent to the annihilation event and the detector in the opposing fan which records the second annihilation photon after it has passed through the patient. This design was modified later by replacing the ring sources radially by continuously rotating "rod" sources. Obviously, the detector block close to the rod source receives a high photon flux rate, causing detector dead time to be a major limitation in this approach.[39] This problem can be relieved by windowing the transmission data so that only events collinear with the known location of the rod are accepted. Scanner manufacturers have adopted this as a standard approach for several years. More recently, some manufacturers have implemented transmission scanning using single-photon sources such as ^{137}Cs ($T_{1/2} = 30.2$ yrs, $E_\gamma = 662$ keV). Transmission data recorded with an external single-photon source can be recorded at higher count rates resulting from the decreased detector dead time. In addition, a ^{137}Cs transmission source produces a higher energy photon and therefore improves object penetration[62] in comparison to a positron-emitting transmission source. The 662 keV photons from ^{137}Cs are less attenuated than the annihilation photons from the PET emission source. The attenuation map generated from transmission data acquired with a ^{137}Cs transmission source must be corrected to account for differences in photon attenuation between the emission and transmission data.

In recent volumetric PET systems like the ECAT ART that operate exclusively in 3D mode, attenuation correction factors are measured with two single-photon collimated point sources of ^{137}Cs capable of producing high-quality scatter-free data with this continuously-rotating partial-ring PET tomograph.[71] This allows transmission data to be acquired with improved counting statistics while drastically diminishing the acquisition time. This scanner is designed around single-photon transmission sources having two sets of 12 slits with an aperture ratio of 15:1 and an axial pitch equal to twice the pitch of the axial crystal ring.[52] A simple mechanism has been devised to produce a "coincidence" event between the detector, which records the transmitted single-photon, and the detector in the opposing fan near the current location of the single-photon source. More recently, a simultaneous emission-transmission scanning system has been developed that reduces contamination of the emission data by the emitted transmission photons using a fast, dedicated, lutetium oxyorthosilicate (LSO)-based reference detector placed close to the collimated coincidence point source used to produce the transmission data.[55]

2.2.1.d Segmentation of transmission data

When radionuclide sources are used to acquire the transmission data, photon statistical noise from the transmission scan can propagate through the reconstruction process, affecting the quality of the reconstructed images. To minimize this effect, long transmission scans are normally acquired to ensure

good statistics at the expense of patient throughput especially in the case of whole-body scanning with low-sensitivity tomographic systems. Alternatively, image segmentation can be applied to delineate different anatomical regions (e.g., lung vs. soft tissue) in the attenuation map. The known attenuation coefficients of these tissues then can be applied to the segmented regions to minimize noise in the resulting attenuation map, with the goal of reducing noise in the associated attenuation-corrected emission tomogram. During the last decade, techniques using transmission image segmentation and tissue classification have been proposed to minimize the acquisition time (less than 3 min) and increase the accuracy of the attenuation correction process, while preserving or even reducing the noise level. The reconstructed transmission image pixels are segmented into populations of uniform attenuation. The classified transmission images are then forward projected to generate new transmission sinograms to be used for attenuation correction of the corresponding emission data. This reduces the noise on the correction maps while still correcting for specific areas of differing attenuation such as the lungs, soft tissue and bone.

The whole process is illustrated in Figure 5. Once the transmission image is segmented, the tissue type corresponding to each label is identified (e.g. lungs, soft tissue, air . . . etc.). The attenuation map is then calculated by weighted averaging combining the segmented and original images.

In a clinical setting, segmentation algorithms must be designed to balance image quality and computational time. The majority of segmentation methods used for attenuation correction fall into one of the following two classes (see chapter 10): histogram-based thresholding techniques[72,73] and fuzzy-clustering based segmentation techniques.[74,75] Threshold approaches use the grey-level histogram counts to distinguish between regions. However, if the geometry of the attenuation map is based solely on the characteristics of the histogram, the technique is most likely to fail in regions where the total number of counts is small (e.g. the skull). Therefore, the performance of these techniques strongly depends on the choice of the thresholds. In

FIGURE 5. Illustration of the segmentation process of a heart/chest phantom transmission images showing from left to right: the original transmission image, FCM segmented image using 5 clusters, merging process to a three- clustered image (the bed is removed here), and the final attenuation map after applying a Gaussian filter and weighted averaging (the low-noise image of the bed acquired and processed off-line is added here).

comparison, fuzzy-clustering based segmentation techniques have demonstrated excellent performance and produced good results as an automated, unsupervised tool for segmenting noisy images in a robust manner. They are iterative procedures that minimize an objective function. As an output, a membership degree is assigned to every voxel with respect to a cluster centre. The number of clusters is generally passed as an input parameter. To automate the process, a cluster validity index can be used to select the optimal number of clusters.[76]

Other interesting approaches to segment noisy transmission data include the use of active contour models,[77] neural networks,[78] morphological segmentation,[79] and hidden Markov modelling.[80] An alternative to segmentation of transmission images with the goal of reducing noise in PET transmission measurements includes Bayesian image reconstruction[81-83] and non-linear filtering.[84,85]

2.2.2 X-ray Transmission Scanning

It is well known that x-ray computed tomography (CT) can provide a patient-specific map of attenuation coefficients that can be used to compensate radionuclide data from PET or SPECT for the effects of photon attenuation.[86,87] This is a natural observation since the CT image inherently represents the three-dimensional spatial distribution of attenuation coefficients of the patient. In addition, CT scanners are widely available and can produce cross-sectional images quickly with low noise and excellent spatial resolution in comparison to transmission images produced with external radionuclide sources. Overall, x-ray transmission techniques offer practical advantages for producing attenuation coefficient maps in terms of high photon output and short procedure times, excellent decay, lack of physical decay with a source that can be turned off between scans and that can be discarded easily at the end of its useful operating life.

Several researchers have investigated the use of x-ray CT to provide a patient-specific map of attenuation coefficients. The first studies of this type were performed by investigators who scanned the patient in separate clinical procedures to acquire the CT and radionuclide data, then relied on image fusion techniques to register the data before the attenuation compensation technique was performed.[43,88,89] These techniques can be applied to imaging studies of the head where the rigidity provided by the skull facilities off-line registration of the x-ray and radionuclide data. However, other regions of the body can bend and flex, thereby complicating image registration. In addition, the heterogeneity of the body, especially in the thorax, accentuates the need for accurate registration when x-ray measurements are used to produce an attenuation map for reconstruction of the radionuclide tomograms. However, the advent of dual-modality imaging systems (see chapter 2) provides a means of acquiring CT and radionuclide data so that they are inherently coregistered. This allows obtaining a patient-specific map of

attenuation coefficients that can be used to compensate the radionuclide data for photon attenuation in all regions of the body. Some early PET/CT systems allowed the operator to obtain transmission data with an external radionuclide source,[49] similar to the technique used with conventional CT where the x-ray subsystem was not available. However, increasingly, dual-modality imaging systems rely exclusively on correlated CT to obtain the attenuation map for reconstruction of the radionuclide tomogram.

While the image data from a CT scanner inherently represents the spatial distribution of linear attenuation coefficients in the patient,[90] the CT data from a specific clinical study cannot be used as an attenuation map for compensation of radionuclide data without accounting for the specific characteristics of CT imaging study. First, the CT projection data are obtained using a polyenergetic x-ray source and are recorded with an array of current-integrating x-ray detectors. Furthermore, as the x-ray beam passes through the patient, the low energy photons are preferentially absorbed in the patient so that the x-ray spectrum transmitted through a thick body part has a higher mean area than that transmitted through thinner regions of the body. This produces the well-known "beam-hardening" artefact in which the raw values of the CT image are lower in the middle of an object than at the periphery, representing the higher mean energy and correspondingly lower attenuation coefficients obtained in thick body parts than those produced for thin regions of the body. All clinical CT scanners incorporate calibrations to correct for spectral changes due to object thickness. Nevertheless, these considerations underscore the need to present the CT image data in a way that is relatively immune to changes in the x-ray energy spectral content. For clinical applications of CT, in order to present the CT image in a way that represents tissue type rather than the measured linear attenuation coefficient, the raw CT data are normalized by calculating each pixel value in terms of "Hounsfield Units" (HU) where

$$HU(x,y) = \frac{\mu_{CT}(x,y) - \mu_w}{\mu_w} \times 1000 \tag{2}$$

where $HU(x,y)$ is the value of the CT scan expressed in Hounsfield units at point (x,y), $\mu_{CT}(x,y)$ is the linear attenuation coefficient obtained from the raw CT scan at the location (x,y), and μ_w is the corresponding value of the linear attenuation coefficient of water. When the CT data are rescaled in terms of Hounsfield units, air corresponds to -1000 HU, water corresponds to 0 HU, bone is represented by values of 1000 to 2000 HU, and fat has a value of approximately -100 HU where the relationship between tissue type and approximate value in Hounsfield units is relatively independent of the x-ray spectrum or other parameters used to generate the CT image. The use of Hounsfield units to present the CT image thereby provides some consistency that facilitates their visual interpretation. However, when CT data from a clinical scanner is used to obtain a patient-specific map of attenuation

coefficients, it therefore is necessary to transform the CT image data so that it is expressed in terms of the linear attenuation coefficient of the corresponding material. Furthermore, this must be done for the energy of the radionuclide photon so that the attenuation data can be used to correct the emission data for photon attenuation.

The first method of calculating an attenuation map from CT requires that the CT image is segmented in a way that parallels the segmentation of transmission data discussed in Section 2.2.1.4. Segmentation is a method in which image processing techniques are used to identify regions of the CT image that represent different material types (e.g., soft-tissue, bone, and lung).[87] After each of these tissue regions is identified, the CT image values for each tissue type are then replaced by the linear attenuation coefficients for the corresponding material at the energy of the radionuclide photon. This method has the advantage that the attenuation map is "noiseless" such that each region of a single material is assigned a single linear attenuation coefficient value. However, the assigned value may be inaccurate, especially in pulmonary regions where the density of the lung is known to vary by as much as 30%.[91]

A second method of transforming the CT data to a map of attenuation coefficients directly transforms the CT image to a map of attenuation coefficients at the energy of the radionuclide photon. This transformation can be performed by acquiring x-ray CT calibration measurements of a calibration phantom containing materials of known composition (and thus of known linear attenuation coefficient at a given photon energy) and from which the CT number (in HU) can be measured directly from the CT scan of the phantom. Some of the earliest work in this area was performed in conjunction with the development of a SPECT/CT system by Blankespoor *et al.*[86] in which transmission data were obtained then reconstructed using x-ray CT, then calibrated using measurements from a phantom with cylindrical inserts containing water, fat-equivalent (ethanol) and bone-equivalent (K_2HPO_4) materials, and iodinated contrast agents. In order to convert the CT numbers in the calibration phantom to units of attenuation coefficient at the photon energy of the radionuclide, the CT numbers extracted from each region were plotted against their known attenuation coefficients at the photon energy of the radionuclide to provide a piece-wise linear calibration curve[86,92] which was stored as a look-up table in the computer used to reconstruct the radionuclide tomograms. During a clinical study, both CT and SPECT data were obtained from the patient and reconstructed using the commercial filtered backprojection software provided by the manufacturer to provide preliminary tomographic data that could be registered spatially. The CT image then is reformatted so that it had the same matrix format, same slice width, and same slice positions as the SPECT data, then was transformed using the calibration look-up table to produce an attenuation map expressed in units of linear attenuation coefficient for the photon energy of the radionuclide image. The resulting attenuation map then was

incorporated with the original radionuclide projection data into an iterative maximum-likelihood expectation-maximization (ML-EM) algorithm to reconstruct the radionuclide data. These techniques have been tested extensively in SPECT/CT studies on phantoms,[93] animals,[92] and patients.[93,94]

A third technique for transforming the CT image to an attenuation map is the hybrid method that combines segmentation and scaling.[44,87] This method uses segmentation to separate regions of the image that represent bone, from those areas which represent other tissue types. Different scale factors then are used to transform the CT values to linear attenuation coefficients in these two regions. The hybrid technique has been applied primarily to PET images, and both bilinear scaling and the hybrid technique have demonstrated that they provide accurate reconstructions when compared to attenuation maps generated with 511 keV positron emitting rod sources.

A SPECT/CT or PET/CT scanner can provide a virtually noiseless map of attenuation coefficients that can be used to correct the radionuclide image for photon attenuation errors. In addition, the transmission data can be acquired more quickly and with significantly better spatial resolution with an x-ray source than with a radionuclide source. While these features are advantageous in clinical studies, several investigators have identified some pitfalls associated with attenuation maps generated with x-ray transmission scanning. One area of concern arises because the x- ray scan is obtained with a polyenergetic bremsstrahlung spectrum with a mean photon energy that generally is different than that of the radionuclide used for the emission scan. The x-ray data therefore must be transformed using one of the techniques described above to obtain an attenuation map that can be used to correct the emission data for photon attenuation. In measurements from a commercial PET/CT scanner, Burger *et al.*[95] found a slight (approximately 3%) but measurable difference in the attenuation map values generated from x-ray transmission data and from ^{68}Ga positron-emitting transmission source (i.e., germanium-68). More noticeable artefacts are associated with respiratory-induced misregistration[96,97] of the emission and transmission data which can lead to a decrease in reconstructed activity at the margins of the lungs and diaphragm, and are seen with attenuation maps generated using both radionuclide and x-ray transmission sources. These arise because the emission scan is acquired over multiple respiratory cycles under shallow-breathing conditions, whereas the CT scan is obtained quickly over just a short interval within one respiratory cycle. Furthermore, the misregistration errors are maximized when the transmission data are acquired with the patient at end-inspiration or with shallow tidal respiration, but are minimized when the transmission data are acquired with the patient at end-expiration[96] A third source of error arises when the patient contains materials that have physical densities and linear attenuation coefficients outside of normal physiological ranges. For example, high-density oral contrast[98] or metal implants in teeth[99] or hip prosthetic devices[100] as well as other high density structures and implants within the body can cause streaks and

"beam-hardening" artefacts in the CT scan or conventional transmission image. In turn, these can lead to artefacts in the PET or SPECT image that mimic areas of increase radionuclide uptake. Fortunately, both intravenous iodinated contrast media[101] and low-density barium contrast agents,[98] typically administered in clinical CT studies, appear to cause only minor artefacts in the transmission and emission data, and appear to be suitable for clinical imaging with PET/CT.

2.2.3 Segmented Magnetic Resonance Imaging

In addition to methods that use radionuclide or x-ray sources to produce attenuation maps for reconstruction of SPECT or PET data, a few studies have addressed the issue of using segmented MR data to construct an attenuation map for attenuation correction purposes in emission tomography[40] and that represent an outgrowth of the many PET/MR image coregistration and fusion algorithms that have been described in the literature.[102] For imaging the brain, the simplest method segments the MRI image by thresholding to create a mask, which delineates the skull and all other tissues but excludes the hollow space of sinus, etc. Every voxel in the mask is then assigned the attenuation coefficient of water ($0.096\,cm^{-1}$). A more robust approach based on registered 3D MRI T1 images has been proposed recently.[41] These images were realigned to preliminary reconstructions of PET data and then segmented using a fuzzy clustering technique by identifying tissues of significantly different density and composition. The voxels belonging to different regions can be classified into bone, brain tissue and sinus cavities. These voxels were then assigned theoretical tissue-dependent attenuation coefficients as reported in the ICRU 44 report.[2] An equivalent method substituting the patient-specific MR images with a coregistered digitized head atlas derived from high-resolution MRI-based voxel head model[103] called inferring-attenuation distributions (IAD) has been proposed by Stodilka *et al.*[104] for brain SPECT and extended later for brain PET imaging. The feasibility of Atlas or template-guided attenuation correction in cerebral 3D PET imaging, where the patient-specific attenuation map is derived by non-linear warping of a transmission template constructed by scanning a representative population of healthy subjects has been recently proposed and is being validated in a clinical environment.[105] Its accuracy depends strongly on the performance of the coregisteration and anatomic standardization techniques applied to the tracer-specific emission template to match the patient data.

The availability of public domain image registration and segmentation software dedicated for brain (e.g. Statistical Parametric Mapping package) facilitates clinical implementation of MR-based attenuation corrections. The recent interest in simultaneous multimodality PET/MRI may motivate future applications of this method to other organs where accurate registration is more difficult to achieve. For example, a prototype small animal PET scanner has been designed with lutetium oxyorthosilicate (LSO) detector blocks of 3.8 cm

ring diameter coupled to three multi-channel PMT's via optical fibres[106] so that the PET detector can be operated within a conventional MRI system. The authors reported no appreciable artefacts caused by the scintillators in the MR images. A second larger (11.2 cm) prototype is being developed for simultaneous PET/MR imaging of mice and rats at different magnetic field strengths.[107] Although the challenges and costs for these devices are substantial, the future potential of multimodality imaging appears to be bright.

A recent study[22] compared the impact of the attenuation map using clinical brain scans corrected for attenuation using both uniform attenuation maps based on manual and automatic contours described in Section 3.1 and nonuniform attenuation maps described in Section 4 including transmission scanning,[53] segmented transmission,[73] coregistered segmented MRI,[41] and the IAD method.[104] This later method was implemented as described by Stodilka *et al.*[104] without any modifications (e.g. adding the bed to the final images). From a purely qualitative analysis, the merits of the more exact methods based on realistic nonuniform attenuation maps are obvious.[22] They produce less visible artefacts, while the approximate methods tend to produce an artefact in which there is a high level of activity along the edge of the image due to overestimation of the head contour on the external slices when using the automatic edge detection method. On the other hand, the quantitative VOI-based analysis of ten patient data sets revealed different performance and statistically significant differences between the different attenuation correction techniques when compared to the gold standard (transmission scanning).

Typical patient brain attenuation maps and corresponding PET images acquired with the ECAT ART camera and reconstructed with measured transmission as well as MRI-guided attenuation and scatter corrections are shown in Figure 6. Both correction methods improve the quality of the images and allow a better definition of brain structures and better overall contrast between gray and white matter compared to the case where no correction is applied; however, the images appear noisier when using transmission-guided attenuation correction. It should be noted that the long acquisition time during transmission scanning significantly improved the signal-to-noise ratio by reducing noise propagation from transmission to emission scanning. In addition, the quantitative analysis showed good agreement between measured transmission-guided and segmented MRI-guided reconstructions with only small differences between the two distributions.

3. Attenuation Correction Techniques in SPECT and PET

In the two-dimensional case, the fundamental relation that links the imaged object $f(x,y)$ and corresponding attenuation map $\mu(x,y)$ to its 1D projection data $p(s,\phi)$ is called the 2D central slice theorem (see chapter 3). The general

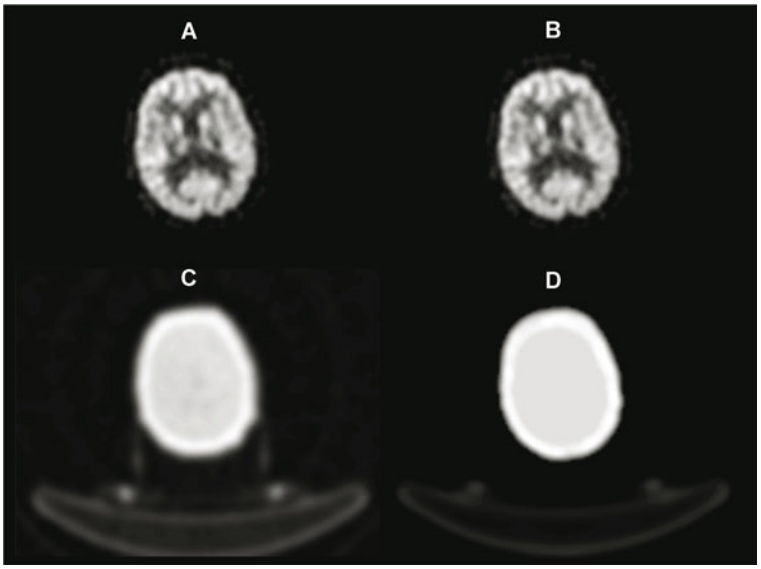


FIGURE 6. Comparison of attenuation maps and PET image slices of a patient study reconstructed using the two different processing protocols. **A.** PET image reconstructed using transmission-guided attenuation and scatter corrections. **B.** PET image reconstructed using segmented MRI-guided attenuation and scatter corrections. **C.** Measured transmission-based attenuation map. **D.** MRI-guided attenuation map (Reprinted with permission from ref.[41]).

equation describing measured projections in term of the radionuclide source distribution inside an attenuating medium is called the attenuated Radon transform and is given in the case of SPECT by:

$$p(s,\phi) = \int_{L(s,\phi)} f(x,y) \exp\left[- \int_{0}^{l(x,y)} \mu(x',y')dl \right] dr \tag{3}$$

where $l(x,y)$ is the distance from the emission point (x,y) in the object to the detector along the line $L(s, \phi)$, ϕ is the angle between the rotating detector plane and the stationary reconstruction plane and $\mu(x',y')$ the attenuation coefficient at position (x',y'). Whereas, in the case of PET, the attenuated Radon transform is given by:

$$p(s,\phi) = \int_{L(s,\phi)} f(x,y)dr \times \exp\left[- \int_{L(s,\phi)} \mu(x,y)dl \right] dr \tag{4}$$

Ideally, one would like to solve the Radon transform exactly to determine or "reconstruct" the radionuclide distribution $f(x,y)$. However, because of the complexity of the equation, no exact analytical reconstruction method

exists to invert the attenuated Radon transform. This is especially true in clinical imaging where the attenuation coefficient distribution is nonuniform.

Transmission-based attenuation correction has been traditionally performed in the case of PET, which started mainly as a research tool where there was greater emphasis on accurate quantitative measurements, and more recently has been applied for SPECT. There are two simple reasons for that: (i) attenuation correction in PET is easy since it requires a simple pre- multiplication of the measured emission data by the corresponding attenuation correction factors, and (ii) the attenuation correction factors are large and quantitation is impossible without attenuation compensation. Interestingly, the magnitude of the correction factors required in PET is far greater than in SPECT. For a given projection, the SPECT attenuation correction factors rarely exceed 10 in virtually all clinical imaging, whereas for PET, they often exceed 100 for some lines of response through the body. Typically, the magnitude of the correction factors ranges from approximately 20 in PET and decreases in SPECT down to 9–10 for ^{201}Tl (69–80 keV), 6–7 for ^{99m}Tc (140 keV) to nearly 3 for ^{18}F (511 keV).[39]

The attenuation factor for a given line of response in PET depends on the total distance travelled by both annihilation photons ($\boldsymbol{a} + \boldsymbol{b}$ in Figure 7) and is independent of the emission point along this line of response. In comparison, SPECT models attenuation processes in which the emitted photon traverses only part of the patient's anatomy before reaching the detector. Figure 7 shows a transmission image along with the attenuation paths for both single-photon and coincidence detection modes. Therefore, correction for attenuation can be performed only for uniform attenuation maps using analytic reconstruction but otherwise requires iterative reconstruction techniques for SPECT, but it is more straightforward and its accuracy limited only by the statistics of the acquired transmission data for PET.

Because attenuation correction in PET is relatively straightforward, only two techniques have emerged. In the first approach, attenuation correction is performed in projection (sinogram) space by multiplying attenuation correction factors (*ACFs*) by emission data. Attenuation correction factors are generated by forward projecting at appropriate angles the attenuation map obtained using one of the methods described in the previous sections. Alternatively, when emission data are reconstructed using an iterative algorithm (see chapter 4), the attenuation correction factors can be used to provide proper statistical weighting to the data as is done in attenuation-weighted OSEM (AWOSEM).[108] It is worth emphasizing that this method attracted the interest of many scanner manufacturers who implemented it in software they supply to end-users.

The problem of photon attenuation in SPECT has proven to be more difficult to solve than for PET, and several types of correction methods have been suggested.[109,110] Nevertheless, recent iterative algorithms converge to a very accurate solution even in SPECT.

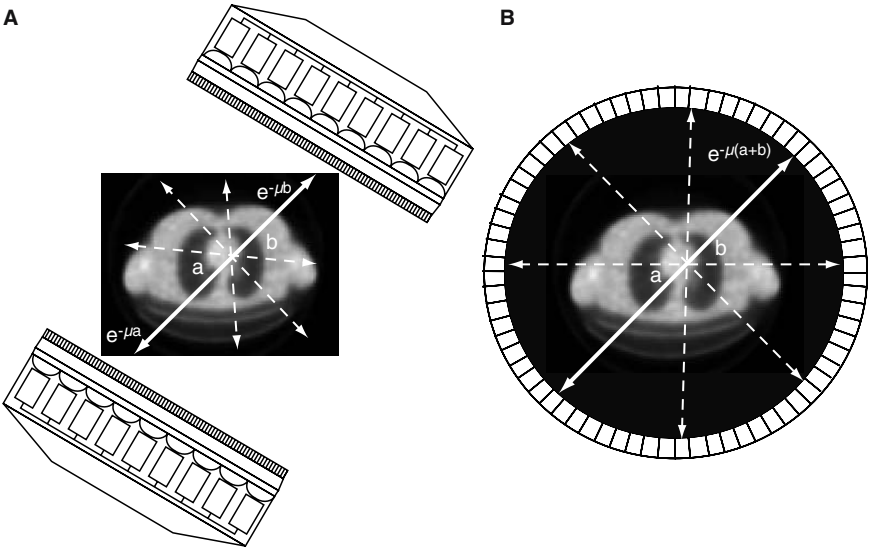


FIGURE 7. Illustration of the main differences between attenuation correction schemes for SPECT and PET on a transmission image of a patient in the thorax region. **A.** In SPECT, the most widely used algorithm on commercial systems calculates the attenuation factor ($exp(-\mu a)$ for uniform attenuating media) for all projection angles and assigns an average attenuation factor to each point within the object along all rays. The procedure can be repeated iteratively and adapted to non-uniform attenuating media. **B.** In PET, the attenuation correction factors are independent of the location of the emission point on the LOR and are therefore given directly by the factor ($exp\{-\mu(a+b)\}$ for uniform attenuating media) for each projection.

3.1. Correction Methods Based on Uniformly Attenuating Medium

Some of the first attenuation correction techniques used clinically in nuclear medicine, especially for SPECT, were designed to correct radionuclide images reconstructed using analytic techniques such as filtered backprojection. However, as noted above, it is difficult, if not impossible, to compensate the SPECT image for nonuniform attenuation using analytic reconstruction techniques. It is possible, however, to correct the radionuclide data assuming that the attenuation is uniform, for example in the head or pelvis (if the bony structure and other nonuniformities are assumed to be negligible). However, the assumption of a uniformly attenuating medium is not suitable for reconstructing emission data from the thorax due to the presence of the lungs. Attenuation correction of radionuclide data using a uniform attenuation map therefore is not appropriate for myocardial perfusion imaging where photon attenuation is considered to be a major source of false positive errors.

In cases where the linear attenuation coefficient is uniform throughout the reconstruction volume, it is possible to reconstruct the radionuclide image with a uniform attenuation map using an analytic reconstruction technique. This can be accomplished using a method developed by Chang,[111] now referred to as the "multiplicative Chang technique". Specifically, if the object has a constant linear attenuation coefficient μ within its borders, then the radionuclide image $f_{AC}(x,y)$ corrected for photon attenuation is given by

$$f_{AC}(x,y) = \frac{f(x,y)}{\frac{1}{M}\sum_{i=1}^{M} \exp(-\mu l_i)} \tag{5}$$

where the radionuclide image $f(x,y)$ is first reconstructed without attenuation correction. The denominator represents the multiplicative correction term for photon attenuation in which M is the total number of projections used to acquire the radionuclide data and l_i is the distance from the point (x,y) to the border of the object in the i^{th} radionuclide projection. When the object being imaged has an extended geometry, the correction term expressed in Eq. 5 overcompensates the image for photon attenuation which results in overcorrection in some parts of the image and undercorrection in other parts. In these cases, the accuracy of the photon attenuation correction can be improved using an iterative technique (referred to as the "iterative Chang technique") in which the initially corrected image $f_{AC}(x,y)$ is reprojected mathematically to form a set of projection data that simulate the radionuclide imaging process in a way that includes *a priori* information about the physical extent and linear attenuation coefficient distribution of the object. The new estimated projections then can be subtracted from the original measured projection data to calculate an error term that can be backprojected using filtered backprojection reconstruction or other reconstruction techniques, which then can be used to correct the radionuclide tomograms initially obtained from filtered backprojection. In spite of its theoretical shortcomings, Chang's method has been used widely in multiple clinical applications and research studies. Other less popular approaches have been described elsewhere[110] and will not be repeated here.

3.2. *Correction Methods Based on Non-Uniformly Attenuating Medium*

Radionuclide images reconstructed using analytic methods such as filtered backprojection generally are reconstructed without attenuation correction, but can incorporate a uniform attenuation map using the techniques described above. However, since most anatomical regions are heterogeneous in their anatomical structure and tissue composition, the radionuclide data inherently contain errors contributed by photon attenuation. In addition, significant errors in the radionuclide image can be contributed by other

physical effects such as scattered radiation and the finite spatial resolution of the imaging system (including the geometrical response of the collimator in the case of SPECT) that are not accounted for in these reconstruction techniques.[112-116] As discussed in the introduction to section 3, PET images can be reconstructed using either analytic or iterative techniques, whereas nonuniform attenuation correction for SPECT generally requires iterative reconstruction.

Several different iterative reconstruction methods have been developed and evaluated for radionuclide imaging. Early techniques used additive correction terms in each iterative step to improve computational speed and included the conjugate gradient weighted least squares (CG-WLS) technique.[117] However, most current reconstruction methods are based on the well-known maximum likelihood expectation-maximization (ML-EM) method.[118,119] This is an iterative reconstruction technique that accounts for the Poisson statistical nature of the radionuclide data, and that can incorporate other physical effects including a projector model that accounts for the geometric response of the collimator. In addition, the ML-EM algorithm has a positivity constraint (i.e., all reconstructed pixel values are non-negative), preserves the integral number of counts in each iterative step, and converges monotonically to a maximum likelihood solution. The mathematical form of the maximum-likelihood expectation-maximization (ML-EM) algorithm generally is expressed as (see chapter 4):

$$f_j^{new} = \frac{f_j^{old}}{\sum_l a_{lj}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old}} \tag{6}$$

where p_i are projection pixel values for counts originating from the object voxel activity concentration f_j, a_{ij} is the transition or system matrix, which represents the likelihood that a photon emitted from object voxel j is detected at detector pixel i, and thereby contains information about the geometry of the imaging system including the collimator design, the characteristics of the detector used to acquire the radionuclide data from the object, and the physical characteristics of the object including photon attenuation and scatter that affect the acquisition process. f_j^{new} and f_j^{old} refer to the current and previous estimates of the reconstructed object, respectively.

The image reconstruction begins with an estimate of the patient radionuclide distribution f_j^0 generally assumed to be uniform within the reconstruction circle of the imaging system for all values of j. Eq. 6 then mathematically projects the radionuclide distribution onto the detector using the properties of the transition matrix and compares the results with the experimentally measured projection data p_i. The algorithm then calculates a correction term that is used to update the estimated image data so that a new iterative estimate can be calculated. The EM algorithm described in Eq. 6 can be accelerated significantly by grouping the projection data into "ordered subsets" the collection of which encompass all of the projection

data measured during the imaging process. In the standard ML-EM algorithm, the EM algorithm is applied to the entire set of projection data. In comparison, in the "ordered-subsets expectation maximization" (OS-EM) algorithm, the EM algorithm then is applied to each of these subsets in turn with the image estimate f_j obtained from processing one subset forming the estimate used as the starting point in Eq. 6 to process the next subset. The iterative OS-EM algorithm can include photon attenuation and other physical effects through their expression in the detection probability matrix a_{ij}, thereby allowing correction of the radionuclide image for the effects of nonuniform photon attenuation.

The direct application of the attenuation correction process can be performed in several ways. In the case of PET, the attenuated Radon transform can be expressed as noted in Eq. 3. Therefore, projection data can be corrected for photon attenuation by using the following expression:

$$p_{AC}(s,\phi) = ACF \times p(s,\phi) = \int_{L(s,\phi)} f(x,y)dr \tag{7}$$

where

$$ACF = \exp\left[\int_{L(s,\phi)} \mu(x,y)dr\right] \tag{8}$$

is the "attenuation correction factor" for the line of response defined by the coincident annihilation photons acquired during the PET imaging process. Since the ACF in PET is defined for lines of response that traverse the entire width of the patient, its value can be determined experimentally using transmission measurements obtained through the patient using an external radionuclide or x-ray source. It therefore is possible to apply Eq. 7 to correct the projection measurements obtained during the PET imaging process, then reconstruct the resulting attenuation-corrected projection data $p_{AC}(s,\phi)$ using either analytic or iterative reconstruction techniques.

For SPECT or PET, the actual process of including nonuniform attenuation in the iterative reconstruction algorithm also can be performed using an attenuated projector-backprojector pair such as the one described by Gullberg *et al.*[120] Referring to Figure 8, Gullberg *et al.*, showed that if f_j and μ_j are the radionuclide concentration and the linear attenuation coefficient respectively at voxel j along a ray at angle ϕ described in the projector model of the radionuclide imaging process, and if r_j and r_{j+1} are the positions of the leading and trailing edge of *voxel j* through which the ray passes enroute to the detector, then the projected value recorded by the detector is given by

$$p(s,\phi) = \int_{L(s,\phi)} f(r)\exp\left[-\int_{r}^{+\infty} \mu(r')dr'\right]dr \tag{9}$$

Eq. 9 also can be expressed as:[120]

$$p(s,\phi) = \sum_j f_j \exp\left[-\int_{r_{j+1}}^{+\infty} \mu(r')dr'\right] \int_{r_j}^{r_{j+1}} \exp\left[-\mu_j(r_{j+1}-r)\right]dr \qquad (10)$$

where the first exponential term represents the attenuation of photons as they emerge from the right side of the pixel toward the detector, and the second exponential term represents the effects of attenuation across the pixel. Finally, Gullberg, *et al.*[120] have shown that the attenuated projection values $p(s,\phi)$ can be expressed as

$$p(s,\phi) = \sum_j f_j F_j(s,\phi) \qquad (11)$$

where

$$F_j(s,\phi) = \begin{cases} \frac{1}{\mu_j}\left\{\exp\left[-\int_{r_{j+1}}^{+\infty}\mu(r)dr\right] - \exp\left[-\int_{r_j}^{+\infty}\mu(r)dr\right]\right\} & \text{if } \mu_j > 0 \\ L_j\left\{\exp\left[-\int_{r_{j+1}}^{+\infty}\mu(r)dr\right]\right\} & \text{if } \mu_j = 0 \end{cases} \qquad (12)$$

and where $L_j = r_{j+1} - r_j$ is the ray length through pixel j having radionuclide concentration f_j and linear attenuation coefficient μ_j (see Figure 8).

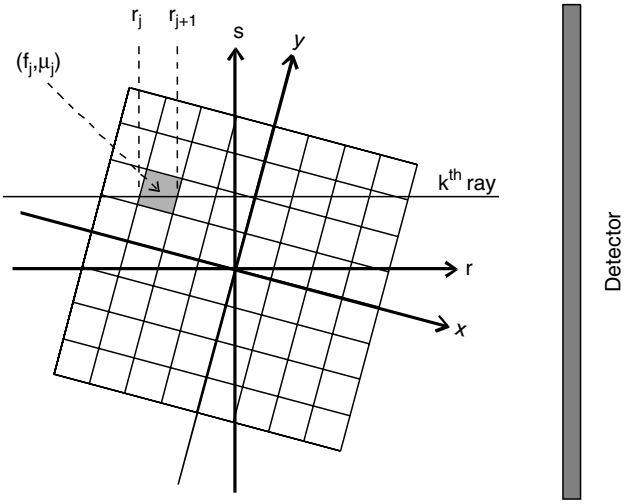


FIGURE 8. Geometry of attenuated projector-backprojector defined for k^{th} ray passing through an object in which the j^{th} voxel has radionuclide concentration f_j and linear attenuation coefficient μ_j.

The astute reader will notice that the expression for $F_j(s,\phi)$ allows the user to program a "bootstrap" operation in which the exponential terms in the expression for $F_j(s,\phi)$ can be used to calculate the subsequent term $F_{j+1}(s,\phi)$ as the ray is indexed through the object to calculate the projection data in a way that facilitates the overall computational process.

4. Dosimetric, Logistic and Computing Considerations

There have been few dose estimates reported in the literature on transmission scanning in emission tomography. Unfortunately, many studies were very superficial and reported very approximate estimates of the maximum absorbed dose at the skin surface using survey meters rather than effective dose equivalent (EDE) values (a quantity, which is suitable for comparing risks of different procedures in nuclear medicine, radiology, and other applications involving ionizing radiation) using anthropomorphic phantoms and suitable dosimeters.[30,121,122] It is worth noting the discrepancy between results reported in the literature even when using the same experimental set-up, scanning source geometries, and dosimeters. The discrepancies may be explained by differences in thermoluminescent dosimeters (TLDs) positioning and other uncontrolled factors. In SPECT, the use of sliding collimated line sources and parallel collimation allows a significant dose reduction when compared with the use of a static collimated line source and fan-beam collimation.[121] Van Laere *et al.*[30] reported an EDE rate of 32 μSv/GBq.h, while only 0.52 μSv/GBq.h has been reported by Almeida *et al.*[121] when using uncollimated and collimated ^{153}Gd transmission line sources, respectively, for typical brain scanning. A scanner manufacturer used low-dose X-ray device that adds a dose that is four-fold lower than that of a state of the art CT. The patient dose for a typical scan ranges from 1.3 mGy at the centre of the CT dose index phantom (32 cm diameter plastic cylinder) to 5 mGy at the surface (skin dose).[46] The radiation dose for another system was 3.4 mGy per slice at the center and 4.3 mGy per slice at the surface of a 16 cm diameter tissue-equivalent phantom.[47] A more recent study compared the difference between ^{68}Ga- and CT-based transmission scanning using a combined PET/CT device in terms of radiation burden to the patients during brain, cardiac and whole-body scans by using a Rando Alderson phantom with TLDs.[123] The authors reported negligible EDEs (<0.26 mSv) when using ^{68}Ga/^{68}Ge rod sources, while EDEs of 8.81 mSv in the high-speed mode and 18.97 mSv in the high-quality mode were measured for whole-body CT-based transmission scans on the Discovery LS system (GE Medical Systems, Milwaukee, WI, USA). It should be pointed out that these results serve only as guides since in most practical situations transmission scanning source geometries have not been characterized for each examination and specific camera using extensive Monte Carlo simulations or experimental measurements with anthropomorphic phantoms. To the best of our knowledge, these estimated numbers

are the only values that are available to end-users. Further efforts need to be devoted to fully characterize transmission scanning in emission tomography.

It is likely that if attenuation correction could be obtained rapidly and accurately, the benefits it adds to whole-body emission tomography would be considered sufficient to offset the increased acquisition time. Fortunately, several methods are under development or already have been implemented to perform attenuation correction more rapidly, more accurately, and with less noise than conventional techniques. Although the earliest emission tomographic studies had transmission images performed before radiotracer administration and emission imaging, such protocols are considerably time-consuming since the transmission and emission imaging sessions are completely separated. Methods that acquire transmission data after tracer injection now are performed more commonly.[124] Such methods save time but may have disadvantages, with potential artefacts if tracer activity is changing in location or intensity during data acquisition. However, a promising approach is the use of segmentation algorithms both to reduce the noise from the attenuation images and to reduce the time of acquisition by computer-aided classification of tissue density into a few discrete categories based on the transmission scan and prior knowledge of expected tissue attenuation coefficients. Indeed, algorithms with only a few minutes of acquisition per level have been developed using ^{68}Ga and ^{137}Cs sources.[75] More rapid data collection using either higher photon flux single-photon or X-ray sources is also promising. Combined SPECT/CT or PET/CT devices acquire transmission data rapidly and simplify the process of registering the emission data to the CT-derived anatomical images.[46,47,49,86] The cost of combined ECT/CT systems may be prohibitive for small nuclear medicine departments and correlation of data acquired from different scanners suffer from the usual problems of working with multi-modality images, namely, the difficulty of accurate coregistration from the different modalities. Nevertheless, dual-modality imaging is gaining in popularity and in acceptance for a number of important clinical applications.

Another aspect that deserves special attention, but which is not discussed in this chapter, is the need for efficient and accurate quality control and quality assurance procedures. Quality control procedures for transmission-emission tomographic systems are needed for insuring a good quality attenuation map and accurate attenuation correction. Currently, quality control for emission-transmission imaging is still undergoing development where only a handful of guidelines have been proposed for routine implementation.

5. Concluding Remarks and Future Prospects

Remarkable achievements and continuing efforts to enhance the performance of radionuclide imaging systems have improved both the visual quality and quantitative accuracy of emission tomography. However, physical

effects such as photon attenuation, scatter, spatial resolution characteristics, and statistical noise still limit the performance of emission tomography. The accuracy and clinical relevance of attenuation correction in emission tomography has been validated using either analytical[125] or Monte Carlo simulations,[29] experimental phantom studies,[75] animal studies,[126] biopsy samples taken after imaging was performed,[127] or more recently using Statistical Parametric Mapping (SPM)-based analyses of functional brain images[128] and receiver operating characteristic (ROC)-based analysis of clinical data.[129] It should be pointed out that the majority of attenuation correction methods described in the literature have been applied primarily to computer-simulated images and simplified experimental arrangements. Some solutions to the problem of attenuation correction are less suitable for routine applications in patients than they are in phantom simulations. The results reported in the literature concerning the quantitative accuracy of emission tomography very much depend on multiple factors including the phantom geometry, source size and distribution, noise characteristics, and the type of scanner. Quantitative data can be achieved within 10% when simulating clinically realistic activity distributions in anthropomorphic phantoms, which would be adequate for the majority of clinical applications. However, the accuracy obtained in phantom studies is unlikely to be reached in clinical investigations and the true clinical feasibility of these methods has yet to be fully investigated.

Considering the difficulties associated with transmission-based attenuation correction and the limitations of current calculated attenuation correction, if they were readily available, ''transmissionless'' attenuation correction might be the method of choice for the foreseeable future as a second best approach in a busy clinical department and remains a focus of many research groups. However, most manufacturers of nuclear medicine instrumentation still rely on transmission-based attenuation correction methods and have not show an interest in sophisticated transmissionless methods for clinical use. Attractive methods using transmissionless techniques to compensate for photon attenuation use either continuous[130-133] or discrete[38,125] consistency conditions. More importantly, the discrete consistency conditions have proved to be useful for estimating the thresholds for segmentation of the transmission image[134] and regularization parameters for attenuation map reconstruction.[135]

Some of these methods are considerably more computationally intensive than conventional techniques[34] and at present, the computational time achieved with common computing facilities available in nuclear medicine departments remains prohibitive, especially for large aperture cameras with fine sampling. However, with the development of more powerful multiple-processor parallel architectures and the potential for using these in conjunction with intelligent algorithms, the challenges of implementing transmissionless attenuation correction techniques may become more tractable. The clinical applicability of this approach remains to be demonstrated.

Acknowledgments. One of the co-authors (HZ) acknowledges the support of Grant SNSF 3152A0-102143 awarded by the Swiss National Science Foundation.

References

1. Berger M. J. and Hubbell, J. H., XCOM: Photon cross sections on a personal computer. National Bureau of Standards (US); Report No. NBSIR 87-3597, 1987.
2. ICRU, Tissue substitutes in radiation dosimetry and measurement. *International Commission on Radiological Units and Measurements*; Report No. 44, 1989.
3. Wackers F. J. T., Should SPET attenuation correction be more widely employed in routine clinical practice? *Eur J Nucl Med* **29:** 412-415 (2002).
4. Ficaro E. P., Should SPET attenuation correction be more widely employed in routine clinical practice? *Eur J Nucl Med* **29:** 409-412 (2002).
5. Hendel R. C., Corbett, J. R., Cullom, S. J. *et al.*, The value and practice of attenuation correction for myocardial perfusion SPECT imaging: a joint position statement from the American Society of Nuclear Cardiology and the Society of Nuclear Medicine. *J Nucl Cardiol* **9:** 135-143 (2002).
6. Bengel F. M., Ziegler, S. I., Avril, N. *et al.*, Whole-body positron emission tomography in clinical oncology: comparison between attenuation-corrected and uncorrected images. *Eur J Nucl Med* **24:** 1091-1098 (1997).
7. Wahl R. L., To AC or not to AC: that is the question. *J Nucl Med* **40:** 2025-2028 (1999).
8. Bedigian M. P., Benard, F., Smith, R. J. *et al.*, Whole-body positron emission tomography for oncology imaging using singles transmission scanning with segmentation and ordered subsets-expectation maximization (OS-EM) reconstruction. *Eur J Nucl Med* **25:** 659-661 (1998).
9. Larsson S. A., Gamma camera emission tomography: development and properties of a multi-sectional emission computed tomography system. *Acta Radiol Suppl* **363:** 1-75 (1980).
10. Ben Younes R., Mas, J. and Bidet, R., A fully automated contour detection algorithm the preliminary step for scatter and attenuation compensation in SPECT. *Eur J Nucl Med* **14:** 586-589 (1988).
11. Wallis J. W., Miller, T. R. and Koppel, P., Attenuation correction in cardiac SPECT without a transmission measurement. *J Nucl Med* **36:** 506-512 (1995).
12. Macey D. J., DeNardo, G. L. and DeNardo, S. J., Comparison of three boundary detection methods for SPECT using Compton scattered photons. *J Nucl Med* **29:** 203-207 (1988).
13. Bergstrom M., Litton, J., Eriksson, L. *et al.*, Determination of object contour from projections for attenuation correction in cranial positron emission tomography. *J Comput Assist Tomogr* **6:** 365-372 (1982).
14. Hosoba M., Wani, H., Toyama, H. *et al.*, Automated body contour detection in SPECT: effects on quantitative studies. *J Nucl Med* **27:** 1184-1191 (1986).
15. Tomitani T., An edge detection algorithm for attenuation correction in emission CT. *IEEE Trans Nucl Sci* **34:** 309-312 (1987).

16. Herbert T. J., Gopal, S. S. and Murphy, P., Fully automated optimization algorithm for determining the 3-D patient contour from photopeak projection data in SPECT. *IEEE Trans Med Imaging* **14:** 122-131 (1995).
17. Madsen M. T., Kirchner, P. T., Edlin, J. P. *et al.*, An emission-based technique for obtaining attenuation correction data for myocardial SPECT studies. *Nucl Med Commun* **14:** 689-695 (1993).
18. Pan T. S., King, M. A., Luo, D. S. *et al.*, Estimation of attenuation maps from scatter and photopeak window single photon-emission computed tomographic images of technetium 99m-labeled sestamibi. *J Nucl Cardiol* **4:** 42-51 (1997).
19. Fleming J. S., Technique for the use of standard outlines for attenuation correction and quantification in SPECT. *Nucl Med Commun* **11:** 685-696 (1990).
20. Weinzapfel B. T. and Hutchins, G. D., Automated PET attenuation correction model for functional brain imaging. *J Nucl Med* **42:** 483-491 (2001).
21. Watabe H., Sato, N., Deloar, H. M. *et al.*, "Acquisition of attenuation map for brain PET study using optical tracking system" *Proc*. IEEE Nuclear Science Symposium and Medical Imaging Conference., San Diego, CA, Vol. 3; *pp* 1458-1461 (2001).
22. Zaidi H., Montandon, M.-L. and Slosman, D. O., Attenuation compensation in cerebral 3D PET: effect of the attenuation map on absolute and relative quantitation. *Eur J Nucl Med Mol Imaging* **31:** 52-63 (2004).
23. Hooper P. K., Meikle, S. R., Eberl, S. *et al.*, Validation of post-injection transmission measurements for attenuation correction in neurological FDG-PET studies. *J Nucl Med* **37:** 128-136 (1996).
24. Licho R., Glick, S. J., Xia, W. *et al.*, Attenuation compensation in 99mTc SPECT brain imaging: a comparison of the use of attenuation maps derived from transmission versus emission data in normal scans. *J Nucl Med* **40:** 456-463 (1999).
25. Van Laere K., Koole, M., Versijpt, J. *et al.*, Non-uniform versus uniform attenuation correction in brain perfusion SPET of healthy volunteers. *Eur J Nucl Med* **28:** 90-98 (2001).
26. Zaidi H. and Montandon, M.-L., Which attenuation coefficient to use in combined attenuation and scatter corrections for quantitative brain SPECT? *Eur J Nucl Med* **29:** 967-969 (2002).
27. Kemp B. J., Prato, F. S., Dean, G. W. *et al.*, Correction for attenuation in technetium-99m-HMPAO SPECT brain imaging. *J Nucl Med* **33:** 1875-1880 (1992).
28. Stodilka R. Z., Kemp, B. J., Prato, F. S. *et al.*, Importance of bone attenuation in brain SPECT quantification. *J Nucl Med* **39:** 190-197 (1998).
29. Arlig A., Gustafsson, A., Jacobsson, L. *et al.*, Attenuation correction in quantitative SPECT of cerebral blood flow: a Monte Carlo study. *Phys Med Biol* **45:** 3847-3859 (2000).
30. Van Laere K., Koole, M., Kauppinen, T. *et al.*, Nonuniform transmission in brain SPECT using 201Tl, 153Gd, and 99mTc static line sources: anthropomorphic dosimetry studies and influence on brain quantification. *J Nucl Med* **41:** 2051-2062 (2000).
31. Montandon M.-L., Slosman, D. O. and Zaidi, H., Assessment of the impact of model-based scatter correction on 18F-[FDG] 3D brain PET in healthy subjects using statistical parametric mapping. *Neuroimage* **20:** 1848-1856 (2003).

32. Censor Y., Gustafson, D., Lent, A. *et al*., A new approach to the emission computerized tomography problem: simultaneous calculation of attenuation and activity coefficients. *IEEE Trans Nucl Sci* **26:** 2275-2279 (1979).
33. Nuyts J., Dupont, P., Stroobants, S. *et al*., Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. *IEEE Trans Med Imaging* **18:** 393-403 (1999).
34. Krol A., Bowsher, J. E., Manglos, S. H. *et al*., An EM algorithm for estimating SPECT emission and transmission parameters from emissions data only. *IEEE Trans Med Imaging* **20:** 218-232 (2001).
35. Kaplan M. S., Haynor, D. R. and Vija, H., A differential attenuation method for simultaneous estimation of SPECT activity and attenuation distributions. *IEEE Trans Nucl Sci* **46:** 535-541 (1999).
36. Dicken V., A new approach towards simultaneous activity and attenuation reconstruction in emission tomography. *Inverse Problems* **15:** 931-960 (1999).
37. Madsen M. T. and Lee, J. R., "Emission based attenuation correction of PET images of the thorax." *Proc*. 1999 IEEE Nuclear Science Symposium, Seattle, WA, Vol. 2; *pp* 967-971 (1999).
38. Bronnikov A. V., Numerical solution of the identification problem for the attenuated Radon transform. *Inverse Problems* **15:** 1315-1324 (1999).
39. Bailey D. L., Transmission scanning in emission tomography. *Eur J Nucl Med* **25:** 774-787 (1998).
40. Rowell N. P., Glaholm, J., Flower, M. A. *et al*., Anatomically derived attenuation coefficients for use in quantitative single photon emission tomography studies of the thorax. *Eur J Nucl Med* **19:** 36-40 (1992).
41. Zaidi H., Montandon, M.-L. and Slosman, D. O., Magnetic resonance imaging-guided attenuation and scatter corrections in three-dimensional brain positron emission tomography. *Med Phys* **30:** 937-948 (2003).
42. Koral K. F., Zasadny, K. R., Kessler, M. L. *et al*., CT-SPECT fusion plus conjugate views for determining dosimetry in iodine-131-monoclonal antibody therapy of lymphoma patients. *J Nucl Med* **35:** 1714-1720 (1994).
43. Fleming J. S., A technique for using CT images in attenuation correction and quantification in SPECT. *Nucl Med Commun* **10:** 83-97 (1989).
44. Kinahan P. E., Townsend, D. W., Beyer, T. *et al*., Attenuation correction for a combined 3D PET/CT scanner. *Med Phys* **25:** 2046-2053 (1998).
45. Nickoloff E. L., Perman, W. H., Esser, P. D. *et al*., Left ventricular volume: physical basis for attenuation corrections in radionuclide determinations. *Radiology* **152:** 511-515 (1984).
46. Bocher M., Balan, A., Krausz, Y. *et al*., Gamma camera-mounted anatomical X-ray tomography: technology, system characteristics and first images. *Eur J Nucl Med* **27:** 619-627 (2000).
47. Patton J. A., Delbeke, D. and Sandler, M. P., Image fusion using an integrated, dual-head coincidence camera with X-ray tube-based attenuation maps. *J Nucl Med* **41:** 1364-1368 (2000).
48. Kalki K., Blankespoor, S. C., Brown, J. K. *et al*., Myocardial perfusion imaging with a combined x-ray CT and SPECT system. *J Nucl Med* **38:** 1535-1540 (1997).
49. Beyer T., Townsend, D. W., Brun, T. *et al*., A combined PET/CT scanner for clinical oncology. *J Nucl Med* **41:** 1369-1379 (2000).
50. Mayneord W. V., The Radiography of the human body with radioactive isotopes. *Br J Radiol* **25:** 517-525 (1952).

51. Kuhl D. E., Hale, J. and Eaton, W. L., Transmission scanning: a useful adjunct to conventional emission scanning for accurately keying isotope deposition to radiographic anatomy. *Radiology* **87:** 278-284 (1966).
52. Watson C. C., Schaefer, A., Luk, W. K. *et al*., Clinical evaluation of single-photon attenuation correction for 3D whole-body PET. *IEEE Trans Nucl Sci* **46:** 1024-1031 (1999).
53. Zaidi H., Laemmli, C., Allaoua, M. *et al*., Optimizing attenuation correction in clinical cerebral 3D PET: which method to use? [abstract]. *J Nucl Med* **42:** P195-196 (2001).
54. Tung C.-H., Gullberg, G. T., Zeng, G. L. *et al*., Non-uniform attenuation correction using simultaneous transmission and emission converging tomography. *IEEE Trans Nucl Sci* **39:** 1134-1143 (1992).
55. Watson C. C., Eriksson, L., Casey, M. E. *et al*., Design and performance of collimated coincidence point sources for simultaneous transmission measurements in 3-D PET. *IEEE Trans Nucl Sci* **48:** 673-679 (2001).
56. Smith R. J., Karp, J. S., Muehllehner, G. *et al*., Singles transmission scans performed post-injection for quantitative whole body PET imaging. *IEEE Trans Nucl Sci* **44:** 1329-1335 (1997).
57. Ficaro E. P., Fessler, J. A., Rogers, W. L. *et al*., Comparison of americium-241 and technetium-99m as transmission sources for attenuation correction of thallium-201 SPECT imaging of the heart. *J Nucl Med* **35:** 652-663 (1994).
58. Bailey D. L., Hutton, B. F. and Walker, P. J., Improved SPECT using simultaneous emission and transmission tomography. *J Nucl Med* **28:** 844-851 (1987).
59. Gilland D. R., Jaszczak, R. J., Greer, K. L. *et al*., Transmission imaging for non-uniform attenuation correction using a three-headed SPECT camera. *J Nucl Med* **39:** 1105-1110 (1998).
60. King M. A., Tsui, B. M. and Pan, T. S., Attenuation compensation for cardiac single-photon emission computed tomographic imaging: Part 1. Impact of attenuation and methods of estimating attenuation maps. *J Nucl Cardiol* **2:** 513-524 (1995).
61. Turkington T. G., Attenuation correction in hybrid positron emission tomography. *Semin Nucl Med* **30:** 255-267 (2000).
62. Karp J. S., Muehllehner, G., Qu, H. *et al*., Singles transmission in volume-imaging PET with a 137Cs source. *Phys Med Biol* **40:** 929-944 (1995).
63. Zaidi H. and Hasegawa, B. H., Determination of the attenuation map in emission tomography. *J Nucl Med* **44:** 291-315 (2003).
64. Cao Z. and Tsui, B. M., Performance characteristics of transmission imaging using a uniform sheet source with parallel-hole collimation. *Med Phys* **19:** 1205-1212 (1992).
65. Ichihara T., Motomura, N., Ogawa, K. *et al*., Evaluation of SPET quantification of simultaneous emission and transmission imaging of the brain using a multi-detector SPET system with the TEW scatter compensation method and fan-beam collimation. *Eur J Nucl Med* **23:** 1292-1299 (1996).
66. Tan P., Bailey, D. L., Meikle, S. R. *et al*., A scanning line source for simultaneous emission and transmission measurements in SPECT. *J Nucl Med* **34:** 1752-1760 (1993).
67. Celler A., Sitek, A., Stoub, E. *et al*., Multiple line source array for SPECT transmission scans: simulation, phantom and patient studies. *J Nucl Med* **39:** 2183-2189 (1998).

68. Beekman F. J., Kamphuis, C., Hutton, B. F. *et al*., Half-fanbeam collimators combined with scanning point sources for simultaneous emission-transmission imaging. *J Nucl Med* **39:** 1996-2003 (1998).
69. Zeng G. L., Gullberg, G. T., Christian, P. E. *et al*., Asymmetric cone-beam transmission tomography. *IEEE Trans Nucl Sci* **48:** 117-124 (2001).
70. Laymon C. M., Turkington, T. G., Gilland, D. R. *et al*., Transmission scanning system for a gamma camera coincidence scanner. *J Nucl Med* **41:** 692-699 (2000).
71. Bailey D. L., Jones, W. F., Brun, T. *et al*., A spiral CT approach to recording accurate single photon transmission data in PET [abstract]. *J Nucl Med* **38:** 113P (1997).
72. Meikle S. R., Dahlbom, M. and Cherry, S. R., Attenuation correction using count-limited transmission data in positron emission tomography. *J Nucl Med* **34:** 143-150 (1993).
73. Xu M., Cutler, P. and Luk, W., An adaptive local threshold segmented attenuation correction method for whole-body PET imaging. *IEEE Trans Nucl Sci* **43:** 331-336 (1996).
74. Bettinardi V., Pagani, E., Gilardi, M. *et al*., An automatic classification technique for attenuation correction in positron emission tomography. *Eur J Nucl Med* **26:** 447-458 (1999).
75. Zaidi H., Diaz-Gomez, M., Boudraa, A. E. *et al*., Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging. *Phys Med Biol* **47:** 1143-1160 (2002).
76. Pal N. R. and Bezdek, J. C., On cluster validity for the fuzzy c-means model. *IEEE Trans Fuzzy Syst* **3:** 370-379 (1995).
77. Tai Y.-C., Lin, K.-P., Dahlbom, M. *et al*., A hybrid attenuation correction technique to compensate for lung density in 3-D total body PET. *IEEE Trans Nucl Sci* **43:** 323-330 (1996).
78. Yu S. K. and Nahmias, C., Segmented attenuation correction using artificial neural networks in positron tomography. *Phys Med Biol* **41:** 2189-2206 (1996).
79. Riddell C., Brigger, P., Carson, R. E. *et al*., The watershed algorithm: a method to segment noisy PET transmission images. *IEEE Trans Nucl Sci* **46:** 713-719 (1999).
80. Anderson J. M. M., Srinivasan, R., Mair, B. A. *et al*., Hidden Markov model based attenuation correction for positron emission tomography. *IEEE Trans Nucl Sci* **49:** 2103-2111 (2002).
81. Mumcuoglu E. U., Leahy, R., Cherry, S. R. *et al*., Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images. *IEEE Trans Med Imaging* **13:** 687-701 (1994).
82. Fessler J. A., Ficaro, E. P., Clinthorne, N. H. *et al*., Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction. *IEEE Trans Med Imaging* **16:** 166-175 (1997).
83. Alenius S., Ruotsalainen, U. and Astola, J., Attenuation correction for PET using count-limited transmission images reconstructed with median root prior. *IEEE Trans Nucl Sci* **46:** 646-651 (1999).
84. Kitamura K., Iida, H., Shidahara, M. *et al*., Noise reduction in PET attenuation correction using non-linear Gaussian filters. *IEEE Trans Nucl Sci* **47:** 994-999 (2000).
85. Demirkaya O., Anisotropic diffusion filtering of PET attenuation data to improve emission images. *Phys Med Biol* **47:** N271-N278 (2002).

86. Blankespoor S. C., Xu, X., Kaiki, K. *et al*., Attenuation correction of SPECT using X-ray CT on an emission-transmission CT system: myocardial perfusion assessment. *IEEE Trans Nucl Sci* **43:** 2263-2274 (1996).
87. Kinahan P. E., Hasegawa, B. H. and Beyer, T., X-ray-based attenuation correction for positron emission tomography/computed tomography scanners. *Semin Nucl Med* **33:** 166-179 (2003).
88. Damen E. M., Muller, S. H., Boersma, L. J. *et al*., Quantifying local lung perfusion and ventilation using correlated SPECT and CT data. *J Nucl Med* **35:** 784-792 (1994).
89. Liu A., Williams, L. E. and Raubitschek, A. A., A CT assisted method for absolute quantitation of internal radioactivity. *Med Phys* **23:** 1919-1928 (1996).
90. McCullough E. C. and Payne, J. T., X-ray-transmission computed tomography. *Med Phys* **4:** 85-98 (1977).
91. Robinson P. J. and Kreel, L., Pulmonary tissue attenuation with computed tomography: comparison of inspiration and expiration scans. *J Comput Assist Tomogr* **3:** 740-748 (1979).
92. Da Silva A., Tang, H., Wong, K. *et al*., Absolute quantification of regional myocardial uptake of 99mTc-sestamibi with SPECT: experimental validation in a porcine model. *J Nucl Med* **42:** 772-779 (2001).
93. Tang T. R., ''A combined X-ray CT-scintillation camera imaging system for measuring radionuclide uptake in tumors.,'' PhD Thesis, University of California, Berkeley and University of California, San Francisco, 1998.
94. Wong K. H., ''Multi-modality imaging for improved staging of prostate cancer.,'' Ph.D Thesis, University of California, 2002.
95. Burger C., Goerres, G., Schoenes, S. *et al*., PET attenuation coefficients from CT images: experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients. *Eur J Nucl Med Mol Imaging* **29:** 922-927 (2002).
96. Osman M. M., Cohade, C., Nakamoto, Y. *et al*., Respiratory motion artifacts on PET emission images obtained using CT attenuation correction on PET-CT. *Eur J Nucl Med Mol Imaging* **30:** 603-606 (2003).
97. Goerres G. W., Burger, C., Kamel, E. *et al*., Respiration-induced attenuation artifact at PET/CT: Technical considerations. *Radiology* **226:** 906-910 (2003).
98. Cohade C., Osman, M., Nakamoto, Y. *et al*., Initial experience with oral contrast in PET/CT: Phantom and clinical studies. *J Nucl Med* **44:** 412-416 (2003).
99. Goerres G. W., Hany, T. F., Kamel, E. *et al*., Head and neck imaging with PET and PET/CT: artefacts from dental metallic implants. *Eur J Nucl Med Mol Imaging* **29:** 367-370 (2002).
100. Goerres G. W., Ziegler, S. I., Burger, C. *et al*., Artifacts at PET and PET/CT caused by metallic hip prosthetic material. *Radiology* **226:** 577-584 (2003).
101. Nakamoto Y., Chin, B. B., Kraitchman, D. L. *et al*., Effects of nonionic intravenous contrast agents in PET/CT imaging: phantom and canine studies. *Radiology* **227:** 817-824 (2003).
102. Hutton B. F., Braun, M., Thurfjell, L. *et al*., Image registration: an essential tool for nuclear medicine. *Eur J Nucl Med* **29:** 559-577 (2002).
103. Zubal I. G., Harrell, C. R., Smith, E. O. *et al*., Computerized 3-dimensional segmented human anatomy. *Med Phys* **21:** 299-302 (1994).
104. Stodilka R. Z., Kemp, B. J., Prato, F. S. *et al*., Scatter and attenuation correction for brain SPECT using attenuation distributions inferred from a head atlas. *J Nucl Med* **41:** 1569-1578 (2000).

105. Montandon M.-L., and Zaidi, H., Atlas-guided non-uniform attenuation correction in cerebral 3D PET imaging. *Neuroimage* **25:** 278-286 (2005).
106. Shao Y., Cherry, S. R., Farahani, K. *et al*., Simultaneous PET and MR imaging. *Phys Med Biol* **10:** 1965-1970 (1997).
107. Slates R., Cherry, S., Boutefnouchet, A. *et al*., Design of a small animal MR compatible PET scanner. *IEEE Trans Nucl Sci* **46:** 565-570 (1999).
108. Michel C., Sibomana, M., Boi, A. *et al*., ''Preserving Poisson characteristics of PET data with weighted OSEM reconstruction'' *Conf. Rec*. of IEEE Nuclear Science Symposium and Medical Imaging Conference, Totonto, ON, Vol. 2; *pp* 1323-1329 (1998).
109. Rosenthal M. S., Cullom, J., Hawkins, W. *et al*., Quantitative SPECT imaging: a review and recommendations by the Focus Committee of the Society of Nuclear Medicine Computer and Instrumentation Council. *J Nucl Med* **36:** 1489-513 (1995).
110. Zaidi H., Quantitative SPECT: Recent developments in detector response, attenuation and scatter correction techniques. *Phys Medica* **12:** 101-117 (1996).
111. Chang L. T., A method for attenuation correction in radionuclide computed tomography. *IEEE Trans Nucl Sci* **25:** 638-643 (1978).
112. Hoffman E. J., Huang, S. C. and Phelps, M. E., Quantitation in positron emission computed tomography: 1. Effect of object size. *J Comput Assist Tomogr* **3:** 299-308 (1979).
113. Huang S. C., Hoffman, E. J., Phelps, M. E. *et al*., Quantitation in positron emission computed tomography: 2. Effects of inaccurate attenuation correction. *J Comput Assist Tomogr* **3:** 804-814 (1979).
114. Zeng G. L., Gullberg, G. T., Tsui, B. M. W. *et al*., Three-dimensional iterative reconstruction algorithms with attenuation and geometric point response correction. *IEEE Trans Nucl Sci* **38:** 693-701 (1991).
115. Galt J. R., Cullom, S. J. and Garcia, E. V., SPECT quantification: a simplified method of attenuation and scatter correction for cardiac imaging. *J Nucl Med* **33:** 2232-2237 (1992).
116. Tsui B. M. W., Zhao, X., Frey, E. *et al*., Quantitative single-photon emission computed tomography: Basics and clinical considerations. *Semin Nucl Med* **24:** 38-65 (1994).
117. Chinn G. and Huang, S. C., A general class of preconditioners for statistical iterative reconstruction of emission computed tomography. *IEEE Trans Med Imaging* **16:** 1-10 (1997).
118. Shepp L. A. and Vardi, Y., Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging* **1:** 113-122 (1982).
119. Lange K. and Carson, R., EM reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomog* **8:** 306-316 (1984).
120. Gullberg G. T., Huesman, R. H., Malko, J. A. *et al*., An attenuated projector-backprojector for iterative SPECT reconstruction. *Phys Med Biol* **30:** 799-816 (1985).
121. Almeida P., Bendriem, B., de Dreuille, O. *et al*., Dosimetry of transmission measurements in nuclear medicine: a study using anthropomorphic phantoms and thermoluminescent dosimeters. *Eur J Nucl Med* **25:** 1435-1441 (1998).
122. Schaefer A., Seifert, H., Donsch, P. *et al*., Radiation exposure to patients caused by single-photon transmission measurement in PET. [German] *Nuklearmedizin* **39:** 204-208 (2000).

123. Wu T.-H., Huang, Y.-H., Lee, J. S. *et al*., Radiation exposure during transmission measurements: comparison between CT- and germanium-based techniques with a current PET scanner. *Eur J Nucl Med Mol Imaging* **31:** 38-43 (2004).
124. Meikle S. R., Bailey, D. L., Hooper, P. K. *et al*., Simultaneous emission and transmission measurements for attenuation correction in whole-body PET. *J Nucl Med* **36:** 1680-1688 (1995).
125. Bronnikov A. V., Reconstruction of attenuation map using discrete consistency conditions. *IEEE Trans Med Imaging* **19:** 451-462 (2000).
126. Da Silva A. J., Tang, H. R., Wong, K. H. *et al*., Absolute in vivo quantitation of myocardial activity. *IEEE Trans Nucl Sci* **47:** 1093-1098 (2000).
127. Weber W. A., Neverve, J., Sklarek, J. *et al*., Imaging of lung cancer with fluorine-18 fluorodeoxyglucose: comparison of a dual-head gamma camera in coincidence mode with a full-ring positron emission tomography system. *Eur J Nucl Med* **26:** 388-395 (1999).
128. Mesina C. T., Boellaard, R., van den Heuvel, O. A. *et al*., Effects of attenuation correction and reconstruction method on PET activation studies. *Neuroimage* **20:** 898-908 (2003).
129. Farquhar T. H., Llacer, J., Hoh, C. K. *et al*., ROC and localization ROC analyses of lesion detection in whole-body FDG PET: effects of acquisition mode, attenuation correction and reconstruction algorithm. *J Nucl Med* **40:** 2043-2052 (1999).
130. Natterer F., Determination of tissue attenuation in emission tomography of optically dense media. *Inverse Problems* **9:** 731-736 (1993).
131. Welch A., Clack, R., Natterer, F. *et al*., Toward accurate attenuation correction in SPECT without transmission measurements. *IEEE Trans Med Imaging* **16:** 532-541 (1997).
132. Welch A., Campbell, C., Clackdoyle, R. *et al*., Attenuation correction in PET using consistency information. *IEEE Trans Nucl Sci* **45:** 3134-3141 (1998).
133. Bromiley A., Welch, A., Chilcott, F. *et al*., Attenuation correction in PET using consistency conditions and a three-dimensional template. *IEEE Trans Nucl Sci* **48:** 1371-1377 (2001).
134. Welch A., Hallett, W., Marsden, P. *et al*., Accurate attenuation correction in PET using short transmission scans and consistency information. *IEEE Trans Nucl Sci* **50:** 427-432 (2003).
135. Panin V. Y., Kehren, F., Hamill, J. *et al*., Application of discrete data consistency conditions for selecting regularization parameters in PET attenuation map reconstruction. *Phys Med Biol* **49:** 2425-2436 (2004).

7
Scatter Correction Strategies in Emission Tomography

H. Zaidi* and K.F. Koral†

1. The Problem of Scatter Detection in Nuclear Imaging

Scatter in nuclear medicine imaging usual refers to Compton scattering in which a gamma ray undergoes an interaction with matter (e.g. patient) whereby it changes direction and looses energy. As the change in direction approaches zero, the energy loss also approaches zero. Coherent scattering is often neglected because of its small contribution to the total cross section for the energy range of interest in nuclear medicine but is included in some simulation studies.[1,2] The scatter counts for which one wishes compensation are those that fall within the photopeak window. In PET or in SPECT, if there is only one emission, they are the ones that don't loose so much energy as to fall below the lower-energy cut-off of the window. In SPECT, if there are multiple gamma-ray emission energies, a higher energy emission can loose the amount of energy necessary to fall within the photopeak window of a lower-energy emission and be an unwanted scatter event.

One must state, and keep in mind, where the scatter occurs when talking about scatter correction. In general, the gamma ray can scatter in the patient, in material outside the patient such as the table or the collimator, in the detector material, or in the material behind the detector. Complicating matters, one may have multiple scatters possibly in the same object, but potentially in two or more places. In considering a scatter-compensation method, one should be aware of which type(s) of scatter is (are) being corrected.

The purpose of scatter correction is usually either improved accuracy of activity quantification, contrast enhancement of the image, or both. The purpose may affect the choice of technique. While it is well accepted by the

*PD Dr H. Zaidi, Geneva University Hospital, Division of Nuclear Medicine, CH-1211 Geneva, Switzerland
†Dr K.F. Koral, Dept. of Radiology, University of Michigan Medical Center, Ann Arbor, MI 48109-0552, USA

nuclear medicine community that the detection of Compton-scattered events degrades image quality, a common question asked by most nuclear medicine physicians is to what extent does scatter affect image interpretation and clinical decision making? Does it reduce diagnostic accuracy? What is the real added value of scatter correction in clinical nuclear imaging? While not being able to answer the questions in detail, most imaging scientists seem to be convinced that scatter correction is a vital component in producing high-resolution, artefact-free, quantitative images.[3] In addition, some recent studies have quantified the effects of a particular scatter correction in a given clinical situation. As an example, Kim *et al.*[4] have shown that the transmission-dependent convolution subtraction scatter correction technique combined with Ordered-Subsets Expectation-Maximization (ML-EM) SPECT reconstruction results in "... significant changes in kinetic parameters..." and thus a 30%–35% increase in the quantitative estimate of the binding potential for a dopamine D2 receptor ligand.

2. Modelling the Scatter Component

In Compton scattering, the photon interacts with an atomic electron and is scattered through an angle θ relative to its incoming direction. Only a fraction of the initial photon energy is transferred to the electron. After Compton (or incoherent) scattering, the scattered photon energy E_s is given by:

$$E_s = 1 + \frac{E}{m_o c^2}(1 - \cos\theta) \tag{1}$$

where E is the energy of the incident photon, m_0 is the rest mass energy of the electron and c is the speed of light. Eq. (1) is derived on the assumption that an elastic collision occurs between the photon and the electron, and that the electron is initially free and at rest. It is evident from Eq. (1) that the maximum energy transfer takes place when the photon is back-scattered 180° relative to its incoming direction and that the relative energy transfer from the photon to the electron increases for increasing photon energies. The expression (1) is plotted in Figure 1a for 511 keV photons and illustrate that rather large angular deviations occur for a relatively small energy loss.

The transport of annihilation photons is described by the integro-differential Boltzmann equation, which is the governing equation for a given monoenergetic, isotropic source of photons and a scattering medium. The differential part describes the generation, propagation and attenuation of photons. The integral part describes the scattering of photons as an integral over direction and energy. The Boltzmann equation describes the evolution of the spatial and angular distribution of photons with respect to time resulting from the generation, propagation, attenuation and scattering of photons within the medium. This equation is difficult to solve because of the scatter integral and the size of the problem when discretised. Many

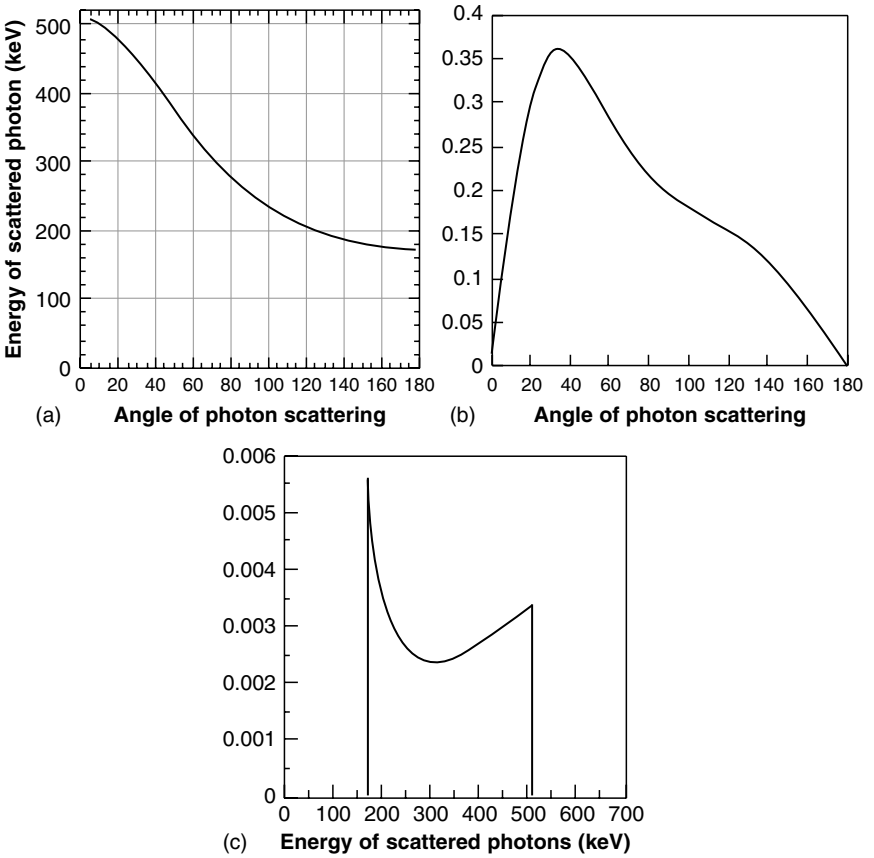


FIGURE 1. (a) The residual energy of a 511 keV photon after Compton scattering through a given angle. The theoretical angular (b) and energy (c) distributions of Compton single-scattered photons having an initial energy of 511 keV according to the Klein-Nishina expression are also shown.

techniques deal with the scatter integral by discretely considering successive orders of scattering.

The probability of Compton scattering is given by the Klein-Nishina equation, which gives the differential scattering cross section $d\sigma/d\Omega$ as a function of scattering angle θ:[5]

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2}(1+\cos^2\theta)\left[\frac{1}{1+\alpha(1-\cos\theta)}\right]^2\left[1+\frac{\alpha^2(1-\cos\theta)^2}{[1+\alpha(1-\cos\theta)](1+\cos^2\theta)}\right] \quad (2)$$

where $\alpha = E/m_o c^2$ and r_e the classical radius of the electron. This equation was extensively used to build appropriate scatter models to correct for this effect in nuclear imaging. Since all unscattered events in PET have 511 keV

before scattering, $\alpha = 1$ for the first scatter. Therefore, the differential Compton scattering cross section relative to that for unscattered annihilation photons is:

$$\frac{d\sigma}{d\Omega} = \frac{r_e^2}{2}\left[\frac{1}{(2-\cos\theta)^2}\right]\left[1+\cos^2\theta+\frac{(1-\cos\theta)^2}{(2-\cos\theta)}\right] \quad (3)$$

The expression (3) gives the differential scattering probability. The integral of this from 0° to any angle gives the integral cross section. The integral gives information about what fraction of scattered photons will be scattered into a cone with a given half-angle. The theoretical angular and energy distributions of Compton-scattered photons are plotted as a function of the scattering angle and energy in Figure 1b and Figure 1c, respectively, for an initial energy of 511 keV. It illustrates relatively that small angle scattering is more likely than large angle scattering and that the most probable scattering angle is located around 35°.

Scattered photons arise from the whole attenuating medium, including the imaging table and the different gamma camera or PET tomograph components. Several studies have reported that the scatter fraction–defined as the ratio between the number of scattered photons and the total number of photons (scattered and unscattered) detected–in PET represents from 35% (brain scanning) to more than 50% (whole-body scanning) of the data acquired in 3D mode, depending on the scanner geometry, the energy window setting, the region to be explored and patient size. In addition to a decrease in the image contrast, events may also appear in regions of the image where there is no activity (e.g. outside the patient). The issue of scatter detection, modelling and correction in emission tomography is addressed in many publications.[6-10] The ideal research tool (gold standard) for scatter modelling and evaluation of scatter correction techniques is the Monte Carlo method.[2] Nevertheless, the complexity and computing requirements of Monte Carlo simulation led to the development of analytic simulation tools based on simplifying approximations to improve speed of operation. For instance, Beekman[11] developed a fast analytic simulator of tomographic projection data taking into account attenuation, distance-dependent detector response, and scatter based on an analytical point spread function (*PSF*) model. It is well accepted that with current computer hardware (unless parallel platforms are used), on-line Monte Carlo calculation is not a viable solution in a clinical environment. However, it is an essential and useful tool for studying the scatter component and developing appropriate scatter models.

The current practice of developing theoretical scatter models involves four different stages: characterisation, development, validation and evaluation.[12]

(i) *Characterisation*. The scatter response function (*srf*) is first studied using a variety of phantom geometries, source locations, scattering media

shape, size and composition as well as imaging system-related parameters (e.g. detector energy resolution) to fully understand and characterize the parameters influencing its behaviour.

(ii) *Development*. From knowledge and insight gained during the characterization step, an appropriate scatter model can be developed. This can be a simple model limited to homogeneous attenuating media or an elaborated one taking into account more complex inhomogeneous media.

(iii) *Validation*. The validation step is the crucial part and involves comparisons between either experimental measurements or Monte Carlo simulation studies and predictions of the developed theoretical model. The second approach is generally given preference for practical reasons owing to ease of modelling and capability of simulations to analyse separately scattered and unscattered components. Again this could be performed using simple phantom geometries (point and line sources in a uniform cylinder) or more complicated anthropomorphic phantoms to mimic clinical situations.

(iv) *Evaluation*. Obviously, evaluation of the theoretical scatter model with respect to the intended use, i.e. scatter correction, constitutes the last step of the whole process. The intrinsic performance of the scatter compensation algorithm based on the developed model as well as its effectiveness in comparison to existing methods is generally recommended.

Accurate simulation of scatter in SPECT/PET projection data is computationally extremely demanding for activity distributions in non-uniform dense media. These methods require information about the attenuation map of the patient. A complicating factor is that the scatter response is different for every point in the object to be imaged. Many investigators used Monte Carlo techniques to study the scatter component or *srf*.[13-18] However, even with the use of variance reduction techniques, these simulations require large amounts of computer time, and the simulation of the *srf* for each patient is impractical. Various methods for tackling this problem have been proposed.

One class of methods uses Monte Carlo simulations[19,20] to compute the transition matrix, which represents the mapping from the activity distribution onto the projections. Monte Carlo simulation can readily handle complex activity distributions and non-uniform media. Unfortunately, hundreds of Gbytes up to several Tbytes of memory are required to store the complete non-sparse transition matrix when the fully 3D Monte Carlo matrix approach is used, and without approximations it can take several weeks to generate the full matrix on a state-of-the-art workstation. In addition, the procedure has to be repeated for each patient. Analytical scatter models, based on integration of the Klein–Nishina (K–N) equation,[21-23] have practical disadvantages, which are similar to those of Monte Carlo-based methods.[24-26]

Another class of methods, which includes anatomy-dependent scatter in the reconstruction, first calculates and stores in tables the scatter responses

of point sources behind slabs for a range of thicknesses, and then tunes these responses to various object shapes with uniform density.[27] This method is referred to as slab-derived scatter estimation (SDSE). A table occupying only a few Mbytes of memory is sufficient to represent this scatter model for fully 3D SPECT reconstruction.[28] A fully 3D reconstruction of a ^{99m}Tc cardiac study based on SDSE can be performed in only a few minutes on a state-of-the-art single processor workstation. A disadvantage of SDSE compared with matrices generated by Monte Carlo simulation or Klein–Nishina integration is that it cannot accurately include the effects of the non-uniform attenuation map of the emitting object. So far, only a few rough adaptations have been proposed to improve the accuracy of this method[29,30] or other similar approaches[31] in non-uniform objects. More recently, Beekman *et al.*[32] reported an accurate method for transforming the response of a distribution in a uniform object into the response of the same distribution in a non-uniform object. However, the time needed to calculate correction maps for transforming a response from uniform to non-uniform objects may be too long for routine clinical implementation in iterative reconstruction-based scatter correction, especially when the correction maps are calculated for all projection angles and each iteration anew. The use of only one order of scatter was sufficient for an accurate calculation of the correction factors needed to transform the scatter response. Since the computation time typically increases linearly with the number of scatter orders, it still remains much shorter than with straightforward Monte Carlo simulation.

Figure 2 shows Monte Carlo simulated scatter distribution functions for a line source located at the centre of the FOV and displaced 5 cm radially

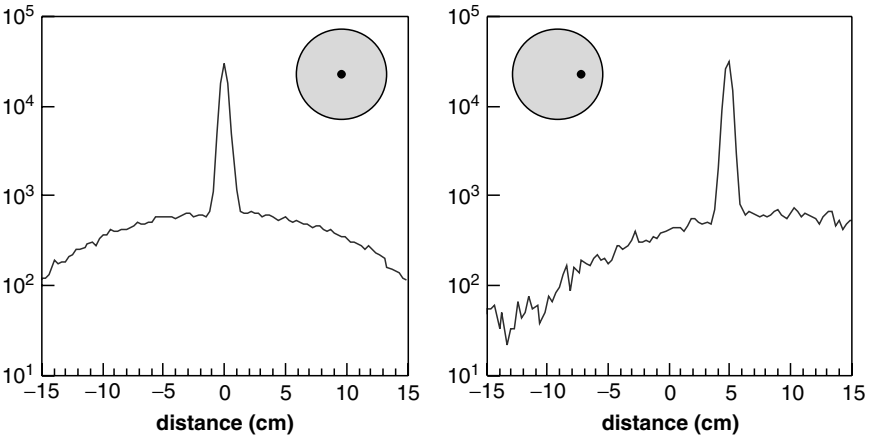


FIGURE 2. Illustration of the scatter response functions simulated using Monte Carlo calculations when the line source is located at the centre of the FOV (left) and displaced 5 cm radially (right). The projections for the line source at the centre are symmetrical and were summed over all projection angles.

off-centre for the ECAT 953B multi-ring PET scanner (CTI PET systems, Knoxville, TN, USA) using the *Eidolon* simulation package.[33] The projections for the line source at the centre are symmetrical and were summed over all projection angles. This is not the case when the source is moved off-centre. For this reason, a profile of a single angular view was used. It can be seen that the projections of a line source placed in a uniform water-filled cylinder are dominated in the wings by the object scatter and in the peak by the unscattered photons. This discussion remains valid when the line source is shifted out of the symmetry centre. The amplitude of the short side of the projection compared to that of the symmetrical case is increased, since the path length of the photons through the phantom becomes shorter, whereas the amplitude of the long side is decreased, due to a longer pathway through the attenuating medium.

3. Scatter Correction Techniques in Planar Imaging

Scatter correction procedures used in planar imaging will be covered within that chapter. However, any scatter correction employed for SPECT that corrects each projection image separately employs a method that can be used in planar imaging, and vice versa. This will generally not be pointed out further in this book. As an example, in 1982 for planar imaging with ^{111}In, Van Reenen *et al.*[34] discussed the idea of employing two energy windows–photopeak and scatter–and a constant, k, to relate the scatter-count component of the total count in the photopeak window to the total count in the scatter window. In 1984, Jaszczak *et al.*[35] popularized this same dual-energy-window approach for SPECT imaging with ^{99m}Tc.

4. Scatter Correction Techniques in SPECT

Koral reviewed SPECT scatter correction methods that employed Monte-Carlo simulation, or used it to carry out testing of the correction method, as of 1998.[9] Some older correction techniques are covered in that review which won't be covered at all here, or will be covered in less detail than there. Scatter correction methods were divided into two broad categories: those without explicit subtraction, and those employing subtraction. That division will not be followed here, but methods will instead be divided into those that employ implicit correction and those that use explicit correction. Implicit correction procedures are those which require nothing beyond satisfying some criterion. Explicit correction methods employ a procedure designed to compensate for scattering. The compensation can be by putting the scatter counts back into the voxel from which the gamma originated, by subtracting the scatter counts and then reconstructing, or by reconstructing unscattered counts in such a way that the existence of scatter counts is accounted for.[3]

4.1 Implicit Correction Methods

One implicit scatter correction involved narrowing the photopeak energy window by moving the lower-energy cut-off energy up so it became less and less likely that scattered gamma rays were accepted. For ^{99m}Tc, this method worked better and better as the window was narrowed, but the number of unscattered counts being accepted was also being reduced.[36] So this approach is usually not employed because one hardly ever is willing to sacrifice the ''good'' counts. Another implicit method simply required system activity calibration to be based on an object with scattering properties that were identical to, or similar to, the case of interest. An example was quantifying the activity of a tumour using a phantom-based camera-collimator calibration that consisted of a known-activity sphere of about the tumour's size placed within a surround that mimicked the patient.[37] The potential advantage was that there wasn't an increase in the noise of the estimate, as there is with explicit scatter correction. The obvious drawback was that it was impractical to simulate each patient exactly. Nevertheless, such an implicit correction based on imaging a known-activity sphere situated within a cylindrical phantom was employed for evaluation of tumour activity in a subset of previously-untreated patients undergoing tositumomab (formerly anti-B1, and also known as Bexxar™) therapy.[38,39] A third implicit correction method involved using lower values for the narrow-beam attenuation coefficient in attenuation correction.[40] This ad hoc method was practical, but seems obviously to have little chance of producing results without bias in particular cases.

4.2 Explicit Correction Methods

Buvat *et al.*[6] published a thorough review of explicit scatter correction methods as of 1994. Narita *et al.*[41] provided a good, brief evaluation of possible methods as of 1996 in their introduction to a paper comparing two methods of SPECT scatter correction. They wrote:

''The most commonly employed scatter correction techniques, due to their ease of implementation, are the convolution subtraction techniques[42] and dual- or triple-energy window (TEW) techniques.[35,43,44] These techniques have recently been evaluated and compared using Monte Carlo simulations. Of the techniques investigated, the TEW technique (with some modification) achieved reasonable quantitative accuracy, does not require careful object and instrument-dependent calibration and does not assume a fixed scatter fraction (k).[45,46] However, due to the small windows used in this technique, noise can be substantially increased in the scatter corrected data. Improvements in accuracy can be achieved by regional spectral analysis,[46] but at the expense of substantially increased complexity. Also there may be difficulty in collecting pixel by pixel spectra on existing gamma cameras.''

The last-mentioned difficulty with multi-window (spectral-analytic) methods restricts some of these methods to certain cameras, or cameras

with experimental acquisition software, or to list-mode acquisition. Below, we comment on all SPECT methods by type, but spend more time on newer techniques.

4.2.1 Approaches based on Convolution or Deconvolution

The convolution or deconvolution approaches used in SPECT are older, and, especially in their simplest forms, have fallen into disfavour compared to other methods, and so a minimum amount of time will be spent discussing them here. They were first introduced in 1984, but in a one-dimensional form that required that the object did not vary in the axial direction.[42] Later, two-dimensional convolution methods took several forms but all used a convolution and subtraction.[14,47,48] They often suffered from the fact that the distribution of all counts was used to approximate the distribution of unscattered counts, although, as described under PET techniques, an iterative approach might help. To correct for the deficiency without using iteration, Floyd *et al*.[49] presented an approach based on true deconvolution in frequency space. In their formalism, the scatter was estimated by a convolution of the non-scatter projection with an exponential scatter function. Solution required inverse Fourier Transformation. This method assumed that the scatter distribution was independent of source location, which is not the case as the source nears the edge of the object, as has been shown in Figure 2. The entire category of scatter correction by convolution or deconvolution will be covered again for PET (see section 5.2).

4.2.2 Dual-Energy-Window (DEW) Approach

One of the early SPECT scatter subtraction techniques, the dual-energy-window approach, assumed that a window immediately below the photopeak window could monitor the number of scatter counts corrupting the total count within the photopeak window.[35] In this study, the k value obtained for ^{99m}Tc from experimental measurement was 0.5 and from simulation 0.57. The lower-window count could either be multiplied by k and then subtracted from that within the photopeak window, pixel by pixel, or the lower-window counts projections could be reconstructed into a scatter image, which would be multiplied by k and then subtracted from the images reconstructed from the total count in the photopeak window. Unfortunately for the method, it was shown by Monte Carlo simulation that the monitoring is deficient in at least some cases.[48,50] As an example of the deficiency in a projection image, if one looks at ^{99m}Tc counts within a circular region of interest corresponding to an on-axis sphere located within a circular cylinder, the number in the lower-energy window when the cylindrical background activity is zero and the sphere activity non-zero is only about 2/3 the number when the cylindrical background activity is non-zero and the sphere activity zero. This is the case when the scatter count total within the 20% photopeak window is the same.[50] One possible solution to the

deficiency of the method involved allowing the scatter multiplier (alias the "*k*" factor) to vary with the assessed non-target background. This was successful for ^{99m}Tc and ^{123}I when the iterative method was employed for the restricted goal of quantification of focal activity rather than the activity distribution throughout the image.[51]

The method, basically as originally introduced, was later applied to ^{131}I SPECT by other groups. Green *et al.*[52] increased the size of the scatter-monitoring window to $\pm$ 18% and then used a *k* value of 0.54. Using the resultant scatter correction, they obtained good lesion-activity estimates in a phantom over a range of background activities. They then compared scatter-corrected SPECT estimated activities for the liver and spleen of twelve patients receiving 2.78 GBq (75 mCi) of an antibody to CEA to those from planar imaging without scatter correction. Koral *et al.* determined that under otherwise identical conditions, the *k* value decreased with gamma energy. They proceeded to apply the *k* value for ^{131}I in their quantification technique (0.75) to accomplish scatter correction for tumour activity quantification for three relapsed lymphoma patients undergoing ^{131}I-labelled MB1 monoclonal antibody therapy and three similar patients undergoing tositumomab therapy.[53]

A variation on the dual-energy-window method was a method called the dual-photopeak-window method in which the normal photopeak window was split into two halves.[44] The similarity to the first method lay in trying to estimate scatter counts from looking at only two count values (in the new method, one for each half of the window). The fall of this method in its most basic form was its need for the location of the photopeak window to be invariant with respect to spatial location in the projection image, and with respect to tomographic angle. This was demonstrated to not be the case for at least some cameras.[54,55]

4.2.3 Multiple-Energy-Window (Spectral-Analytic) Approaches

The idea of using multiple energy windows was introduced in 1988.[56] The multiple-energy-window methods assume that given enough windows above and below the photopeak window, one can estimate either the complete energy spectrum of the scattered counts, or at least the integral of that spectrum from the lower-energy cut-off of the photopeak window to the upper-energy cut-off of that window. These methods input more information to potentially get more accurate answers. Subtraction of the scattered counts, pixel by pixel, is still the goal.

The triple-energy-window method is the simplest approach.[43] It uses a window immediately above the photopeak and one immediately below the photopeak. It further assumes that the total scatter counts is accurately approximated when a straight line is used for the shape of the scatter spectrum (Note that the true shape may be different without loss of accuracy as long as the integral from lower-energy cut-off to upper-energy cut-off is

the same for both functions.) This method has the advantage of simplicity. It can suffer from excessive noise if the two monitoring windows are too narrow. In the case of ^{99m}Tc, the method degenerates into a version of the dual energy window approach, because the upper monitor window is usually assumed to have zero counts.[45]

In most of the other multiple-energy-window approaches, a given range of the energy spectrum for the detected gamma rays is divided into equal width sub-windows. In one of these methods, a scatter-free spectrum is assumed to be known from measurements. In this method, the spectrum of all counts obtained for a given pixel is assumed to be comprised of a sum of two terms. The first term is the scatter-free spectrum times a constant. The second term is a scatter spectrum which is assumed to have a particular functional form, with parameters to be determined. The scatter estimation problem then becomes fitting the measured data with the two component spectrum.[56,57] As with the dual-photopeak-window method, it would be handy in this method for the scatter-free spectrum to be invariant with respect to spatial location in the projection image and with respect to tomographic angle. However, with this method it is obvious that, in principle, the scatter-free spectrum can equally well be measured as a function of those parameters.

In the later, more-refined version of the spectral-fitting method,[57] a spatial regularization allows for each pixel in a projection image to have its scatter component estimated while taking into account the estimate for adjacent pixels. A characteristic of the method is that, due to the need for low-noise spectra, pixels may need to be grouped before analysis.

In another one of the multiple-window methods, weights for each energy window are determined from an optimisation procedure based on a specific task.[58,59] In the case of brain imaging with ^{99m}Tc, the task may be both accurate lesion and non-lesion activity concentration. The weights, which have positive and negative values and which tend to vary pseudo-sinusoidaly over the energy range, are then combined with the measured spectrum to produce the estimate of primary-count total.[58] It seems clear that the main assumption here is that the same set of weights are optimum, or near optimum, for the case of interest as they were for the "training" cases.

Another multi-window approach that uses training in an even more definitive way is scatter estimation using artificial neural networks. These were introduced for nuclear medicine scatter correction by means of a Monte Carlo simulation study in 1993.[60] Quantification accuracy for ^{99m}Tc activity of 5% was achieved. Later, the neural networks were employed for scatter correction or for scatter and attenuation correction.[61] The radionuclide was again ^{99m}Tc but both Monte Carlo simulation and experimental data were employed. Recently, a neural network approach was employed to estimate the photopeak count total for each of the three major emissions of ^{67}Ga, using Monte-Carlo simulation.[62] In this study, the network was composed of a 37-channel input layer, an 18-node hidden layer, and a three-channel output layer. Each node was connected to every node of

the succeeding layer. The effect of each input was regulated through a specific weight. The values for the weights were established during a "learning" phase. The 37 input channels accepted count totals from sub-windows of the energy spectrum that were of unequal width and covered the energy range from 60 to 370 keV. The output channels yielded the ratio of primary to total counts for each of the three major emissions of ^{67}Ga. The analysis was carried out for a given pixel in a projection independently from the result for other pixels.

The entire learning set was divided into two components. One component was used to set the weights through an error-back-projection algorithm. The other component was used "... for testing global convergence." The exact meaning of global convergence and the way that testing was done are unclear. The method yielded a bias for all structures that was much improved compared to that with standard windows (19% compared to 85%). The question remaining is again how well a network trained with a given set of spectra can handle spectra from a wide variety of cases. A recent study has compared the neural network method, using separate networks for each of the three gamma emissions of ^{67}Ga, to the window optimisation method for the task of tumour detection and activity estimation.[59] A very brief summary is that they both performed well.

4.2.4 Method Requiring a Transmission Measurement

Recently, a method using a transmission measurement to provide additional, hopefully-pertinent information about the scatter in a particular imaging situation was proposed. It was introduced by Meikle *et al.*[63] in 1994 and is called the transmission-dependent convolution subtraction (TDCS) method. It uses a bigger scatter correction in regions of the projection image where the transmission through the patient is small. It was developed for ^{99m}Tc and ^{201}Tl and tested for heart imaging.[41,64] It draws upon earlier approaches, and is basically an iterative approach although sometimes only one iteration is used. It also takes the geometric mean of conjugate views, relies on a convolution, uses a ratio of scattered events divided by total events, k_{ST}, and employs a depth-dependent build-up factor, $B(d)$. The k_{ST} and the $B(d)$ are both variable across the two-dimensional projection image. In a version that employs more than one iteration, k_{ST} seems to be used in place of the ratio of scatter to primary events, k_{SP}. The transmission is defined for a narrow-beam attenuation coefficient, although its broad-beam experimental measurement may only approximate this condition. The iteration is designed to correct for the fact that the observed projection image is used to approximate the true scatter-free image. (It is the scatter-free image that is wanted to carry out the scatter estimation, but it is initially not available).

Narita *et al.*[41] made several small changes to the original method. First of all, they did use 10 iterations whereas only 1 was originally employed. Secondly, they used a scatter function that was the sum of an exponential

plus a Gaussian, whereas only an exponential was originally employed. Lastly, they averaged the dependence of k_{ST} on transmission factor from two quite different activity distributions, one of which is similar to the single dependence originally employed. They compared results from their version of TDCS to the results from TEW scatter correction and concluded that their method produced a much smoother scatter estimate, and that the resulting signal-to-noise was better than with TEW correction.[41] The smoothness of the estimate is shown in Figure 3.

The original method was also recently applied to ^{123}I brain imaging for a phantom and for six patients.[65] In the research, the original equation for k_{ST} as a function of transmission was modified to include a constant additive term to account for septal penetration due to gamma rays with energies greater than 500 keV. The authors found that TDCS provided "... an acceptable accuracy..." in the estimation of the activity of the striatum and of the occipital-lobe background. Moreover, parameter values averaged over six collimators from three different SPECT cameras yielded "... minimal differences..." among the collimators so new users might not have to calibrate their collimator-camera system. Simple Gaussian modelling of the probability of photon scattering through a given angle incorporated within the framework of a 2D projector/backprojector pair of a ML-EM reconstruction algorithm[66] was also reported.

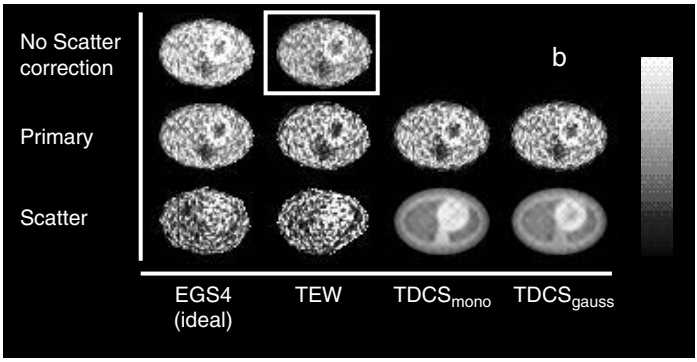


FIGURE 3. Images reconstructed under various conditions for a slice through a chest phantom. The top row shows reconstructions without scatter correction, but with attenuation correction. The left image of the pair used narrow beam μ values and the right image broad-beam μ values. The middle row compares reconstructions of the primary counts at the left with reconstructions of the complete projections employing one of three scatter-compensation methods. Either the triple-energy window method (TEW) or the transmission-dependent convolution subtraction method (TDCS) is employed. Subscript "mono" refers to the use of a mono-exponential in the model and "gauss" to a Gaussian function. In the third row, reconstruction of only scatter counts is shown at the left, while the last three images shows the estimated scatter image from the three methods. Most striking is the smoothness of the scatter estimate with TDCS. (Reprinted from ref.[41] with permission).

4.2.5 Approaches based on Iterative Reconstruction

The original iterative method for SPECT scatter correction was called inverse Monte Carlo when first introduced in 1986.[19] As indicated in the comments in section 1, this method requires large computer storage capacity. So much so, that when the first testing of it was carried out with Maximum-Likelihod Expectation-Maximization (ML-EM) reconstruction in 1991,[67] the authors employed phantoms in which the activity did not vary with distance along the axis of rotation.[68] Improvements in computer speed and recent advances to accelerate reconstruction[69] have led to renewed interest in algorithms which incorporate scatter as part of the emission model.[20] Iterative reconstruction-based scatter compensation has received considerable attention during the last decade. Some of these studies[26-28,70] indicate that when scatter is modelled in iterative reconstruction, higher contrast and lower noise can be obtained than when scatter subtraction is used.

As discussed in chapter 4 of this book, the normal approach for implementation of a scatter model is to incorporate the scatter estimation directly in the transition matrix (a_{ij}) of the maximum-likelihood expectation-maximization (ML-EM) or its accelerated version, the ordered-subsets expectation-maximization (OS-EM) equation, in which case it becomes considerably larger than for the case where only attenuation and geometric factors are included. Because of the larger transition matrix, computation is slow (scatter is essentially recalculated for each iteration). Efficiency has been improved by utilising a dual matrix approach[71] in which scatter is incorporated in the forward projection step only. In this case, the projector includes attenuation, detector blurring and scatter, while the backprojector only includes attenuation and detector blurring. One of the requirements of this method is computing the scatter response function at each point in the attenuator for all projection views and each iteration anew. To avoid slow computation, the correction factors could be calculated only once or alternatively a few times only, given that the calculated scatter component does not change much after having carried out the first few iterations of accelerated statistical reconstruction.[32] Thus, the scatter estimate can be kept either as a constant term in all or only in later iterations instead of modifying the scatter estimate in each iteration.[72,73] In this way, a constant pre-calculated (using one of the methods described above) scatter component can be introduced in the denominator, i.e. the forward projection step of the ML-EM equation (see Chapter 4).

$$f_j^{new} = \frac{f_j^{old}}{\sum_l a_{lj}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old} + \hat{s}} \tag{4}$$

where ($\hat{s}$) is the scatter estimated on all projections.

An efficient algorithm for scatter estimation was recently described in which the spatial scatter distribution is implemented as a spatially invariant convolution for points of constant depth in tissue.[72] The scatter estimate is weighted by a space-dependent build-up factor based on the measured

attenuation in the body. Consequently, $(\hat{s})$ can be estimated once, based on an estimate of the activity distribution obtained from the conventional ML-EM equation with the final scatter-corrected result obtained by proceeding with Eq. (4). This approach was validated by Monte Carlo simulation studies of a realistic thorax phantom. This approach has been recently extended by Beekman *et al*.[26] by implementing a computationally efficient approach to preserve the main advantages of statistical reconstruction while achieving a high accuracy through modelling the scatter component in the projector using Monte Carlo-based calculation of low-noise scatter projections of extended distributions, thus, completely avoiding the need for massive transition matrix storage. Figure 4 illustrates the advantages of

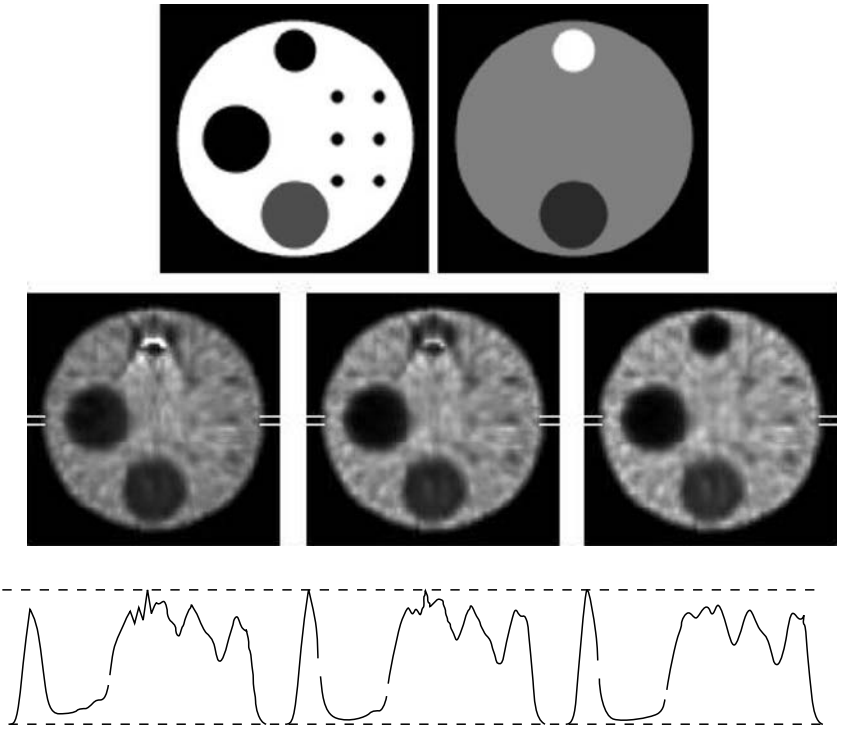


FIGURE 4. From top to bottom: representative slice of cylindrical phantom used in simulation studies showing activity distribution with six small lesions (left), and nonuniform attenuation distribution (right). The lesions provide the opportunity to assess contrast with different approaches. Reconstructed images and horizontal profiles using ten iterations of dual matrix ordered subset without scatter correction (left), and with scatter correction integrated in the projector using an approximate scatter *PSF* model (centre), and using Monte Carlo simulation combined with convolution-based forced detection (MCS-CFD) as a reprojector. The scatter-induced artefacts in and around the high-density cylinder are removed with the latter. (Reprinted from ref.[26] © 2002 IEEE, with permission).

this approach compared to simple *PSF*-based scatter modelling in the projector. Whether the accuracy of modelling achieved in simulated phantoms could be attained in clinical studies still needs to be shown to guarantee applicability of the approach in a clinical setting.

5. Scatter Correction Techniques in PET

Over the last two decades, many methods have been developed for the purpose of reducing the resultant degradation of image contrast and loss of quantitative accuracy in PET due to contribution from scattered events. The main difference among the correction methods is the way in which the scatter component in the selected energy window is estimated. The most reliable method to determine the actual amount of scatter in the image is physical modelling of the scatter process to resolve the observed energy spectrum into its unscattered and scattered components. By observing how accurately a scatter correction algorithm estimates the amount and distribution of scatter under conditions where it can be accurately measured or otherwise independently determined, it is possible to optimise scatter correction techniques. A number of scatter correction algorithms for PET have been proposed in the literature. They fall into four broad categories[3,10,74]:

- Energy window-based approaches,
- Convolution/deconvolution-based approaches,
- Approaches based on direct estimation of scatter distribution.
- Statistical reconstruction-based scatter compensation approaches.

Different versions of the above methods have been successfully implemented for 3D PET and are briefly discussed below.

5.1 Energy Window-based Approaches

A DEW method was originally developed for SPECT[35] as mentioned above. The developments of 3D acquisition mode and improvements in the detector energy resolution in PET have allowed the implementation of scatter correction based on the analysis of energy spectra. Several groups investigated the potential of acquiring data in two,[75,76] three[77] and multiple[78] energy windows to develop corrections for scattering in 3D PET.

Two variants of dual-energy window techniques have been proposed: methods estimating the scatter component in the photopeak window from the events recorded in a lower energy window placed just below the photopeak and methods estimating the unscattered component in the photopeak window from the unscattered counts recorded in a high energy window in the upper portion of the photopeak. The dual energy window (DEW) technique[75] belongs to the former while the estimation of trues method (ETM)[76] belongs to the latter.

In the practical implementation of the DEW method on the ECAT 953B scanner, detected coincidence events are assigned to the upper energy window when both photons deposit energy between 380 keV and 850 keV, or to the lower energy window when one or both photons deposit within 200 keV and 380 keV.[75] Both energy windows are assumed to contain object scattered and unscattered events. Based on data collected in the two energy windows and scaling parameters derived from measurements of the ratios of counts from line sources due to unscattered (measurements in air) and scattered events (measurements in a head-sized phantom), two equations containing four unknown parameters are solved to estimate the unscattered component in the acquisition energy window.

The estimation of trues method[76] consists in acquiring data simultaneously in two energy windows: a high window with a lower energy threshold higher than 511 keV and a regular acquisition window including the higher window. Therefore, both windows have the same upper energy threshold (UET) value. The hypothesis of this method is that the number of unscattered coincidences recorded in a given energy range depends on the energy settings of the window and the angle of incidence of the annihilation photons on the detector face. Hence, the unscattered component in the high-energy window can be related to the unscattered coincidences in the standard wider window through a function of the energy settings, the radial position in the sinogram for a given line of response (LOR) and the axial opening for a given radial position. This calibrating function is in principle independent of the source distribution. The unscattered component in the wide energy window can thus be calculated and subsequently subtracted from the data recorded in the regular window to produce a scattered sinogram. The unscattered component in the regular window is then obtained by smoothing that sinogram and subtracting it from the data recorded in the standard window.

The triple energy window method[77] was suggested as an extension of the DEW technique. Coincidence events are recorded in three windows: two overlapping windows having the same UET settings (450 keV) located below the photopeak window and a regular window centred on the photopeak and adjacent to the low windows. A calibrating function that accounts for the distribution of scattered coincidences at low energies is obtained by calculating the ratio of the coincidence events recorded in both low energy windows for the scanned object and for a homogeneous uniform cylinder. The scatter component in the standard acquisition window is then estimated from the calibrating function and the narrower low energy window.

The multispectral method is based on the acquisition of data in a very large number (typically 256) of windows of the same energy width (16×16 energy values for the two coincident photons). The spatial distribution of scattered and unscattered components in each window can be adjusted using simple mono-exponential functions.[78] The method presents the advantage of increased sensitivity due to smaller rejection of detected coincidence events

(~10%) since the position of the scattered photons in the detectors can be restored. The statistical noise in each window and the necessity to accommodate the acquisition hardware with the required electronic boards to allow acquisition in multiple windows remain the major obstacles for implementation of the method on commercial PET scanners.

5.2 Convolution-Deconvolution based Approaches

Techniques based on convolution or deconvolution estimate the distribution of scatter from the standard photopeak data. The scatter fraction (*SF*) which gives an indication about the expected amount of scatter and the scatter response function (*srf*) which defines the spatial distribution of scatter in the photopeak data are usually the two parameters required for estimation of scatter component and need to be determined *a priori*. A pure additive model of the imaging system is generally assumed where the recorded data are composed of an unscattered and a scattered component plus a noise term due to statistical fluctuations, and can be written in the following form:

$$p_o = p_u + p_s + \eta \tag{5}$$

where p_o are the observed data, p_u and p_s are the unscattered and scattered components respectively, and η is the noise term. The problem to be addressed consists in estimating the unscattered distribution (p_u) from the measured data (p_o) contaminated by scatter, or alternatively estimate the scattered component (p_s) and then derive p_u. The proposed methods differ in the way the scatter function is defined.

The convolution-subtraction (CVS) technique developed for 3D PET[79] operates directly on projection data (pre-reconstruction correction) and is based on the assumption that the convolution of the source distribution with the *srf* gives an estimate of the scatter component. Two assumptions have been made: the stationary and nonstationary assumptions. In the stationary assumption, the scatter is assumed to be analytically defined and not dependent on the object, activity distribution, etc. The nonstationary assumption overcomes this problem by taking into consideration the dependence of scatter upon source locations, object size, detector angle, etc. Using the stationary assumption, the scatter component can be related to the unscattered component by the convolution relation:

$$p_s = p_u \otimes srf \tag{6}$$

where $\otimes$ denotes the convolution operator. The scatter estimate (p_s) is then subtracted from the measured data (p_o) after scaling by a appropriate scatter fraction (*SF*). The process can be repeated iteratively with each step using the previous estimate of the scatter-free distribution as input to the scatter estimation.[79] The observed data are used as a first approximation to the unscattered distribution (p_u) and the process is repeated iteratively with each

step using the previous estimate of the scatter-free distribution as input to the scatter estimation:

$$\hat{p}_u^{(n)} = p_o - SF\left(\hat{p}_u^{(n-1)} \otimes srf\right) \tag{7}$$

where the '^' indicates that the parameter is a rough estimate of the scatter and n is the iteration number. The rationale is that with each iteration, the input to the scatter estimation step more closely approximates p_u. It was also suggested to use a damping factor to prevent oscillation in the result.

The convolution-subtraction approach can also be applied to reconstructed images (post-reconstruction). In this case, the scatter estimates are reconstructed and then subtracted from the non-corrected reconstructed images of the acquired data.[80]

There is a continuing interest in developing non-stationary convolution-subtraction scatter correction techniques, which overcome the inherent limitations of the stationary approach by taking into consideration the dependence of scatter upon source locations, object size, detector angle, etc. Different methods of non-stationary deconvolution have been proposed in the literature for SPECT[14] and 2D PET imaging;[81] the extension of such models for 3D PET should in principle be straightforward.

Scattered photons degrade the point-spread function (*PSF*) of the PET camera; the long tails of the *PSF* are mainly due to scatter. Thus, deconvolution methods, which correct the images for the *PSF* will also, implicitly correct for scatter. In general, the *PSF* will act as a low-pass filter. Deconvolution will restore the high-frequency contents of the signal emitted by the object, at least as long as they have not yet been completely removed by *PSF*. Not only the high frequencies in the signal are restored, but also the high-frequency noise is amplified, which in turn can degrade the image again. Therefore, the restoration filter is often combined with a low-pass filter that balances that image improvement by deconvolution and its degradation due to amplification of noise. Well-known examples of these filters are the Wiener and Metz filters. Links *et al.*[82] studied the use of two-dimensional Fourier filtering to simultaneously increase quantitative recovery and reduce noise. The filter is based on the inversion of the scanner's measured transfer function, coupled with high-frequency roll-off. In phantom studies, they found improvements in both "hot" and "cold" sphere quantification. Fourier-based image restoration filtering is thus capable of improving both accuracy and precision in PET.

The curve fitting approach is based on the hypothesis that detected events assigned to LORs outside of the source object must have scattered and that the scatter distribution corresponds to a low-frequency component that is relatively insensitive to the source distribution. Estimation of the unscattered component can thus be performed in three successive steps: (i) fitting the activity outside the source object with an analytical function (e.g. Gaussian), (ii) interpolating the fit inside the object, and (iii) subtracting the scatter

component from the observed data.[83] The accuracy of this class of scatter correction methods depends on how accurately the scatter component can be estimated. The appropriate choice of a set of fitting parameters, which should be optimised for each PET scanner and for different radioactivity and attenuation coefficients distributions, is the dominant factor.

5.3 Approaches based on Direct Calculation of Scatter Distribution

This class of methods assume that the distribution of scattered events can be estimated accurately from the information contained in the emission and transmission data. For the majority of detected scattered events, only one of the two annihilation photons undergoes a single Compton interaction. The rationale of most of these methods is that the overall scatter distribution can be computed from the single scatter distribution ($\sim$75% of detected scattered events) and that this latter can be scaled to model the distribution of multiple-scattered events.[84] The multiple scatter distribution is generally modelled as an integral transformation of the single scatter distribution. The same approach can also be applied to scatter estimation in transmission imaging.[85,86] Monte Carlo simulation studies of various phantom geometries demonstrated the accuracy of this method for fully 3D PET imaging by direct comparison of analytic calculations with Monte Carlo estimates.[15]

The model-based scatter correction method developed by Ollinger[84] uses a transmission scan, an emission scan, the physics of Compton scatter, and a mathematical model of the scanner in a forward calculation of the number of events for which one photon has undergone a single Compton interaction. Parameterization of a fast implementation of this algorithm has been recently reported.[87] A single-scatter simulation technique for scatter correction where the mean scatter contribution to the net true coincidence data is estimated by simulating radiation transport through the object was also suggested and validated using human and chest phantom studies.[23] Watson[88] reported on a new numerical implementation of the single-scatter simulation algorithm, which is faster than the previous implementation, currently requiring less than 30 sec execution time per bed position for an adult thorax. The normalization problem was solved and multiple scatter partially taken into account. However, the above methods do not correct for scatter from outside the field-of-view. This effect can be directly taken into account by acquiring short, auxiliary scans adjacent to the axial volume being investigated. This implicitly assumes that the distribution of scatter from outside the FOV has the same shape as that of scatter from inside the FOV. These extra data are naturally available in whole-body imaging. However, this method is impractical for isotopes with a short half-life or rapid uptake relative to the scanning interval. It has also been shown that the attenuation map to be used as input for estimation of the scatter

distributions can be derived from magnetic resonance images for brain PET scanning.[89] Contribution of scatter from outside the FOV remains a challenging issue that needs to be addressed carefully in whole-body imaging especially with large axial FOV 3D PET scanners.

The true scatter component (being an experimentally impossible measurement) can be accurately estimated using rigorous Monte Carlo simulations. Given a known radioactive source distribution and density of the object, Monte Carlo techniques allow detected events to be classified into unscattered and scattered events and the scatter component to be determined in an independent way. However, the source and scattering geometry is generally not known in clinical studies. In their Monte Carlo-based scatter correction (MCBSC) method, Levin *et al.*[24] used filtered-backprojection reconstructions to estimate the true source distribution. This input image is then treated as a 3D source intensity distribution for a photon-tracking simulation. The number of counts in each pixel of the image is assumed to represent the isotope concentration at that location. The image volume planes are then stacked and placed at the desired position in the simulated scanner geometry assuming a common axis. The program then follows the history of each photon and its interactions in the scattering medium and traces escaping photons in the block detectors in a simulated 3D PET acquisition. The distributions of scattered and total events are calculated and sorted into their respective sinograms. The scatter component is equal to the difference between measured data and the scaled and smoothed scatter component. To reduce the calculation time, coarser sampling of the image volume over regions equivalent to 4×4 pixels was adopted assuming that the Compton scatter distribution varies slowly over the object. For obvious reasons, the implemented method does not correct for scatter from outside the field-of-view and further refinements of the technique were required to take this effect into account. A modified version of this approach was therefore suggested.[90] The data sets were pre-corrected for scatter and the reconstructed images are then used as input to the Monte Carlo simulator.[33] This approach seems reasonable for a more accurate estimation of the true source distribution. Faster implementations of similar approaches have also been described elsewhere.[25]

5.4 Iterative Reconstruction-based Scatter Correction Approaches

Some scatter compensation methods incorporate scatter in the transition matrix or point-spread function during iterative reconstruction. It has been shown that this can lead to high quantitative accuracy[19,27] and improved signal-to-noise ratio in the reconstructed images.[26,28] Development of scatter models that can be incorporated in statistical reconstruction such as OSEM for PET continues to be appealing; however, implementation must be

efficient to be clinically applicable. It is worthwhile to point out that, with one exception[70] most of the research performed in this field is related to SPECT imaging as described in section 4.2.5. The preliminary results obtained using a fast implementation of the single-scatter simulation algorithm[88] were not satisfactory spurring further research to incorporate a more accurate model taking into account multiple scatter to yield significant improvements compared to conventional subtraction-based approaches. Further development and validation of this class of algorithms in whole-body 3D PET are still required.

More recently, a new technique for scatter correction in 3D PET called statistical reconstruction-based scatter correction was proposed.[18] The method is based on two hypotheses: (i) the scatter distribution consists mainly of a low-frequency component in the image, (ii) the low-frequency components will converge faster than the high-frequency ones in successive iterations of statistical reconstruction methods. This non-uniform convergence property is further emphasized and demonstrated by Fourier analysis of the ML-EM algorithm[91] and successive iterations of inverse Monte Carlo-based reconstructions.[19] The low-frequency image is estimated using one iteration of the OSEM algorithm. A single iteration of this algorithm resulted in similar or better performance results as compared to 4 iterations of the CVS method.[90]

6. Strategies for Evaluation of Scatter Correction

Similar to what has been said in Chapter 4 about validation of image reconstruction procedures, evaluation of scatter correction algorithms is inherently difficult and sometimes unconvincing. There is a clear need for guidelines to evaluate correction techniques and other image processing issues in PET. Most of the algorithms developed so far have been evaluated using either simulated or experimentally measured phantom studies, in addition to qualitative evaluation of clinical data.[90] Modelling and simulation of nuclear imaging is best done with phantom models that match the gross parameters of an individual patient. Recent three- and four-dimensional computer phantoms seek a compromise between ease of use, flexibility and accurate modelling of populations of patient anatomies, and attenuation and scatter properties and biodistributions of radiopharmaceuticals in the patients. Modelling of the imaging process has been improved by more accurate simulation of the physics and instrumentation involved in the process. Monte Carlo software packages, especially those developed specifically for nuclear medicine and with different performance characteristics, have been found useful in the modelling work. The combination of realistic computer phantoms and accurate models of the imaging process allows simulation of nuclear imaging data that are ever closer to actual patient data. Simulation techniques will find an increasingly important role in

the future of nuclear medicine research, especially scatter modelling and correction, in light of further development of realistic computer phantoms, accurate modelling of projection data and computer hardware. However, cautions must be taken to avoid errors in the simulation process and verification via comparison with experimental and patient data is crucial.[2]

Evaluation of scatter correction has been extended more recently to objective assessment of image quality using Receiver Operating Characteristics (ROC) analysis based on human or computer observers.[92] An interesting approach in comparative evaluation studies for functional brain imaging would be to carry out voxel-based statistical analysis using Statistical Parametric Mapping (SPM).[93] A recent investigation of the impact of model-based scatter correction on spatial distribution of ^{18}F-[FDG] in reconstructed brain PET images of healthy subjects using this kind of analysis demonstrated that significant differences in ^{18}F-[FDG] distribution arise when images are reconstructed with and without explicit scatter correction for some cerebral areas.[94] Figure 5 illustrates the areas with significant changes in brain metabolism obtained by comparing distributions with and without explicit scatter correction normalised using the same ^{18}F-[FDG] template. The results obtained when comparing scatter corrected to not scatter corrected images suggest that ^{18}F-[FDG] distribution decreases significantly in the frontal gyri, in addition to the middle temporal and postcentral gyri. On the other hand, the distribution increases in the cerebellum, thalamus, insula, the brainstem, temporal lobe and the frontal

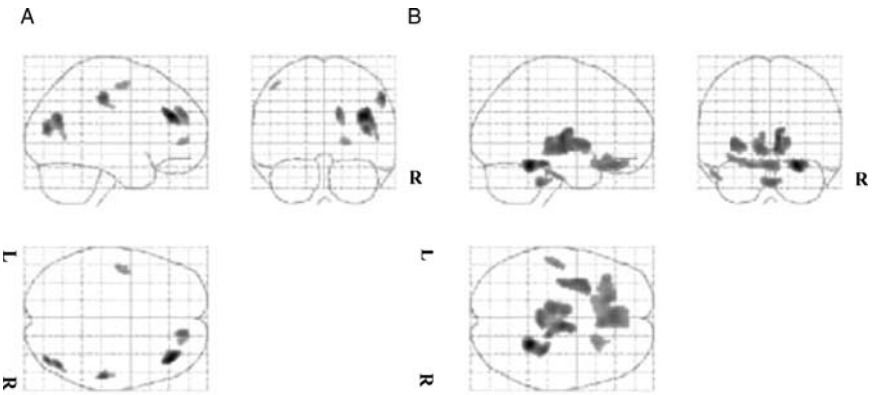


FIGURE 5. Statistical parametric maps resulting from the comparison of images reconstructed using model-based scatter correction and those corrected for attenuation using an effective linear attenuation coefficient without explicit scatter correction normalised using the same ^{18}F-[FDG] template showing areas of significant regional decreases (**A**) and increases (**B**) in brain metabolism. (Reprinted from ref.[94] with permission).

cortex. This needs to be considered for adequate interpretation of ^{18}F-[FDG] 3D brain PET images after applying scatter correction associated with software upgrades or modifications of processing protocols.

In a clinical environment, evaluation of scatter correction is further hampered by the multiplicity of the medical purposes for which the corrections may be studied. For any specific medical task, the evaluation should ideally be based on the performance of human observers. However, this is costly and complex, since a reasonable number of experienced observers should be used to analyse many images under carefully controlled conditions, etc. One severe limitation of performing psychophysical experiments for evaluation of image correction techniques is that it is time consuming and costly. Furthermore, for optimisation of reconstruction algorithms in which possible parameter settings suffer a combinatorial explosion, human psychophysical studies are simply not viable.

Another method to assess the effect of scatter is to investigate the errors in tracer kinetic parameters estimation after scatter compensation. Very few papers addressed this issue during the last decade using SPECT[4] and 3D brain PET data.[83,95] It has been shown that errors of between 10 and 30% can typically result if 3D PET studies are not corrected for scatter.[83] Further investigation using different tracers and different applications using both SPECT and PET data are necessary to fully characterise the effect of scatter correction on tracer kinetic parameters estimation.

There is no single figure of merit that summarises algorithm performance, since performance ultimately depends on the diagnostic task being performed. Well-established figures of merit known to have a large influence on many types of task performance are generally used to assess image quality.[96] With a few exceptions, most papers dealing with the evaluation of scatter correction techniques compare relative concentrations within different compartments of a given phantom with the background compartment serving as a reference. This approach possibly obscures what is actually going on, does not necessarily reflect the accuracy of the correction procedure and might bias the evaluation procedure. Therefore attempts should be made to put such results into absolute terms.

As an example of a comparative study, reconstructions of Monte Carlo simulated 3D digital Hoffman brain phantom without corrections, after applying attenuation correction alone and when using different scatter correction algorithms described above are shown in Figure 6.[90] All the methods improve the contrast compared to the case where no correction is applied. In particular, the low count regions and structures are better recovered after scatter compensation. The CVS and MCBSC techniques tend to overcorrect while SRBSC undercorrects for scatter in most regions of this phantom.

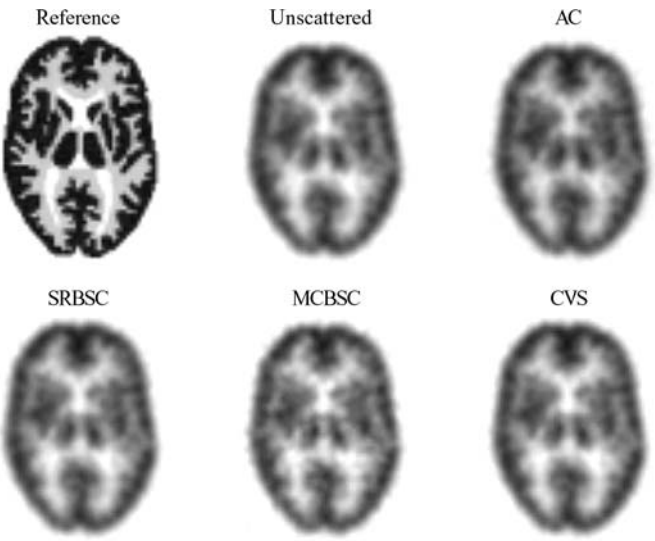


FIGURE 6. Reconstructed images of Monte Carlo simulated Hoffman 3D brain phantom. From top left, clockwise: the reference image used as input to the Monte Carlo simulations, reconstructed image of unscattered events only, attenuation correction only (AC), convolution-subtraction (CVS), statistical reconstruction-based scatter correction (SRBSC), and finally Monte Carlo-based scatter correction (MCBSC). (Reprinted from ref.[90] With permission).

7. Concluding Remarks and Future Prospects

Much research and development has been concentrated on the scatter compensation required for quantitative emission tomography. Increasingly sophisticated scatter correction procedures are under investigation, particularly those based on accurate scatter models, and iterative-based scatter compensation approaches.[3] Monte Carlo methods give further insight and might in themselves offer a possible correction procedure.

The major manufacturers of dedicated SPECT cameras and PET tomographs supply scatter correction software to end-users, whereas the gamma camera-based PET market is still suffering in this respect. As the future of this technology does not seem to be bright, very few efforts have been spent in this area. However, it is expected that commercial software for accurate SPECT/PET quantitation using computationally efficient algorithms will be available in the near future. The scatter correction issue in 3D SPECT/PET is an area of considerable research interest and many research groups are very active in this field, leading the nuclear medicine community to forecast a promising progress during the next few years.

Acknowledgments. One of the co-authors (HZ) acknowledges the support of Grant SNSF 3152A0-102143 awarded by the Swiss National Science Foundation. KFK acknowledges the support of Grant R01 CA87955 awarded by the National Cancer Institute, United States Department of Health and Human Services. The contents are solely the responsibility of the authors and do not necessarily represent the official view of the sponsoring agencies.

References

1. Rosenthal M. S. and Henry, L. J., Scattering in uniform media (for correction of nuclear medicine images). *Phys Med Biol* **35:** 265-274 (1990).
2. Zaidi H., Relevance of accurate Monte Carlo modeling in nuclear medical imaging. *Med Phys* **26:** 574-608 (1999).
3. Zaidi H. and Koral, K. F., Scatter modelling and compensation in emission tomography. *Eur J Nucl Med Mol Imaging* **31:** 761-782 (2004).
4. Kim K. M., Watabe, H., Shidahara, M. *et al*., ''Impact of scatter correction in the kinetic analysis of a D2 receptror ligand SPECT study.'' IEEE Medical Imaging Conference, San Diego, Vol. 3; *pp* 1509-1512 (2001).
5. Klein O. and Nishina, Y., Über die Streuung von Strahlung durch freie Elektronen nach der neuen relativistischen Quantendynamik von Dirac Z. *Z Phys* **52:** 853-868 (1929).
6. Buvat I., Benali, H., Todd_Pokropek, A. *et al*., Scatter correction in scintigraphy: the state of the art. *Eur J Nucl Med* **21:** 675-694 (1994).
7. Links J. M., Scattered photons as ''good counts gone bad:'' are they reformable or should they be permanently removed from society? *J Nucl Med* **36:** 130-132 (1995).
8. Msaki P., Bentourkia, M. and Lecomte, R., Scatter degradation and correction models for high-resolution PET. *J Nucl Med* **37:** 2047-2049 (1996).
9. Koral K. F., ''Monte Carlo in SPECT scatter correction'' in: *Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging*., edited by M Ljungberg, S-E Strand, and M A King (Institute of Physics Publishing, Bristol 1998), *pp* 165-181.
10. Zaidi H., Scatter modelling and correction strategies in fully 3-D PET. *Nucl Med Commun* **22:** 1181-1184 (2001).
11. Beekman F. J. and Viergever, M. A., Fast SPECT simulation including object shape dependent scatter. *IEEE Trans Med Imaging* **14:** 271-282 (1995).
12. Ljungberg M., ''The Monte Carlo method applied in other areas of SPECT imaging'' in: *Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging*., edited by M Ljungberg, S-E Strand, and M A King (Institute of Physics Publishing, Bristol, 1998), *pp* 207-220.
13. Floyd C. E., Jaszczak, R. J., Harris, C. C. *et al*., Energy and spatial distribution of multiple order Compton scatter in SPECT: a Monte Carlo investigation. *Phys Med Biol* **29:** 1217-1230 (1984).
14. Ljungberg M. and Strand, S. E., Scatter and attenuation correction in SPECT using density maps and Monte Carlo simulated scatter functions. *J Nucl Med* **31:** 1560-1567 (1990).

15. Barney J. S., Rogers, J. G., Harrop, R. *et al*., Object shape dependent scatter simulations for PET. *IEEE Trans Nucl Sci* **38:** 719-725 (1991).
16. Frey E. C. and Tsui, B. M. W., Parameterization of the scatter response function in SPECT imaging using Monte Carlo simulation. *IEEE Trans Nucl Sci* **37:** 1308-1315 (1990).
17. Adam L. E., Karp, J. S. and Brix, G., Investigation of scattered radiation in 3D whole-body positron emission tomography using Monte Carlo simulations. *Phys Med Biol* **44:** 2879-2895 (1999).
18. Zaidi H., Reconstruction-based estimation of the scatter component in positron emission tomography. *Ann Nucl Med Sci* **14:** 161-171 (2001).
19. Floyd C. E., Jaszczak, R. J., Greer, K. L. *et al*., Inverse Monte Carlo as a unified reconstruction algorithm for ECT. *J Nucl Med* **27:** 1577-1585 (1986).
20. Buvat I., Lazaro, D. and Breton, V., "Fully 3D Monte Carlo reconstruction in SPECT: proof of concept and is that worthwhile?" *Conf. Proc*. of the VIIth International Meeting on Fully Three-dimensional Image Reconstruction in Radiology and Nuclear Medicine, 29 June-4 July 2003, Saint-Malo, France. Available on CD-ROM.
21. Riauka T. A., Hooper, H. R. and Gortel, Z. W., Experimental and numerical investigation of the 3D SPECT photon detection kernel for non-uniform attenuating media. *Phys Med Biol* **41:** 1167-1189 (1996).
22. Jonsson C. and Larsson, S. A., A spatially varying compton scatter correction for SPECT utilizing the integral Klein-Nishina cross section. *Phys Med Biol* **46:** 1767-1783 (2001).
23. Watson C. C., Newport, D., Casey, M. E. *et al*., Evaluation of simulation-based scatter correction for 3-D PET cardiac imaging. *IEEE Trans Nucl Sci* **44:** 90-97 (1997).
24. Levin C. S., Dahlbom, M. and Hoffman, E. J., A Monte Carlo correction for the effect of Compton scattering in 3-D PET brain imaging. *IEEE Trans Nucl Sci* **42:** 1181-1188 (1995).
25. Holdsworth C. H., Levin, C. S., Janecek, M. *et al*., Performance analysis of an improved 3-D PET Monte Carlo simulation and scatter correction. *IEEE Trans Nucl Sci* **49:** 83-89 (2002).
26. Beekman F. J., de Jong, H. W. and van Geloven, S., Efficient fully 3-D iterative SPECT reconstruction with Monte Carlo-based scatter compensation. *IEEE Trans Med Imaging* **21:** 867-877 (2002).
27. Frey E. C. and Tsui, B. M. W., A fast projector-backprojector pair modeling the asymmetric, spatially varying scatter response function for scatter compensation in SPECT imaging. *IEEE Trans Nucl Sci* **40:** 1192-1197 (1993).
28. Beekman F. J., Kamphuis, C. and Frey, E. C., Scatter compensation methods in 3D iterative SPECT reconstruction: a simulation study. *Phys Med Biol* **42:** 1619-1632 (1997).
29. Frey E. C. and Tsui, B. M. W., Modeling the scatter response function in inhomogeneous scattering media for SPECT. *IEEE Trans Nucl Sci* **41:** 1585-1593 (1994).
30. Beekman F. J., den Harder, J. M., Viergever, M. A. *et al*., SPECT scatter modelling in non-uniform attenuating objects. *Phys Med Biol* **42:** 1133-1142 (1997).
31. Wells R. G., Celler, A. and Harrop, R., Analytical calculation of photon distributions in SPECT projections. *IEEE Trans Nucl Sci* **45:** 3202-3214 (1998).

32. Beekman F. J., de Jong, H. W. and Slijpen, E. T., Efficient SPECT scatter calculation in non-uniform media using correlated Monte Carlo simulation. *Phys Med Biol* **44:** N183-192 (1999).
33. Zaidi H., Herrmann Scheurer, A. and Morel, C., An object-oriented Monte Carlo simulator for 3D positron tomographs. *Comput Meth Prog Biomed* **58:** 133-145 (1999).
34. van Reenen P. C., Lotter, M. G., Heyns, A. D. *et al*., Quantification of the distribution of 111In-labelled platelets in organs. *Eur J Nucl Med* **7:** 80-84 (1982).
35. Jaszczak R. J., Greer, K. L., Floyd, C. E. *et al*., Improved SPECT quantification using compensation for scattered photons. *J Nucl Med* **25:** 893-900 (1984).
36. Koral K. F., Clinthorne, N. H. and Rogers, W. L., Improving emission-computed-tomography quantification by Compton-scatter rejection through off-set window. *Nucl Instr Meth Phys Res* **A242:** 610-614 (1986).
37. Koral K. F., Dewaraja, Y. and Lin, S., ''131I tumor quantification: a new background-adaptive method'' IEEE Medical Imaging Conference, Vol. 2; *pp* 1155-1159 (1997).
38. Koral K. F., Dewaraja, Y., Li, J. *et al*., Initial results for hybrid SPECT–conjugate-view tumor dosimetry in 131I-anti-B1 antibody therapy of previously untreated patients with lymphoma. *J Nucl Med* **41:** 1579-1586 (2000).
39. Koral K. F., Dewaraja, Y., Li, J. *et al*., Update on hybrid conjugate-view SPECT tumor dosimetry and response in 131I-tositumomab therapy of previously untreated lymphoma patients. *J Nucl Med* **44:** 457-464 (2003).
40. Harris C. C., Greer, K. L., Jaszczak, R. J. *et al*., Tc-99m attenuation coefficients in water-filled phantoms determined with gamma cameras. *Med Phys* **11:** 681-685 (1984).
41. Narita Y., Eberl, S., Iida, H. *et al*., Monte Carlo and experimental evaluation of accuracy and noise properties of two scatter correction methods for SPECT. *Phys Med Biol* **41:** 2481-2496 (1996).
42. Axelsson B., Msaki, P. and Israelsson, A., Subtraction of Compton-scattered photons in single-photon emission computerized tomography *J Nucl Med* **25:** 490-494 (1984).
43. Ogawa K., Harata, Y., Ichihara, T. *et al*., A practical method for position-dependent Compton-scatter correction in single-photon emission CT. *IEEE Trans Med Imaging* **10:** 408-412 (1991).
44. King M. A., Hademenos, G. J. and Glick, S. J., A dual-photopeak window method for scatter correction. *J Nucl Med* **33:** 605-612 (1992).
45. Ljungberg M., King, M. A., Hademenos, G. J. *et al*., Comparison of four scatter correction methods using Monte Carlo simulated source distributions. *J Nucl Med* **35:** 143-151 (1994).
46. Buvat I., Rodriguez-Villafuerte, M., Todd-Pokropek, A. *et al*., Comparative assessment of nine scatter correction methods based on spectral analysis using Monte Carlo simulations. *J Nucl Med* **36:** 1476-1488 (1995).
47. Msaki P., Axelsson, B., Dahl, C. M. *et al*., Generalized scatter correction method in SPECT using point scatter distribution functions. *J Nucl Med* **28:** 1861-1869 (1987).
48. Ljungberg M., Msaki, P. and Strand, S.-E., Comparison of dual-window and convolution scatter correction techniques using the Monte-Carlo method *Physics in Medicine and Biology* **35:** 1099-1110 (1990).

49. Floyd C. E., Jr., Jaszczak, R. J., Greer, K. L. *et al.*, Deconvolution of Compton scatter in SPECT. *J Nucl Med* **26:** 403-408 (1985).
50. Luo J. Q., Koral, K. F., Ljungberg, M. *et al.*, A Monte Carlo investigation of dual-energy-window scatter correction for volume-of-interest quantification in 99mTc SPECT. *Phys Med Biol* **40:** 181-199 (1995).
51. Luo J. Q. and Koral, K. F., Background-adaptive dual-energy-window correction for volume-of-interest quantification in 99m-Tc SPECT. *Nucl Instr Meth Phys Res* **A353:** 340-343 (1994).
52. Green A. J., Dewhurst, S. E., Begent, R. H. *et al.*, Accurate quantification of 131I distribution by gamma camera imaging. *Eur J Nucl Med* **16:** 361-365 (1990).
53. Koral K. F., Zasadny, K. R., Kessler, M. L. *et al.*, CT-SPECT fusion plus conjugate views for determining dosimetry in iodine-131-monoclonal antibody therapy of lymphoma patients. *J Nucl Med* **35:** 1714-1720 (1994).
54. Swailem F. M., Koral, K. F. and Rogers, W. L., Count-based monitoring of Anger-camera spectra–local energy shifts due to rotation. *Med Phys* **18:** 565-567 (1991).
55. Koral K. F., Luo, J. Q., Ahmad, W. *et al.*, Changes in local energy spectra with SPECT rotation for two anger cameras. *IEEE Trans Nucl Sci* **42:** 1114-1119 (1995).
56. Koral K. F., Wang, X. Q., Rogers, W. L. *et al.*, SPECT Compton-scattering correction by analysis of energy spectra. *J Nucl Med* **29:** 195-202 (1988).
57. Wang X. and Koral, K. F., A regularized deconvolution-fitting method for Compton-scatter correction in SPECT. *IEEE Trans Med Imaging* **11:** 351-360 (1992).
58. Moore S. C., Kijewski, M. F., Muller, S. P. *et al.*, Evaluation of scatter compensation methods by their effects on parameter estimation from SPECT projections. *Med Phys* **28:** 278-287 (2001).
59. El Fakhri G., Kijewski, M. F., Maksud, P. *et al.*, The effects of compensation for scatter, lead x-rays, and high-energy contamination on tumor detectability and activity estimation in Ga-67 imaging. *IEEE Trans Nucl Sci* **50:** 439-445 (2003).
60. Ogawa K. and Nishizaki, N., Accurate scatter compensation using neural networks in radionuclide imaging. *IEEE Trans Nucl Sci* **40:** 1020-1025 (1993).
61. Maksud P., Fertil, B., Rica, C. *et al.*, Artificial neural network as a tool to compensate for scatter and attenuation in radionuclide imaging. *J Nucl Med* **39:** 735-745 (1998).
62. El Fakhri G., Moore, S. C. and Maksud, P., A new scatter compensation method for Ga-67 Imaging using artificial neural networks. *IEEE Trans Nucl Sci* **48:** 799-804 (2001).
63. Meikle S. R., Hutton, B. F. and Bailey, D. L., A transmission-dependent method for scatter correction in SPECT. *J Nucl Med* **35:** 360-367 (1994).
64. Narita Y. and Iida, H., [Scatter correction in myocardial thallium SPECT: needs for optimization of energy window settings in the energy window-based scatter correction techniques] *Kaku Igaku* **36:** 83-90 (1999).
65. Kim K. M., Varrone, A., Watabe, H. *et al.*, Contribution of scatter and attenuation compensation to SPECT images of nonuniformly distributed brain activities. *J Nucl Med* **44:** 512-9 (2003).
66. Welch A., Gullberg, G. T., Christian, P. E. *et al.*, A transmission-map-based scatter correction technique for SPECT in inhomogeneous media. *Med Phys* **22:** 1627-1635 (1995).

67. Bowsher J. E. and Floyd, C. E., Jr., Treatment of Compton scattering in maximum-likelihood, expectation-maximization reconstructions of SPECT images. *J Nucl Med* **32:** 1285-1291 (1991).
68. Koral K. F., Correction for patient Compton scattering-current status. *J Nucl Med* **32:** 1291-1293 (1991).
69. Hudson H. M. and Larkin, R. S., Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* **13:** 601-609 (1994).
70. Werling A., Bublitz, O., Doll, J. *et al*., Fast implementation of the single scatter simulation algorithm and its use in iterative image reconstruction of PET data. *Phys Med Biol* **47:** 2947-2960 (2002).
71. Kamphuis C., Beekman, F. J., van Rijk, P. P. *et al*., Dual matrix ordered subsets reconstruction for accelerated 3D scatter compensation in single-photon emission tomography. *Eur J Nucl Med* **25:** 8-18 (1998).
72. Hutton B. F. and Baccarne, V., Efficient scatter modelling for incorporation in maximum likelihood reconstruction. *Eur J Nucl Med* **25:** 1658-1665 (1998).
73. Kadrmas D. J., Frey, E. C., Karimi, S. S. *et al*., Fast implementations of reconstruction-based scatter compensation in fully 3D SPECT image reconstruction. *Phys Med Biol* **43:** 857-873 (1998).
74. Bailey D. L., "Quantitative procedures in 3D PET" in: *The theory and practice of 3D PET*, edited by B Bendriem and D W Townsend (Kluwer Academic Publishers, The Netherlands, 1998) *pp* 55-109.
75. Grootoonk S., Spinks, T. J., Sashin, D. *et al*., Correction for scatter in 3D brain PET using a dual energy window method. *Phys Med Biol* **41:** 2757-2774 (1996).
76. Bendriem B., Trebossen, R., Frouin, V. *et al*., "A PET scatter correction using simultaneous acquisitions with low and high lower energy thresholds" Proc. IEEE Med. Imag. Conf., San Francisco, CA, Vol. 3; *pp* 1779-1783 (1993).
77. Shao L., Freifelder, R. and Karp, J. S., Triple energy window scatter correction technique in PET. *IEEE Trans Med Imaging* **4:** 641-648 (1994).
78. Bentourkia M., Msaki, P., Cadorette, J. *et al*., Energy dependence of scatter components in multispectral PET imaging. *IEEE Trans Med Imaging* **14:** 138-145 (1995).
79. Bailey D. L. and Meikle, S. R., A convolution-subtraction scatter correction method for 3D PET. *Phys Med Biol* **39:** 411-424 (1994).
80. Lercher M. J. and Wienhard, K., Scatter correction in 3D PET. *IEEE Trans Med Imaging* **13:** 649-657 (1994).
81. Bentourkia M. and Lecomte, R., Energy dependence of nonstationary scatter subtraction–restoration in high resolution PET. *IEEE Trans Med Imaging* **18:** 66-73 (1999).
82. Links J. L., Leal, J. P., Mueller-Gartner, H. W. *et al*., Improved positron emission tomography quantification by Fourier-based restoration filtering. *Eur J Nucl Med* **19:** 925-932 (1992).
83. Cherry S. and Huang, S. C., Effects of scatter on model parameter estimates in 3D PET studies of the human brain. *IEEE Trans Nucl Sci* **42:** 1174-1179 (1995).
84. Ollinger J. M., Model-based scatter correction for fully 3D PET. *Phys Med Biol* **41:** 153-176 (1996).
85. Wegmann K., Adam, L.-E., Livieratos, L. *et al*., Investigation of the scatter contribution in single photon transmission measurements by means of Monte Carlo simulations. *IEEE Trans Nucl Sci* **46:** 1184-1190 (1999).

86. Celler A., Axen, D., Togane, D. *et al.*, Investigation of scatter in SPECT transmission studies. *IEEE Trans Nucl Sci* **47:** 1251-1257 (2000).
87. Wollenweber S. D., Parameterization of a model-based 3-D PET scatter correction. *IEEE Trans Nucl Sci* **49:** 722-727 (2002).
88. Watson C. C., New, faster, image-based scatter correction for 3D PET. *IEEE Trans Nucl Sci* **47:** 1587-1594 (2000).
89. Zaidi H., Montandon, M.-L. and Slosman, D. O., Magnetic resonance imaging-guided attenuation and scatter corrections in three-dimensional brain positron emission tomography. *Med Phys* **30:** 937-948 (2003).
90. Zaidi H., Comparative evaluation of scatter correction techniques in 3D positron emission tomography. *Eur J Nucl Med* **27:** 1813-1826 (2000).
91. Pan T. S. and Yagle, A. E., Numerical study of multigrid implementations of some iterative image reconstruction algorithms. *IEEE Trans Med Imaging* **10:** 572-588 (1991).
92. Sankaran S., Frey, E. C., Gilland, K. L. *et al.*, Optimum compensation method and filter cutoff frequency in myocardial SPECT: a human observer study. *J Nucl Med* **43:** 432-438 (2002).
93. Friston K., Holmes, A., Worsley, K. *et al.*, Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping* **2:** 189-210 (1995).
94. Montandon M.-L., Slosman, D. O. and Zaidi, H., Assessment of the impact of model-based scatter correction on ^{18}F-[FDG] 3D brain PET in healthy subjects using statistical parametric mapping. *Neuroimage* **20:** 1848-1856 (2003).
95. Townsend D., Price, J., Mintun, M. *et al.*, "Scatter correction for brain receptor quantitation in 3D PET" in: *Quantification of Brain Function Using PET*, edited by Myers R, Cunningham V, Bailey DL and Jones T (Academic Press, San Diego, CA 1996), *pp* 76-81.
96. Kinahan P. E. and Karp, J. S., Figures of merit for comparing reconstruction algorithms with a volume-imaging PET scanner. *Phys Med Biol* **39:** 631-642 (1994).

8
Correction for Partial Volume Effects in Emission Tomography

O.G. Rousset* and H. Zaidi†

1. Introduction

Despite 25 years of continual improvement of the physical characteristics of positron emission tomography (PET) and single-photon emission computed tomography (SPECT) instruments, including crystal technology, improved electronics, and faster computers, PET and SPECT are still plagued with relatively low spatial resolution compared to anatomy-oriented imaging devices such as magnetic resonance (MR) imaging or x-ray computed tomography (CT). Further, in order to accurately explore cell metabolism, reproducibility and sensitivity of the data analysis procedures must at least match the subtle changes occurring in metabolism that one tries to investigate. While it is of primary importance to compensate for physical effects such as interaction of photons with matter resulting in their attenuation (see Chapter 6) and scattering (see Chapter 7), geometry-dependent interactions between the imaging system and the distribution of radioactivity in the field-of-view must also be accounted for. This includes correction methods used to account for collimator response in SPECT (see Chapter 5). Despite all the efforts aimed at improving the quality and meaningfulness of emission tomography (ET), there remains the need to correct for limited resolution (or *partial volume*) effects if one wants to obtain absolute image quantification.[1,2]

Partial volume effects (PVE's) have been shown to result in large bias in the estimate of regional radioactivity concentration, both in experimental phantom and simulation studies. Partial volume is usually addressed in the context of "small" regions, i.e., with dimensions smaller than around 2-4 times the full-width at half-maximum (FWHM) of the scanner's point-spread function (*PSF*). It is rather hard indeed to find for example a single

*Dr O.G. Rousset, Department of Radiology, Johns Hopkins Medical Institutions, Baltimore, MD 21287 USA
†PD Dr H. Zaidi, Division of Nuclear Medicine, Geneva University Hospital, CH-1211 Geneva, Switzerland

brain structure that can elude partial volume given the spatial resolution of current emission tomography systems.[3]

In the presence of tissue atrophy, such as in the case of Dementia of Alzheimer Type (DAT), the signal is diluted even further since the scanner's resolution cell contains even less cortical grey matter. What is important to the clinician and research investigator is the ability to distinguish the loss of radioactivity due to increase in partial volume effects from the true loss of tissue function. In the case of dynamic studies of the heart and brain, time-varying contrasts between the target region and its surrounding lead to distortion of true tissue time-activity curves and subsequent underestimation of physiological parameters.

The purpose of this chapter is to introduce basic notions and describe correction methods for partial volume effects in emission tomography as well as their evaluation strategies.

2. The Problem of Partial Volume Effect

The general problem of partial volume effect was first introduced in the context of the early generation of X-ray CT and PET scanners.[4] With crystal size soaring at over 2 cm thick, substantial part of the anatomical region under study would only partially occupy the imaging volume resolution cell. In emission tomography, this phenomenon of *partial volume* came to include both the loss of detection sensitivity in the axial direction (slice thickness), as well as in-plane resolution effect resulting from in-plane detector response and filtering of the projection prior to backprojection.

The limited spatial resolution of emission tomography depends on a number of factors from various sources: i) physical, such as positron range and non-colinearity of emitted annihilation photons (in PET), scattered radiation (in SPECT and PET), ii) instrumental, such as scintillation crystal size, their detection efficiency and geometrical arrangement, or collimator geometry (in SPECT) and iii) methodological, such as the choice of parameters for image acquisition and reconstruction processes. Finally, even if this will not influence the resolution of the system *per se*, the choice of the regions-of-interest (RoI) for extracting regional activity concentrations will further modulate the degree of recovery of actual tracer distribution.

Due to the imperfections of the imaging system, the response of the scanner to a point source of activity, or *point-spread function* (*PSF*) , will show a bell shape. The full width of the *PSF* taken at half the maximum of the profile (or full-width at half-maximum—FWHM) is generally taken as the measure of spatial resolution (Figure 1).

If we consider the simple one-dimensional case where an object possesses one of its dimensions smaller that 2-3 times the spatial resolution along that particular direction, a profile through the produced image will show an underestimation of true tracer concentration (Figure 2). The signal profile of this

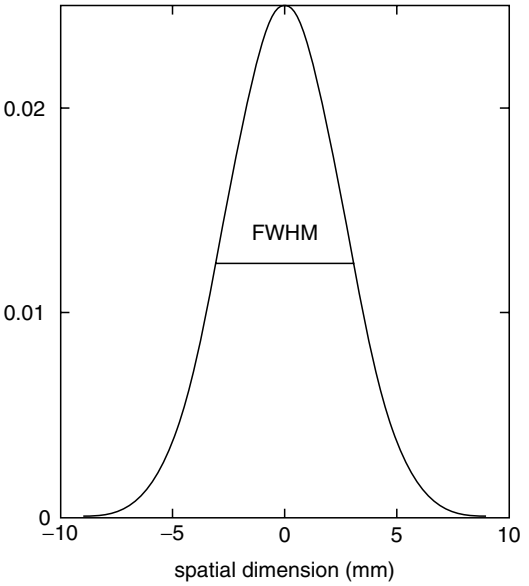


FIGURE 1. Mono-dimensional Gaussian function representing typical spatial response function, or point-spread function (*PSF*), of modern emission tomographs. Spatial resolution is usually given in terms of full-width-at-half-maximum (FWHM) of the *PSF*, which has been chosen equal to 6 mm in this illustration.

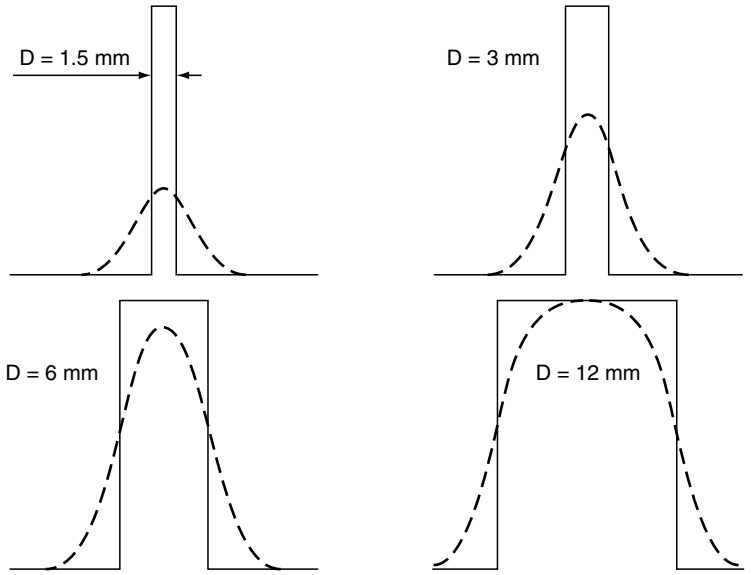


FIGURE 2. Schematic illustration of the partial volume effect. The observed signal (dashed line) is underestimated compared to true profiles (solid lines) depending on object size D with respect to the scanner's spatial resolution (chosen as equal to 6 mm (FWHM) in this example). We see clearly that in order for the object D to exhibit 100% of true original activity, its dimension needs to be greater than $2\times$FWHM ($D > 12$ mm).

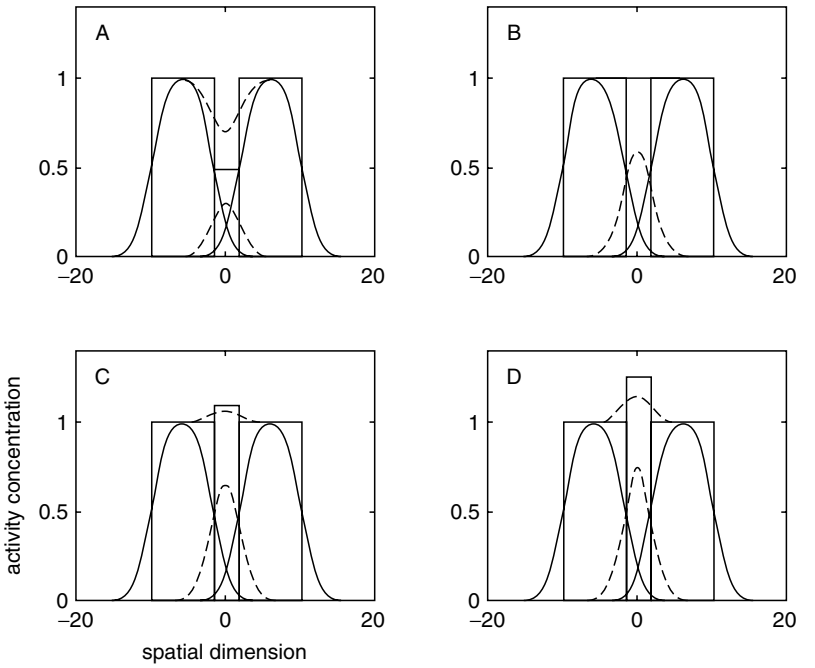


FIGURE 3. More general definition of partial volume effect(s). When a small object of width D (D=6 mm in this illustration) is surrounded by background activity, the observed signal will result from the addition of the target object signal (dashed line) plus a contamination component referred to as "spill-over" or "spill-in". When the activity in the target object is smaller than in the surrounding, we observe an overestimation of its activity (A). In the absence of image contrast, i.e., if the target activity is the same as in the surrounding (B), the observed signal will be equal to the true signal as activity spilled-out is exactly compensated by activity spill-in. When the activity in the target object is above that of the background, the observed signal is underestimated (C and D).

isolated object is further distorted by the presence of neighbouring activity (Figure 3). The goal of partial volume correction (PVC) is to be able to account for both the loss of signal due to the limited size of the object with respect to resolution, and the signal contamination coming from the rest of the image.

3. Theoretical Considerations: Object-Image Relationship

For a non-linear system, with a spatially variant point-spread function $h(r)$, the resulting image $g(r)$ of a distribution of activity $f(r)$ can be written as:

$$g(r) = \int_{\Re} h(r, r', f(r'))\, dr' \tag{1}$$

where r and r' are the vectors in image and object space respectively. For a linear system we have:

$$g(r) = \int_{\Re} h(r, r')\, f(r') dr' \tag{2}$$

where $h(r, r')$ represents the system's *PSF*. Those two last equations represent the most general description of the image formation process, where the *PSF* is a function of the spatial coordinates in both spaces (object and image), and is referred to as spatially variant. In the case where the response to a point source of activity is the same for all the points in the object space, the *PSF* is said to be spatially invariant. In this case, h only depends on the difference in coordinates r-r' since the response to a point source depends only on the relative distance between these 2 points in the image plane. For a spatially invariant, non-linear system, Eq. (2) becomes:

$$g(r) = \int_{\Re} h(r - r', f(r')) dr' \tag{3}$$

Whereas for a spatially invariant, linear system:

$$g(r) = \int_{\Re} h(r - r') f(r')\, dr' \tag{4}$$

We recognize here the convolution integral: the image is equal to the convolution of the object distribution with the *PSF* of the imaging system. The *PSF* transfers the information from the object space to the image space and incorporate all the geometrical infidelities of the imaging system. It is therefore not surprising that the attempt to compensate for those degradations be referred to as a "deconvolution" process. It is worth emphasizing here that this term should in theory only be applicable in situation where the *PSF* is a spatially-invariant function.

4. The General Problem of Image Restoration

If we concentrate on the case where the imaging system is linear and spatially invariant, we have seen (Eq. 4) that the object $f(r)$ and the image $g(r)$ are linked by the convolution integral that can be simply written as:

$$g(r) = h(r) \otimes f(r) \tag{5}$$

where $\otimes$ represents the convolution operator. In Fourier space, the convolution becomes a simple multiplication of the Fourier transforms $H(u)$ and $F(u)$ of the functions $h(r)$ and $f(r)$, respectively, (convolution theorem).We can then write:

$$G(u) = H(u) \times F(u) \tag{6}$$

where u represents the spatial frequency corresponding to the spatial coordinate r. The relationship between a function and its Fourier transform being given by:

$$F(u) = \int_{\Re} f(r) \times e^{-2\pi i u r} dr \tag{7}$$

and

$$f(r) = \int_{\Re} F(u) \times e^{2\pi i u r} du \tag{8}$$

Equation (8) indicates that the object f can be separated into its spatial frequency components u and how these elements can be recombined to recover the original object f. This representation in Fourier space contains the same information than in real space, only in a different form. It is then clear that the Fourier transform H of the *PSF* h corresponds to the fraction of the object distribution component of spatial frequency u that is transferred to the image distribution at the same spatial frequency. H regulates the transfer of information for each spatial frequency, and is often termed the *modulation transfer function* (*MTF*) . If the imaging system was perfect (i.e., $H = 1$), the image would be a perfect representation of the object, and h would be a Dirac function (*PSF* infinitely narrow). Unfortunately, for every imaging system, and in particular for PET and SPECT, we observe a dispersion of the system *PSF* which corresponds to a decrease of the *MTF* magnitude with increasing spatial frequency. It is therefore clear that there is a loss of spatial information for the high spatial frequencies. If one thinks of recovering this information by performing a direct deconvolution, it becomes clear that we need to invert Eq. (6):

$$F(u) = G(u)/H(u) \tag{9}$$

thus

$$f(r) = \int_{\Re} \frac{G(u)}{H(u)} e^{2\pi i u} du \tag{10}$$

The practical use of this simple procedure is not very reasonable if we consider that H decreases in magnitude with increasing spatial frequency. It is at those high frequencies that the image becomes dominated by noise, that the deconvolution suggested in Eq. (10) would amplify to unacceptable levels.

In fact, the image formation process given in Eq. (2) could present the false impression that object and image spaces are only connected by geometrical transformations. In practice, the images are contaminated by a whole variety of noise of various origins. A more accurate description of the image formation process is hence the following for a linear system:

$$g(r) = \int_{\Re} h(r, r')\, f(r')dr' + \eta(r) \tag{11}$$

where η represents the noise distribution in the image (assumed to be additive here). There is no unique solution to this equation, and the problem is said to be "ill-posed", meaning that a small perturbation in the image can lead to large errors in the solution. Image restoration consists in finding the best solution to Eq. (11) while limiting noise amplification resulting from the correction process. If we consider that the image can be written as the effect of an operator on the object, plus a noise component, we can then write Eq. (11) in the form:

$$g = Hf + \eta \tag{12}$$

If we apply the inverse operator, we obtain:

$$H^{-1}g = H^{-1}Hf + H^{-1}\eta \tag{13}$$

If we define a measurement of the object as being the inverse of H applied to g, then:

$$\hat{f} = f + H^{-1}\eta \tag{14}$$

This equation indicates that the "processed" image $\hat{f}$ is equal to the real object plus a term representing the noise amplification. If the operator H is singular, this image can not even be obtained, and even if H is only slightly ill-conditioned, the second term of Eq. (14) becomes predominant and will invalidate the correction method.

5. Theoretical Activity Recovery of Known Size Objects

Typical *PSF* of emission tomography systems can be approximated by a Gaussian function of a few millimetres FWHM. The partial volume effect on spherical objects of inner radii varying from 4 to 15.5 mm, for emission tomography systems with spatial resolution varying between 2 and 10 mm FWHM, are illustrated in Figure 4. It is worth noticing that the smallest sphere (4 mm) is hardly visible when using a system with similar spatial resolution and not detectable at all for systems with lower spatial resolution (> 4 mm).

By definition, a normalized Gaussian aperture function of the spatial function *r(x,y,z)* can be expressed as:

$$h(r) = \frac{(2\pi)^{-3/2}}{\sigma_{S_x}\sigma_{S_y}\sigma_{S_z}} \times \exp\left(-\frac{1}{2}\left[\frac{x^2}{\sigma_{S_x}^2} + \frac{y^2}{\sigma_{S_y}^2} + \frac{z^2}{\sigma_{Sz}^2}\right]\right) \tag{15}$$

where σ_{Sx}, σ_{Sy}, and σ_{Sz} represent the standard deviation in the x, y, and z directions. This symmetric function has a maximum value of $1/\sqrt{2\pi\sigma_{S_x}}$ in 1-D, $1/\sqrt{2\pi\sigma_{S_x}\sigma_{Sy}}$ in 2-D, and $1/\sqrt{2\pi\sigma_{S_x}\sigma_{Sy}\sigma_{Sz}}$ in 3-D.

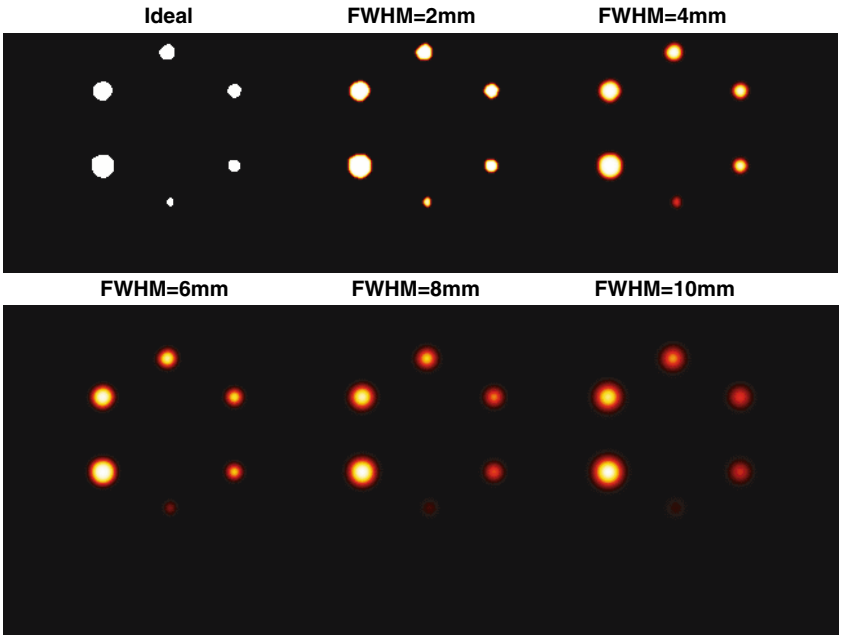


FIGURE 4. Illustration of the partial volume effect using simulated spheres of 4.0, 7.5, 8.5, 11.0, 13.0, and 15.5 mm inner radii for imaging systems with spatial resolution FWHM varying between 2 and 10 mm.

If we consider the Gaussian function definition given in Eq. (15) and the definition of the spatial resolution in terms of the FWHM of the scanner's *PSF*, we can relate the FWHM to the standard deviation σ as follows:

$$\exp\left(-\frac{1}{2}\left(\frac{FWHM}{2\sigma}\right)^2\right) = 1/2 \Leftrightarrow \frac{FWHM}{2\sigma} = \sqrt{2\ln 2} \approx 1.18 \qquad (16)$$

If the Gaussian aperture function presented in Eq. (15) is convolved with a Gaussian object of standard deviations σ_{Ox}, σ_{Oy}, and σ_{Oz} along x, y, and z, the result can be directly assessed after Fourier transform and multiplication in the frequency domain, and can be expressed as:

$$g(r) = \frac{\sigma_{O_x}\sigma_{O_y}\sigma_{O_z}}{\sqrt{\left(\sigma_{O_x}^2+\sigma_{Sx}^2\right)\left(\sigma_{O_y}^2+\sigma_{Sy}^2\right)\left(\sigma_{O_z}^2+\sigma_{Sz}^2\right)}} \times \exp\left(-\frac{1}{2}\left[\frac{x^2}{\sigma_{O_x}^2+\sigma_{Sx}^2}+\frac{y^2}{\sigma_{O_y}^2+\sigma_{Sy}^2}+\frac{z^2}{\sigma_{O_z}^2+\sigma_{Sz}^2}\right]\right) \qquad (17)$$

The results for 1-D and 2-D objects can be readily derived from the previous equation by dropping the z and y variables respectively. For

instance, the maximum recovery coefficient (RC) for a 2-D Gaussian object of identical standard deviation along x and y (i.e., $\sigma_{Ox} = \sigma_{Oy} = \sigma_O$), becomes:

$$RC = \frac{(\sigma_O/\sigma_S)^2}{1 + (\sigma_O/\sigma_S)^2} \tag{18}$$

We can predict analytically the maximum RC values in the case of objects of simple geometrical shape. In the case of a bar of width D that is the only dimension to suffer from partial volume effects, the recovery coefficient can be expressed by the convolution integral:

$$\begin{aligned} RC(x/\sigma) &= \int_{-D/2}^{D/2} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{x^2}{\sigma^2}\right) dx \\ &= erf\left(\frac{D/2}{\sqrt{2}\sigma}\right) \end{aligned} \tag{19}$$

where

$$erf\left(\frac{u}{\sqrt{2}\sigma}\right) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\omega^2}{2\sigma^2}\right) d\omega \tag{20}$$

and represents the error function.[5] The calculation is also possible in 2-D in the case of a cylinder of elliptical section whose length is large with respect to the resolution along the z dimension.[6] For an elliptical section given by:

$$x^2/b^2 + y^2/c^2 = R^2 \tag{21}$$

then

$$\begin{aligned} &RC(R/\sigma) = 1 - \exp(-R^2/2\sigma^2) \\ &RC(D/FWHM) = 1 - \exp(-\ln 2 \times D^2/FWHM^2) \quad \because D = 2R \end{aligned} \tag{22}$$

In the case where a sphere of radius R is centred on an isotropic Gaussian function of standard deviation σ, the maximum value for the 3-D recovery coefficient that one can obtain in the image is given by:

$$RC_{3D} = \frac{1}{(2\pi)^{3/2}\sigma^3} \int_{-R}^{R} dz \int_{-R}^{R} dy \int_{-R}^{R} dx \left(\exp\left(-\frac{x^2 + y^2 + z^2}{2\sigma^2}\right)\right) \tag{23}$$

The result of this integral is proposed in the work published by Kessler *et al.*[7] and is illustrated in figure 5 along with the results obtained in 1-D and 2-D. This calculation was expended to the more general case where the sphere's centre is shifted by an amount Z_p with respect to the centre of the Gaussian aperture function:

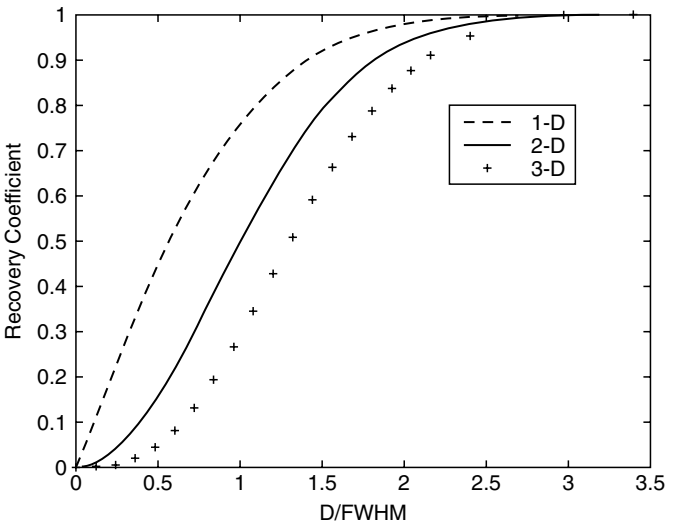


FIGURE 5. Maximum theoretical recovery coefficients (RC) resulting from the computation of the convolution integrals for various geometries: bar (1-D), cylinder (2-D), and sphere (3-D), plotted as a function of their spatial characteristic D normalized to the image resolution in terms of its FWHM.

$$RC(R/\sigma, Z_p/\sigma) = erf\left(\frac{R - Z_p}{\sqrt{2}\sigma}\right) - erf\left(\frac{-R - Z_p}{\sqrt{2}\sigma}\right) - \frac{1}{\sqrt{2\pi}}\frac{\sigma}{Z_p} \exp\left(-\frac{R^2 + Z_p^2}{2\sigma^2}\right) \times \left[\exp\left(\frac{R \times Z_p}{\sigma^2}\right) - \exp\left(-\frac{R \times Z_p}{\sigma^2}\right)\right] \tag{24}$$

For every combination of object and impulse function (*PSF*) , a whole range of *RC* values are possible, with *RC* reaching a maximum when the object and the *PSF* have maximum overlap.[8]

It should be noted here that the *RC* defined by Hoffman *et al.*[9] characterizes only one of the two aspects of the partial volume effect. It represents in fact the fraction of TRUE activity contained in the measurement in the absence of activity other than that present in the object (''cold'' background). This concept does not take into account the presence of a ''warm'' surrounding medium that contaminates the measurements of the ''hot'' spots. We can nevertheless mention the introduction of the concept of contrast recovery coefficient (*CRC*)[7] that reflects the rate of recovery that lies above the surrounding medium (BKG), i.e., *CRC=(obs-BKG)/(TRUE-BKG)*. This parameter is only justified in the case where the BKG is not itself subject to PVE's, and is of known and uniform activity concentration.

6. Partial Volume Correction Strategies

Attempts to compensate for partial volume effects date back to the time where they were first pointed out as a serious limitation in quantitative analysis. The reference paper is probably that of Hoffman *et al.*[9] where he proposed the computation of recovery coefficients from the known geometry of the objects being imaged and their position relative to the centre of a given PET slice. From the same group, Mazziotta then predicted the recovery coefficients for numerous brain structures based on their approximation by a series of non-overlapping spheres.[10] Although the partial volume phenomenon was first addressed in the context of "hot" objects in a "cold" background, emphasizing on the apparent loss of radioactivity due to the small object size with respect to the spatial resolution of the system,[9] it became obvious that it was necessary not only to account for activity "spilling-out" of the "hot" region, but that "spill-in" from the surrounding usually "warm" area should also be accounted for in the regional measurements.[7] Several authors attempted some sort of partial volume correction by applying the recovery coefficients described by Hoffman as well as those derived from Kessler's formulation.[11,12]

Some correction methods require only the original emission data. These include methods making all the necessary corrections for physical effects at the projection level, such as in the method proposed originally by Huesman *et al.*[13] for scatter correction. Iterative reconstruction techniques that incorporate a blurring model in their iterative scheme have been proposed to compensate for the inherent blurring of SPECT.[14,15] This aspect is covered in detail in Chapter 5. There has also been a great deal of search for image processing tools that would restore, or at least visually enhance, the noisy images obtained in emission tomography. Those can be regrouped into the general class of filters used during image reconstruction (low-pass filtering), and those used post-reconstruction for the purpose of restoration filtering. The latter include methods such as Wiener's filtering widely used in evoked potential, and investigated as a way of suppressing noise while maintaining the signal content of time-activity curves.[16,17] In SPECT, depth-dependent blurring can be corrected by first back-projecting each projection and then applying a depth-dependent Wiener filter row by row.[18] Finally, another approach that does not require additional data is based on the computation of correction factors during mathematical modelling of kinetic data, such as regional cerebral blood flow (CBF) measurement with PET, both in the heart[19] and brain.[20]

A distinct class of correction methods require the definition of the various objects being imaged in addition to the characterization of the scanner's *PSF*. These include anatomy-based post-reconstruction correction methods that make use of concomitant high-resolution structural information from

MR imaging or CT.[21-27] Being the most popular methods, those will be described in details later in this section.

Finally, empirical methods based on the derivation of correction factors from experiments with physical test objects remain an active way of characterizing partial volume effects.[28,29]

6.1 *Correction Methods based on Physical Phantom Experiments*

Such empirical methods mimic the object (radioactivity distribution) by a series of simple geometrical shapes (cylinders, spheres) to derive correction factors for actual anatomical structure that can be approximated by a simple shape or a linear combination of simple geometrical shapes. Some researchers have proposed to describe the effect of scatter counts and partial volume on the *PSF* by a series of functions with parameters derived from physical phantoms experiments.[28] This method requires the availability of a realistic physical phantom of the organ or tissue of interest, although empirical rules allow to deriving the *PSF* of an arbitrary object from the *PSF* of a known object.

Also, the estimation of the arterial input function required for the absolute quantification of fast dynamic studies of tracer pharmacokinetics, has been proposed based on its estimation from a large artery after correction for PVE's. Such correction factors can be derived from imaging cylindrical distribution of activity of various diameters and for various levels of background activity.[29] This method is based on Kessler's formulation of the hot spot recovery coefficient (*HSRC*) or ratio of image activity concentration to the true activity concentration in a "hot" isolated region, and cold spot recovery coefficient (*CSRC*) or ratio of image activity in "cold" spot to true background concentration.[7,30] In this case, in the presence of the 2-component system (target+background), according to the formulation of Kessler, the observed estimate of activity within the target region (e.g., arterial vessel) can be expressed as:

$$t_{target} = HSRC \times T_{target} + CSRC \times T_{background} \tag{25}$$

with $T_{background}$ representing the true radioactivity concentration of the background region, and t_{target} is the target radioactivity concentration observed with the imaging system. Under the condition that the target region is totally surrounded by the background region, we have the relationship:

$$HSRC + CSRC = 1 \tag{26}$$

By substituting *CSRC* from Eq. 26 in Eq. 25, we can estimate the corrected activity within the target region:

$$T_{target} = \frac{1}{HSRC}(t_{target} - T_{background}) + T_{background} \tag{27}$$

We can note that the *HSRC* is generally simply referred to as the recovery coefficient. In the present case, the application of this approach assumes that the background activity can be accurately assessed directly from the emission image. In other words, the background is assumed not to be affected by partial volume. This method also assumes that the size of e.g., the vessel, can be accurately estimated. This is a crude approach in its assumption on large surrounding homogeneous background, but has the advantage of only requiring the estimation of the *RC* of the target region. Similarly, recovery coefficients derived from phantom studies have been proposed to correct for partial volume effects in SPECT myocardial perfusion studies.[31] A similar approach was adopted to demonstrate the feasibility of such correction in a clinical setting where the anatomical structures can be approximated by simple geometrical objects.[30] The study concluded that a recovery correction is feasible only for PET data down to lesions of size ~1.5 × FWHM.

However, this kind of approach represents a real technical challenge for more complex organs where anatomical variability or the presence of anatomical abnormalities cannot be addressed by a single physical model. While physical models of the heart in various sizes are available (Data Spectrum, Hillsborough, NC), there is no existing brain phantom that can reproduce the brain circumvolutions or structures smaller than about 10 ml in volume. However, even in the absence of gross abnormalities such as tissue atrophy, normal anatomical variability especially that found in the human brain would not favour the use of a single physical model for correction purposes. However, some authors have for instance proposed a method to overcome the problem of axial under-sampling of emission tomographs based on phantom experiments.[32] Their method make use of a human basal ganglia phantom to validate their assumption according to which the intensity profile in the axial direction can be accurately assessed by Gaussian fitting in order to derive correction factors that compensate for the non-uniform axial response function of the scanner, making the signal independent from axial positioning of the head.

6.2 Reconstruction-based Correction Methods

Such methods may or may not require the availability of supplementary structural information from e.g., MR imaging. Because of the inherent degradation of spatial resolution during filtered backprojection, there is a great deal of research for statistical reconstruction approaches that would overcome this problem of resolution loss while dealing with the intrinsic Poissonian nature of the noise. For example, iterative reconstruction methods such as those derived from the Maximum-Likelihood Expectation-Maximization scheme[33] can incorporate all kind of information such as scatter or spatial variation of the system's *PSF* to compensate for non-stationary resolution (see Chapter 4). Those methods have been explored both in SPECT[14,34-36] and PET.[37]

Some other iterative reconstruction methods not only incorporate some model of the scanner's *PSF*, but can also incorporate *a priori* anatomical information from e.g., MRI, in an attempt to compensate for partial volume effect, both in SPECT[38] and PET.[39,40] However those methods still fail to account for resolution effects in the axial direction, which remains the predominant source of partial volume effects.

Another class of correction schemes, based on quantification directly in projection space, is derived from the early work of Huesman *et al.*[13] and was extended to include not only scatter but also spatial resolution effects, both in the heart[41] and tumour imaging.[42] The major advantage of this approach is its ability to derive accurate figure of regional variance. However this kind of methods cannot account for partial volume effects of objects with cross-section about the size of the intrinsic resolution of the scanner, and has yet to be extended in 3-D.

6.3 Post-reconstruction Correction Methods

It becomes increasingly common, if not systematic, to have access to both the functional information from emission tomography, together with its underlying anatomy defined from high-resolution, structure-oriented, scanning devices such as MR or CT. The combined use of anatomical and functional data allows for a more accurate identification and definition of regions used for the assessment of regional tracer concentration. This trend is also reflected in the development of PET/CT and PET/MRI devices that allow access to both types of information simultaneously, thus avoiding the problem of inter-modality spatial realignment errors (see Chapter 9).

Such correction method consists in solving the imaging equation (Eq. 2). Due to the stochastic and Poissonian form of the signal, it is not possible to simply deconvolve the image with the inverse of the *PSF* without unbearable noise magnification (see Chapter 4).

In order to reduce the number of unknowns in the imaging equation, i.e., minimizing the noise propagation issue, it is necessary to perform a data reduction, i.e., a segmentation of the anatomical information provided by CT or MR imaging. It is therefore assumed that each "segment" of the "activity distribution model" represents a distinct and homogeneous activity distribution. If we consider that the activity distribution $f(r)$ consists of N tissue components of true activity T_i, the resulting image $g(r)$ for a linear system (spatially invariant or not) can be written as the imaging equation:

$$\begin{aligned} g(r) &= \sum_{i=1}^{N} \int_{D_i} h(r, r')\, f_i(r)\, dr + \eta(r) \\ \text{with } & \bigcup_{i=1}^{N} f_i(r) = f(r) \end{aligned} \tag{28}$$

There is no assumption here on the noise characteristics of the image. Although Poissonian in essence, the final noise $\eta(r)$ in the image includes also components from attenuation, scatter, or other type of corrections during data acquisition and reconstruction. Since the partial volume solutions proposed here do not call for the explicit characterization of a noise model, we will omit the noise component in the subsequent equations.

6.3.1 Pixel-guided Approach

If one seeks to recover the entire true activity distribution $f_i(r)$ within tissue component D_i, there are N unknowns, the true image distribution of activity of each tissue component of the tracer distribution model, but only one equation, the emission image $g(r)$. Several authors proposed to solve this equation, first with $N=1$ to compensate for signal dilution in spaces void of activity.[43] Compensating for dilution of radioactivity in non-active tissues such as cerebro-spinal fluid is more important in the case of tissue atrophy, as the decrease of metabolism seen with increased tissue atrophy might be confounded by the loss of signal consequent to increased partial volume effect. For example, the decline in blood flow observed with PET has been shown not to decline with age after correcting for tissue atrophy using this method.[44] These techniques make use of an anatomical mask defined from MRI or CT, and by assigning pixels corresponding to the cerebral tissue (i.e., grey and white matter) a value of 1, and the space occupied by non-cerebral tissue being kept at 0. This is equivalent to defining an anatomical mask $f_1(r)$ as follows:

$$f_1(r) = \begin{cases} 1 & r = cerebral\ tissue \\ 0 & r = non - cerebral \end{cases} \tag{29}$$

As discussed earlier (Eq. 5), if we consider the point-spread function h of the imaging system as being a spatially invariant function, the equation becomes a simple convolution. The next step of this correction method hence consists in convolving the binarized volume $f_1(r)$ at the scanner's resolution. The PET image in divided by this "low resolution" mask in order to obtain images corrected for dilution of radioactivity in non-active spaces. Hence, the very approximate of the true activity distribution is given by:

$$f(r) = \frac{f_1(r)g(r)}{f_1(r) \otimes h(r)} \tag{30}$$

This equation being derived from the approximation:

$$\frac{f(r)}{f(r) \otimes h(r)} \approx \frac{f_1(r)}{f_1(r) \otimes h(r)} \tag{31}$$

It can be seen that the approximation proposed in Eq. (31) is only justified when $f(r)$ possesses the same spatial characteristics as those defined by the anatomical mask $f_i(r)$. This approach ignores the difference in tracer uptake between grey and white matter.

An extension to a more heterogeneous distribution of the tracer was later proposed, this time with a more realistic mask that makes the distinction between grey and white matter to account for white matter activity contribution to measurements of grey matter activity concentration.[22] We have this time the problem that the number of unknowns is now equal to $N = 2$, while the number of equations has not changed (only one). The way these authors overcame this problem is by transforming one of the unknowns into a known variable. They assume that the true white matter activity can be accurately measured from a large remote white matter area in the emission image (Figure 6). This new mask can be defined as follows:

$$f(r) = \begin{cases} f_1(r) & r = \textit{Grey matter} \\ f_2(r) & r = \textit{White matter} \\ f_3(r) & r = \textit{background \& CSF} \end{cases} \tag{32}$$

Where e.g., $f_1(r) = 1$ for pixels identified as Grey matter, and $f_1(r) = 0$ elsewhere. The tissue components Grey, White, and $BKG + CSF$ are obtained by segmentation of the MR images realigned with the PET image volume. Figure 6 illustrates the general principle of MR-guided PVC in functional brain imaging.

By considering that the radioactivity concentrations in white matter and CSF spaces are constant, and that those components do not suffer themselves from partial volume thanks to their important dimensions with respect to the imaging system resolution, the PET image can be written as:

$$\begin{aligned} g(r) &= f(r) \otimes h(r) \\ &= [T_1 f_1(r) + T_2 f_2(r) + T_3 f_3(r)] \otimes h(r) \\ &= T_1 f_1(r) \otimes h(r) + [T_2 f_2(r) + T_3 f_3(r)] \otimes h(r) \end{aligned} \tag{33}$$

where T_2 and T_3 are *known* constants representing the true activity of the white matter and background plus CSF, respectively. After rearranging the previous equation, one can write:

$$T_1(r) = \frac{g(r) - T_2 f_2(r) \otimes h(r) - T_3 f_3 \otimes h(r)}{f_1(r) \otimes h(r)} \tag{34}$$

White matter activity T_2 is considered as being uniform throughout the brain whose activity concentration is considered as being accurately estimated from a measurement in the pons, assumed to elude partial volume due to its large cross-sectional size. This value is assigned to the low-resolution white matter image $f_2(r) \otimes h(r)$, and then subtracted from the PET image $g(r)$. These authors proposed to do the same with background + CSF compartment, although one might wonder whether the measured value for that compartment is not the result of spillover from adjacent tissue plus noise. However, for increased realism, and to make a distinction in grey matter tracer uptake, these authors extended this elimination-substitution

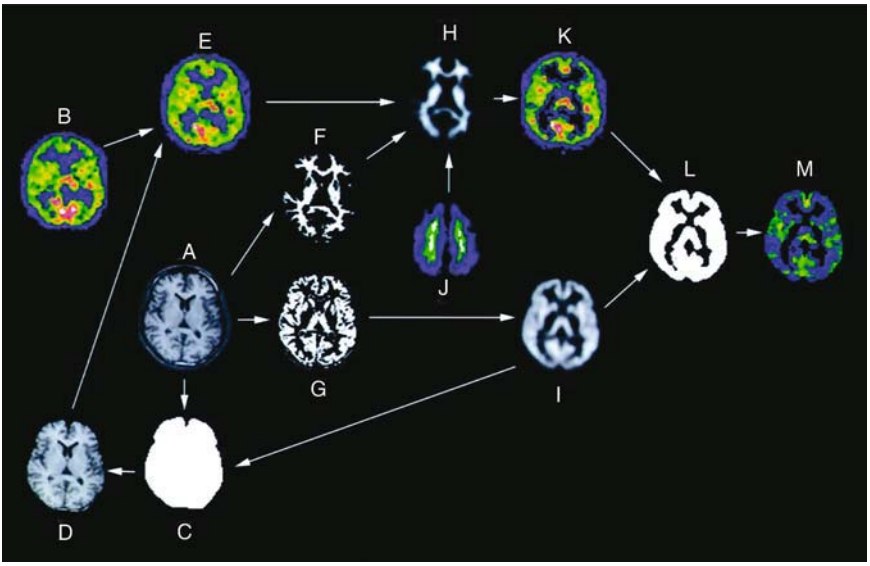


FIGURE 6. Schematic representation of the different steps required for MRI-guided correction for partial volume effect illustrating the original MR (A) and SPECT (B) images, binary mask for whole brain (C); MRI after scalp editing (D); original SPECT image coregistered to MRI (E); MR image segmented into white matter (F) and grey matter (G); White matter SPECT image (H) simulated from convolved white matter MR image; convolved grey matter MR image (I); white matter MR images (J); grey matter SPECT images (K) obtained by subtraction of simulated white matter SPECT image from original SPECT image coregistered to MRI; binary mask for grey matter (L) applied to finally obtain grey matter SPECT image corrected for partial volume effect (M). (Reprinted from ref.[45] with permission).

scheme one step further by incorporating a distinct volume-of-interest (VoI) such as the amygdale.[46] Their method consists first in solving the problem of white matter contribution by using Eq. (33). Subsequently, true cortical activity concentration is measured from the corrected image given in Eq. (34). The corrected image for the amygdala is then given by:

$$T_{VOI}(r) = \frac{g(r) - T_1 f_1(r) \otimes h(r) - T_2 f_2(r) \otimes h(r) - T_3 f_3 \otimes h(r)}{f_{VOI}(r) \otimes h(r)} \tag{35}$$

The "true" cortical value T_1 derived from the "white matter and CSF-corrected image" given in Eq. (34), must satisfy the following criteria:

(i) it must be representative of the true cortical grey activity that actually contaminates the VoI;
(ii) it must be sufficiently far from the VoI to avoid integrating some of its signal.

We can see that those two criteria might rapidly introduce some conflict with increasing tracer heterogeneity, and would be for example violated if one seeks to account for cross-contamination between e.g., the caudate nucleus and the Putamen since they are lying so close to each other.

6.3.2 Model-based Approach

Another type of post-reconstruction correction methods is the model-based optimization method developed by Chen *et al.*[42] to simultaneously recover the size and the activity concentration of small spheroids thus improving estimates of lesion activity in clinical oncology when object size is unknown (Figure 7). The algorithm is based on a 3D spatially varying object size- and contrast-dependent Gaussian model of the system PSF. A match index is then used to estimate the best model parameters. The authors report a reduction in the activity error by 11%-63% compared to the error obtained without correction. Moreover, the accuracy of the simple RC method is dependent on object-to-background ratio (OBR) and the data used for estimating fitting parameters. Large errors were reported for small spheroids, which are obviously very sensitive to OBR variation and noise. A modified version of the algorithm described above combined with an extension to non-spherical objects was recently proposed.[47] The method is being improved currently allowing the quantification of lung tumours with smallest radii with improved convergence properties towards the true model.

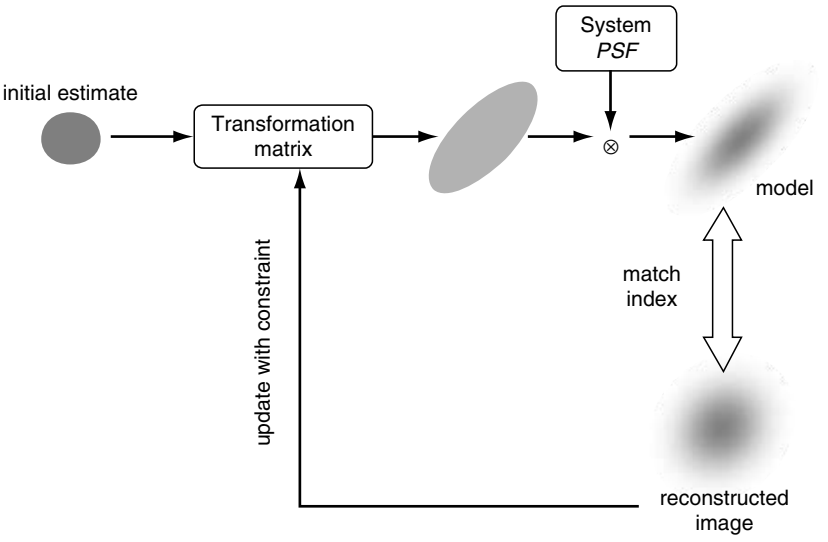


FIGURE 7. Description of the model-based method for partial volume correction in clinical oncology.[42]

6.3.3 Geometric Transfer Matrix (GTM) Approach

If we directly compute the effect of signal degradation due to limited spatial resolution on the mean regional concentration within a limited region of space, or region-of-interest (RoI[1]), we can obtain as many equations as there are unknowns. This is the basis for the method proposed by Rousset *et al.*[23,24,48] which allows for the calculation of corrected estimates without *a priori* knowledge on any activity level. For instance, the observed activity t_j within tissue component D_j from a given RoI_j is given by:

$$t_j = \sum_{i=1}^{N} \omega_{ij} T_i \tag{36}$$

where T_i represents the true tracer concentration within tissue component i. The weighting factors ω_{ij} represent the fraction of true activity T_i from tissue i that is integrated in the measurement t_j from RoI_j of volume v_j. They can be expressed as:

$$\omega_{ij} = \frac{1}{v_j} \int_{ROI_j} RSF_i(r)dr \tag{37}$$

where $RSF_i(r)$ represents the regional spread function of tissue i and corresponds to the response of the scanner to the distribution of activity D_i:

$$RSF_i = \int_{D_i} h(r, r')dr' \tag{38}$$

The weighting factors ω_{ij} constitute the geometric transfer matrix (GTM) and express the distortions introduced by the limited intrinsic spatial resolution of the scanner, as well as smoothing introduced during image back-projection, and further modulation during extraction of regional tracer concentration (RoI analysis). Both the effect of type and size of filter used during FBP[49] as well as definition of RoI[50,51] have been shown to introduce bias in parameter estimates.

In practice, these partial volume factors are computed from simulation of the noise-free *RSF* images and sampling with a user-defined set of RoIs. The number of RoIs must be equal to the number of tissue components identified in the tracer model in order to provide a full-rank GTM. In that case, the diagonal terms of the GTM represent the tissue self-recovery *RC* (or "spill-out") while off-diagonal terms ($\omega_{ij}, j \pm i$) represent the spill-over, or "spill-in", factors. If the number of RoIs is greater than N, the problem is over-constrained and can be solved by ordinary linear least square regression. The same set of RoIs must be used to extract observed values from the

[1]sample of image pixels for average radioactivity concentration computation

actual PET images in order to obtain the vectors t_j (tissue time-activity curves). The set of linear equations can then be solved for the true tracer concentration in each region by inverting the GTM matrices and multiplying by the observed regional values.[23,24,48]

This method was initially developed based on the simulation of sinogram data, in order to reproduce very accurately the effect of finite sampling and backprojection. This further allows for the incorporation of Poisson noise at the detector level for more realistic computer simulations of radioligand uptake for validation purposes.[23,48] It has since been shown that instead of simulating the *RSF* images from sinogram data, one can achieve similar levels of accuracy by creating the *RSF* images by direct convolution of the individual D_i maps with a kernel representative of the spatial resolution of the PET image.[52] This method can be seen as the image-based variant of the GTM method, and is much faster to implement without apparent sacrifice in data accuracy. Nonetheless, the sinogram approach has the advantage of being independent of the filter used during image reconstruction since it uses the same filter and FBP algorithm as the original PET data.

7. Performance Evaluation of Partial Volume Correction

Like any other correction algorithm, PVC methods must be validated and their limit tested before being applied to clinical or research data. This includes the assessment of the accuracy of the algorithm and its sensitivity towards methodological considerations such as simulation of the system's *PSF*, MR-ET image registration and MR image segmentation.

Absolute measurement of accuracy is only attainable in phantom and simulation studies. Accuracy is usually given as a percentage of the true activity, or can be expressed by the apparent recovery coefficient (*ARC*), which represents the apparent (observed or partial volume corrected) regional radioactivity concentration to true activity ratio.[23,48,52] Precision, or data reproducibility, is a more subjective measure of performance and will depend to a great extent on the level of automation of data analysis and correction. Methods that do not require human intervention may have a great level of reproducibility, but particular attention needs to be paid to accuracy of unsupervised methods.

7.1 Registration and Segmentation Errors

Post-reconstruction partial volume correction methods requiring additional anatomical information from e.g., MR imaging, such as the pixel-based approach or the GTM approach, rely on the spatial realignment of functional and anatomical images. Usually, a rigid-body transformation is used, with spatial transformation parameters being derived from the minimization of the distance between homologous points found in both modalities, or increasingly based on some similarity criterion between the two images.

These include methods such as Woods' popular minimization of the variance of ratio between the 2 images to be registered.[53] Other similarity criteria, such as the maximum cross-correlation54 or mutual information,[55] are also popular methods for image realignment.[56] The final image resolution, signal-to-noise ratio, and contrast in the image, all condition the success of the registration (see chapter 9).

The accuracy of the partial volume correction method will depend in part on the degree of accuracy in the realignment of the anatomical images with the emission image of tracer distribution. This has been investigated for both the pixel-based method[22,46,57,58] as well as for the GTM approach.[24,52,58,59] For the GTM approach, it is interesting to note that errors introduced during mis-registration only affect the observed estimates, and does not modify the coefficients of the GTM matrix. As a consequence, the registration error effect on the corrected estimates is of the same magnitude as the effect of mis-registration on the observed estimates due to poor RoI placement.[59] Those errors have been found to have relatively little impact (<2% of true value for typical 1-2 mm mis-registration error) on the final accuracy of the corrected estimates.[52,59] As for errors in segmentation of the tissue components of the system, they have been found to be of greater significance with for example a 5% decrease in Caudate Nucleus *ARC* if a 25% error in total volume is made.[52] However, it has been shown that the effect of the segmentation error was limited to the mis-segmented region.[52] Overall, it appears that the success of the segmentation of the structural information provided by e.g., MR images, has a higher impact on the accuracy of the corrected estimates,[60] compared to the influence of image co-registration, although some authors recently suggested that mis-registration errors have the strongest impact on data accuracy and precision.[58] This recent finding is also in contradiction with the conclusion achieved in the case of the performance evaluation of the method proposed by Müller-Gärtner.[22,57] The accuracy of this latter method further depends upon the accuracy in measurement of background (i.e., white matter) activity concentration. This error has been evaluated as being in the order of 5% error in grey matter (GM) PET estimate for a 20% error in white matter tracer concentration.[22]

As for the overall performance, i.e., in the absence of major sources of registration or segmentation errors, partial volume corrected estimates have been found to be typically within 5-10% of true tracer concentration with a standard deviation of a few percent in both phantom and simulation studies.[24,48,52,58,59,61]

7.1.1 Tissue Homogeneity

Segmentation errors can be thought as a more general problem of tissue heterogeneity. Indeed, the major limiting factor of those methods is primarily the assumption made about the homogeneity of tracer distribution in

each identified, or segmented, tissue components. Aston and colleagues have proposed a test of inhomogeneity based on Krylov subspace iteration that can test for hypothesis on homogeneous tracer distribution.[61] This would however only be effective in the case where the specified noise model is sufficiently accurate, which still remains a difficult endeavour.

However, if known, the tissue mixture of each identified component can be used to compute the various regional spread functions (Eq. 35). This is achieved by the use of statistical probabilistic anatomical maps (SPAM's)[62] that represent the probability of each tissue class (e.g., GM, WM, CSF) to be present in a given location of a standardized, or stereotaxic space.

7.1.2 Data Precision

Even if the registration, segmentation, or other homogeneity errors were inexistent, like any type of "deconvolution" procedure, compensating for partial volume effects will result in some degradation of the precision of the processed data. The variance associated with the corrected estimates can be estimated by explicit study of the covariance nature of the correction method. The degradation of the coefficient of variation (std/mean) after/before PVC can be seen as the noise magnification factor (*NMF*) resulting from the correction.[24] Maximum theoretical values of the NMF can be easily predicted for the GTM method and have shown to be in good agreement with experimental values derived from a brain phantom experiment.[24] Variance associated with the correction has been shown to only slightly increase after partial volume correction[23,24,48,52] suggesting the applicability of GTM-PVC to dynamic emission studies.[63,64]

7.2 Simulation Studies

Computer simulations are a powerful way to explore the limits of correction algorithms since they give a complete control to the operator. By reproducing realistic noise characteristics associated with emission tomography, one can demonstrate the usefulness of PVC algorithms and their applicability for a wide range of realistic situations. Another powerful application of computer simulation is the replication of methodological flaws and their consequence on image quantitation. These include simulation of erroneous definition of the scanner's *PSF*, registration errors between MR/CT and PET/SPECT, segmentation errors, and assumptions made about homogeneity of tracer distribution.

Computer simulation of emission tomography images would usually consist in reproducing the 4-D spatial distribution of the tracer in a digitized model of the organ of interest, and mimicking the geometry and other physical effects specific to the scanner.

7.2.1 Simulation of Objects

Simulation can consist in using mathematical models of simple shaped objects such as spheres or ellipses. For example, Mullani[65] designed a special phantom to measure the quantitative recovery of myocardial emission tomography images as a function of the size of the object and the angulation of a 1-cm-thick simulated myocardium inclined with respect to the image plane. Muzic *et al.*[41] used a set of 5 different elliptical models as part of their validation for the proposed PVC method applied to myocardial PET. This mathematical model represented the porcine torso including the lungs, myocardium, ventricular cavity, and soft tissue background.

Early digital phantoms were derived from post-mortem brain slices and used to evaluate the signal-to-noise performance with decreasing detecting crystal size.[66-68] A digitized version of the Data Spectrum sphere phantom was used for the evaluation of the Müller-Gärtner method.[57]

More recently, the development of classification and segmentation techniques applied to MR or CT images have allowed for the creation of numerical anatomical models of the spatial distribution of the parameter of interest. Digitized models can be a carefully segmented single individual MR volume[69] or models derived from the average of repeated high-resolution images of the same subject.[70] Atlases derived from tissue probability maps have become a popular way for automatic image segmentation,[71] and can be applied to individual data sets to create customized numerical phantoms. Simulated MR images can also be used to investigate the impact of MR image quality on segmentation accuracy[60,72] and could be used to study its subsequent effect on the accuracy of partial volume corrected estimates.

Those objects can be subject to local or global deformations to simulate a variety of situations. For example, the effect of tissue atrophy can be simulated to estimate the effect of a reduced brain structure volume on expected observed estimates.[73] Segmentation error can be simulated by erosion or dilation of the structure of interest and assess the resulting effect on the corrected estimates.[52,58]

7.2.2 Simulation of Image Formation Process

The effect of finite spatial resolution is usually reproduced either at the projection level, or directly in image space. The latter allows for a much faster implementation but does not offer the possibility of adding appropriate noise characteristics that are hard to define in image space. Adding appropriate noise to the image data is necessary for assessing the accuracy and precision of the correction technique in the presence of noise conditions comparable to that encountered in clinical practice. On the other hand, simulation of sinogram data allows for incorporation of the scatter component and the effect of attenuation at the detector level. This provides an opportunity to reproduce the propagation of Poisson noise encountered in a

more realistic way. Further, filtered backprojection type of image reconstruction is still widely used despite the forthcoming of reliable iterative reconstruction schemes for emission tomography (see Chapter 4). Smoothing the projection before projecting them back onto the image grid will get rid of a lot of the high-frequency noise, and will modulate both the magnitude of the signal—further degrading spatial resolution—and the characteristics of the noise.

Monte Carlo simulations represent a powerful way to reproduce the physics occurring in emission tomography (see Chapter 11). Similar to popular analytical simulators, simulated PET or SPECT images derived from Monte Carlo simulations can be used for validation of partial volume correction methods. In particular, Monte Carlo-based generation of projection data typical of [^{18}F]-dopa and [^{11}C]-raclopride uptake have been simulated using a digitized brain phantom[69] to assess the recovery capability of the GTM algorithm.[52]

7.2.3 Simulation of Tracer Kinetics

Simulation of the dynamic uptake of PET tracers such as neuroreceptor ligands, allows for testing the PVC capability to recover known distributions, both in terms of raw activity concentrations, and in terms of physiological parameters extracted from the kinetic modelling process. Tissue time-activity curves (TAC's) can be derived from individual emission data, corrected for partial volume effects using the GTM approach, and then fitted with a mathematical model normally used for extracting meaningful physiological parameters. The fitting curves can then be taken as true TAC input, assigned to the various tissue components identified from anatomical source, and processed through the simulator. This provides a convenient way for studying for example the effect of cortical grey matter heterogeneity on subcortical PVC estimates for various levels of heterogeneity and for varying contrast conditions.[74] More generally, this allows testing the performance of the PVC algorithms in noisy conditions similar to those expected or seen in clinical and research setting.

7.3 Experimental Phantom Studies

As discussed at the beginning of the previous section, physical phantoms can be used to directly derive correction factors. Those phantoms are limited in terms of their applicability due to the limitation in complexity one can achieve in reproducing the organ under study. However, physical phantom experiments are mostly used for validation purposes and represent the *sine qua non* step for proper acceptance of the technique. Physical phantoms of various degree of complexity can be used to demonstrate that the PVC algorithm works well for known distributions of radioactivity. They include simple geometrical objects with one, two, or all three dimensions suffering

from PVE for which the analytical expressions of quantitative recovery are available (cf. Eqs. 19-24). Validation is usually sought out by first making sure the partial volume correction method works for simple geometrical shapes such as cylinders or spheres. For example, a set of 5 spheres of various sizes was used for the validation of the simulated *PSF* used in the GTM-PVC algorithm.[24] Validation is further carried out with phantoms of increased complexity, with possible multiple isotope experiments for simulating time-varying contrast.[52] Realistic phantoms containing several independent compartments surrounded by a realistic medium, such as the brain phantom used to validate the GTM approach[24,48] represents more realistic imaging conditions. Those models are suitable for neuroreceptor studies where the tracer accumulates specifically in the striatum. For tracers diffusing more homogeneously throughout the cortex, an anthropomorphic phantom (STEPBRAIN) separating the cortex from white matter has been proposed.[75]

7.4 Clinical and Research Studies

7.4.1 Brain Imaging

Because of limitations of spatial resolution, quantitative PET measurements of cerebral blood flow, glucose metabolism and neuroreceptor binding, are influenced by partial-volume averaging among neighbouring tissues with differing tracer concentrations.[1] Partial volume effects are important for describing the true functional contribution of nuclear medicine images.[76] Decomposition of these images into a functional and structural component is necessary for studies of healthy ageing and pathophysiology, as well as for assessing clinical patients.[3,63]

Several algorithms have been proposed to improve positron emission tomography quantification by combining anatomical and functional information. The anatomical information could also be used to build an attenuation map for attenuation and scatter correction purposes.[77] The precision of these methods when applied to real data depends on the precision of the manifold correction steps, such as *PSF* modelling, magnetic resonance imaging-positron emission tomography registration, tissue segmentation, or background activity estimation. A good understanding of the influence of these parameters thus is critical to the effective use of the algorithms.[20,24,78] It has been shown that a two-compartment approach is better suited for comparative PET/SPECT studies, whereas the three-compartment algorithm is capable of greater accuracy for absolute quantitative measures.[57]

Several neuroimaging studies have attempted to verify the neurophysiologic correlates of age-related changes in the brain. In the early 1980s, using the ^{133}Xe inhalation method, researchers investigated cerebral blood flow and reported significant reductions with age. Since these pioneering studies, which suffered from poor spatial resolution and other limitations, the advent

of PET has provided neuroscientists with more sophisticated tools for the quantitative measurements of brain physiology.[25] Cerebral volume loss resulting from healthy aging processes can cause underestimation of PET physiologic measurements, despite great improvement in scanner resolution.[79] Thus, the failure to account for the effect of partial-volume averaging of brains with expanded sulci has contributed to the confounding results in functional imaging studies of aging. After partial volume correction, no CBF decline with age in healthy individuals is described.[44,80]

It is therefore expected that an important contribution to PET and SPECT imaging of the brain will be obtained by enhanced reconstruction algorithms incorporating resolution recovery techniques. An example of such method applicable to clinical data is a data-driven automated deconvolution approach.[81] Promising results to achieve such goal have been obtained by several authors e.g. using probabilistic MRI segmentation, subsequent binarization and convolution to obtain dispersion coefficients.[52]

Although different statistical mapping methods may yield grossly similar patterns of hypometabolism or hypoperfusion, the extent, severity, and peak location of metabolic changes can be inconsistent. Deformation accuracy appears to be more prone to atrophy.[82] Accurate estimates of striatal uptake and BP in ^{123}I brain SPECT are feasible with PVC, even with small errors in registering SPECT with anatomic data or in segmenting the striata.[64] In various pathological conditions, PVC can assess the functional contribution to pathology. Reduced glucose metabolism measured by PET in DAT is not simply an artefact due to an increase in CSF space induced by atrophy, but reflects a true metabolic reduction per gram of tissue.[83] Also in epileptic foci, hypometabolism is larger than a mere atrophy effect.[84]

It has been demonstrated that cerebral atrophy could not solely account for the loss of tissue function seen in DAT.[85] Conversely, when the same type of dilution correction is applied in the case of the study of normal aging, partial volume correction annihilates the significant decrease in cerebral blood flow in normal aging commonly reported before atrophy correction.[44] This is in disagreement with several earlier reports of decreased metabolism with normal aging.[86,87]

Brain perfusion studies performed in patients with probable DAT showed that rCBF was decreased in the parahippocampal gyrus but not in the hippocampus after pixel-based partial volume correction.[88] It was also demonstrated that the apparent decrease in uptake of a muscarinic cholinergic antagonist seen in temporal lobe epilepsy was due to a decrease in hippocampal volume rather than a decrease in receptor concentration.[89] This correction was based on global scaling by total measured volume of hippocampus from MRI.

Dopamine transporters have been shown to be markedly reduced in Lesh-Nyhan disease, partial volume correction only accentuating this finding.[73] Dopa-decarboxylaze activity has been shown to be greatly reduced in patients with Parkinson disease compared to normal controls.[63]

Studies of the dopaminergic system with SPECT have been shown to greatly benefit from partial volume correction using a striatal phantom modelling the basal ganglia[90] and a numerical model of the RSD brain phantom.[64] The latter study has shown that the bias in binding potential can be reduced from 50% to about 10%.

7.4.2 Myocardial Imaging

Several investigators have proposed to compute recovery coefficients and spillover factors from geometric measurement directly from the image[91] or derived from phantom studies.[92] In both cases, count recovery and spillover factors are derived by convolving the *PSF* in one dimension with an analytical model of the heart, comprised of a centre circle (blood pool or radius R) surrounded by an annulus (myocardium or thickness d), both of uniform activity. In this case, the *PSF* being considered as spatially invariant, the recovery coefficient for the blood pool corresponds to that of a 2-D cylindrical object or diameter $2R$, and can be expressed as (see Eq. 22):

$$F_{BB} = 1 - e^{-R_2/2s^2} \tag{39}$$

Similarly, the contribution of the blood pool to the myocardium F_{BM}, the self recovery of muscle tissue F_{MM}, and the contribution of the myocardial tissue to the blood compartment F_{MB} can be computed using the convolution integrals over the myocardium/ventricle model. They can be derived using formulas recalled in section 5. The original work of Henze and colleagues[91] proposed to extract the myocardium thickness and left ventricle diameter directly from PET using specific markers of the 2 regions. Herrero *et al.*, derived computed recovery and spillover factors using the same analytical procedure, but computed the dimensions of the left ventricle and myocardium wall from the dimensions of a standard heart phantom[92] (Data Spectrum, Hillsborough, NC).

Based on the assumption that partial volume errors are the same for transmission and emission images, it was demonstrated that the activity per gram of extravascular tissue can be estimated by dividing the perfusion regional data by extravascular density for the same region.[36] It is worth emphasizing that no convincing evidence was found of thickness above the partial volume limit in a large sample of 75 normotensive and 25 hypertensive patients.[93] Therefore it is likely that relations between myocardial count increases and wall thickening are similar throughout the cardiac cycle, even in patients with left ventricular hypertrophy. Using PET and $^{13}NH_3$-ammonia for the quantification of myocardial blood flow, Nuyts *et al.*[94] have obtained recovery coefficients of 59% for the myocardial wall and 86% for the blood pool in animal experiments. In addition, spillover from the blood pool into the myocardial was ~14%.

More recently, some authors have proposed to use correlated anatomical images to derive correction maps for partial volume in myocardial perfusion

studies. For instance, the group from UCSF has used CT images to define regional myocardial attenuation and count recovery maps of the porcine heart.[27] This method is similar to the pixel-by-pixel approach used for simple atrophy correction in brain PET,[43,95] although it does not provide corrected maps of the myocardium perfusion, but rather an average regional value.

7.4.3 Oncology Imaging

Despite the widespread applications of nuclear imaging (especially PET) in oncology, only a limited number of studies investigated the PVE problem in tumour imaging. The effects of object shape, size and background on SPECT ^{131}I activity quantification without detector response compensation has been reported in a detailed Monte Carlo study.[96] The activity quantification was carried out using a constant calibration factor and no PVC. It has been shown that the bias increases significantly with decreasing lesion size as a result of the increased spill-out of counts. More importantly, the bias for cylindrical lesions is consistently higher than for spherical lesions because spill-out is more significant for non-spherical objects. The bias also depends significantly on tumour-to-background ratio (TBR) because of the spill-in effect.

The simplest approach relies on the use of pre-calculated recovery coefficients for more reliable estimate of the standardized uptake value in pulmonary lesions.[97,98] The bias affecting TBR estimates owing to PVE is lesion size dependent. The generally accepted criterion is that PVC is required if the lesion size is less than 2-3 times the spatial resolution (FWHM) of the imaging system when the parameter of interest is the maximum voxel value within a particular VoI. In fact, Soret *et al.*[99] have demonstrated that when the parameter of interest is the average count density, the bias introduced by the PVE could exceed 10% even for lesions $\sim$6 times the FWHM depending on the true TBR.

Notwithstanding the known limitations, small animal imaging with a clinical dual-modality PET/CT scanner has been shown to be feasible for oncological imaging where the high resolution CT could be used for more precise localization of PET findings in addition to PVC through size-dependent recovery coefficient correction.[100] An aspect which deserves further attention is that robust models have also been developed to correction for partial volume effect in helical CT.[101]

8. Summary

Partial volume correction of emission tomography data remains a priority for accurate image quantitation. The ability to compensate for partial volume effects usually requires to: (1) characterize the point-spread function of the imaging system, (2) characterize the tissue components that participate in

the uptake and metabolism of the tracer, and (3) characterize the resolution effects in terms of correction factors or maps.

The imaging system *PSF* can be accurately measured by phantom experiments and is easily modelled by simple mathematical functions such as Gaussian or a series of Gaussian functions depending on whether the scatter component is included at that level or not. With the growing availability of multimodality integration and processing software, additional anatomical information provided by MR or CT imaging can be used to segment the distribution of the tracer into functionally meaningful regions. By studying the interaction of the system's *PSF* with a model of the tracer distribution, the contribution of each distinct region to the emission image can be computed. These interactions can be modelled using sophisticated Monte Carlo simulation, analytical projectors, or a simple convolution with a defined resolution kernel in image space.[102]

In practice self-recovery (*RC*) and spillover (or "spill-in") factors can be easily extracted and processed through any given set of user-defined *a priori* or *a posteriori* RoIs,[23,24,48] or *RC* maps and spill-in contributions can be derived to create images of the corrected radioactivity distribution in cortical grey matter.[22,25] With increased image resolution, the limit in size of regions that PVC can be applied to will constantly be pushed further, and it will become increasingly important to be able to account for cross-contamination of activity between small adjacent regions for which the assumptions made during the application of the pixel-based method[22] will be largely violated. The GTM approach[24] possesses the formidable advantage of not requiring any assumption of any tracer level at any time. The extraction of the GTM matrix is a one-time process, and does not require any additional computation for dynamic data, i.e., the same inverse of the GTM matrix is applied to all the data points since it expresses geometric interaction between the imaging system and the tissue component, independent of the contrast present in the image. Comparative assessment of PVC strategies is an important part of the validation step.[52,58,103]

Those methods proved to be sufficiently accurate to be applied in a growing number of research studies, if one considers that the number of publications related to the effect of the application of PVC algorithms to research data has grown significantly in the past few years. PVC is now a powerful and reliable tool that should become systematically used in research or clinical studies involving the use of emission tomography. It is expected that improvement in all aspects of the prerequisite for accurate partial volume correction are still required, especially for what concerns the quality of anatomo-functional mapping needed for accurate quantitation of cell-specific function and metabolism.

Acknowledgments. OR was supported by the National Institute of Neurological Disorders and Stroke (NINDS) under grant number

R01NS38927 and by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) under grant number R01AA12839. HZ was supported by the Swiss National Science Foundation under grant SNSF 3152A0-102143.

References

1. Zaidi H. and Sossi, V., Correction for image degrading factors is essential for accurate quantification of brain function using PET. *Med Phys* **31:** 423-426 (2004).
2. Links J. M., Lesion detectability and quantitative recovery with FDG-PET. *Eur J Nucl Med* **26:** 681-682 (1999).
3. Fazio F. and Perani, D., Importance of partial-volume correction in brain PET studies. *J Nucl Med* **41:** 1849-1850 (2000).
4. Brooks R. A. and Di Chiro, G., Slice geometry in computer assisted tomography. *J Comput Assist Tomogr* **1:** 191-199 (1977).
5. Gautschi W., "Error functions and Fresnel integrals." in: *Handbook of Mathematical Functions*, edited by Abramowitz M and Segun I A Dover, New-York, (1968), *pp* 295-330.
6. Abramowitz M. and Segun, I. A., *Handbook of Mathematical Functions*, Dover publications, New-York, (1965).
7. Kessler R. M., Ellis, J. R. and Eden, M., Analysis of emission tomographic scan data: limitations imposed by resolution and background. *J Comput Assist Tomogr* **8:** 514-522 (1984).
8. Hoffman E. J., Huang, S. C., Plummer, D. *et al*., Quantitation in positron emission computed tomography: 6. effect of nonuniform resolution. *J Comput Assist Tomogr* **6:** 987-999 (1982).
9. Hoffman E. J., Huang, S. C. and Phelps, M. E., Quantitation in positron emission computed tomography: 1. Effect of object size. *J Comput Assist Tomogr* **3:** 299-308 (1979).
10. Mazziotta J. C., Phelps, M. E., Plummer, D. *et al*., Quantitation in positron emission computed tomography: 5. Physical–anatomical effects. *J Comput Assist Tomogr* **5:** 734-743 (1981).
11. Herscovitch P., Auchus, A. P., Gado, M. *et al*., Correction of positron emission tomography data for cerebral atrophy. *J Cereb Blood Flow Metab* **6:** 120-124 (1986).
12. Wahl L. M., Asselin, M. C. and Nahmias, C., Regions of interest in the venous sinuses as input functions for quantitative PET. *J Nucl Med* **40:** 1666-1675 (1999).
13. Huesman R. H., A new fast algorithm for the evaluation of regions of interest and statistical uncertainty in computed tomography. *Phys Med Biol* **29:** 543-552 (1984).
14. Tsui B. M., Hu, H., Gilland, D. *et al*., Implementation of simulataneous attenuation and detector response in SPECT. *IEEE Trans Nucl Sci* **35:** 778-783 (1988).
15. Zeng G. L., Gullberg, G. T., Bai, C. *et al*., Iterative reconstruction of fluorine-18 SPECT using geometric point response correction. *J Nucl Med* **39:** 124-130 (1998).

16. King M. A. and Miller, T. R., Use of a nonstationary temporal Wiener filter in nuclear medicine. *Eur J Nucl Med* **10:** 458-461 (1985).
17. Links J. M., Jeremy, R. W., Dyer, S. M. *et al*., Wiener filtering improves quantification of regional myocardial perfusion with thallium-201 SPECT. *J Nucl Med* **31:** 1230-1236 (1990).
18. Links J. M., Becker, L. C., Rigo, P. *et al*., Combined corrections for attenuation, depth-dependent blur, and motion in cardiac SPECT: a multicenter trial. *J Nucl Cardiol* **7:** 414-425 (2000).
19. Iida H., Kanno, I., Takahashi, A. *et al*., Measurement of absolute myocardial blood flow with H215O and dynamic positron-emission tomography. Strategy for quantification in relation to the partial-volume effect. *Circulation* **78:** 104-115 (1988).
20. Iida H., Law, I., Pakkenberg, B. *et al*., Quantitation of regional cerebral blood flow corrected for partial volume effect using O-15 water and PET: I. Theory, error analysis, and stereologic comparison. *J Cereb Blood Flow Metab* **20:** 1237-1251 (2000).
21. Chawluk J., Alavi, A., R, D. *et al*., Positron emission tomography in aging and dementia: effect of cerebral atrophy. *J Nucl Med* **28:** 431-437 (1987).
22. Muller-Gartner H. W., Links, J. M., Prince, J. L. *et al*., Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects. *J Cereb Blood Flow Metab* **12:** 571-583 (1992).
23. Rousset O., Ma, Y., Kamber, M. *et al*., 3D simulations of radiotracer uptake in deep nuclei of human brain. *Comput Med Imaging Graph* **17:** 373-379 (1993).
24. Rousset O. G., Ma, Y. and Evans, A. C., Correction for partial volume effects in PET: principle and validation. *J Nucl Med* **39:** 904-911 (1998).
25. Meltzer C. C., Kinahan, P. E., Greer, P. J. *et al*., Comparative evaluation of MR-based partial-volume correction schemes for PET. *J Nucl Med* **40:** 2053-2065 (1999).
26. Labbe C., Froment, J. C., Kennedy, A. *et al*., Positron emission tomography metabolic data corrected for cortical atrophy using magnetic resonance imaging. *Alzheimer Dis Assoc Disord* **10:** 141-170 (1996).
27. Da Silva A. J., Tang, H. R., Wong, K. H. *et al*., Absolute quantification of regional myocardial uptake of 99mTc-sestamibi with SPECT: experimental validation in a porcine model. *J Nucl Med* **42:** 772-779 (2001).
28. Chen C. H., Muzic, R. F., Nelson, A. D. *et al*., A nonlinear spatially variant object-dependent system model for prediction of partial volume effects and scatter in PET. *IEEE Trans Med Imaging* **17:** 214-227 (1998).
29. Brix G., Bellemann, M. E., Hauser, H. *et al*., [Recovery coefficients for the quantification of the arterial input functions from dynamic PET measurements: experimental and theoretical determination]. *Nuklearmedizin* **41:** 184-190 (2002).
30. Geworski L., Knoop, B. O., de Cabrejas, M. L. *et al*., Recovery correction for quantitation in emission tomography: a feasibility study. *Eur J Nucl Med* **27:** 161-169 (2000).
31. Blankespoor S. C., Xu, X., Kaiki, K. *et al*., Attenuation correction of SPECT using X-ray CT on an emission-transmission CT system: myocardial perfusion assessment. *IEEE Trans Nucl Sci* **43:** 2263-2274 (1996).
32. Sossi V., Buckley, K. R., Snow, B. J. *et al*., Recovery of the human striatal signal in a slice oriented positron emission tomograph. *J Nucl Med* **34:** 481-487 (1993).

33. Shepp L. A. and Vardi, Y., Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imag* **1:** 113-122 (1982).
34. Hutton B. F. and Lau, Y. H., Application of distance-dependent resolution compensation and post-reconstruction filtering for myocardial SPECT. *Phys Med Biol* **43:** 1679-1693 (1998).
35. Pretorius P. H., King, M. A., Pan, T. S. *et al*., Reducing the influence of the partial volume effect on SPECT activity quantitation with 3D modelling of spatial resolution in iterative reconstruction. *Phys Med Biol* **43:** 407-420 (1998).
36. Hutton B. F. and Osiecki, A., Correction of partial volume effects in myocardial SPECT. *J Nucl Cardiol* **5:** 402-413 (1998).
37. Brix G., Doll, J., Bellemann, M. E. *et al*., Use of scanner characteristics in iterative image reconstruction for high-resolution positron emission tomography studies of small animals. *Eur J Nucl Med* **24:** 779-786 (1997).
38. Calvini P., Vitali, P., Nobili, F. *et al*., Enhancement of SPECT reconstructions by means of coregistered MR data. *IEEE Trans Nucl Sci* **48:** 750-755 (2001).
39. Kim H. J., Zeeberg, B. R. and Reba, R. C., Compensation for three-dimensional detector response, attenuation and scatter in SPECT grey matter imaging using an iterative reconstruction algorithm which incorporates a high-resolution anatomical image. *J Nucl Med* **33:** 1225-1234 (1992).
40. Ouyang X., Wong, W., Johnson, V. *et al*., Incorporation of correlated structural images in PET image-reconstruction. *IEEE Trans Med Imaging* **13:** 627-640 (1994).
41. Muzic R. F., Chen, C. H. and Nelson, A. D., A method to correct for scatter, spillover, and partial volume effects in region of interest analysis in PET. *IEEE Trans Med Imaging* **17:** 202-213 (1998).
42. Chen C. H., Muzic, R. F., Nelson, A. D. *et al*., Simultaneous recovery of size and radioactivity concentration of small spheroids with PET data. *J Nucl Med* **40:** 118-130 (1999).
43. Videen T. O., Perlmutter, J. S., Mintun, M. A. *et al*., Regional correction of positron emission tomography data for the effects of cerebral atrophy. *J Cereb Blood Flow Metab* **8:** 662-670 (1988).
44. Meltzer C. C., Cantwell, M. N., Greer, P. J. *et al*., Does cerebral blood flow decline in healthy aging? A PET study with partial-volume correction. *J Nucl Med* **41:** 1842-1848 (2000).
45. Matsuda H., Ohnishi, T., Asada, T. *et al*., Correction for partial-volume effects on brain perfusion SPECT in healthy men. *J Nucl Med* **44:** 1243-1252 (2003).
46. Meltzer C. C., Zubieta, J. K., Links, J. M. *et al*., MR-based correction of brain PET measurements for heterogeneous gray matter radioactivity distribution. *J Cereb Blood Flow Metab* **16:** 650-658 (1996).
47. Schoenahl F. and Zaidi, H., ''Towards optimal model-based partial volume effect correction in oncological PET.'' *Proc*. IEEE Nuclear Science Symposium and Medical Imaging Conference, Oct. 19-22, Rome, Italy, (2004) *in press*
48. Rousset O. G., Ma, Y., Marenco, S. *et al*., ''In vivo correction for partial volume effects in PET: accuracy and precision.'' in: *Quantification of brain function using PET*, edited by R. Myers, Cunningham, V., Bailey, DL, Jones, T Academic Press, San Diego, (1996), *pp* 158-165.
49. Spinks T., Guzzardi, R. and Bellina, C. R., Measurement of resolution and recovery in recent generation positron tomographs. *Eur J Nucl Med* **15:** 750-755 (1989).

50. Kuwert T., Ganslandt, T., Jansen, P. *et al.*, Influence of size of regions of interest on PET evaluation of caudate glucose consumption. *J Comput Assist Tomogr* **16:** 789-794 (1992).
51. Kuwert T., Morgenroth, C., Woesler, B. *et al.*, Influence of size of regions of interest on the measurement of uptake of 123I-alpha-methyl tyrosine by brain tumours. *Nucl Med Commun* **17:** 609-615 (1996).
52. Frouin V., Comtat, C., Reilhac, A. *et al.*, Correction of partial volume effect for PET striatal imaging: fast implementation and study of robustness. *J Nucl Med* **43:** 1715-1726 (2002).
53. Woods R. P., Mazziotta, J. C. and Cherry, S. R., MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* **17:** 536-546 (1993).
54. Collins D. L., Neelin, P., Peters, T. M. *et al.*, Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr* **18:** 192-205 (1994).
55. Maes F., Vandermeulen, D. and Suetens, P., Medical image registration using mutual information. *Proceedings of the IEEE* **91:** 1699-1722 (2003).
56. Robbins S., Evans, A., Collins, D. *et al.*, Tuning and comparing spatial normalization methods. *Lecture Notes in Computer Science* **2879:** 910- 917 (2003).
57. Strul D. and Bendriem, B., Robustness of anatomically guided pixel-by-pixel algorithms for partial volume effect correction in positron emission tomography. *J Cereb Blood Flow Metab* **19:** 547-559 (1999).
58. Quarantelli M., Berkouk, K., Prinster, A. *et al.*, Integrated software for the analysis of brain PET/SPECT studies with partial-volume-effect correction. *J Nucl Med* **45:** 192-201 (2004).
59. Slifstein M., Mawlawi, O. and Laruelle, M., "Partial volume effect correction: Methodological consideration." in: *Physiological Imaging of the Brain with PET.*, edited by SB Hansen A. Gjedde, GM Knudsen, OB Paulson Academic Press, San Diego, (2001), *pp* 67-75.
60. Zaidi H., Ruest T., Schoenahl F., Montandon M- L., Comparative assessment of brain MR image segmentation algorithms and their impact on partial volume correction in PET. (2004) *submitted*
61. Aston J. A., Cunningham, V. J., Asselin, M. C. *et al.*, Positron emission tomography partial volume correction: estimation and algorithms. *J Cereb Blood Flow Metab* **22:** 1019-1034 (2002).
62. Evans A., Marrett, S., Neelin, P. *et al.*, Anatomical mapping of functional activation in stereotactic coordinate space. *Neuroimage* **1:** 43-53 (1992).
63. Rousset O. G., Deep, P., Kuwabara, H. *et al.*, Effect of partial volume correction on estimates of the influx and cerebral metabolism of 6-[(18)F]fluoro-L-dopa studied with PET in normal control and Parkinson's disease subjects. *Synapse* **37:** 81-89 (2000).
64. Soret M., Koulibaly, P. M., Darcourt, J. *et al.*, Quantitative accuracy of dopaminergic neurotransmission imaging with 123I SPECT. *J Nucl Med* **44:** 1184-1193 (2003).
65. Mullani N. A., A phantom for quantitation of partial volume effects in ECT. *IEEE Trans Nucl Sci* **36:** 983-987 (1989).
66. Phelps M. E., Huang, S. C., Hoffman, E. J. *et al.*, An analysis of signal amplification using small detectors in positron emission tomography. *J Comput Assist Tomogr* **6:** 551-565 (1982).

67. Hoffman E., "Positron Emission Tomography: principles and quantitation." in: *Positron Emission Tomography and Autoradiography*, edited by Mazziotta JC and Schelbert HR Phelps ME Raven Press, (1986), *pp* 237-286.
68. Hoffman E. J., Cutler, P. D., Digby, W. M. *et al.*, 3-D phantom to simulate cerebral blood flow and metabolic images for PET. *IEEE Trans Nucl Sci* **37:** 616-620 (1990).
69. Zubal I. G., Harrell, C. R., Smith, E. O. *et al.*, Computerized 3-dimensional segmented human anatomy. *Med Phys* **21:** 299-302 (1994).
70. Holmes C. J., Hoge, R., Collins, L. *et al.*, Enhancement of MR images using registration for signal averaging. *J Comput Assist Tomogr* **22:** 324-333 (1998).
71. Collins D., Holmes, C., Peters, T. *et al.*, Automatic 3-D model-based neuroanatomical segmentation. *Human Brain Mapping* **3:** 190-208 (1995).
72. Shan Z. Y., Yue, G. H. and Liu, J. Z., Automated histogram-based brain segmentation in T1-weighted three-dimensional magnetic resonance head images. *Neuroimage* **17:** 1587-1598 (2002).
73. Wong D. F., Harris, J. C., Naidu, S. *et al.*, Dopamine transporters are markedly reduced in Lesch-Nyhan disease in vivo. *Proc Natl Acad Sci U S A* **93:** 5539-5543 (1996).
74. Rousset O. G., Ma, Y., Wong, D. F. *et al.*, "Pixel-versus region-based partial volume correction in PET." in: *Quantitative Functional Imaging with Positron Emission Tomography*., edited by RE Carson, Daube-Witherspoon, ME, Herscovitch, P Academic Press, San Diego, (1998), *pp* 67-75.
75. Alfano B., Brunetti, A., Prinster, A. *et al.*, "STEPBRAIN: a stereolithographed phantom of the brain for nuclear medicine: computed tomography, and magnetic resonance applications. (abstract)" Annual meeting of the Radiological Society of North America, Chicago, USA, (2003).
76. Szabo Z., Links, J. M., Seki, C. *et al.*, Scatter, spatial resolution, and quantitative recovery in high resolution SPECT. *J Comput Assist Tomogr* **16:** 461-467 (1992).
77. Zaidi H., Montandon, M.-L. and Slosman, D. O., Magnetic resonance imaging-guided attenuation and scatter corrections in three-dimensional brain positron emission tomography. *Med Phys* **30:** 937-948 (2003).
78. Law I., Iida, H., Holm, S. *et al.*, Quantitation of regional cerebral blood flow corrected for partial volume effect using O-15 water and PET: II. Normal values and gray matter blood flow response to visual activation. *J Cereb Blood Flow Metab* **20:** 1252-1263 (2000).
79. Braem A., Chamizo Llatas, M., Chesi, E. *et al.*, Feasibility of a novel design of high-resolution parallax-free Compton enhanced PET scanner dedicated to brain research. *Phys Med Biol* **49:** 2547-2562 (2004).
80. Van Laere K. J. and Dierckx, R. A., Brain perfusion SPECT: age- and sex-related effects correlated with voxel-based morphometric findings in healthy adults. *Radiology* **221:** 810-817 (2001).
81. Mignotte M., Meunier, J., Soucy, J.-P. *et al.*, Comparison of deconvolution techniques using a distribution mixture parameter estimation : application in SPECT imagery. *J Electron Imag* **11:** 11-25 (2002).
82. Ishii K., Willoch, F., Minoshima, S. *et al.*, Statistical brain mapping of 18F-FDG PET in Alzheimer's disease: validation of anatomic standardization for atrophied brains. *J Nucl Med* **42:** 548-557 (2001).

83. Ibanez V., Pietrini, P., Alexander, G. E. *et al*., Regional glucose metabolic abnormalities are not the result of atrophy in Alzheimer's disease. *Neurology* **50:** 1585-1593 (1998).
84. Knowlton R. C., Laxer, K. D., Klein, G. *et al*., In vivo hippocampal glucose metabolism in mesial temporal lobe epilepsy. *Neurology* **57:** 1184-1190 (2001).
85. Meltzer C. C., Zubieta, J. K., Brandt, J. *et al*., Regional hypometabolism in Alzheimer's disease as measured by positron emission tomography after correction for effects of partial volume averaging. *Neurology* **47:** 454-461 (1996).
86. Melamed E., Lavy, S., Bentin, S. *et al*., Reduction in regional cerebral blood flow during normal aging in man. *Stroke* **11:** 31-35 (1980).
87. Martin A. J., Friston, K. J., Colebatch, J. G. *et al*., Decreases in regional cerebral blood flow with normal aging. *J Cereb Blood Flow Metab* **11:** 684-689 (1991).
88. Matsuda H., Kanetaka, H., Ohnishi, T. *et al*., Brain SPET abnormalities in Alzheimer's disease before and after atrophy correction. *Eur J Nucl Med Mol Imaging* **29:** 1502-1505 (2002).
89. Weckesser M., Hufnagel, A., Ziemons, K. *et al*., Effect of partial volume correction on muscarinic cholinergic receptor imaging with single-photon emission tomography in patients with temporal lobe epilepsy. *Eur J Nucl Med* **24:** 1156-1161 (1997).
90. Koole M., Laere, K. V., de Walle, R. V. *et al*., MRI guided segmentation and quantification of SPECT images of the basal ganglia: a phantom study. *Comput Med Imaging Graph* **25:** 165-172 (2001).
91. Henze E., Huang, S. C., Ratib, O. *et al*., Measurements of regional tissue and blood-pool radiotracer concentrations from serial tomographic images of the heart. *J Nucl Med* **24:** 987-996 (1983).
92. Herrero P., Markham, J. and Bergmann, S. R., Quantitation of myocardial blood flow with H2 15O and positron emission tomography: assessment and error analysis of a mathematical approach. *J Comput Assist Tomogr* **13:** 862-873 (1989).
93. Nichols K., DePuey, E. G., Friedman, M. I. *et al*., Do patient data ever exceed the partial volume limit in gated SPECT studies? *J Nucl Cardiol* **5:** 484-490 (1998).
94. Nuyts H., Maes, A., Vrolix, M. *et al*., Three-dimensional correction for spillover and recovery of myocardial PET images. *J Nucl Med* **37:** 767-774 (1996).
95. Meltzer C. C., Leal, J. P., Mayberg, H. S. *et al*., Correction of PET data for partial volume effects in human cerebral cortex by MR imaging. *J Comput Assist Tomogr* **14:** 561-570 (1990).
96. Dewaraja Y. K., Ljungberg, M. and Koral, K. F., Monte Carlo evaluation of object shape effects in iodine-131 SPET tumor activity quantification. *Eur J Nucl Med* **28:** 900-906 (2001).
97. Vesselle H., Schmidt, R. A., Pugsley, J. M. *et al*., Lung cancer proliferation correlates with [F-18]Fluorodeoxyglucose uptake by positron emission tomography. *Clin Cancer Res* **6:** 3837-3844 (2000).
98. Menda Y., Bushnell, D. L., Madsen, M. T. *et al*., Evaluation of various corrections to the standardized uptake value for diagnosis of pulmonary malignancy. *Nucl Med Commun* **22:** 1077-1081 (2001).
99. Soret M., Riddell, C., Hapdey, S. *et al*., Biases affecting the measurements of tumor-to-background activity ratio in PET. *IEEE Trans Nucl Sci* **49:** 2112-2118 (2002).

100. Tatsumi M., Nakamoto, Y., Traughber, B. *et al*., Initial experience in small animal tumor imaging with a clinical positron emission tomography/computed tomography scanner using 2-[F-18]Fluoro-2-deoxy-D-glucose. *Cancer Res* **63:** 6252-6257 (2003).
101. Hsieh J., Nonlinear partial volume artifact correction in helical CT. *IEEE Trans Nucl Sci* **46:** 743-747 (1999).
102. Zaidi H., Relevance of accurate Monte Carlo modeling in nuclear medical imaging. *Med Phys* **26:** 574-608 (1999).
103. Koole M., Van de Walle, R., Van Laere, K. *et al*., Study of the quantification of FBP SPECT images with a correction for partial volume effects. *IEEE Trans Nucl Sci* **49:** 69-73 (2002).

9
Image Registration Techniques in Nuclear Medicine Imaging

B.F. Hutton*, M. Braun[§] and P. Slomka[†]

1. Introduction

Nuclear medicine has a long tradition of incorporating quantitative analysis in its diagnostic procedures. Until recently, the analysis was based on radionuclide images as the sole input although the importance of the complementary information available from other modalities or from earlier scans has long been recognized. Indeed, qualitative correlation between images, based on anatomical expertise, has always been part of the repertoire of the nuclear medicine clinician. However, spatial distortion between images, caused largely by differences in posture and acquisition technique, prevented the extension of these techniques to quantitative analysis. Recent advances in image registration software and hardware have made it increasingly possible to utilize that complementary image information in a clinical setting. Reporting stations now often provide access to data from multiple modalities and permit various forms of fused display for viewing spatially registered images. Recently introduced dual-modality systems (described in Chapter 2 of this volume) provide a hardware approach to spatial registration that is of particular value in those cases where numerical algorithms fail. However, software approaches remain essential to the solution of many registration problems in clinical practice. In this chapter, we define the registration problem, introduce various algorithmic approaches to registering images, and discuss issues of clinical application with emphasis on nuclear medicine. A more extensive coverage of image registration is

*Prof. B.F. Hutton, Institute of Nuclear Medicine, University College London, London W1E 6BT, UK; also Centre for Medical Radiation Physics, University of Wollongong, NSW 2522, Australia
[§]Dr M. Braun, Department of Applied Physics, University of Technology, Sydney, NSW 2007, Australia
[†]Dr P. Slomka, Department of Imaging, Cedars Sinai Medical Center and the School of Medicine, University of California, Los Angeles, CA, USA

provided by Hajnal *et al.*[1] There are also useful reviews covering various aspects of medical image registration.[2-16]

2. Registration Problem

The precise form of the registration problem depends on the application. Figure 1 attempts to define the registration problem in a general form that takes account of various applications. The starting point is an underlying baseline anatomy, containing all anatomical structures that can be imaged. The baseline anatomy is exemplified by an atlas representation. The end-point is a pair of images. The process that links the starting point to each of the final images is complex. A number of the stages of this process introduce spatial distortions. Other stages have a corrupting influence on the image data. The registration problem can be posed as the search for a relative deformation between image domains given the two images.

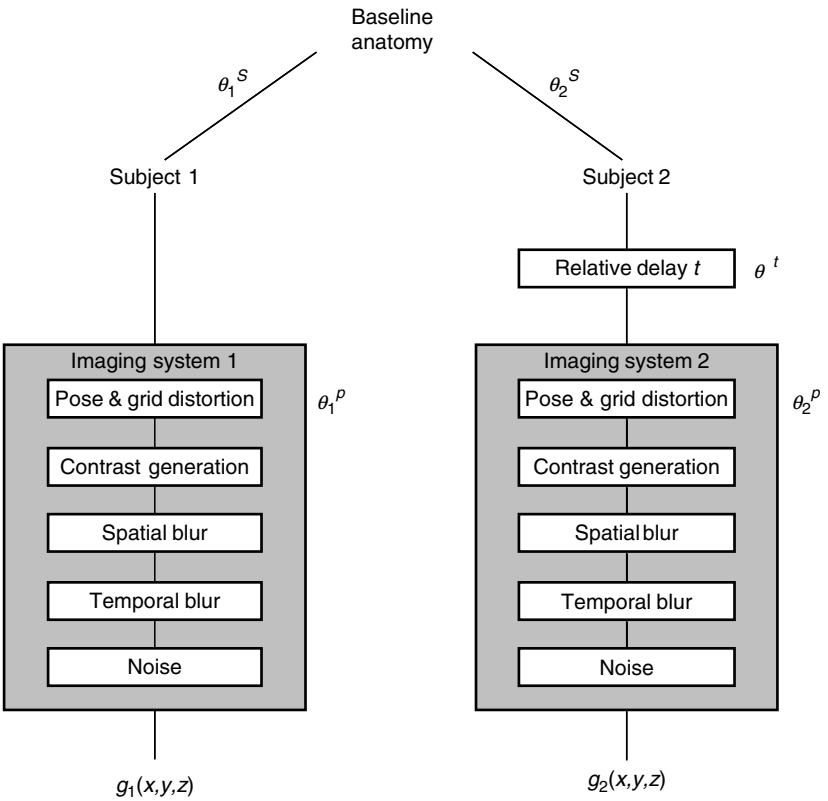


FIGURE 1. A generic representation of the registration problem.

2.1 Contributions to Deformation

The anatomy of each individual subject is treated as an instantiation of the baseline anatomy. Individual variation in size, shape and position of anatomical structures from the baseline is captured by deformation θ^s. In order to preserve the generality of the description, the deformation will be represented by a displacement field, in which a displacement vector is associated with every point in the spatial domain. In inter-subject registration, the aim is to capture deformation $\theta^s = \theta_2^s - \theta_1^s$. The implicit assumption that the entire anatomical structure of one subject can be obtained by a spatial mapping from another subject is not always valid. For example, the presence of gender-specific organs will clearly violate the assumption. Other normal anatomical variations (e.g. the variable number of brain sulci) may also present difficulties.

If $\theta_2^s = \theta_1^s$, the registration becomes an intra-subject problem, and a significant source of anatomical deformation is the relative delay t between the times the two images were acquired. Depending on the time scale of the delay, different morphometric changes may be evident. For example, on a relatively short time scale, bladder filling will affect not merely the size, shape and position of the bladder but also of nearby organs. On a longer scale, normal developmental changes in children, ageing processes, and changes associated with reproductive cycle, may contribute to the deformation θ^t. In a clinical setting, principal sources of deformation are the pathological process (e.g. tumours, oedema) and treatment (e.g. surgical excision).

In both inter- and intra-subject applications, registration aims to bridge the morphometric differences so that images can be compared on a common spatial grid. However, where the primary focus is on elucidating rather than bypassing the morphometric differences, the deformation map found by the registration process can provide the basis for quantifying these differences.

As the anatomical information is transformed into an image, a further deformation θ^p takes place. There are two principal contributions to the deformation that occur in the imaging process. The subject's pose (e.g. prone/supine, arms up/down, breath hold) will depend on the imaging modality and the type of examination. Each variation can give rise to significant and complex deformations. In addition, even where the same type of pose is assumed, replication of the pose in fine detail cannot be assumed (except where extrinsic locating devices are used, such as a stereotactic frame in the imaging of the brain.[17,18] The other contribution comes from the distortion of the spatial grid intrinsic to the imaging system. For example, in magnetic resonance imaging (MRI), the spatial grid is distorted by the non-uniformity of the steady magnetic field and gradient nonlinearity.[19]

The registration problem therefore amounts to finding the deformation $\Delta\theta = \Delta\theta^s + \Delta\theta^t + \Delta\theta^p$. In inter-subject studies, the effect of a delay between

image acquisitions can be built into the inter-subject deformation, and we are left with $\Delta\theta = \Delta\theta^s + \Delta\theta^p$. In intra-subject studies, the deformation becomes $\Delta\theta = \Delta\theta^t + \Delta\theta^p$. In studies on the same subject carried out in relatively quick succession (e.g. in dual systems described in Chapter 2 of this volume) or where no significant deformations occur between image acquisitions, the deformation becomes $\Delta\theta = \Delta\theta^p$. In dual-modality systems ($\Delta\theta$ is usually assumed to be zero.

2.2 Corrupting Influences

Because the registration algorithm acts on final images, it is subject to the corrupting influences of the imaging system. Figure 1 depicts the effect of an imaging system in a simplified form. The heart of the imaging process is the mapping of tissue-specific properties to regions of the image space (referred to as contrast generation in Figure 1). In x-ray computed tomography (CT), the property mapped is related to x-ray attenuation. In MRI, it is usually free proton density modulated by spin-lattice and spin-spin relaxation, and in nuclear medicine, it is the uptake of the radiopharmaceutical. Contrast is generated where the property values are significantly different. Because of this fundamental difference between image modalities, image intensity in one modality is not necessarily equal to that in another modality and we cannot achieve registration by simply matching intensity values.

Consider a multi-modality registration of single photon emission computed tomography (SPECT) and CT of the brain. With the appropriate choice of a radiopharmaceutical (e.g. ^{99m}Tc labelled HMPAO) the SPECT image will show a strong contrast between grey matter and white matter. However, x-ray attenuation of the two tissues is so similar that there is no appreciable contrast between them in the CT image. Thus registration of grey matter in SPECT with grey matter in CT will not be possible, although registration of the two brain images may still succeed by relying on the registration of other contrasting tissues. On the other hand, radionuclide images generally manifest poor anatomical contrast, with few landmarks available for registration purposes.

The finite resolution of each imaging system imposes a spatial blur on image data. This can have a significant effect on the displacement field, as illustrated in Figure 2. In (a), the fine-scale displacement field is uniform (characteristic of rigid-body translation). It is likely that, following blurring in which the 2×2 image block is replaced by one pixel, the displacement will remain largely unchanged. On the other hand, a displacement field characterized by rapid change, such as the radial expansion in Figure 2(b), is likely to be severely affected by spatial blurring. Indeed, since the resulting pixel can only have a single displacement vector, that vector cannot capture accurately deformations other than uniform translation. The direction and magnitude of the vector will be determined by the details of the image and the registration algorithm. Figure 2(c) illustrates this point further. In order

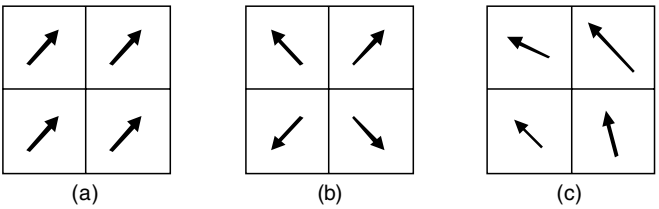


FIGURE 2. Examples of fine scale displacement fields used in considering the effect of spatial blur: (a) uniform translation, (b) radial expansion, and (c) counter-clockwise rotation.

to represent a uniform rotation exactly, the displacement field would ideally be defined with infinitesimally small pixels. The poorer the resolution, the coarser is the approximation to uniform rotation.

In inter-modality and intra-modality applications, where images are acquired with different resolution, the undesirable effects on the displacement field may be reduced by blurring the higher resolution image to match the resolution of the other image although noise reduction is also an important consideration.[20] In practice, a number of registration algorithms adopt a multi-scale approach, which allows for controlled blurring of both images. Note, however, that multi-scale properties of the displacement field are less well understood than those of the image itself. Resolution effects are further complicated by the fact that imaging systems are, at best, only approximately shift invariant and the resolution varies from one part of the imaged volume to another. An example where the approximation breaks down is the position-dependent resolution in SPECT. Where possible, such effects should be corrected prior to the application of the registration algorithm.

Temporal blur occurs where the duration of image acquisition is of the order of, or longer than, the time constant of any movement of tissues. In modalities such as CT and fast MRI, acquisition can be completed with breath hold, so that only cardiac motion contributes significantly to data blurring (non-blur artefacts may also occur). In slower acquisition modalities (e.g. SPECT), respiratory motion, peristaltic intestinal motion, and involuntary patient's motion, in addition to cardiac motion, can all contribute to temporal blur. The effect on the registration process depends on the application. In intra-modality problems, where the temporal effects in both images are similar, the principal consequence is that the desired deformation is defined with greater uncertainty, leading to a possible diminution of registration accuracy. In inter-modality applications, the temporal blurring may be very different in the two images, impeding the registration process. Unlike the case of spatial blur, we cannot bring the two images to a common temporal scale with a uniform spatial blur.

Noise can have a significant effect in that the random fluctuation of intensity values may confuse the registration algorithm. Nuclear medicine images are generally noisier than those obtained with other modalities and this can be

a significant factor in registration problems. An additional complication is that the noise level in those images depends on local mean intensity.

Thus the aim of the registration algorithm is to determine the relative deformation between image domains given images $g_1(x,y,z)$ and $g_2(x,y,z)$, each subject to the corrupting influences of contrast generation, spatial and temporal blur, as well as noise.

3. Generic Registration Algorithm

Almost all registration algorithms follow the iterative scheme shown in Figure 3. One can distinguish between the deformable floating image and the fixed reference image. The floating image is first deformed using some initial values of the deformation parameters. The core of the algorithm is the evaluation of the similarity between the deformed floating image and the reference image. An iterative optimization algorithm is employed to find the deformation that maximizes similarity. The output is the optimum deformation defined by a set of parameters. The similarity measure may be evaluated for a part of the image at a time, instead of the entire image.

4. Types of Deformation

So far, the deformation θ was represented by a displacement vector field, where deformation at each point r in the image is given by $\Delta \boldsymbol{r} = (\Delta x, \Delta y, \Delta z)^T$. There are two major issues arising from this representation:

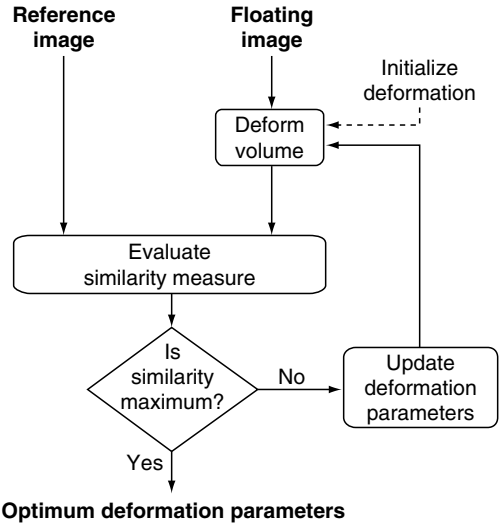


FIGURE 3. A flowchart for image registration.

computational cost and lack of local coherence in the vector field. For an image of size N^3, finding the deformation requires solving N^3 independent local registration problems, each yielding 3 parameters. This is computationally expensive. Moreover, it is reasonable to expect that a deformation of tissue at one point affects deformation at adjacent points. Several models have been proposed to express this correlation, including thin plate spline,[21] elastic body spline[22] and viscous flow.[23,24] Smoothing with, for example, a median filter,[25] also serves to constrain the displacement field based on adjacent values.

An extreme case of inter-voxel correlation is expressed by the affine transformation model,[26] which permits only global translations, rotations, scaling along each axis and shearing deformations. In this case, only one registration problem needs to be solved, yielding 12 parameters embedded in a transformation matrix. For each location (x, y, z), the transformed coordinates are given by

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \Delta x \\ a_{21} & a_{22} & a_{23} & \Delta y \\ a_{31} & a_{32} & a_{33} & \Delta z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{1}$$

where the translation vector is given by $[\Delta x, \Delta y, \Delta z]^T$, and the affine transformation is given by the 9 coefficients a_{ij}. The displacement vector for point (x, y, z) is $(x' - x, y' - y, z' - z)$.

The rigid-body model represents a somewhat more severe case of inter-voxel correlation. It is a special case of the affine model, with shearing transformations removed and scaling independently pre-determined. The registration algorithm needs to determine only 6 parameters (3 orthogonal rotations and 3 translations). Much of the early registration work was confined to intra-subject brain applications, where the confinement of densely packed tissues to the interior of the skull makes a rigid-body model usually an adequate approximation. The scaling coefficients would be obtained separately from the knowledge of acquisition parameters.

Apart from exploiting inter-voxel dependence, we can reduce the computational cost by solving for $\boldsymbol{\Delta r}$ at a subset of locations obtained by subsampling the image space. The registration algorithm will then need to be invoked $(N/p)^3$ times, where p is the subsampling factor in each dimension. The displacement at remaining locations can be obtained by interpolation. The interpolating function should be consistent with the adopted physical model.

In a computationally attractive coarse-to-fine approach,[27-29] the displacement field is estimated on a sparse grid first, then on progressively finer grids (Figure 4). This approach is also expected to make the registration algorithm more robust.

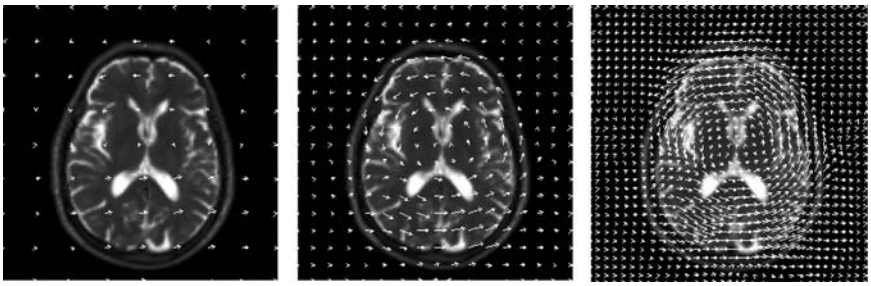


FIGURE 4. Three steps of a coarse-to-fine scheme. First (left), the coarsely sampled displacement field is estimated using a large image block. Then the displacement field is refined using successively smaller image blocks and finer sampling.

5. Similarity Measures

The criterion that drives the registration algorithm is known as the similarity measure (alternative formulations refer to the cost function or the objective function). The maximum of the similarity measure should correspond to the closest alignment between the two images being registered. Let $A(\mathbf{r})$ and $B(\mathbf{r})$ be two such images defined over respective image domains Ω_A and Ω_B, respectively. The registration process can be expressed as the search for the deformation θ that maximizes the similarity measure S between the reference image A and the deformed floating image B,

$$\theta_{\max} = \arg\max_{\theta} S(A(r), B(\theta_r)) \tag{2}$$

where θ_r constructs the deformed image domain. The similarity measure may be evaluated over subsets of image domains rather than the entire domains.

Many similarity measures have been proposed. Broadly, these can be categorized as intensity-based and feature-based measures. The feature-based measures are not applied to the original images. Instead, the images are segmented in some way so that certain spatial features can be extracted. These features may be points, lines, or surfaces. Their spatial characteristics, such as the position of a landmark point or the shape of a surface, are then compared in the similarity measure.

In the simplest case, an intensity-based measure makes a direct pixel-by-pixel comparison of intensity values. More commonly, some characteristics derived from the intensity data are compared. In some instances, the intensity characteristics are derived from spatially constrained regions. Such cases straddle the border between measures based on spatial features and those based on intensity but will be classified as intensity-based measures.

For a particular application, the choice of the similarity measure will be guided by the characteristics of the image modalities and the nature of the expected deformation.

5.1 Measures based on Spatial Features

One can categorize similarity measures according to the type of feature used to construct the measures, as follows.

5.1.1 Landmark Measures

In this type of registration, the similarity measure acts on a set of points extracted from each image. Each point gives the location of a landmark, a small feature (artificial or natural) that is clearly present in both images. The similarity measure would typically relate to the average distance between the corresponding landmarks in the two images. For example, the similarity measure can be defined as the (negative of the) average distance between corresponding landmarks,

$$S(A(r),B(\theta_r)) = -\frac{1}{L}\sum_{j=1}^{L}\left\| r_{A_j} - \theta r_{B_j} \right\| \tag{3}$$

where j indexes pairs of corresponding landmarks and $\|.\|$ denotes the (usually Euclidean) norm. Alternatively, the similarity measure is given in a least squares form:

$$S(A(r),B(\theta_r)) = -\sum_{j=1}^{L}\left\| r_{A_j} - \theta r_{B_j} \right\|^2 \tag{4}$$

If θ is a rigid body transformation, this is the so-called Procrustes metric, which is well suited to this subset of registration problems.[30]

One of the challenges of these algorithms is to establish the correct pairing of landmarks in the two images. Aside from manual identification, a common automatic approach is to find, for each landmark in one image, the closest landmark in the other image. Such approaches usually require that the misalignment be small. A version of this, with iterative reassignment of closest pairs, forms a part of the popular iterative closest point (ICP) algorithm[31] used in surface matching. Another factor influencing the success of landmark-based registration is the density of landmarks. Both the total number and the disposition of landmarks must be adequate to model accurately the deformation of the image domain. In practice, the number of landmarks is much smaller than the number of voxels, leading to a fast algorithm, which is an important advantage of landmark-based methods.

In accordance with their origin, landmarks can be classified as extrinsic, anatomical, or geometrical.

5.1.2 Extrinsic Markers

Fiducial markers, visible in the two images to be registered, are attached to the body. This is more difficult to accomplish where different modalities are involved but dual-purpose external markers that are visible in MRI or CT on

the one hand, and in positron emission tomography (PET) or SPECT on the other, have been widely used. Another requirement is that the markers be rigidly attached to the body. This may be difficult to accomplish depending on what is meant by the "body". If the markers are attached to the skin but the organs of interest are in the body interior, the markers may not manifest the deformation being sought. For example, because the skin can move relative to the skull, skin markers may not be adequate for detection of intracranial misalignment. In patients about to undergo neurosurgery, a rigid stereotactic frame is sometimes attached directly to the skull and offers a set of landmarks that accurately captures skull misalignments.[17,18] However, such an invasive procedure cannot be justified by the requirements of image registration alone.

Extrinsic markers play an important role in assessing the accuracy of registration methods.[32,33] Since such assessments depend on accurate localization of the markers, the design of the markers and the localization technique must minimize errors and those errors need to be quantified.[30,33,34] The principal disadvantage of extrinsic markers is that they cannot be used in retrospective studies. A potentially important limitation of extrinsic markers is that, being exterior to the body, they may not accurately capture deformations of the body interior. Additionally, the sparseness of the markers limits the registration to global rigid-body or affine transformations.

5.1.3 Anatomical Landmarks

Markers based on distinctive anatomical features avoid the retrospectivity and invasiveness of the extrinsic markers (e.g. Evans *et al.*[35]). However, identification of anatomical landmarks requires considerable operator interaction. Small sets of anatomical landmarks have been used to constrain an unrelated registration algorithm.[35] The lack of anatomical definition and low resolution of nuclear medicine images makes the use of anatomical landmarks problematic.

5.1.4 Geometrical Landmarks

Corners, local curvature extrema and other geometric features are intrinsic landmarks that can be identified automatically.[36] They appear to be particularly useful in serial studies of high resolution modalities. However, in nuclear medicine images, the low signal-to-noise ratio and the relatively poor resolution impedes the use of geometrical landmarks.

5.2 Boundary Measures

In manual registration, the user usually attempts to align corresponding boundary elements (surfaces or edges). The process includes identification of the boundary elements and establishing anatomical correspondence between corresponding elements in the two images, and finally the alignment.

Algorithms based on boundary measures emulate this process but with a high degree of automation. A principal challenge in boundary-based algorithms is the requirement to extract the appropriate surfaces or edges. The similarity measure quantifies an average ''distance'' between the corresponding surfaces. If surface extraction is carried out in just one of the images, the average distance is computed between the surface detected in one image and a set of points in the other image. If both surfaces are represented by sets of points, the iterative closest point (ICP) method provides a popular optimization algorithm for image registration.[31] The algorithm seeks a least squares minimum distance between data sets but allows the correspondence between points to be reassigned at each iteration based on the current closest point.

As was the case with landmark measures, the process of segmenting spatial features removes the dependence on voxel values, making these measures well suited to inter-modality registration.

There is a wide range of techniques for matching based on surface information, utilizing either characteristics of the surface or edge. These may be extensions of the geometrical landmarks referred to earlier, such as the crest lines,[37,38] surface points[39] or more general models (e.g. Gee *et al.*[40]). Until recently, one of the most popular algorithms for intra-subject registration of functional and anatomical images was that of Pelizzari and co-workers.[39] A small number of points was defined on the brain surface imaged with the lower resolution modality, treating this as a 'hat' to be fitted to the 'head', a set of contours defined by the higher resolution modality. Problems can occur in automating the segmentation step that defines the brain outline as opposed to the skull or scalp and in finding a unique fit, independent of head shape. Although designed specifically for neurological applications, similar approaches have been developed for matching surfaces of internal organs.[41,42] The segmentation step needs to be designed with the particular part of the anatomy in mind. Other algorithms rely indirectly on edge information (e.g. maximisation of zero-crossings in difference images[43]). However, most of these techniques rely on the images having high resolution so that edges are well defined. A detailed review of surface registration techniques can be found in ref.[12]

Since organ boundary descriptors define the shape of the organ, boundary measures can be considered as shape measures. More specific shape measures can be extracted from segmented regions. One example is the principal axes method.[44,45] The method requires a computation of moments up to order two in the two images to be registered. In the simplest (rigid-body) case, the centres of mass (counts) are determined and the images are translated so that the centres of mass coincide. The rotation is then determined from the orientation of the principal axes. The method can be extended easily to incorporate affine transformations. The algorithm needs to visit a large number of voxels but it is usually not iterative, so the method is reasonably fast. A drawback is that the principal axes measures are sensitive to differences in the field of view in the two images.

5.3 Intensity Measures

These measures make explicit use of the intensity values or the statistical parameters of their histogram. Removing the explicit dependence on the spatial features and the use of all voxels frees the intensity measures from restrictions on the allowable transformations. In principle, transformations with a large number of degrees of freedom could be sought. However, restrictions re-appear where the algorithm acts on a small subset of the image. The simplest similarity measures are those that directly compare intensity values between the two images, voxel by voxel. Such measures will only be suitable for intra-modality registration. When images arise from different modalities, with unrelated contrast mechanisms, such simple measures are clearly inadequate. We distinguish the following types of similarity measure.

5.3.1 Minimum Intensity Difference

The simplest of these voxel-by-voxel measures is the (negative of) the absolute or squared intensity differences.[9] The sum of squared differences,

$$S(A(r),B(\theta_r)) = -\sum_{j=1}^{L} \left\| r_j - B(\theta r_j) \right\|^2 \tag{5}$$

can be shown to be optimal if images differ only due to Gaussian noise.[46] However, it is sensitive to bias when some voxel pairs manifest large intensity differences. It is assumed that voxel values in registered images are similar, an assumption that is only true for similar studies in the same modality (e.g. paired studies in the same individual using the same radionuclide) (Figure 5). Nevertheless the sum of absolute differences has been demonstrated to be widely applicable, even when there are quite large differences between studies.[9]

5.3.2 Cross Correlation

Another common similarity measure is the cross-correlation. In this case, it is possible to relax the condition of the similarity of values so as to apply the technique for both intra- and inter-modality registration problems.[47-49]

5.3.3 Minimum Variance

The underlying assumption for minimum variance measures is that an image consists of largely homogenous (small variance) regions that differ substantially (large variance) from one another. If the two images are correctly registered, a homogeneous region in one image will correspond to a homogeneous region in the other image, even though the mean intensity values will be, in general, unrelated. This is the basis for some popular inter-modality implementations. Woods *et al.*[50] performed a simple

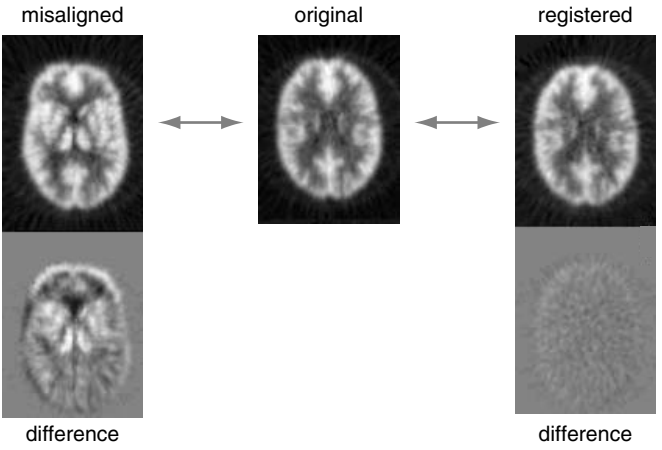


FIGURE 5. The absolute difference image is substantial when the alignment is poor (left) but is reduced as the alignment improves (right). The sum of the difference image can be used to guide the registration process.

intensity-based segmentation that did not account for any correlation among neighbouring voxels. Ardekani *et al.*[51] performed segmentation of a MRI or CT image using K-means clustering, followed by connected component analysis, to define a large set of ''homogeneous'' regions. This segmentation was then imposed (induced) on the second lower resolution image (SPECT or PET image). Both these approaches are well suited for registration of SPECT/ PET data with other modalities since segmentation need only be applied to the higher resolution anatomical image. Note that although the measure is extracted from intensity values, those values come from pre-segmented regions and, in this sense, straddle the boundary between intensity-based measures and those based on spatial features.

5.3.4 Mutual Information

Similarity measures arising from information theoretic approaches have recently become popular. These include minimization of histogram dispersion,[52] maximization of mutual information[53-56] or maximization of the correlation ratio.[25,57] These approaches are attractive as they are generally applicable to both intra- and inter-modality registration classes, with the underlying assumption being that there is a correlation between groups of voxels that have similar values. Although the techniques would appear to be less constrained than approaches where some spatial relationship between like-voxels is implied, they have been demonstrated to be widely applicable for matching both similar and dissimilar images. The basis for histogram based approaches is summarised in Figure 6. See also the useful discussion in the review by Hill *et al.*[14]

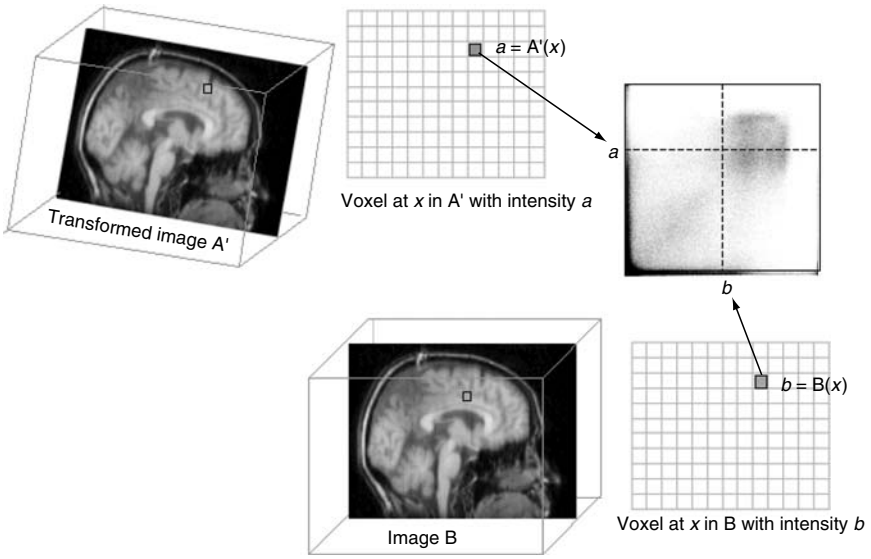


FIGURE 6. A joint intensity histogram is formed from the intensities in two images. If a voxel at x has intensity a in transformed image A′, and b in image B, this pair of voxels will contribute one count to the joint histogram at (a, b). The final value of the histogram at (a,b) will be the count of all locations x where $a = A'(x)$ and $b = B(x)$.

In a recent assessment of voxel similarity measures, Holden *et al.*[58] have demonstrated clear preference for those based on entropy measures including mutual information. To reduce problems that occur when the two images to be registered do not represent the same volume of the body, mutual information can be normalized.[54,59] Lau *et al.*[25] have demonstrated an advantage of a symmetric version of the correlation ratio compared to normalised mutual information when severe sub-sampling is used to improve speed. Recently, a number of groups have suggested further spatial constraints to assist registration such as edge information,[60] regional labelling,[61] or identification of parts of the image volume corresponding to bone where rigid-body transformation applies.[62] Entropy measures, and particularly mutual information, have found widespread use in nuclear medicine applications.

6. Display of Registered Data

Fused images need to be displayed in an integrated fashion, which allows simultaneous evaluation of multiple modalities or multiple image-sets both in 2D and in 3D. There are several imaging techniques, which can assist the user in viewing such combined images.

6.1 Synchronized Cursor

The simplest but very effective technique of imaging is the use of a synchronized dual cursor, which can be seen on multiple images at the same time. Although the images are not displayed in an integrated fashion with this technique, the cursors allow identification of corresponding features. The main advantage of using the synchronized cursors and linked side-by-side displays is that no information is lost during the display and features present at the same location in the two images will not obscure, and interfere with each other.

6.2 Fused Display

Medical images (CT, MRI, PET, SPECT) are typically monochrome with pseudo-colour lookup tables. Most often, anatomical images such as CT or MRI are displayed using the greyscale only. On the other hand, functional information derived from PET and SPECT can be effectively depicted using colour lookup tables. To combine the colour functional images with the greyscale anatomical images, various display techniques can be applied. When using older type 8-bit displays, it is not possible to mix or combine colours at a pixel level; however, it is possible to interleave the colour and greyscale pixel in grid-like fashion. When screen pixels are small, such interleaved pixels appear as a superimposition of colour with greyscale with the degree of transparency proportional to the interleave step. The disadvantage of such techniques is that the display effects depend on the screen resolution and various aliasing artefacts often occur when changing the effective image resolution, for example, when images are zoomed. In addition, image information may be lost if images need to be sub-sampled for the display purposes. When using 24-bit display, which is currently standard on most workstations, it is possible to perform colour blending with the degree of apparent transparency. Instead of interleaving the pixels, the intensities are combined and a new colour value is computed for each pixel, which includes the contribution from both images. Such fused images have higher quality. However, it is possible that bright structures on CT or MRI images can artificially increase the perceived intensity of the functional image. Therefore, fused images, although impressive, need to be viewed with caution and the review of original un-fused images is always recommended.

Other techniques include an interactive navigation of a roving window on the anatomical images, which reveals a sub-image containing the functional information. In a checkerboard display, small sub-images containing data from different modalities are interleaved. If serial images of the same modality are to be compared, for example stress/rest images of the heart or ictal and inter-ictal brain images, subtracted display can also be utilized to visualize changes.[45,63] Fused display cannot be used with serial images of the same modality since such images can be meaningfully compared only when the same colour lookup table is applied to both sets.

6.3 3D visualization

Fused data can also be visualized in 3D using a combination of all the above techniques. 3D volume rendered images can be created for anatomical images and 2D orthogonal or oblique cut-planes can be used to reveal the corresponding functional information. Combination of various techniques can be very effective. For example, oncological PET images are best displayed using the 3D Maximum Intensity Projection (MIP) technique whereas MRI is probably best displayed using a volume rendering technique or texture mapped 2D slices. 3D displays can also include surface-based displays shown at a certain threshold sometimes with colour-coded functional information. These techniques have been proven to be very useful in multimodal cardiac,[64] or brain visualization[65] (Figure 7). The multi-modality displays and even image registration can be implemented utilizing the latest computer graphics hardware innovations, such as hardware-based 3D or 2D texture mapping, which can accelerate 3D navigation, image blending and image registration.[66,67] The hardware-based graphics display is typically accelerated by the order of magnitude in comparison to the software-based rendering. Often 3D and 2D images are linked together with the help of synchronized cursors to create comprehensive display pages.

7. Approaches to Validation

Validating registration is not a trivial task, particularly in the case of non-rigid algorithms. Performance of algorithms can be task-dependent and, to some extent, system-dependent. Rigorous testing therefore ideally requires proof of performance in a range of situations. Although validation in

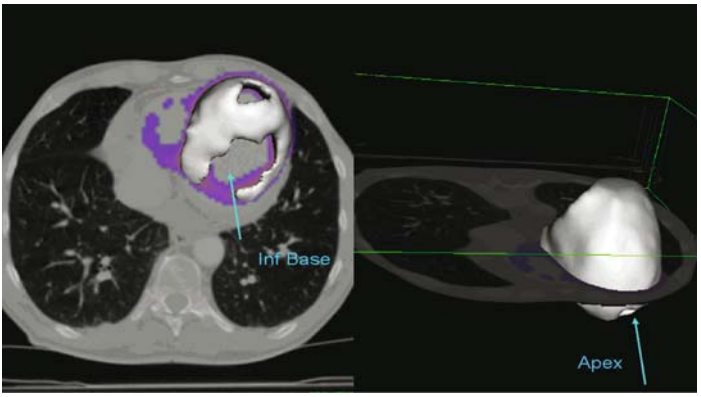


FIGURE 7. An example of a cardiac fusion display combining the iso-surface derived from one modality (^{18}F-FDG PET scan demonstrating infarction), with co-registered slice images from another modality (Electron Beam Computed Tomography).

realistic clinical data is desirable, it is rarely feasible, since 'ground truth' is usually unknown; but alternatives based on simulation never quite reproduce the clinical situation. Interpretation of parameters used to assess registration can be misleading, e.g. errors in transformation parameters can conceal the true clinical impact of misregistration. Only a limited number of studies adopt rigorous validation methods that demonstrate not only accuracy and precision, but also robustness (usability with different degrees of mis-registration) and flexibility (applicability to a range of situations).

7.1 Rigid-body Registration

Possibly the most rigorous studies have involved the use of external fiducial markers, visible using multiple modalities.[32,33] In these studies, the fiducials can be used to provide high quality registration and can be digitally removed so that alternative approaches can be independently tested in comparison with the known transformation. Acquiring these studies is non-trivial and, unfortunately, rarely involves all the modalities of potential interest. Nevertheless, these studies have been used to demonstrate accuracy of approximately half a voxel. A limitation of these studies is that the measured errors are fiducial registration errors (FRE), although the average registration error for the area of interest, normally referred to as the target registration error (TRE) can be derived.[30,68]

An alternative is to use multiple registrations to demonstrate consistency in the registration result;[69] e.g. for three image sets A, B, and C, the direct transformation to register A to C can be compared with the combined transformation in registering A to B and subsequently B to C. A basic assumption in this approach is that the errors in registration can be equally attributed to each of the transformations, an assumption that may not hold in general.

7.2 Non-rigid Registration

Validation of non-rigid registration is particularly problematic. Although visual assessment of clinical data can provide a qualitative evaluation (e.g. Meyer *et al.*[70]), it is difficult to guarantee that correct registration has been achieved. Simulation can be used (e.g. Pluim *et al.*[71]; Lau *et al.*[72]), provided the applied distortion is independent of the registration method and provided the range of applied distortions is similar to that encountered in clinical practice. Validation normally involves comparison of the displacement vector field that results from the registration with the known displacement vector field that was applied in the simulation. Finding an adequate description of the range of displacement errors continues to be a topic of research interest.

8. Clinical Applications of Image Registration

8.1 Neurology

Multi-modality image registration involving PET or SPECT has been utilized in several neurological applications. One particularly useful clinical example of multi-modality imaging is the combination of MRI and SPECT in imaging of epilepsy. SPECT radiotracer injection can be rapidly performed during the onset of seizures, enabling ictal imaging, which can localize the seizure foci. Such ictal images can be co-registered with MRI to perform image-guided resection of these regions.[63,73] Although SPECT discerns seizure foci, MRI information is needed for the navigation and anatomical correlation during the surgical procedure. In addition, ictal images acquired during seizures can be co-registered and subtracted from baseline (inter-ictal) images creating difference images which greatly enhance the visualization of foci.[63,74,75] Another useful application of SPECT and MRI image fusion is the guidance of SPECT region definition as required for image quantification. In one study, analysis of SPECT scans of Alzheimer's patients was assisted by the delineation of anatomical regions of interest derived from co-registered MRI,[76] MRI was used in a similar fashion in a neuro-receptor study to define true anatomical borders of basal-ganglia.[77] Figure 8 illustrates fusion of MRI with ^{123}I-IBZM SPECT. MRI was also helpful in localizing functional activation detected with SPECT.[78] Although functional MRI (fMRI) techniques can be used to localize functional

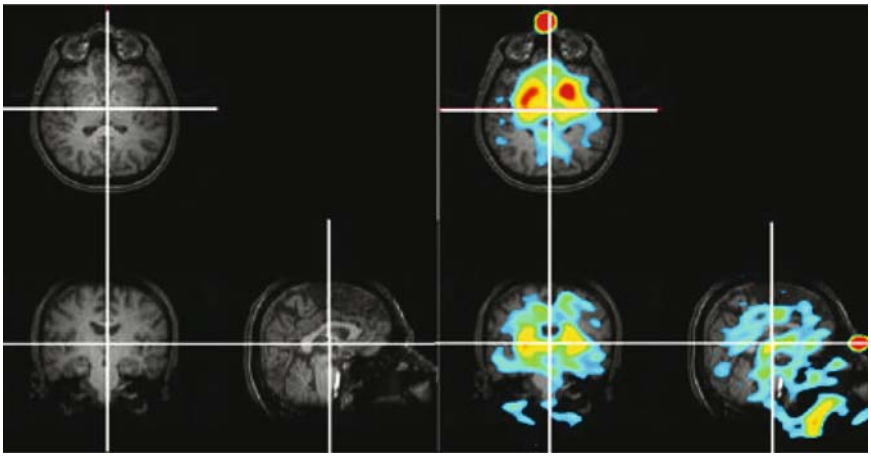


FIGURE 8. Image fusion of MRI with ^{123}I-IBZM SPECT neuro-receptor images of the basal ganglia. High resolution MRI may allow the identification of the anatomic boundaries of the basal ganglia. MRI could be used in such cases to define precisely the regions of interest for the purposes of SPECT quantification or for partial volume correction.

activation areas, they are susceptible to motion artefacts. In addition, activation tasks must be performed inside the scanner, which is not required in SPECT imaging. Co-registration of functional MRI with PET activation studies could be performed to maximize the information in pre-surgical planning.[79] PET activation images were especially helpful in the presence of fMRI artefacts. In a more basic research studies, it has been shown that the co-registration of brain PET with MRI could be used for partial volume correction of PET data enhancing the quantitative accuracy for measurements of cerebral blood flow, glucose metabolism and neuro-receptor binding.[80,81]

Brain SPECT and PET images of several patients can also be co-registered to common spatial coordinates and/or image templates with linear and non-linear techniques to create intra-modality, inter-patient reference databases for either group,[82] or individual comparisons.[83-85] This approach allows for automatic assessment of significance of local perfusion in comparison to the normal population. These registration-based brain PET and SPECT quantification techniques have been applied to a wide variety of clinical problems including Alzheimer's disease,[86-90] analysis or neuro-activation of PET[82,91] and SPECT studies,[92,93] head injuries,[94] and depression.[95,96] Similar database techniques have been also applied to the analysis of functional neuroreceptor SPECT studies,[92,97-99] which is more challenging since the tracer is accumulated solely in the basal ganglia. Image registration-based quantification techniques do not require definition of anatomical regions for individual patients, are observer-independent and have the potential to reliably detect subtle changes. Although such methods are fully automated, their performance depends on the quality of population- and equipment-specific databases; research is being conducted to evaluate the suitability and interchangeability of such normal databases.[100,101]

In addition to multi-modality registration and inter-patient registration of PET and SPECT brain neuro-images, some investigators have applied intra-modality serial registration techniques to assess serial changes in neuro-activation[102] in stroke imaging[103] and in a unique study of twin data comparison to examine chronic fatigue syndrome.[104] Serial registration techniques may be more sensitive than techniques that rely on the registration to the normal databases since serial images are obtained from the same patient and are highly similar in the absence of physiological change.

8.2 Cardiology

Algorithms for image registration have been applied to cardiac nuclear medicine images primarily to perform automated image reorientation and image quantification of cardiac images by registration and comparison to normal templates.[45,105-108] Some of the template-based systems utilizing image registration have been successfully validated in larger groups of patients with high success rate.[107,109] Image registration algorithms have

also been applied for intra-modality serial or sequential studies such as stress and rest SPECT and cardiac PET images.[9,110,111] Transmission scans have been utilized for serial PET registration.[112] Image registration was also applied to improve the quality of cardiac images by ''motion-tracking'' and registration of ECG gated cardiac PET images to obtain de-blurred static image of myocardial perfusion.[113] In a preliminary study, similar techniques were applied to obtain ''motion frozen'' perfusion images from gated SPECT.[114] Although shown to be very effective in estimation of changes, serial/sequential cardiac registration techniques are currently not typically used in commercial quantification packages.

The registration of four-dimensional ECG-gated multi-modality images of the heart has not been widely investigated. Cardiac registration via thoracic data has been proposed for PET/SPECT registration,[115] and for SPECT/CT registration,[116,117] and in a clinical application for the registration of ^{18}F-FDG PET data with electron beam CT.[118] Automated registration based on a combination of left ventricular segmentation with the ''head and hat'' algorithm[39] was developed.[41] Manual fusion has been proposed for electron beam CT and N-13 PET,[119] for MRI and PET,[120] and for CT and In-111 SPECT thrombus imaging.[121] In an effort to combine the vascular tree obtained from X-ray angiography with SPECT perfusion, knowledge-based vascular tree models and left ventricular segmentation were applied.[64] In related work, a prototype system was developed for combined X-ray angiography with PET/SPECT.[122] A technique for using transmission scans for PET and MRI alignment has been presented.[123] In a preliminary study, an automated 4D registration of cardiac MRI with myocardial perfusion SPECT images was presented, based on the extension of mutual information and motion-based segmentation of cardiac MRI.[124] In that study, the correlating motion of the heart was used as additional information utilized by the image registration algorithm. Most of the published studies in multi-modality cardiac registration describe initial preliminary experiences with various methodologies without extensive validation.

Although currently multi-modality fusion techniques are not applied in clinical application of cardiac imaging, there are several potential applications, which could prove to be clinically practical and useful. Future nuclear cardiology applications may include matching of perfusion defects defined by SPECT and PET imaging with the location of stenosis obtained with emerging coronary CT angiography techniques. Since both the perfusion defects and stenotic lesions are often depicted with poor quality, spatial correlation of these findings could potentially improve synergistically sensitivity and specificity of these techniques. Another potential application in nuclear cardiology is the quantification of mismatch from PET-SPECT multi-modality studies where ^{18}F-FDG viability PET findings need to be directly compared with perfusion SPECT scans.[125] Such techniques are proposed for the fusion of 2D X-ray angiography with SPECT,[64] but are not yet clinically utilized due to difficulties with aligning the 2D X-ray

projection data with 3D SPECT. The availability of fully tomographic 3D CT angiography techniques may allow for practical implementations. Another potential application is the use of standalone CT or MRI data for cardiac SPECT[126] or brain PET[127] attenuation correction, respectively, since SPECT and PET transmission imaging methods add cost and complexity to the imaging procedures and have relatively poor quality. There are also preliminary studies which explore the possibility of depicting vulnerable plaque in the coronary arteries, aorta and carotids with ^{18}F-FDG and correlating it with vessel anatomy obtained by CT.[128] Although these results have been demonstrated with the use of the hybrid PET/CT scanner, similar studies could be performed using software image registration and fusion.

8.3 Oncology

The applications of image fusion in oncology have been rapidly expanding with the introduction of dedicated hybrid SPECT/CT[129,130] and PET/CT scanners.[131] Recent reviews summarize comprehensively these applications for PET/CT[132] and SPECT/CT scanners.[133] Although most of the current clinical applications are reported with the use of hybrid devices, without using software image registration, most of these applications could be performed with software fusion. For brain images, the image registration can be also performed with an aid of stereo-tactic frames or external fiducial markers,[14] described in section 5.3. However, in thoracic or abdominal images common in oncological imaging, changes in the external body shape may not correspond to the shifts of internal organs; therefore, these external reference systems are not practical. Furthermore, the necessary set-up of the external reference systems during both imaging sessions can add to the complexity and time of the imaging procedures.

Although software fusion has been initially developed for the brain imaging and original software fusion applications were reported in the brain oncology area,[39] software-based registration of other organs for oncological imaging has been pursued with good results. The multi-modality registration of CT and PET, or SPECT of the thorax has been approached as a rigid-body problem using surface-based techniques based on chamfer matching,[134] lung segmentation,[135,136] interactively-defined homologous markers,[137] and mutual information maximization.[138] Automated software-based SPECT/CT fusion has been proposed for Prostascint™ pelvis images using vascular tree information visualized on both scans.[139] Linear registration techniques can be suitable for whole-body imaging if patient body configuration and breathing conditions are similar between the separate scans (Figure 9). If PET images are acquired with a different breathing pattern and with different body configuration than CT, or MRI, then rigid-body image registration is typically not sufficient.

To correlate images of the thorax and abdomen, non-rigid approaches may be required in combination with linear registration to correct for

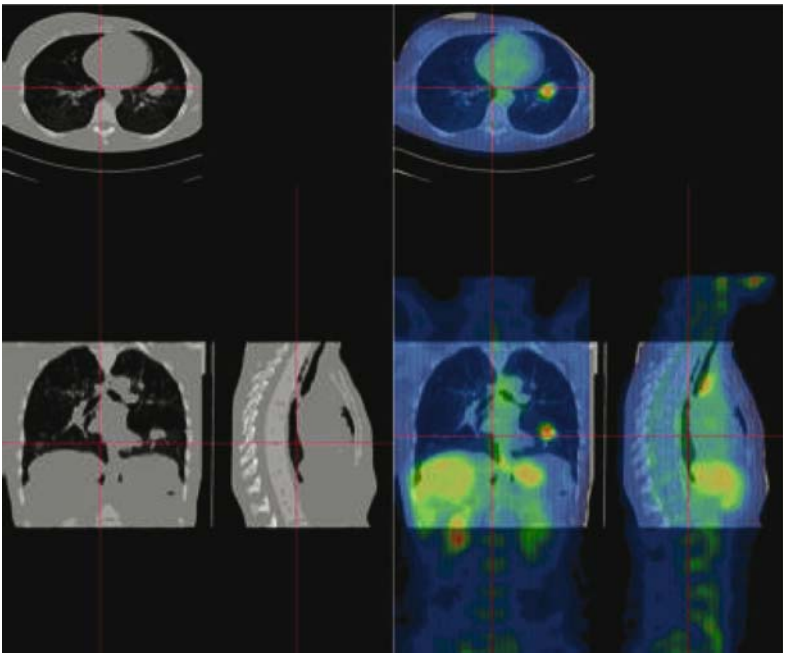


FIGURE 9. An example of linear image registration applied to the fusion of thoracic CT with whole-body ^{18}F-FDG PET. Both scans were acquired with similar breathing patterns.

changes in body configuration, breathing differences, or internal organ displacements. Non-linear approaches based on thin-plate spline (TPS) interpolation[70] or piecewise linear approaches[138,140] have been proposed to compensate for soft tissue deformations. Image warping is typically accomplished by energy minimizing interpolation such as thin-plate-splines,[21] or B-splines.[141] Such non-linear correction may be very difficult to determine from functional PET images alone (section 7.2). Therefore one possibility is to use auxiliary transmission images, which are almost always acquired during PET imaging session to enable attenuation correction and absolute quantification of PET radiopharmaceutical uptake.[116,142,143] These transmission images are not affected by the potential mismatch between functional and anatomical information, therefore they can provide a robust estimate of the nonlinear image transformation. Recently, a practical, fully automated algorithm for non-linear 3D image registration of whole-body PET and CT images, which compensates for the non-linear deformation due to the breath-hold imaging technique and variable body configuration, has been proposed.[143] In general, the nonlinear surface-based registration techniques remain sensitive to initial errors in the image segmentation phase, while the computational cost of volume-based iterative approach can be prohibitive. The principal remaining difficulties are the selection of suitable

deformation models, the validation of the non-linear registration algorithms, and the computational cost. The full automation of the non-linear registration techniques is of great importance, since the number of adjustable image parameters can be large.

8.3.1 Radiotherapy Applications

Image registration and fusion enables the use of sensitive functional modalities such as PET as a guide for radiation therapy planning. Software image fusion has been applied with success to augment radiation treatment planning systems with PET or SPECT information. PET- or SPECT-guided treatment plans have been reported for brain lesions,[144] lung cancer,[70,145-148] head and neck cancers,[149] and abdominal cancer.[150] SPECT ^{99m}Tc MAA lung perfusion images have also been incorporated in the treatment plans to allow avoidance of highly functioning lung tissue.[151] Although software fusion performs very well in brain imaging, PET or SPECT may not offer significant "added-value" for this application and most often MRI/CT brain fusion is utilized. The most significant changes in the definitions of the gross tumour volumes when adding PET registered data have been reported for the lung cancer radiotherapy planning. PET, in many cases, can identify additional lesions or allow dose escalation by identifying necrotic portions of tumours, which do not need to be irradiated. PET has much higher sensitivity and specificity in lung cancer imaging, but cannot be used alone for planning of the radiation therapy, which is based on simulation CT. Combined PET/CT images allow interactive definition of gross tumour volumes (GTV) using synchronized multi-modality displays. The thoracic PET/CT software fusion for radiotherapy planning seems to be a practical application since the impact of using fused images is high and despite non-linear deformations, sufficient image detail is present on both CT and PET images to allow accurate image registration. More precise functionally-based treatment planning techniques will increase in importance with concurrent advances in the radiation delivery methods such as intensity-modulated radiotherapy (IMRT) and image guided brachytherapy.[152]

9. Future Directions

9.1 Hardware vs. Software Image Fusion

Software image fusion can be challenging to perform on a routine clinical basis since it requires excellent DICOM connectivity, compatibility of imaging protocols and cooperation between various imaging departments. These difficulties may be bypassed with the use of hybrid PET/CT scanners.[131] Software fusion, however, may offer greater flexibility. For example, it is possible to perform PET-MRI image fusion with registration software;[32,153] MRI is superior to CT for oncological brain imaging and recent

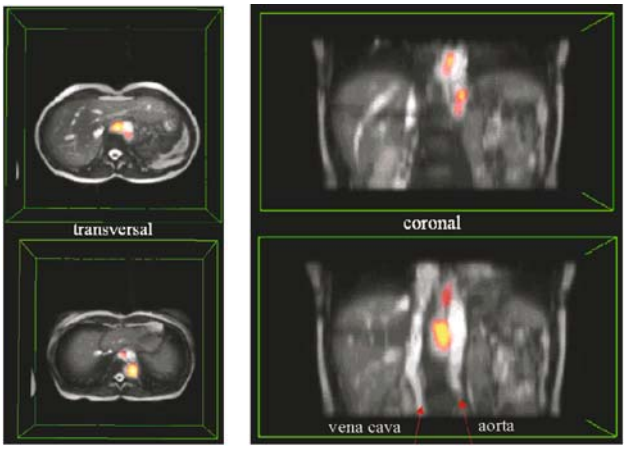


FIGURE 10. Image fusion of MRI and ^{18}F-FDG PET in a patient with malignant pheochromocytoma and abdominal metastases. In the transverse slices (left), initial infiltration of the vertebra can be detected. The coronal slices show inter-aorto-caval metastases. This example demonstrates that soft tissue contrast/outline of vascular structures by MRI may be needed to accurately localize lesions in the abdomen. (Data courtesy of R.P. Baum, Zentralklinik, Bad Berka).

advances in MRI have allowed fast thoracic and abdominal imaging with excellent contrast.[154] An example of PET-MRI fusion is shown in Figure 10. There are efforts underway to use MRI as the primary modality for radiotherapy planning. Furthermore, without the availability of dedicated PET/CT scanner in the radiotherapy department, it may be more feasible to match retrospectively standalone PET scans to the simulation CT scans performed in the course of the radiation therapy in the radiation oncology department. The accuracy of image registration obtained on hybrid PET/CT may be higher since images are acquired close in time, without changing patient position; However, software fusion may be sufficient for the radiation therapy planning since tumour volume definitions are derived interactively with large error margins especially in lung cancer patients.[146]

Software registration techniques can be applied for the follow-up assessment and comparison of serial scans during the course of therapy and presumably for correction of the motion and mismatch of data acquired on the PET/CT machine. Perhaps one of the most difficult applications for software image fusion is multi-modality imaging of the head and neck where significant non-linear mismatches can occur due to different position of arms and/or head. Abdominal and thoracic areas are also often deformed but many significant landmarks are present which can aid non-linear software algorithms to achieve accurate superimposition. In the abdominal and pelvic region, the required accuracy of fusion may be very high due to possible false-positive uptake in the bowels and urinary tract adjacent to the

possible sites of malignant lesions such as ovaries. In these cases, hardware-based PET/CT image fusion may perform better than software-based techniques. If MRI was also indicated for better tissue contrast, software registration of MRI with PET/CT images could be performed. Sometimes, it may be necessary to acquire images in different body configurations for specific purposes, even if images were acquired on a hybrid PET/CT system. For example, CT imaging may need to be performed in deep inspiration due to superior image quality and some benefits in radiation treatment planning.[155,156] Such imaging techniques would require non-linear software registration even if images were obtained on the same scanner. Most likely, a combination of software and hardware-based techniques will be used in a clinical practice with software techniques prominent in brain, and thoracic imaging and in image-guided radiotherapy areas as well as for modality combinations other than PET/CT. The biggest challenge for the proliferation of software fusion into the clinical environment is true compatibility and connectivity between various diagnostic and therapeutic imaging systems. The reasons for selecting hardware or software-based approaches may be more logistical and economical than based on the capabilities of each technology. Despite rapid proliferation of hybrid PET/CT systems, there will be a large installed base of standalone PET scanners, for which software fusion will be required; already in some departments software fusion is used routinely, especially in conjunction with PET scans.[157]

9.2 Potential Applications for Registration

The full potential of registered medical images has yet to be realised. To a large extent, the utility of complementary information has not been exploited. For various reasons, the acceptance of image registration has been relatively slow. Problems in achieving efficient data transfer, choice of software algorithm and lack of sufficiently well documented validation have all contributed to this. However, the introduction of dual-modality imaging has very quickly changed this situation and also is likely to impact on the demand for software registration solutions. There are a number of potential applications of registered data beyond simple visualisation of fused images; these include the incorporation of anatomical information in reconstruction or quantification and motion correction. The wider availability of accurately registered images is likely to increase the popularity of these techniques.

9.2.1 Incorporation of Anatomical Information

The most obvious use of anatomical data is in attenuation correction, particularly for use in the thorax where there is non-homogeneous attenuation. However, knowledge of the anatomy can be utilised for other purposes. For example, if the activity distribution can be assumed to be uniform

within anatomically distinct regions, determined by segmentation of MRI or CT, an estimate of partial volume loss can be derived by smoothing the segmented study to the same resolution as PET or SPECT and measuring the resultant signal dilution. Care needs to be taken to account for the spillover from tissues adjacent to the region of interest, especially if the surrounding activity is uncertain. This form of correction can be applied on a region or voxel basis (e.g.[80,158,159]).

A technique that has shown promise, but failed to attract widespread use, involves the direct incorporation of anatomical information in reconstruction of emission data. The well defined anatomical edges can be used to constrain the reconstruction, discouraging smooth boundaries.[160-162] More recent refinements[163] are discussed further in Chapter 4. In addition, the use of direct Monte Carlo simulation was recently demonstrated, based on measured attenuation coefficients, to directly estimate scatter in emission reconstruction. Novel approaches to improving the speed of Monte Carlo estimation make such an approach feasible.[164]

9.2.2 Motion Correction

As resolution and the degree of quantitative correction improves in SPECT and PET, the importance of motion correction increases. Motion of either the subject or some internal structure can be regarded as a specific intra-modality registration problem (motion tracking is effectively a form of constrained registration). Depending on the type of motion, correction of multiple acquired images can be achieved by either rigid-body or non-rigid registration to one reference image (usually taken from the sequence). These techniques have been applied to rapid dynamic sequences[165] and gated studies.[113] Correction of motion during the frequently lengthy, emission tomography acquisition is more complex, although fully automated correction based on 3D registration has been demonstrated[166,167] as discussed further in chapter 4.

9.2.3 Atlas-based and Inter-subject Analysis

Registering individual studies to a standard atlas or reference study is not trivial, especially for areas other than brain. However, availability of a standard atlas permits inter-subject comparison based on pre-defined regions and offers potential for automatically defining tissue signatures that can be used to classify subjects (e.g. based on shape, texture or size). The wider availability of registered data should encourage more widespread use of complementary information, which should aid probabilistic approaches to tissue characterisation.

Acknowledgments. We are indebted to Dennys Yiu Hon Lau for assistance with some of the figures.

References

1. Hajnal J. V., Hill, D. L. G. and J., H. D., *Medical Image Registration*, CRC Press, Boca Raton, (2001).
2. van den Elsen P. A., Pol, E.-J. D. and Viergever, M. A., Medical image matching-a review with classification. *IEEE Eng Med Biol* **12:** 26-39 (1993).
3. Maurer C. R. and Fitzpatrick, J. M., "A review of medical image registration." in: *Interactive Image-Guided Neurosurgery*, edited by R. J. Maciunas American Association of Neurological Surgeons, Parkridge, IL, (1993), *pp* 17-44.
4. Weber D. A. and Ivanovic, M., Correlative image registration. *Semin Nucl Med* **24:** 311-323 (1994).
5. Lavallee S., "Registration for computer-integrated surgery: methodology, state of the art." in: *Computer-integrated surgery, Technology and Clinical Applications*, edited by S. Lavallee R. H. Taylor, et al. MIT Press, Cambridge, (1996), *pp* 77-97.
6. McInerney T. and Terzopoulos, D., Deformable models in medical image analysis: a survey. *Med Image Anal* **1:** 91-108 (1996).
7. Hawkes D. J., Algorithms for radiological image registration and their clinical application. *J Anat* **193:** 347-361 (1998).
8. Maintz J. B. and Viergever, M. A., A survey of medical image registration. *Med Image Anal* **2:** 1-36 (1998).
9. Eberl S., Kanno, I., Fulton, R. R. *et al.*, Automated interstudy image registration technique for SPECT and PET. *J Nucl Med* **37:** 137-145 (1996).
10. Lester H. and Arridge, S. R., A survey of hierarchical non-linear medical image registration. *Pattern Recognition* **32:** 129-149 (1999).
11. Thurfjell L., Pagani, M., Andersson, J. L. *et al.*, Registration of neuroimaging data: implementation and clinical applications. *J Neuroimaging* **10:** 39-46 (2000).
12. Audette M. A., Ferrie, F. P. and Peters, T. M., An algorithmic overview of surface registration techniques for medical imaging. *Med Image Anal* **4:** 201-217 (2000).
13. Viergever M. A., Maintz, J. B., Niessen, W. J. *et al.*, Registration, segmentation, and visualization of multimodal brain images. *Comput Med Imaging Graph* **25:** 147-151 (2001).
14. Hill D. L., Batchelor, P. G., Holden, M. *et al.*, Medical image registration. *Phys Med Biol* **46:** R1-45 (2001).
15. Hutton B. F., Braun, M., Thurfjell, L. *et al.*, Image registration: an essential tool for nuclear medicine. *Eur J Nucl Med* **29:** 559-577 (2002).
16. Hutton B. F. and Braun, M., Software for image registration: algorithms, accuracy, efficacy. *Semin Nucl Med* **33:** 180-192 (2003).
17. Greitz T., Bergstrom, M., Boethius, J. *et al.*, Head fixation system for integration of radiodiagnostic and therapeutic procedures. *Neuroradiology* **19:** 1-6 (1980).
18. Peters T., Davey, B., Munger, P. *et al.*, Three-dimensional multimodal image-guidance for neurosurgery. *IEEE Trans Med Imaging* **15:** 121-128 (1996).
19. Maudsley A. A., Simon, H. E. and Hilal, S. K., Magnetic field measurement by NMR imaging. *J Phys E Sci Instrum* **17:** 216-220 (1984).
20. Thurfjell L., Lau, Y. H., Andersson, J. L. *et al.*, Improved efficiency for MRI-SPET registration based on mutual information. *Eur J Nucl Med* **27:** 847-856 (2000).
21. Bookstein F. L., Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans Pattern Anal Machine Intell* **11:** 567-585 (1989).

22. Davis M. H., Khotanzad, A., Flamig, D. P. *et al.*, A physics-based coordinate transformation for 3-D image matching. *IEEE Trans Med Imaging* **16:** 317-328 (1997).
23. Christensen G. E., Rabbitt, R. D. and Miller, M. I., 3D brain mapping using a deformable neuroanatomy. *Phys Med Biol* **39:** 609-618 (1994).
24. Bro-Nielsen M. and Gramkow, C., "Fast fluid registration of medical images." in: *Lecture Notes in Computer Science.*, edited by K. H. Hohne and R. Kikinis Springer-Verlag, Hamburg, (1996), *pp* 267-276.
25. Lau Y. H., Braun, M. and Hutton, B. F., Non-rigid image registration using a median-filtered coarse-to-fine displacement field and a symmetric correlation ratio. *Phys Med Biol* **46:** 1297-1319 (2001).
26. Collins D. L., Neelin, P., Peters, T. M. *et al.*, Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr* **18:** 192-205 (1994).
27. Maes F., Vandermeulen, D. and Suetens, P., Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Med Image Anal* **3:** 373-386 (1999).
28. Studholme C., Hill, D. L. and Hawkes, D. J., Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Med Phys* **24:** 25-35 (1997).
29. Thevenaz P., Ruttimann, U. E. and Unser, M., A pyramid approach to subpixel registration based on intensity. *IEEE Trans Imag Proc* **7:** 27-41 (1998).
30. Fitzpatrick J. M., West, J. B. and Maurer, C. R., Predicting error in rigid-body point-based registration. *IEEE Trans Med Imaging* **17:** 694-702 (1998).
31. Besl P. J. and McKay, N. D., A method for registration of 3D shapes. *IEEE Trans Patt Anal Mach Intell* **14:** 239-256 (1992).
32. West J., Fitzpatrick, J. M., Wang, M. Y. *et al.*, Comparison and evaluation of retrospective intermodality brain image registration techniques. *J Comput Assist Tomogr* **21:** 554-566 (1997).
33. Barnden L., Kwiatek, R., Lau, Y. *et al.*, Validation of fully automatic brain SPET to MR co-registration. *Eur J Nucl Med* **27:** 147-154 (2000).
34. Wang M. Y., Maurer, C. R., Fitzpatrick, J. M. *et al.*, An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head. *IEEE Trans Biomed Eng* **43:** 627-637 (1996).
35. Evans A. C., Marrett, S., Torrescorzo, J. *et al.*, MRI-PET correlation in three dimensions using a volume-of-interest (VOI) atlas. *J Cereb Blood Flow Metab* **11:** A69-78 (1991).
36. Thirion J., New feature points based on geometric invariants for 3D image registration. *Int J Comput Vision* **18:** 121-137 (1996).
37. Monga O. and Benayoun, S., Using partial derivatives of 3D images to extract typical surface features. *Comput. Vision Image Understanding* **61:** 171-189 (1995).
38. Gueziec A. and Ayache, N., Smoothing and matching of 3D space curves. *Int J Comput Vision* **12:** 79-104 (2004).
39. Pelizzari C. A., Chen, G. T., Spelbring, D. R. *et al.*, Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. *J Comput Assist Tomogr* **13:** 20-26 (1989).
40. Gee J. C., Barillot, C., Briquer, L. *et al.*, "Matching structural images of the human brain using statistical and geometrical images features" in: *Visualization*

in Biomedical Computing, edited by R. Robb SPIE Press, Bellingham, Washington, (1994), *pp* 191-201.
41. Faber T. L., McColl, R. W., Opperman, R. M. *et al.*, Spatial and temporal registration of cardiac SPECT and MR images: methods and evaluation. *Radiology* **179:** 857-861 (1991).
42. Scott A. M., Macapinlac, H. A., Divgi, C. R. *et al.*, Clinical validation of SPECT and CT/MRI image registration in radiolabeled monoclonal antibody studies of colorectal carcinoma. *J Nucl Med* **35:** 1976-1984 (1994).
43. Venot A., Lebruchec, J. F. and Roucayrol, J. C., A new class of similarity measures for robust image registration. *Comput Vis Graph Imag Proc* **28:** 176-184 (1984).
44. Alpert N. M., Bradshaw, J. F., Kennedy, D. *et al.*, The principal axes transformation–a method for image registration. *J Nucl Med* **31:** 1717-1722 (1990).
45. Slomka P. J., Hurwitz, G. A., Stephenson, J. *et al.*, Automated alignment and sizing of myocardial stress and rest scans to three-dimensional normal templates using an image registration algorithm. *J Nucl Med* **36:** 1115-1122 (1995).
46. Viola P. A., ''Alignment by maximization of mutual information,'' PhD thesis, Massachusetts Institute of Technology, 1995.
47. Junck L., Moen, J. G., Hutchins, G. D. *et al.*, Correlation methods for the centering, rotation, and alignment of functional brain images. *J Nucl Med* **31:** 1220-1226 (1990).
48. Rizzo G., Pasquali, P., Gilardi, M. C. *et al.*, ''Multi-modality biomedical image integration: use of a cross-correlation technique.'' Proc. IEEE Eng Med Biol Soc, Vol. 13; *pp* 219-220 (1991).
49. Andersson J. L., Sundin, A. and Valind, S., A method for coregistration of PET and MR brain images. *J Nucl Med* **36:** 1307-1315 (1995).
50. Woods R. P., Mazziotta, J. C. and Cherry, S. R., MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* **17:** 536-546 (1993).
51. Ardekani B. A., Braun, M., Hutton, B. F. *et al.*, A fully automatic multimodality image registration algorithm. *J Comput Assist Tomogr* **19:** 615-623 (1995).
52. Hill D. L. G., Studholme, C. and Hawkes, D. J., ''Voxel similarity measures for automated image registration.'' in: *Visualization in Biomedical Computing*, edited by R. Robb SPIE Press, Bellingham, Washington, (1994), *pp* 205- 216.
53. Collignon A., Maes, F., Delaere, D. *et al.*, ''Automated multi-modality image registration based on information theory.'' in: *Information Processing in Medical Imaging*, edited by C. Barillot Y. Bizais, R. Di Paolo Kluwer Academic, Dordrecht, (1995), *pp* 263-274.
54. Maes F., Collignon, A., Vandermeulen, D. *et al.*, Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* **16:** 187-198 (1997).
55. Viola P. and Wells III, W. M., Alignment by maximization of mutual information. *Int J Comput Vision* **24:** 137-154 (1997).
56. Wells W. M., Viola, P., Atsumi, H. *et al.*, Multi-modal volume registration by maximization of mutual information. *Med Image Anal* **1:** 35-51 (1996).
57. Roche A., Malandain, G., Pennec, X. *et al.*, ''The correlation ratio as a new similarity measure for mutimodal image registration.'' in: *Lecture Notes in Computer Science. Proc. MICCAI'98.*, edited by A. Colchester W. M. Wells, and S. Delp Springer-Verlag, Berlin, (1998), *pp* 1115-1124.

58. Holden M., Hill, D. L., Denton, E. R. *et al*., Voxel similarity measures for 3-D serial MR brain image registration. *IEEE Trans Med Imaging* **19:** 94-102 (2000).
59. Studholme C., Hill, D. L. G. and Hawkes, D. J., An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recogn* **32:** 71-86 (1999).
60. Pluim J. P., Maintz, J. B. and Viergever, M. A., Image registration by maximization of combined mutual information and gradient information. *IEEE Trans Med Imaging* **19:** 809-814 (2000).
61. Studholme C., Hill, D. L. G. and Hawkes, D. J., ''Incorporating connected region labelling into automated image registration using mutual information.'' Proceedings of the IEEE workshop on Mathematical Methods in Biomedical Image Analysis, San Francsico Ca, *pp* 23-31 (1996).
62. Little J. A., Hill, D. L. G. and Hawkes, D. J., Deformations incorporating rigid structures. *Comput Vis Imag Understanding* **66:** 223-232 (1997).
63. Zubal I. G., Spencer, S. S., Imam, K. *et al*., Difference images calculated from ictal and interictal technetium-99m-HMPAO SPECT scans of epilepsy. *J Nucl Med* **36:** 684-689 (1995).
64. Aguade S., Candell-Riera, J., Faber, T. L. *et al*., [Unified three-dimensional images of myocardial perfusion and coronary angiography]. *Rev Esp Cardiol* **55:** 258-265 (2002).
65. Minoshima S., Frey, K. A., Koeppe, R. A. *et al*., A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *J Nucl Med* **36:** 1238-1248 (1995).
66. Hastreiter P. and Ertl, T., ''Integrated registration and visualization of medical image data.'' Proceedings Computer Graphics International, *pp* 78-85 (1998).
67. Enders F., Strengert, M., Iserhardt-Bauer, S. *et al*., ''Interactive volume rendering of multi-modality cardiac data with the use of consumer graphics hardware.'' SPIE 2003 Medical Imaging: Image visualization, San Diego, CA, Vol. 5029; *pp* 119-128 (2003).
68. West J. B. and Fitzpatrick, J. M., ''The distribution of target registration error in rigid-body, point-based registration.'' in: *Lecture Notes in Computer Science. Proc. IPMI 1999*, edited by A. Kuba et al. Springer-Verlag, Berlin, (1999), *pp* 460-465.
69. Woods R. P., Grafton, S. T., Holmes, C. J. *et al*., Automated image registration: I. General methods and intrasubject, intramodality validation. *J Comput Assist Tomogr* **22:** 139-152 (1998).
70. Meyer C. R., Boes, J. L., Kim, B. *et al*., Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. *Med Imag Anal* **1:** 195-206 (1997).
71. Pluim J. P. W., Maintz, J. B. A. and Viergever, M. A., Mutual information matching in multiresolution contexts. *Image Vision Comput* **19:** 53-62 (2001).
72. Lau Y. H., ''Application of joint intensity algorithms to the registration of emission tomography and anatomical images'' PhD thesis, University of Technology, Sydney, 2004.
73. Murphy M., O'Brien, T. J., Morris, K. *et al*., Multimodality image-guided epilepsy surgery. *J Clin Neurosci* **8:** 534-538 (2001).
74. O'Brien T. J., O'Connor, M. K., Mullan, B. P. *et al*., Subtraction ictal SPET co-registered to MRI in partial epilepsy: description and technical validation of the method with phantom and patient studies. *Nucl Med Commun* **19:** 31-45 (1998).

75. Vera P., Kaminska, A., Cieuta, C. *et al*., Use of subtraction ictal SPECT co-registered to MRI for optimizing the localization of seizure foci in children. *J Nucl Med* **40:** 786-792 (1999).
76. Julin P., Lindqvist, J., Svensson, L. *et al*., MRI-guided SPECT measurements of medial temporal lobe blood flow in Alzheimer's disease. *J Nucl Med* **38:** 914-919 (1997).
77. Barthel H., Muller, U., Wachter, T. *et al*., [Multimodal SPECT and MRT imaging data analysis for an improvement in the diagnosis of idiopathic Parkinson's syndrome] *Radiologe* **40:** 863-869 (2000).
78. Zifko U. A., Slomka, P. J., Young, G. B. *et al*., Brain mapping of median nerve somatosensory evoked potentials with combined 99mTc-ECD single-photon emission tomography and magnetic resonance imaging. *Eur J Nucl Med* **23:** 579-582 (1996).
79. Krings T., Schreckenberger, M., Rohde, V. *et al*., Functional MRI and 18F FDG-positron emission tomography for presurgical planning: comparison with electrical cortical stimulation. *Acta Neurochir* **144:** 889-899; discussion 899 (2002).
80. Muller-Gartner H. W., Links, J. M., Prince, J. L. *et al*., Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects. *J Cereb Blood Flow Metab* **12:** 571-583 (1992).
81. Meltzer C. C., Kinahan, P. E., Greer, P. J. *et al*., Comparative evaluation of MR-based partial-volume correction schemes for PET. *J Nucl Med* **40:** 2053-65 (1999).
82. Friston K. J., Frith, C. D., Liddle, P. F. *et al*., Comparing functional (PET) images: the assessment of significant change. *J Cereb Blood Flow Metab* **11:** 690-699 (1991).
83. Houston A. S., Kemp, P. M. and Macleod, M. A., A method for assessing the significance of abnormalities in HMPO brain SPECT images. *J Nucl Med* **35:** 239-244 (1994).
84. Minoshima S., Koeppe, R. A., Frey, K. A. *et al*., Anatomic standardization: linear scaling and nonlinear warping of functional brain images. *J Nucl Med* **35:** 1528-1537 (1994).
85. Slomka P., Reid, B., Hurwitz, G. and Stephenson, J. "Automated template-based quantification of brain SPECT." in: *A Textbook of SPECT in Neurology in Psychiatry*, edited by Dierckx RA De Deyne PP, Alavi A, Pickut BA John Libbey, London, (1997), *pp* 507-512.
86. Kogure D., Matsuda, H., Ohnishi, T. *et al*., Longitudinal evaluation of early Alzheimer's disease using brain perfusion SPECT. *J Nucl Med* **41:** 1155-1162 (2000).
87. Radau P. E. and P.J, S., "Constrained localized warping reduced registration errors due to lesions in functional neuroimages." Proc SPIE Medical Imaging, San Diego, CA, Vol. 4322; *pp* 588-601 (2001).
88. Radau P. E., Slomka, P. J., Julin, P. *et al*., Evaluation of linear registration algorithms for brain SPECT and the errors due to hypoperfusion lesions. *Med Phys* **28:** 1660-1668 (2001).
89. Huang C., Wahlund, L. O., Svensson, L. *et al*., Cingulate cortex hypoperfusion predicts Alzheimer's disease in mild cognitive impairment. *BMC Neurol* **2:** 9 (2002).
90. Song E. K. and Kolecki, P., A case of pulmonary artery dissection diagnosed in the Emergency Department. *J Emerg Med* **23:** 155-159 (2002).

91. Schoenahl F., Montandon, M.-L., Slosman, D., Zaidi H., "Assessment of the performance of SPM analysis in PET neuroactivation studies: A Monte Carlo investigation." *Conf. Proc.* of the VIIth International Meeting on Fully Three-dimensional Image Reconstruction in Radiology and Nuclear Medicine, 29 June-4 July 2003, Saint-Malo, France, Available on CD-ROM.
92. Lahorte P., Vandenberghe, S., Van Laere, K. *et al.*, Assessing the performance of SPM analyses of spect neuroactivation studies. Statistical Parametric Mapping. *Neuroimage* **12:** 757-64 (2000).
93. Van Laere K. J., Versijpt, J., Koole, M. *et al.*, Experimental performance assessment of SPM for SPECT neuroactivation studies using a subresolution sandwich phantom design. *Neuroimage* **16:** 200-216 (2002).
94. Stamatakis E. A., Wilson, J. T. and Wyper, D. J., Analysis of HMPAO SPECT scans in head injury using Statistical Parametric Mapping. *Behav Neurol* **12:** 29-37 (2000).
95. MacHale S. M., Lawlie, S. M., Cavanagh, J. T. *et al.*, Cerebral perfusion in chronic fatigue syndrome and depression. *Br J Psychiatry* **176:** 550-556 (2000).
96. Bonte F. J., Trivedi, M. H., Devous, M. D. *et al.*, Occipital brain perfusion deficits in children with major depressive disorder. *J Nucl Med* **42:** 1059-1061 (2001).
97. Radau P. E., Linke, R., Slomka, P. J. *et al.*, Optimization of automated quantification of 123I-IBZM uptake in the striatum applied to parkinsonism. *J Nucl Med* **41:** 220-227 (2000).
98. Van Laere K., Koole, M., D'Asseler, Y. *et al.*, Automated stereotactic standardization of brain SPECT receptor data using single-photon transmission images. *J Nucl Med* **42:** 361-375 (2001).
99. Bosman T., Van Laere, K. and Santens, P., Anatomically standardised 99mTc-ECD brain perfusion SPET allows accurate differentiation between healthy volunteers, multiple system atrophy and idiopathic Parkinson's disease. *Eur J Nucl Med Mol Imaging* **30:** 16-24 (2003).
100. Van Laere K., Koole, M., Versijpt, J. *et al.*, Transfer of normal 99mTc-ECD brain SPET databases between different gamma cameras. *Eur J Nucl Med* **28:** 435-449 (2001).
101. Van Laere K. J. and Dierckx, R. A., Brain perfusion SPECT: age- and sex-related effects correlated with voxel-based morphometric findings in healthy adults. *Radiology* **221:** 810-817 (2001).
102. Barber D. C., Tindale, W. B., Hunt, E. *et al.*, Automatic registration of SPECT images as an alternative to immobilization in neuroactivation studies. *Phys Med Biol* **40:** 449-463 (1995).
103. Siddique M. S., Fernandes, H. M., Wooldridge, T. D. *et al.*, Reversible ischemia around intracerebral hemorrhage: a single-photon emission computerized tomography study. *J Neurosurg* **96:** 736-741 (2002).
104. Lewis D. H., Mayberg, H. S., Fischer, M. E. *et al.*, Monozygotic twins discordant for chronic fatigue syndrome: regional cerebral blood flow SPECT. *Radiology* **219:** 766-773 (2001).
105. Declerck J., Feldmar, J., Goris, M. L. *et al.*, Automatic registration and alignment on a template of cardiac stress and rest reoriented SPECT images. *IEEE Trans Med Imaging* **16:** 727-737 (1997).

106. Goris M. L., Pace, W. M., Maclean, M. *et al*., Three-dimensional quantitative analysis of scintigraphic tomographic images after elastic transformation to a template. *Surg Technol Int* **IX:** 81-90 (2000).
107. Slomka P. J., Radau, P., Hurwitz, G. A. *et al*., Automated three-dimensional quantification of myocardial perfusion and brain SPECT. *Comput Med Imaging Graph* **25:** 153-164 (2001).
108. Peace R. A., Staff, R. T., Gemmell, H. G. *et al*., Automatic detection of coronary artery disease in myocardial perfusion SPECT using image registration and voxel to voxel statistical comparisons. *Nucl Med Commun* **23:** 785-794 (2002).
109. Slomka P. J., Hurwitz, G. A., St_Clement, G. *et al*., Three-dimensional demarcation of perfusion zones corresponding to specific coronary arteries: application for automated interpretation of myocardial SPECT. *J Nucl Med* **36:** 2120-2126 (1995).
110. Hoh C. K., Dahlbom, M., Harris, G. *et al*., Automated iterative three-dimensional registration of positron emission tomography images. *J Nucl Med* **34:** 2009-2018 (1993).
111. Slomka P., Aboul-Enein, F., Berman, D. *et al*., Direct quantification of ischemia from registered stress-rest myocardial perfusion scans without normal databases. [abstract] *J Nucl Med* **44:** 202P (2003).
112. Bacharach S. L., Douglas, M. A., Carson, R. E. *et al*., Three-dimensional registration of cardiac positron emission tomography attenuation scans. *J Nucl Med* **34:** 311-321 (1993).
113. Klein G. J., Reutter, B. W. and Huesman, R. H., Non-rigid summing of gated PET via optical flow. *IEEE Trans Nucl Sci* **44:** 1509-1512 (1997).
114. Slomka P. J., Baddredine, M., Baranowski, J. *et al*., "Motion frozen" quantification and display of myocardial perfusion gated SPECT [abstract]. *World J Nucl Med* **1:** S 23 (2002).
115. Gilardi M. C., Rizzo, G., Savi, A. *et al*., Correlation of SPECT and PET cardiac images by a surface matching registration technique. *Comput Med Imaging Graph* **22:** 391-398 (1998).
116. Dey D., Slomka, P. J., Hahn, L. J. *et al*., Automatic three-dimensional multimodality registration using radionuclide transmission CT attenuation maps: a phantom study. *J Nucl Med* **40:** 448-455 (1999).
117. Kashiwagi T., Yutani, K., Fukuchi, M. *et al*., Correction of nonuniform attenuation and image fusion in SPECT imaging by means of separate X-ray CT. *Ann Nucl Med* **16:** 255-261 (2002).
118. Slomka P. J., Dey, D., Przetak, C. *et al*., Nonlinear image registration of thoracic FDG-PET and CT. [abstract] *J Nucl Med* **42:** 11P (2001).
119. Mullani N., Brandt, M., Strite, D. *et al*., Superimposition of EBCT determined coronary calcium deposits onto myocardial PET perfusion images by Rubidium-82 and Nitrogen-13 Ammonia for assessment of flow limiting defects. [abstract] *Clin Positron Imaging* **3:** 148 (2000).
120. Sinha S., Sinha, U., Czernin, J. *et al*., Noninvasive assessment of myocardial perfusion and metabolism: feasibility of registering gated MR and PET images. *AJR Am J Roentgenol* **164:** 301-307 (1995).
121. Suga K., Nishigauchi, K., Kume, N. *et al*., Superimposition of In-111 platelet SPECT and CT/MR imaging in intracardiac thrombus. *Clin Nucl Med* **21:** 595-601 (1996).

122. Schindler T. H., Magosaki, N., Jeserich, M. *et al*., 3D assessment of myocardial perfusion parameter combined with 3D reconstructed coronary artery tree from digital coronary angiograms. *Int J Card Imaging* **16:** 1-12 (2000).
123. Makela T., Clarysse, P. and Lotjonen, J., ''A New method for the registration of cardiac PET and MR images using deformable model based segmentation of the main thorax structures.'' MICCA2001, Vol. LNCS 2208; *pp* 557-564 (2001).
124. Aladl U. E., Dey, D. and Slomka, P. J., ''Four dimensional multi-modality image registration applied to gated SPECT and gated MRI.'' Medical Imaging 2003: Image Processing, San Diego, CA, Vol. 5032; *pp* 1166-1175 (2003).
125. Beanlands R. S., Ruddy, T. D., deKemp, R. A. *et al*., Positron emission tomography and recovery following revascularization (PARR-1): the importance of scar and the development of a prediction rule for the degree of recovery of left ventricular function. *J Am Coll Cardiol* **40:** 1735-1743 (2002).
126. Takahashi Y., Murase, K., Mochizuki, T. *et al*., Segmented attenuation correction for myocardial SPECT. *Ann Nucl Med* **18:** 137-143 (2004).
127. Zaidi H., Montandon, M.-L. and Slosman, D. O., Magnetic resonance imaging-guided attenuation and scatter corrections in three-dimensional brain positron emission tomography. *Med Phys* **30:** 937-948 (2003).
128. Dunphy M., Freiman, A., Larson, S. M. *et al*., Detecting F-18 FDG in the coronary arteries, aorta,carotids and illiac vessels; comparison to vascular calcification [abstract]. *J Nucl Med* **44:** 58P (2003).
129. Bocher M., Balan, A., Krausz, Y. *et al*., Gamma camera-mounted anatomical X-ray tomography: technology, system characteristics and first images. *Eur J Nucl Med* **27:** 619-627 (2000).
130. Patton J. A., Delbeke, D. and Sandler, M. P., Image fusion using an integrated, dual-head coincidence camera with X-ray tube-based attenuation maps. *J Nucl Med* **41:** 1364-1368 (2000).
131. Beyer T., Townsend, D. W., Brun, T. *et al*., A combined PET/CT scanner for clinical oncology. *J Nucl Med* **41:** 1369-1379 (2000).
132. Cohade C. and Wahl, R. L., Applications of positron emission tomography/computed tomography image fusion in clinical positron emission tomography-clinical use, interpretation methods, diagnostic improvements. *Semin Nucl Med* **33:** 228-237 (2003).
133. Keidar Z., Israel, O. and Krausz, Y., SPECT/CT in tumor imaging: technical aspects and clinical applications. *Semin Nucl Med* **33:** 205-218 (2003).
134. Cai J., Chu, J. C., Recine, D. *et al*., CT and PET lung image registration and fusion in radiotherapy treatment planning using the chamfer-matching method. *Int J Radiat Oncol Biol Phys* **43:** 883-891 (1999).
135. Yu J. N., Fahey, F. H., Harkness, B. A. *et al*., Evaluation of emission-transmission registration in thoracic PET. *J Nucl Med* **35:** 1777-1780 (1994).
136. Yu J. N., Fahey, F. H., Gage, H. D. *et al*., Intermodality, retrospective image registration in the thorax. *J Nucl Med* **36:** 2333-2238 (1995).
137. Wahl R. L., Quint, L. E., Cieslak, R. D. *et al*., ''Anatometabolic'' tumor imaging: fusion of FDG PET with CT or MRI to localize foci of increased activity. *J Nucl Med* **34:** 1190-1197 (1993).
138. Erdi A. K., Hu, Y. C. and Chui, C. S., ''Using mutual information (MI) for automated 3D registration in the pelvis and thorax region for radiotherapy treatment planning.'' Medical Imaging 2000: Image Processing, San Diego, CA, Vol. 3979; *pp* 416-425 (2000).

139. Hamilton R. J., Blend, M. J., Pelizzari, C. A. *et al*., Using vascular structure for CT-SPECT registration in the pelvis. *J Nucl Med* **40:** 347-351 (1999).
140. Tai Y.-C., Lin, K. P., Hoh, C. K. *et al*., Utilization of 3-D elastic transformation in the registration of chest X-ray CT and whole body PET. *IEEE Trans Nucl Sci* **44:** 1606-1612 (1997).
141. Rueckert D., Sonoda, L. I., Hayes, C. *et al*., Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging* **18:** 712-721 (1999).
142. Skalski J., Wahl, R. L. and Meyer, C. R., Comparison of mutual information-based warping accuracy for fusing body CT and PET by 2 methods: CT mapped onto PET emission scan versus CT mapped onto PET transmission scan. *J Nucl Med* **43:** 1184-1187 (2002).
143. Slomka P. J., Dey, D., Przetak, C. *et al*., Automated 3-dimensional registration of stand-alone (18)F-FDG whole-body PET with CT. *J Nucl Med* **44:** 1156-1167 (2003).
144. Hamilton R. J., Sweeney, P. J., Pelizzari, C. A. *et al*., Functional imaging in treatment planning of brain lesions. *Int J Radiat Oncol Biol Phys* **37:** 181-188 (1997).
145. Nestle U., Walter, K., Schmidt, S. *et al*., 18F-deoxyglucose positron emission tomography (FDG-PET) for the planning of radiotherapy in lung cancer: high impact in patients with atelectasis. *Int J Radiat Oncol Biol Phys* **44:** 593-597 (1999).
146. Erdi Y. E., Rosenzweig, K., Erdi, A. K. *et al*., Radiotherapy treatment planning for patients with non-small cell lung cancer using positron emission tomography (PET). *Radiother Oncol* **62:** 51-60 (2002).
147. Mah K., Caldwell, C. B., Ung, Y. C. *et al*., The impact of (18)FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study. *Int J Radiat Oncol Biol Phys* **52:** 339-350 (2002).
148. Schmidt S., Nestle, U., Walter, K. *et al*., Optimization of radiotherapy planning for non-small cell lung cancer (NSCLC) using 18FDG-PET. *Nuklearmedizin* **41:** 217-220 (2002).
149. Nishioka T., Shiga, T., Shirato, H. *et al*., Image fusion between 18FDG-PET and MRI/CT for radiotherapy planning of oropharyngeal and nasopharyngeal carcinomas. *Int J Radiat Oncol Biol Phys* **53:** 1051-1057 (2002).
150. Mutic S., Malyapa, R. S., Grigsby, P. W. *et al*., PET-guided IMRT for cervical carcinoma with positive para-aortic lymph nodes-a dose-escalation treatment planning study. *Int J Radiat Oncol Biol Phys* **55:** 28-35 (2003).
151. Munley M. T., Marks, L. B., Scarfone, C. *et al*., Multimodality nuclear medicine imaging in three-dimensional radiation treatment planning for lung cancer: challenges and prospects. *Lung Cancer* **23:** 105-114 (1999).
152. Perez C. A., Bradley, J., Chao, C. K. *et al*., Functional imaging in treatment planning in radiation therapy: a review. *Rays* **27:** 157-173 (2002).
153. Nelson S. J., Day, M. R., Buffone, P. J. *et al*., Alignment of volume MR images and high resolution [18F]fluorodeoxyglucose PET images for the evaluation of patients with brain tumors. *J Comput Assist Tomogr* **21:** 183-191 (1997).
154. Heidemann R. M., Ozsarlak, O., Parizel, P. M. *et al*., A brief review of parallel magnetic resonance imaging. *Eur Radiol* **13:** 2323-2337 (2003).

155. Hanley J., Debois, M. M., Mah, D. *et al*., Deep inspiration breath–hold technique for lung tumors: the potential value of target immobilization and reduced lung density in dose escalation. *Int J Radiat Oncol Biol Phys* **45:** 603-611 (1999).
156. Mah D., Hanley, J., Rosenzweig, K. E. *et al*., Technical aspects of the deep inspiration breath–hold technique in the treatment of thoracic cancer. *Int J Radiat Oncol Biol Phys* **48:** 1175-1185 (2000).
157. Przetak C., Baum, R. and Slomka, P., Image fusion raises clinical value of PET. *Diagnostic Imaging Europe* **5:** 10-15 (2001).
158. Rousset O. G., Ma, Y., Léger, G. *et al*., ''Correction for partial volume effects in PET using MRI-based 3-D simulations of individual human brain metabolism.'' in: *Quantification of Brain Function. Tracer Kinetics and Image Analysis in Brain PET*, edited by R. Myers, Cunningham, V., Bailey, DL, Jones, T Elsevier Science Publishers B. V., Amsterdam, (1993), *pp* 113-123.
159. Labbe C., Froment, J. C., Kennedy, A. *et al*., Positron emission tomography metabolic data corrected for cortical atrophy using magnetic resonance imaging. *Alzheimer Disease and Associated Disorders* **10:** 141-70 (1996).
160. Gindi G., Lee, M., Rangarajan, A. *et al*., Bayesian reconstruction of functional images using anatomical information as priors. *IEEE Trans Med Imaging* **12:** 670-680 (1993).
161. Ardekani B. A., Braun, M., Hutton, B. F. *et al*., Minimum cross-entropy reconstruction of PET images using prior anatomical information. *Phys Med Biol* **41:** 2497-2517 (1996).
162. Som S., Hutton, B. F. and Braun, M., Properties of minimum cross-entropy reconstruction of emission tomography with anatomically based prior. *IEEE Trans Nucl Sci* **45:** 3014-3021 (1998).
163. Baete K., Nuyts, J., Van Paesschen, W. *et al*., Anatomical-based FDG-PET reconstruction for the detection of hypo-metabolic regions in epilepsy. *IEEE Trans Med Imaging* **23:** 510-519 (2004).
164. Beekman F. J., de Jong, H. W. and van Geloven, S., Efficient fully 3-D iterative SPECT reconstruction with Monte Carlo-based scatter compensation. *IEEE Trans Med Imag* **21:** 867-877 (2002).
165. Picard Y. and Thompson, C. J., Motion correction of PET images using multiple acquisition frames. *IEEE Trans Med Imaging* **16:** 137-144 (1997).
166. Hutton B. F., Kyme, A. Z., Lau, Y. H. *et al*., A hybrid 3-D reconstruction/registration algorithm for correction of head motion in emission tomography. *IEEE Trans Nucl Sci* **49:** 188-194 (2002).
167. Kyme A. Z., Hutton, B. F., Hatton, R. L. *et al*., Practical aspects of a data-driven motion correction approach for brain SPECT. *IEEE Trans Med Imaging* **22:** 722-729 (2003).

10
Image Segmentation Techniques in Nuclear Medicine Imaging

A.O. Boudraa* and H. Zaidi†

1. Introduction

Recent advances in nuclear medicine imaging systems design has resulted in significant improvements in the areas of anatomical, functional, and dynamic imaging procedures. With these developments, computer-aided diagnosis is becoming a reality. These computer-based tools allow physicians to understand and diagnose human disease through virtual interaction. The role of medical imaging is not limited to visualization and inspection of anatomic structures, but goes beyond that to patient diagnosis, advanced surgical planning, simulation, and radiotherapy planning. In order to identify and quantify the information gathered from the images, regions (also called regions of interest – RoIs) must be precisely delineated and separated out for processing. This process, called *image segmentation*, is the most critical step for automated analysis in medical imaging.[1-3]

Segmentation is the first essential and important step of low level vision. This process partitions the image I into non-intersecting subregions, $R_1, R_2, \ldots, R_C$ such that each region is homogeneous and the union of two non-adjacent regions is homogeneous. A point x_k in a region R_i is connected to x_l if there is a sequence $\{x_k, \ldots, x_l\}$ such that x_j and x_{j+1} are connected and all the points are in R_i. R_i is a connected region if the set of points in R_i has the property that every pair of points is connected. Formally segmentation can be defined as follows:

(a) $\bigcup_{i=1}^{C} R_i = I$;
(b) R_i is a connected region, $i = 1, 2, \ldots, C$;
(c) $R_i \bigcap R_j = \emptyset \ \forall i, j, i \neq j$;
(d) $P(R_i) = \text{TRUE}$ for $i = 1, 2, \ldots, C$;
(e) $P(R_i \cup R_j) = \text{FALSE}$ for $i \neq j$.

*Dr A.O. Boudaa, Ecole Navale, Département Signal, Lanvéoc Poulmic BP600, F 29240 Brest, France
†PD Dr H. Zaidi, Geneva University Hospital, Division of Nuclear Medicine, 1211 Geneva, Switzerland

where P ( ) is a uniformity (homogeneity) predicate and C is the number of regions. A set of regions satisfying conditions (a) through (e) is known as a *partition*. Condition (a) indicates that the segmentation must be complete; i.e. every pixel must be in a region. Condition (b) requires that points in a region must be connected. Condition (c) indicates that the regions must be disjoint. Condition (d) deals with the properties that must be satisfied by the pixels in a segmented region; for example $P(R_i) = \text{TRUE}$ if all pixels in R_i have the same grey level. Finally, condition (e) indicates that regions R_i and R_j are different in the sense of predicate P ( ). In the standard approach, most segmentation algorithms assume that a given pixel or voxel belongs to a single cluster or tissue type (condition (c)). However, the complexity of tissue boundaries causes many pixels to contain a mixture of surrounding tissues (partial volume effect). Thus, image ambiguity within pixels is due to the possible multi-valued levels of brightness in the image. This indeterminacy is due to inherent vagueness rather than randomness. Since image regions (particularly nuclear medicine images) are not always crisply defined, it becomes more convenient and appropriate to view them as fuzzy subsets of the image.[4-7] Thus, the condition (c) is in general not verified:

$$(c') \bigcap_i R_i \neq \emptyset \tag{1}$$

The fuzzy subsets are characterized by the fuzzy membership degree associated to each pixel to belong to one of them. One popular method for assigning multi-subset membership value to pixels, for segmentation or other image processing tasks is the Fuzzy C-Means (FCM).[8]

2. Image Segmentation in Medical Imaging

Medical image segmentation is becoming an increasingly important image processing step for a number of clinical applications including: (a) identification of RoIs such as lesions to measure their volume and thus assess response to therapy;[5] (b) detection of the left ventricle (LV) cavity to determine the ejection fraction;[4,6] (c) volume visualization and quantification of organ uptake[9,10] or uptake defect of the tracer in the myocardium;[7] (d) study of motion or conduction abnormalities of the heart;[11] and (e) attenuation correction in emission tomographic imaging.[12,13] All subsequent interpretation tasks like feature extraction, object recognition, and classification depend largely on the quality of the segmentation output. The level to which the segmentation is carried depends on the problem being solved. That is, segmentation should stop when the regions of interest for a specific application have been isolated. For example, in automated detection of brain lesions using PET images,[5] interest lies in accurate delineation of the lesions with the objective of assessing the true size and the tracer uptake of these lesions. In this case, brain structures such as grey matter and white

matter are oversegmented and thus not well delineated. Segmentation in medical imaging is generally considered as a very difficult problem. This difficulty mainly arises due to the sheer size of the data sets coupled with the complexity and variability of the anatomic organs. The situation is worsened by the shortcomings of these imaging modalities, such as sampling artefacts, noise, and low contrast which may cause the boundaries of anatomical structures to be indistinct and disconnected. Thus, the main challenge of the segmentation process is to accurately extract the contours of the organ or RoI and separate it out from the rest of the data sets.

3. Segmentation Techniques in Nuclear Medicine Imaging

PET and SPECT images give information about the volume distribution of biologically significant radiotracers. The interest lies in the diagnosis of abnormalities in the biochemical process of interest, e.g. metabolic activity. Unfortunately, these images are inherently noisy and provide less quantitative information about the qualitative features discerned by experienced observers. In order to identify and quantify nuclear medicine image information, image regions must be delineated.[14] Compared to MR images, where brain structures such as CSF or grey matter can relatively be easily identified,[15] segmentation of nuclear medicine images is more difficult. Indeed, in addition to the inherently poor spatial resolution and SNR, there is additional bias introduced by physical degrading effects such as scatter, attenuation, partial volume effect, and patient motion during scanning. Several approaches have been proposed for image segmentation,[16,17] but only a few of them can satisfactorily be applied to nuclear medicine images. Image segmentation algorithms generally are based on one of two basic properties of intensity values: *discontinuity* and *similarity*.[17] In the first category, the basic approach is to partition an image based on abrupt changes in intensity, e.g. edges.[18] The main approaches in the second category are based on partitioning an image into regions that are similar according to a set of predefined criteria.[17] Thresholding, clustering, region growing are examples of methods belonging to this category. Segmentation is based on sets of feature that can be extracted from the images such as pixel intensities, which in turn can be used to calculate other features such as edges, and texture. Many segmentation approaches of nuclear images use the grey scale values of the pixels. In this case, pixel intensity input come from a single image. Single image segmentation methods are described in the following sections.

3.1 Thresholding Approaches

Thresholding is one of the simplest and most popular techniques for image segmentation. It can be performed based on global information (e.g. grey level histogram of the entire image) or using local information (e.g. co-occurrence matrix) of the image.[16] Automatic thresholding is important in applications where the speed or the physical conditions prevent interactive human selection of thresholds.[19] If only one threshold is used for the entire image, the process is called *global thresholding*. On the other hand, when the image is partitioned into several subregions and a threshold determined for each of the subregions, it is referred to as *local thresholding*. Techniques of thresholding can also be classified as *bilevel thresholding* and *multithresholding*. In *bilevel thresholding*, the histogram of the image is usually assumed to have one valley between two peaks, respectively. There are various solutions to the problem of locating the intensity threshold that ideally represents the bottom of this sometimes elusive histogram valley. The following strategy is an example of a thresholding method to obtain automatically the optimum threshold value:

1. Select an initial estimate for threshold T;
2. Segment the image using T. This will provide two regions of pixels: R_1 consisting of all pixels with grey level values $>$T and R_2 consisting of pixels with values $<$T;
3. Compute the average grey level values μ_1 and μ_2 for the pixels regions R_1 and R_2;
4. Compute a new threshold value;

$$T = (\mu_1 + \mu_2)/2 \tag{2}$$

5. Repeat Step 2 through 4 until the mean values μ_1 and μ_2 in successive iterations do not change.

Note that the threshold T can be specified using a heuristic technique based on visual inspection of the histogram but this approach is operator-dependent. If the image is noisy, the selection of the threshold is not trivial. Thus, more sophisticated methods have been proposed. Let $p(t) = n_p/N_I$ be the normalized histogram – brightness probability density function (*pdf*). N_I is the number of pixels in the observed image, I, and $n_p(t)$ is the number of pixels of I having grey level value t. We assume that the distribution of grey values in each region of I follows a Gaussian distribution. The observed histogram is a mixture of grey values of pixels of the object(s) and pixels of the background. The histogram may be considered as an estimate of the brightness *pdf*, say, p(t). Then p (t) is a mixture *pdf* given by $p(t) = P_1 \times p_1(t) + P_2 \times p_2(t)$ where $p_1(t) = \sum_{i=0}^{t} p(i)$ and $p_2(t) = 1 - p_1(t)$

are the *pdf*s of the *object* and the *background*, respectively, and P_1 and P_2 are their *a priori* probabilities such that $P_1 + P_2 = 1$. For a Gaussian case, p (t) is written as follows:

$$p(t) = \sum_{i=1}^{2} \frac{P_i}{\sigma_i \sqrt{2\pi}} \exp\left[\frac{(t-\mu_i)^2}{2\sigma_i^2}\right] \tag{3}$$

where μ_i and σ_i are the *a priori* mean values and standard deviations of the two distribution probabilities ($\mu_1 < \mu_2$). Thus, the mixture has five unknown parameters. If all the parameters are known, the optimal threshold is easily determined. A threshold, T, may be defined so that all pixels with grey values below T are considered *object* and those above T are considered *background* pixels. The probability of erroneously classifying a *background* point as an *object* point is $Er_1(T) = \int_{-\infty}^{T} p_2(z)dz$. Similarly, the probability of classifying an *object* point as *background* point is $Er_2(T) = \int_{T}^{+\infty} p_1(z)dz$. The overall probability of erroneously classifying a pixel from one of the two distributions as a pixel from the other distribution, is simply: $Er(T) = P_1 \times Er_1(T) + P_2 \times Er_2(T)$. The threshold for which $Er(T)$ is minimum is given by:

$$\begin{aligned} &dEr(T)/dT = 0 \Rightarrow P_1 \times p_1(T) = P_2 \times p_2(T) \Rightarrow \alpha T^2 + \beta T + \gamma = 0 \\ &\alpha = \sigma_1^2 - \sigma_2^2, \beta = 2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2), \gamma = \sigma_1^2\mu_2^2 - \sigma_2^2\mu_1^2 + 2\sigma_1^2\sigma_2^2 \log\left(\frac{\sigma_2 P_1}{\sigma_1 P_2}\right) \end{aligned} \tag{4}$$

The determination of parameters is not a simple matter and often requires numerical solution of a number of simultaneous non-linear equations. If the variances are assumed to be equal, T is given by: $T = (\mu_1 + \mu_2)/2 + \sigma_2/(\mu_1 - \mu_2)\log(P_2/P_1)$. If $P_1 = P_2$, the optimal threshold is simply the average of the means: $T = (\mu_1 + \mu_2)/2$. In this case, the above proposed strategy is found. A well known method for *global* thresholding using the histogram is Otsu's method.[20] This method chooses the optimal threshold by maximizing the *a posterior* between-class variance: $\sigma_B^2(t) = p_1(t)[\mu_1(t) - \mu_0]^2 + p_2(t)[\mu_2(t) - \mu_0]^2 = \frac{[\mu_0 p_1(t) - \mu_1(t)]^2}{p_1(t)[1 - p_1(t)]}$. The optimal threshold T is found by a sequential search for the maximum of $\sigma_B^2(t)$ for values of t where $0 < p_1(t) < 1$. Otsu's method gives satisfactory results when the number of pixels in classes are similar. Reddi *et al.*[21] proposed a method based on the same assumption by selecting a threshold T so that the *a posterior* between-class variance of *object* and *background* regions is maximized. They have shown that the between-class variance $\sigma_B^2(t)$ has a unique maximum and established the condition for this maximum.[21] Thus, $\sigma_B^2(t)$ can be written as follows:

$$\sigma_B^2 = \left(\sum_{z=0}^{t} zp(z)\right)^2 \times (p_1(t))^{-1} + \left(\sum_{z=t+1}^{N-1} zp(z)\right)^2 \times (p_2(t))^{-1} - \mu_0^2 \tag{5}$$

Setting $\partial\sigma_B^2(t)/\partial t = 0$, the following relation is found: $\mu_1(T) + \mu_2(T) = T$, where $\mu_1(T)$ and $\mu_2(T)$ are the mean values below and above the threshold

T. μ_0 is the average brightness. Finding the value of T that satisfies the above relation can be accomplished by an exhaustive sequential search, giving the same result as Otsu's method. Starting with threshold $T = \mu_0$, fast convergence is obtained. Kittler and Illingworth[22] proposed a method that minimizes the Kullback-Leibler (KL) distance which measures the discrepancy D between the probability distribution of the histogram and its approximation:

$$D = (1 + \log 2\pi)/2 - \sum_{i=1}^{2} p_i(t) \log p_i(t) + \left(p_1(t) \log \sigma_1^2 + p_2(t) \log \sigma_2^2\right)/2 \quad (6)$$

$p_i(T)$ is the estimated probability that pixel value falls in the i^{th} region. The threshold T should be chosen to minimize the information measure H. The method of Pun[23,24] is based on the maximum entropy criterion and has been examined in detail by Kapur *et al.*[25] It relies on the assumption that the threshold T corresponds to the maximum value of the sum of the two within-region entropies: $T = \arg\max_t [H_1(t) + H_2(t)]$ where $H_i(t)$ is the Shannon entropy of the i^{th} region. The $H_i(t)$ are given by:

$$H_1(t) = -\sum_{g=0}^{t} \frac{p(g)}{p_1(t)} \log \frac{p(g)}{p_1(t)} \text{ and } H_2(t) = -\sum_{g=t+1}^{N-1} \frac{p(g)}{p_2(t)} \log \frac{p(g)}{p_2(t)} \quad (7)$$

Thresholding is the most common approach used in segmenting SPECT images.[9,26-31] Figure 1(a) shows a transverse PET cardiac slice through the LV of a resting patient. The thresholding result using the method proposed by Reddi *et al.*[21] (T = 104) is shown in Fig. 1(b).

When the image is composed of several regions (clusters) one needs several thresholds for segmentation. Let the image be described by a

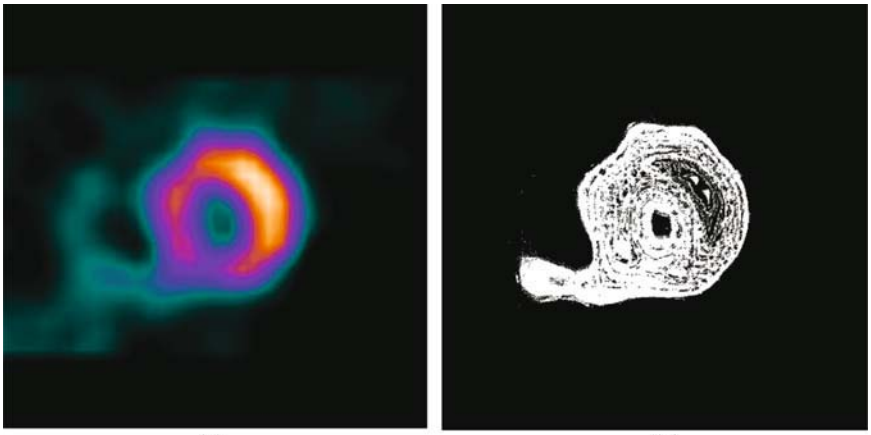


FIGURE 1. Illustration of Reddi *et al.*[21] thresholding method. (a) PET cardiac image. (b) Thresholded image.

two-dimensional function I (x, y), where (x, y) denotes the spatial coordinates and I (x, y) the feature value of pixel at location (x,y). In *multithresholding* one tries to get a set of thresholds $\{t_1, t_2, \ldots, t_K\}$ such that all pixels with $I(x, y) \in [t_i, t_{i+1}], i = 0, 1, \ldots, K$ constitute the i^{th} region type $t_0 = 0$, $t_{K+1} = G - 1$. G is the number of grey levels. There are various methods for *multithresholding*. The method of Reddi *et al.*[21] is fast and stands for a version extended to *multithresholding* of Otsu's method.[20] Wang and Haralick[32] presented a recursive technique for automatic multiple threshold selection. The local properties of the image (edges) are used to find the thresholds. The method of Carlotto[33] determines thresholds by handling the information derived from the changes of zero-crossing in the second derivatives. The method of Papamarkos and Gatos[34] is based on a combination of hill-clustering algorithm and an appropriate linear programming approximate technique. Recently, Fleming and Alaamer[31] have proposed case-specific volume-of-interest segmentation rules. The iterative algorithm calculates local context sensitive thresholds along radial profiles from the centre of gravity of the object.

Figure 2 shows an illustration of the multithresholding method of Reddi *et al.*[21] Figure 2(a) is obtained with two thresholds. Note that the image is not well segmented because the pattern of tracer uptake is not discriminated. Figure 2(b) shows the segmentation result of five thresholds where the pattern of tracer uptake is detected. Figure 2(c) shows the result of eight threshold values corresponding to an over-segmentation of the image.

3.2 Region Growing Approaches

The main drawback of histogram-based region segmentation is that the histogram provides no spatial information. Region growing approaches exploit the important fact that pixels which are close to each other have

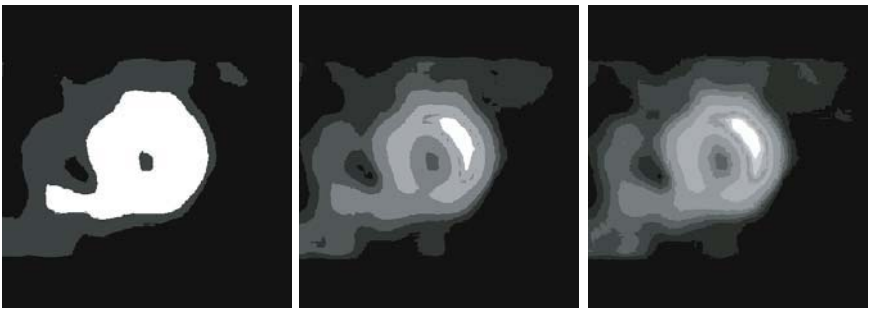


FIGURE 2. Illustration of the Reddi *et al.*[21] multithresholding method. Segmentation of the PET cardiac image shown in Fig. 1(a) with (a) two, (b) five and (c) eight thresholds, respectively.

similar grey values. Region growing is a bottom-up procedure which starts with a set of seed pixels. The aim is to grow a uniform, connected region from each seed. The user selects a pixel or a set of pixels and a region grows out from that (these) seed(s) until some criteria for stopping growth are met. A pixel is added to a region if and only if:

- it has not been assigned to another region;
- it is a neighbour of that region;
- the new region created by addition of the new pixel is still uniform.

The algorithm can be summarised by the following pseudo-code:

Let I be an image, and $R_1, R_2, \ldots, R_C$ a set of regions each consisting of a single seed pixel

Repeat
  For $l = 1, \ldots C$
    For each pixel k at the border R_l
      For all neighbours of pixel k
        Let x,y be the neighbour's coordinates
          Let μ_l be the mean grey level of pixels in R_l
            If neighbour unassigned and $|I(x,y) - \mu_l| \leq \varepsilon$
              Add neighbour to R_l
            Update μ_l
Until no more pixels are being assigned to regions.

The main assumption of the region growing approach is that regions are nearly constant in image intensity. Region growing has been used by Van Laere *et al.*[35] in the anatomic standardization and comparison technique with normal templates of brain SPECT studies. An automated interpretation of myocardial perfusion SPECT using region growing has been reported by Slomka *et al.*[36] The study results show that template-based region growing is a robust technique for volumetric quantification and localization of abnormal regions.

3.3 Classifiers

During the segmentation process, labels are affected to regions. Classification is a pattern recognition technique that deals with association of classes with tissue types, e.g. white matter, grey matter, CSF, ... etc. This step is also called *labelling*. Labels association can be performed by an operator or by an interpreting physician. Thus, pre-segmented images called *training data* are required. Classification seeks to partition *feature space* derived from the image using known labels. The feature space is the range of p-dimensional feature vectors formed at each pixel. The *features* could include pixel intensity, the gradient at a given pixel and so on. The simplest form of classifiers is the Nearest-Neighbour Classifier (NNC) where each pixel is classified in the

same class as the training datum with closest intensity. Examples of classifiers are the k-Nearest Neighbour (k-NN), the Nearest Mean Classifier (NMC),[37] Fisher Linear Discriminant (FLD),[38] Parzen classifier[39] and Support Vector Machine (SVM).[40] The k-NN classifier is the generalization of the NNC, where the pixel is classified according to the majority of the k closest training data. In Parzen window, classification is made according to the majority vote within a predefined window of the feature space centred at the unlabeled pixel (mapped to feature space). SVM is a supervised learning method which is suitable for high dimensional data.[40] Given training examples labelled either ''yes'' or ''no'', a maximum margin hyperplane splits the ''yes'' and ''no'' training examples such that the distance from the closest examples (the margin) to the hyperplane is maximized. The use of the maximum margin hyperplane is motivated by statistical learning theory, which provides a probabilistic test error bound that is minimized when the margin is maximized. If there exists no hyperplane that can split the ''yes'' and ''no'' examples, an SVM will choose a hyperplane that splits the examples as cleanly as possible, while still maximizing the distance to the nearest cleanly split examples. NMC calculates the mean of the training vectors for each class while searching and classification are done only against the mean vector for each class. FLD classifier projects high-dimensional data onto a line and performs classification in this one-dimensional space. The projection maximizes the distance between the means of the two classes while minimizing the variance within each class. FLD and NMC separate the object (image) with one hyperplane in feature space into two classes. The discriminant function of NMC is defined as follows:

$$f_{NMC}(x) = (x - \mu_2)^t(x - \mu_2) - (x - \mu_1)^t(x - \mu_1) \tag{8}$$

where x is the object to be classified and μ_1 and μ_2 are the means of the feature vectors in the training set for the classes ''1'' and ''2'', respectively. It spans up an equidistant hyperplane between both class means. Contrarily to NMC, the FLD classifier takes both class means and covariance, A, into account (assumed to be common to both classes):

$$f_{FLD}(x) = [x - (\mu_1 + \mu_2)/2]^t A^{-1}(\mu_1 - \mu_2) \tag{9}$$

When A is the identity matrix, NMC and FLD are equivalent. If the number of training samples is smaller than p, A is singular and pseudo FLD (PFLD) is formed by replacing A by its pseudo inverse, $(A^tA)^{-1}A^t$. Stoeckel *et al.*[41] proposed a method for classifying a single SPECT HMPAO image as being an image of a normal subject or a probable Alzheimer patient. The aim is to provide a tool to assist the interpretation of SPECT images for the diagnosis of dementia of Alzheimer type (DAT). PFLD and NMC are used to assign one of the possible class labels (DAT or normal) to a SPECT image given as an input.[41] The success rates of classification is estimated using a classification error, *leave one out*.[42] As reported by the authors, PFLD outperforms

NMC, which can be explained by the fact that PFLD takes the covariance into account, and thus the shape of the classes in feature space.[41] The differences in success rates between normal and DAT images might be explained by the presence of nearly two times more normal than DAT images.

3.4 Clustering-based Approaches

The limitation of thresholding techniques is that they only apply to a single-band image, such as greyscale image or a single band of a multi-band image. Data clustering is another effective technique for image segmentation. The advantage of clustering is that it can be applied to a multi-band image such as a colour image or image composed of multiple feature layers. The main disadvantage of clustering techniques is that the appropriate number of clusters should be determined beforehand.[43-45] Clustering algorithms essentially perform the same function as classifiers without the use of training data. Thus, they are called *unsupervised* methods. In order to compensate for the lack of training data, clustering methods iterate between segmenting the image and characterizing the properties of each class. In a sense, clustering methods train themselves using the available data. *Unsupervised* methods include the C-means,[46] the FCM algorithm,[8] and the expectation-maximization (EM) algorithm.[47] *Unsupervised* clustering has shown promise for tumour detection in PET imaging, but algorithm initialization is an important step for reliable clustering results and for reduction of computation time.[5] FCM and C-means are the most commonly used algorithms for nuclear medicine[4-6,48] and transmission[13] image segmentation.

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a finite data set and $C \geq 2$ an integer; n is the number of data points; x_k is a p-dimensional feature vector. R^{Cxn} denote the set of all real $C \times n$ matrices. A fuzzy C-partition of X is represented by a matrix $U = [\mu_{ik}] \in R^{Cxn}$, the entries of which satisfy

$$\begin{aligned} &\mu_{ik} \in [0,1]\ 1 \leq i \leq C;\ 1 \leq k \leq n \\ &\sum_{i=1}^{C} \mu_{ik} = 1;\ 1 \leq k \leq n \\ &0 < \sum_{k=1}^{n} \mu_{ik} < n;\ 1 \leq i \leq C \end{aligned} \tag{10}$$

U can be used to describe the cluster structure of X by interpreting μ_{ik} as the degree of membership of x_k to cluster i. Good partitions U of X may be defined by the minimization of the FCM objective functional:[8]

$$J_m(U,V;X) = \sum_{k=1}^{n} \sum_{i=1}^{C} (\mu_{ik})^m \, \| x_k - v_i \|_A^2 \tag{11}$$

where $m \in [1, +\infty)$ is a weighting exponent called the fuzzifier, $V = (v_1, v_2, \ldots v_c)$ is the vector of the cluster centres. $\| x \|_A = \sqrt{x^T Ax}$ is any inner product norm where A is any positive definite matrix. Approximate optimization of J_m by the FCM algorithm is based on iteration through the following necessary conditions for its local extrema.

FCM theorem.[8] Assume $m \geq 1$ and

$\| x_k - v_i \|_A^2 > 0$; $1 \leq i \leq C$; $1 \leq k \leq n$, (U, V) may minimize J_m only if:

$$\mu_{ik} = \left[\sum_{j=1}^{C} \left(\frac{\| x_k - v_i \|_A}{\| x_k - v_j \|_A} \right)^{\frac{2}{(m-1)}} \right]^{-1} \tag{12}$$

$$v_i = \frac{\sum_{k=1}^{n} (\mu_{ik})^m x_k}{\sum_{k=1}^{n} (\mu_{ik})^m} \tag{13}$$

The FCM algorithm consists of iterations alternating between Eqs. (12) and (13). This algorithm converges to either a local minimum or a saddle point of J_m.[8] Boudraa *et al.*[4] have segmented nuclear cardiac images with FCM using as features the spatial information (x, y) and grey level value I (x, y) (in this case $p = 3$). Recently, a fast version of the FCM was proposed.[49] This version is based on one dimensional attribute such as the grey-level. Let H_S be the histogram of I with G-levels. Each pixel has a feature that lies in the discrete set X. In the new formulation, FCM minimizes the following functional, which is very similar to that of Bezdek:[8]

$$J_m(U,V;X) = \sum_{g=0}^{N-1} \sum_{i=1}^{C} \left(\mu_{ig} \right)^m . H_S(g) . \| g - v_i \|_A^2 \tag{14}$$

The FCM only operates on the histogram and hence is faster than the conventional version which processes the whole data set. Thus, the computation of the membership degrees of $H_S(g)$ pixels is reduced to that of only one pixel with g as grey level value. The algorithm is outlined in the following steps:

⟨*FCM*0⟩ Find the maximum and the minimum values of H_S: G_{min} and G_{max}

⟨*FCM*1⟩ Fix the number of clusters C, $2 \leq C \leq G_{max}$, and the threshold ε;

⟨*FCM*2⟩ Find the number of occurrences, $H_S(g)$, of the level g; $g = G_{min}, \ldots, G_{max}$

⟨*FCM*3⟩ Initialize the membership degrees μ_{il} using the $(G_{max} - G_{min} + 1)$ grey levels such that: $\sum_{i=1}^{c} \mu_{ig} = 1$, $g = G_{min}, \ldots, G_{max}$

⟨*FCM*4⟩ Compute the centroid v_i as follows:

$$v_i = \frac{\sum_{g=G_{min}}^{G_{max}} \left(\mu_{ig}\right)^m . H_S(g).g}{\sum_{g=G_{min}}^{G_{max}} \left(\mu_{ig}\right)^m . H_S(g)} \quad i = 1, 2, \ldots, C \tag{15}$$

⟨*FCM5*⟩ Update the membership degrees

$$\tilde{\mu}_{ig} = \left[\sum_{j=1}^{c} \left(\frac{\| g - v_i \|_A}{\| g - v_j \|_A}\right)^{\frac{2}{(m-1)}}\right]^{-1} \tag{16}$$

⟨*FCM6*⟩ Compute the defect measure

$$E = \sum_{i=1}^{C} \sum_{g=G_{min}}^{G_{max}} \left|\tilde{\mu}_{ig} - \mu_{ig}\right| \tag{17}$$

If ($E > \varepsilon$)
{
$\mu_{ig} \leftarrow \tilde{\mu}_{ig}$
goto <*FCM*4>
}
⟨*FCM7*⟩ Defuzzification process.

Figure 3 illustrates an automated outlining of the LV contour in gated radionuclide ventriculography. The FCM is run at the left anterior oblique (LAO) projection providing the best separation of the left and right (RV) ventricles. The number of clusters is estimated during the clustering process using a cluster validity index.[6,11] The clustering is performed with ε set to 10^{-3} (Fig. 3b). The cluster formed by the LV and the RT (RT = RV + Atria) is separated from the background (Fig. 3c) followed by a labelling analysis of the connected components LV and RT. Figure 3d) shows that the LV is well separated. Since the LV cluster has the greatest area of the cardiac

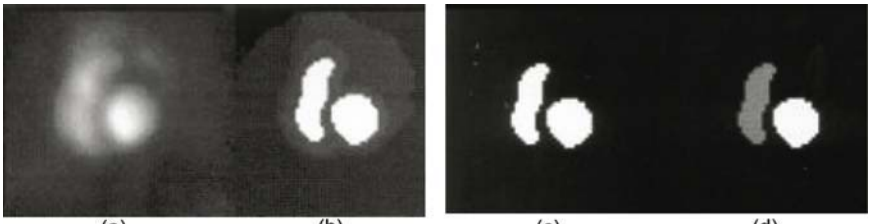


FIGURE 3. LV edge detection by FCM in LAO projection. (a) ED frame. (b) Fuzzy clustering result of the ED frame. (c) Separation of the LV and RT clusters from the background. (d) Labelling of the connected components LV and RT. (Reprinted with permission from ref.[6]).

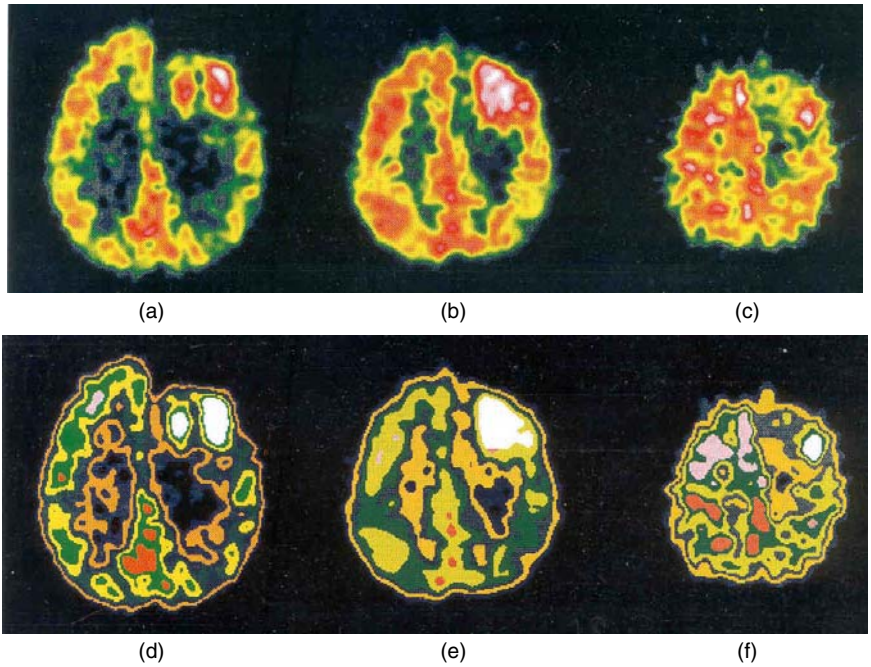


FIGURE 4. Segmentation of FDG-PET images of a patient obtained two weeks after surgery. (a)–(c) are the three slices where the tumour is evidenced by the CT. (d)–(f) are the corresponding segmented images. (Reprinted with permission from ref.[5]).

images sequence in End-diastolic (ED) phase, a framing operation is performed to obtain automatically, a ''box'' enclosing the LV cluster. Thus, the fuzzy clustering of the remaining images of the cardiac cycle is narrowed to this ''box'' with C set to 2 (LV and background).[6]

Figure 4 shows an application of the FCM to localization and quantitation of brain lesions using PET imaging.[5] The CT scans have shown that the brain tumour is only evidenced in three slices (Figs. 4a–c). The FCM is started with an overspecified number of clusters ($C = 20$), believed by the experts to make up the image being segmented. This value leads to an oversegmentation of soft tissues but reduces the chance that the lesion(s) is(are) clustered into classes that contain soft tissues. The fuzzy clustering is followed by a merging process.[5] The brain tumour is hypermetabolic. A careful examination of PET images (Figs. 4a–c) and the corresponding segmented images (Figs. 4d–f) shows that there is a good agreement between the two sets of images and that the tumour is correctly delineated. By comparing the segmented image and the original one, one may easily compute the area (in pixels) of the selected RoI and its total count using the corresponding label affected by the FCM. It has been noticed that the larger is the abnormality, the easier is its detection, even in slices of poor contrast. Nevertheless, a

small tumour which is not readily discerned from its surrounding region (because it has almost the same glucose metabolism) necessitates for its detection, a clustering with high precision.

It should be mentioned that FCM is sensitive to noise and outliers. To handle the problem of noise, the Possibilistic C-Means (PCM)[50] or Maximum Entropy Principle-based Fuzzy Clustering (MEPFC)[51] algorithms can be used. PCM assumes that the membership function of a point in a fuzzy set (or cluster) is absolute, i.e. it is an evaluation of a degree of typicality not depending on the membership values of the same point in other clusters. By contrast, clustering approaches including FCM and MEPFC impose a probabilistic constraint, according to which the sum of the membership values of a point in all the clusters must be equal to one. The PCM algorithm treats each cluster independently of the rest and thus can provide several identical prototypes while completely failing to detect some of other clusters. Furthermore PCM is heavily dependent on initialisation and its ''bandwidth parameter'', η, is difficult to estimate. It was also reported that this algorithm has the undesirable tendency to produce coincidental clusters.[52] Also, MEPFC depends heavily on initialisation and often generates coincidental clusters since the objective function of this algorithm is separable. In addition the ''scale parameter'', β, depends on the data set and usually has to be determined experimentally.[12] Even if FCM is more sensitive to noise than PCM and MEPFC, it does not require parameters such as η and β to run. Furthermore, FCM is less sensitive to initialisation than both PCM and MEPFC. For slightly noisy data, the performance of FCM is not disturbed. However, for high noisy data the performance of FCM is affected.

3.5 Edge Detection

Segmentation can be performed through edge detection of various image regions. Edges are formed at intersection of two regions where there are abrupt changes in grey level intensity values. Edge detection works well on images with good contrast between regions. However, the detection is limited in regions with low contrast. Furthermore, it is difficult to find correlation between the detected edges of the RoIs. Since edges are local features, they are determined based on local information. There are different types of differential operators such as Roberts gradient, Sobel gradient, Prewitt gradient and the Laplacian operator.[17] While the first operators are called first difference operator, Laplacian is a second difference operator. An edge is marked if a significant spatial change occurs in the second derivative. The Laplacian of a 2-D function I(x, y) is defined by $G_r(x, y) = -\nabla^2\{I(x, y)\}$ where the Laplacian operator is defined as $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$.$G_r(x, y)$ exhibits a sign change at the point of inflection of I(x,y). The zero crossing of $G_r(x, y)$ indicates the presence of an edge. The digital second derivative form encountered most frequently in practice is given by:

$$\begin{aligned} \partial^2 I(x, y)/\partial x^2 &= I(x+1, y) + I(x-1, y) - 2I(x, y) \\ \partial^2 I(x, y)/\partial y^2 &= I(x, y+1) + I(x, y-1) - 2I(x, y) \end{aligned} \tag{18}$$

The digital implementation of the 2-D Laplacian is obtained by summing the two components:

$$\nabla^2 I(x, y) = I(x+1, y) + I(x-1, y) + I(x, y+1) + I(x, y-1) - 4I(x, y) \tag{19}$$

The Laplacian generally is not used in its original form for edge detection for several reasons:[17] it is not orientation invariant, and typically is unacceptably sensitive to noise owing to the small size of its template. The magnitude of the Laplacian produces doubles, an undesirable effect because it complicates segmentation. A good edge detector should be a filter with the following two features. First, it should be a differential operator, taking either a first or second partial derivative of the image. Second, it should be capable of being tuned to act at any desired scale, so that large filters can be used to detect blurry shadow edges, and small ones to detect sharply focused details.[16] An operator that satisfies the above conditions is the Laplacian of Gaussian (LoG) in which Gaussian-shaped smoothing is performed prior to application of the Laplacian.[53] Its formulation can be expressed in the continuous domain as follows:

$$G_r(x, y) = -\nabla^2 (I(x, y) \otimes H_\sigma(x, y)) = I(x, y) \otimes \left(-\nabla^2 (H_\sigma(x, y))\right) \tag{20}$$

where $H(x, y) = -\nabla^2 (H_\sigma(x, y)) = \frac{1}{\pi\sigma^4}\left[1 - \frac{x^2+y^2}{2\sigma^2}\right] \exp\left\{\frac{x^2+y^2}{2\sigma^2}\right\}$ is the impulse function of the operator. The position of the edges is present at zero-crossing in $G_r(x, y)$. $H_\sigma(x, y)$ is a Gaussian smoothing filter where σ is the spread of the Gaussian and controls the degree of smoothing. A discrete domain version of the LoG operator can be obtained by sampling the continuous domain of the impulse function $H(x, y)$ over the $W_F \times W_F$ window. In order to avoid deleterious truncation effects, the size of the array should be set such that $W_F = 3b$ or greater, where $b = 2\sqrt{2\sigma}$ is the width of the central excitatory region of the LoG operator. In practice, $G_r(x, y)$ is calculated by multiplying the original image by a template generated by sampling the impulse function $H(x, y)$. For the computation of $G_r(x, y)$, many researchers proposed approximation schemes to speed up the procedure.[53,54] As a result of various sources of error, it has been shown that zero crossing do not always lie at pixel sample points. Sub-pixel accuracy of LoG detector was discussed in ref.[54] To make LoG edge detection robust, zero-crossing can be combined with other measures to judge whether an edge exist or not.

Another popular algorithm is the Canny edge detector.[18] This operator was designed to be an optimal edge detector. Canny's approach is based on optimizing the trade-off between two performance criteria:

- good edge detection: the probabilities of failing to mark real edge points and marking false edge points should be low;

- good edge localization: the positions of edge points marked by the edge detector should be as close as possible to the real edge.

The optimization can be formulated by maximizing a function that is expressed in terms of:

- the signal-to-noise ratio of the image;
- the localization of the edges;
- a probability that the edge detector only produces a single response to each actual edge in an image.

The Canny edge detection algorithm is based on the following steps:

1) smooth the image with a Gaussian filter;
2) compute the gradient magnitude and orientation using finite-difference approximations for the partial derivatives;
3) apply nonmaxima suppression to the gradient magnitude;
4) perform hysteresis thresholding algorithm to detect and link edges.

The Canny edge detector approximates the operator that optimizes the product of signal-to-noise ratio and localization.

3.5.1 Smoothing

Let I(x, y) denote the image. Convolution of I(x, y) with $H_\sigma(x, y)$ gives an array of smoothed data: $S(x, y) = H_\sigma(x, y)I(x, y)$.

3.5.2 Gradient Calculation

Firstly, the gradient of the smoothed array S(x, y) is used to produce the x and y partial derivatives P(x, y) and Q(x, y), respectively:

$$\begin{aligned} P(x, y) &= [S(x, y+1) - S(x, y) + S(x+1, y+1) - S(x+1, y)]/2 \\ Q(x, y) &= [S(x, y) = [S(x, y) - S(x+1, y) + S(x, y+1) - S(x+1, y+1)]/2 \end{aligned} \tag{21}$$

The partial derivatives are computed by averaging the finite differences over the 2×2 square. From the standard formulas for rectangular-to-polar conversion, the magnitude and orientation of the gradient can be computed as:

$$M(x, y) = \sqrt{P^2(x, y) + Q^2(x, y)},\ \theta(x, y) = \tan^{-1}(Q(x, y), P(x, y)) \tag{22}$$

3.5.3 Nonmaxima Suppression (edge thinning)

Once the rate of intensity change at each point in the image is found, edges must be placed at the points of maxima; or rather non-maxima must be suppressed. A local maximum occurs at a peak in the gradient function M(x, y), or alternatively where the derivative of the gradient function is set to zero. However, in this case we wish to suppress non-maxima perpendicular

to the edge direction, rather than parallel to (along) the edge direction, since we expect continuity of edge strength along an extended contour. This assumption creates a problem at corners. Instead of performing an explicit differentiation perpendicular to each edge, another approximation is often used. Each pixel in turn forms the centre of a nine pixel neighbourhood. By interpolation of the surrounding discrete grid values, the gradient magnitudes are calculated at the neighbourhood boundary in both directions perpendicular to the centre pixel. If the pixel under consideration is not greater than these two values (i.e. non-maximum), it is suppressed.

3.5.4 Edge Thresholding

The Canny operator works in a multi-stage process. First of all the image is smoothed by Gaussian convolution Then a simple 2-D first derivative operator (somewhat like the Roberts Cross) is applied to the smoothed image to highlight regions of the image with high first spatial derivatives. Edges give rise to ridges in the gradient magnitude image. The algorithm then tracks along the top of these ridges and sets to zero all pixels that are not actually on the ridge top so as to give a thin line in the output, a process known as *non-maximal suppression*. The tracking process exhibits hysteresis controlled by two thresholds: T_1 and T_2, with $T_1 > T_2$. These two thresholds can be calculated based on noise estimates in the image, but they often are set by hand. Tracking can only begin at a point on a ridge higher than T_1. Tracking then continues in both directions out from that point until the height of the ridge falls below T_2. This hysteresis helps to ensure that noisy edges are not broken up into multiple edge fragments. The effect of the Canny operator is determined by three parameters: σ, T_1, and T_2. Increasing σ reduces the detector's sensitivity to noise at the expense of losing some of the finer details in the image. The error localization in the detected edges also increases slightly as σ is increased.

Figure 5 shows an example of Canny detector applied to blood pool cardiac image. Note that the LV edges are well delineated (Fig. 5f). Edge detection methods have been proposed to extract the edges of the LV in gated cardiac nuclear medicine images,[55,56] and for volume quantitation in SPECT.[28,57,58] Edge detection of gated SPECT images has been performed to estimate LV volume and left myocardial volume in hypertrophic cardiomyopathy patients.[59] Canny detector is used to segment SPECT images in order to analyze abnormal functional asymmetry of the prefrontal boys with attention deficit hyperactivity disorder.[60] Since nuclear medicine images are inherently noisy, the performance of edge detection algorithms is affected. In most cases, these algorithms are not used on their own for segmentation, but coupled with other segmentation algorithms to solve a particular segmentation problem. For example, Dai *et al.*[56] combined the Laplacian operator with mean field annealing method to delineate the LV edges.

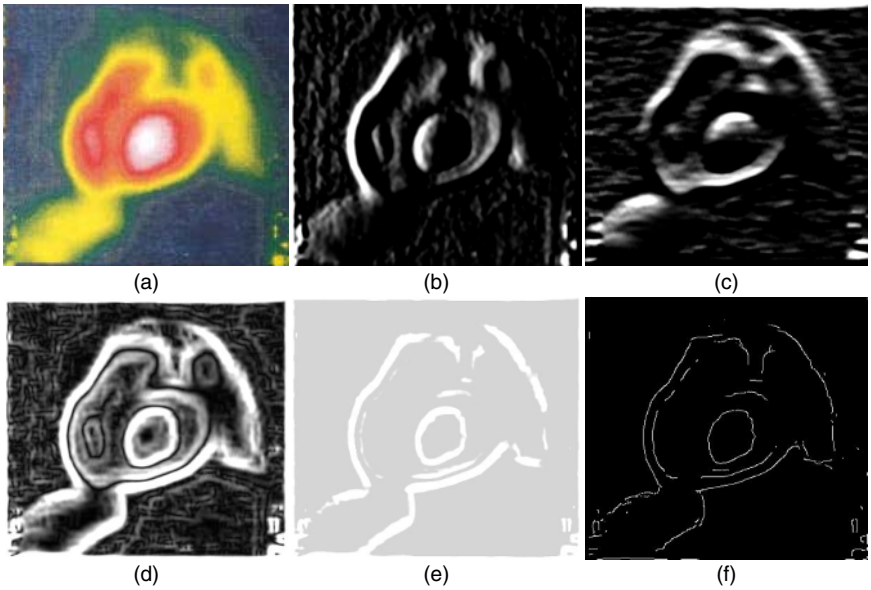


FIGURE 5. Canny edge detector applied to blood pool cardiac image. (a) Initial image. (b) Edge detection along X-axis direction. (c) Edge detection along Y-axis. (d) Norm of the gradient. (e) The norm of the gradient after thresholding. (f) Thinning.

3.6 Markov Random Field Models

Thresholding methods or clustering techniques such as FCM and C-means work well on noise-free images, with slow spatial variation in intensity, and non textured images. These methods typically do not take into account the spatial characteristics of the image, which render them sensitive to noise and thus particularly less reliable. Furthermore, spatial feature is an important piece of information for segmentation. It is advantageous to statistically model the noise and any texture which is random by nature. One may also take advantage of two-dimensional spatial ergodicity to average the effects of noise.[61] If a region is spatially ergodic then a pixel and its neighbours will have similar statistical properties. A natural way to incorporate spatial correlation into a segmentation process is to use Markov random fields (MRFs). MRFs are a class of statistical models that describe contextual constraints. They can be interpreted as a generalization of Markov chain models, which describe (unidirectional) temporal constraints. Assume that an image is defined on $M \times N$ lattice Ω indexed by (i, j) so that $\Omega = \{(i, j); 1 \leq i \leq M, 1 \leq j \leq N\}$. Each element of Ω is called *site*. The sites in Ω are related one to another via a *neighbourhood system*, which is defined as $N_v = \{N_{ij}, (i, j) \in \Omega\}$, where $N_{ij} \subseteq \Omega$ is the set of neighbouring sites (i, j), $(i, j) \in N_{ij}$ and $(i, j) \in N_{i'j'} \Leftrightarrow (i', j') \in N_{ij}$. The sites together with N_v

form an undirected graph, which is used to define the contextual constraints between site labellings. A site often represents a point or a region in the Euclidean space such as an image pixel or an image feature such as a corner point or a line segment. Let $Y = \{Y_{ij} = y_{ij};\, (i,j) \in \Omega\}$ be the observed grey level image where pixels take values from the set $\Lambda = \{1, 2, \cdots, L-1\}$. The observation image itself can be denoted by y. The image segmentation problem involves assigning to each pixel a class label taking a value from the set $B = \{1, 2, \cdots, C\}$ where C is the number of classes. Let $X = \{X_{ij};\, (i,j) \in \Omega\}$ be a family of random variables defined on the set Ω, in which each random variable X_{ij}, associated to the site $(i,j) \in \Omega$, takes a label value x_{ij} in B. Thus a label is an event that may happen to a site. The family X is called random field (or labels field). The notation $X_{ij} = x_{ij}$ denotes that the event takes the value x_{ij} and the notation $(X_{11} = x_{11}, \cdots, X_{MN} = x_{MN})$ denotes the joint event and is also a labelling of the sites in Ω in terms of the labels in B. A *configuration* of X, corresponding to a realization of the field, is noted by $x = \{x_{11}, x_{12}, \cdots, x_{MN}\}$. The probability that a random variable X_{ij} takes the value x_{ij} is denoted $P(X_{ij} = x_{ij})$. A random field X is said to be a MRF on Ω with respect to N_{ij} if

$$P(x_{ij}) > 0 \;\forall\; (i,j) \in \Omega \text{ and } P\left(x_{ij} | X_{\Omega - \{(i,j)\}}\right) = P\left(x_{ij} | X_{N_{ij}}\right) \tag{23}$$

where $X_{\Omega-\{(i,j)\}} = \{x_{kl};\, k = 1, 2, \cdots, i-1, i+1, \cdots, M;\, l = 1, 2, \cdots, j-1, j+1, \cdots, N\}$. $\Omega - \{(i,j)\}$ denotes the set of all sites excluding (i,j). $X_{\Omega-\{(i,j)\}}$ and $X_{N_{ij}}$ denote the random variables associated with the sites $\Omega - \{(i,j)\}$ and N_{ij}, respectively. A first order neighbourhood consists of the four nearest pixel sites; a second order neighbourhood consists of the eight nearest pixel sites, and so on. A clique is a set of one or more sites such that each site in the clique is a neighbourhood of all other sites in the clique. The segmentation problem can now be simply as follows. Given the corrupted data y, find an estimate $\tilde{x}$ of the true labelling configuration x^* (observed under ideal conditions). This inverse problem is ill-posed. Prior information is often very useful in formulating a regularized process so that optimal results can be obtained by solving the regularized problem.[62] A well known approach to this problem is the Bayesian estimation that incorporates prior information through an *a priori* distribution of the random field X, P (X). The Bayesian estimation can be formulated as a maximum *a posteriori* (MAP) estimation that maximizes the posterior probability or likelihood of $\tilde{x}$ given y. In particular if P ( ) is an appropriate probability measure, then one would like to find the estimate $\tilde{x}$ which maximizes $P(X = \tilde{x}|Y = y)$. The Bayes theorem is given by the following relation: $P(X = \tilde{x}|\, Y = y) = P(Y = y|X = \tilde{x}).P(X = \tilde{x})/P(Y = y)$ where $P(Y = y|X = \tilde{x})$ is the conditional probability of Y given the field X. Since $P(Y = y)$ is independent of $\tilde{x}$, we can maximize the relation $P(X = \tilde{x}|Y = y) \propto P(Y = y|X = \tilde{x}).\, P(X = \tilde{x})$ or $\log[(X = \tilde{x}|Y = y)] \propto \log[(Y = y|X = \tilde{x})] + \log[(X = \tilde{x})] + \log[(X = \tilde{x})]$. It

should be pointed out that the difficulty in maximizing this equation lies in the fact that it is a joint log-likelihood for all the image data. It does not simply describe the likelihood of a single pixel. In particular, the maximizing $\tilde{x}$ is one of $C^{M\times N}$ possibilities. The solution is given by

$$\hat{x} = \arg\max_{x\in X} \{\log P[(y|\tilde{x})] + \log[P(\tilde{x})]\} \tag{24}$$

The first term of Eq. 24 is the likelihood, telling how the data are obtained from the labels while the second is the prior probability of a given labelling. In practice, the first term forces fidelity to the data while the second penalizes unlikely trough labelling. Note that, without the *priori* distribution P(x), the estimation scheme becomes a maximum likelihood estimation (MLE). MLE is often used as the initial estimate in the iterative MAP estimation. For practical use, a means to specify the conditional probabilities is required. From the above equation, computation of the prior probability of the class and the likelihood probability of the observation is needed. Since X is assumed as a realization of a MRF, according to the Hammerseley-Clifford theorem,[63] the probability density of X is given by a Gibbs distribution having the form:[62,64]

$$P(X) = Z^{-1}\exp(-U(X)/T) = Z^{-1}\exp\left(-\sum_{c\in C_L} V_c(X)/T\right) \tag{25}$$

where Z is a normalizing constant, called the partition coefficient. T stands for "temperature" and controls the degree of peaking in the probability density, i.e. the larger the value, the larger is the peaking. U ( ) in Eq. 25 is an energy function, composed of potentials of the cliques $V_c(\ )$, which specifies the degree of penalty imposed on the neighbours. A clique potential specifies the admissibility of a specific labelling of the nodes of a clique. The value of $V_c(\ )$ depends on the local configuration of clique c. The set of cliques C_L comprises the completely connected subgraphs of the MRF graph. An example of potential functions is given by:

$$V_c(X) = \begin{cases} \beta & \text{if } x_{ij} = x_{i'j'}\ (i,j),\ (i',j') \in C_L \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

Assuming that the noise in the image, is additive, Gaussian, independent, and with zero mean and variance σ^2, $y_i = x_i + \eta$, the conditional density probability of $(y_i - x_i)$ is simply the probability of the noise

$$P(y_{ij}|x_{ij}) = \left(1/\sigma\sqrt{2\pi}\right)\exp\left(-(y_{ij} - x_{ij})^2/\sigma^2\right) \tag{27}$$

Assuming the noise is independent, we find for the entire image $P(y|x) = \Pi_{(i,j)\in\Omega} P(y_{ij}|x_{ij})$. Thus,

$$P(y|x) = \left(\Pi_{(i,j)\in\Omega}\sigma\sqrt{2\pi}\right)^{-1}\exp(-U(y|x)) \tag{28}$$

where $U(y|x) = \Sigma_{(i,j)\in\Omega}(y_i - x_i)^2/(2\sigma^2)$ is the likelihood energy. It is easy to show that $\log[P(x|y)] \propto -U(x|y)$ where $U(x|y) = U(y|x) + U(x) + \text{Const}$ (where Const is a constant) is the posterior energy. Finally, the MAP estimation is equivalent to minimizing the posterior energy function:

$$\hat{x} = \arg\min_{x\in X} \{U(y|x) + U(x)\} \tag{29}$$

Although mathematically simple, finding a global minimum for MAP estimation, given all the configurations x is a difficult task. Therefore, optimal solutions are usually computed using iterative optimization methods. A well known method is the iterated conditional modes (ICM) algorithm of Besag,[63] which uses the "greedy" strategy in the iterative local minimization and convergence is guaranteed after only a few iterations. Given the image y and labels $x^k_{\Omega-\{(i,j)\}}$ ICM sequentially updates each $x^{(k)}_{ij}$ into $x^{(k+1)}_{ij}$ by minimizing $U\left(x_{ij}|y, x_{\Omega-\{(i,j)\}}\right)$, the conditional posterior probability with respect to x_{ij} (k is the iteration number).

The MRF model has been applied to segmentation of both PET and SPECT images as a pre-process for classification and restoration.[65-69] For example, MRF model was used to segment brain tissues, such as striatum, GM and WM, in dynamic FDOPA-PET studies. The goal is to classify the tissues according to their physiological functions.[69] Markovian segmentation of 3D brain SPECT images was also performed to find the support of the objects to be restored using a 3D blind deconvolution technique.[67] The aim of this restoration is to improve the spatial and inter-slice of SPECT volumes for easy clinical interpretation. In another study, 3D SPECT brain images were segmented into three classes (CSF, WM, GM) using 3D MRF.[68] Feature vectors are extracted from the segmented SPECT volumes to classify brains into two classes, namely "healthy brains" and "diseased brains" (i.e., brains with possible cerebrovascular disease).

3.7 Artificial Neural Networks

Classical segmentation methods such as edge detection or deformable models discussed in section 3.8 often require considerable user expertise. For any artificial vision application, one desires to achieve robustness of the system with respect to random noise and failure of process. Moreover, a system can probably be made artificially intelligent if it is able to emulate some aspects of the human information processing.[16] Another important requirement is to have the output in real time. Artificial Neural Networks (ANNs) are attempts to achieve these goals.

ANNs are massively connected and parallel networks of elementary processors.[70-81] Each processor simulating biological learning is capable of performing elementary computation. Learning is achieved through the adaptation of weights (synaptic coefficients) assigned to the connections between processors. The massive connectionist architecture usually makes

the system robust while the parallel processing enables the system to produce output in real time.[16] ANNs are widely used in segmentation as a classifier. The weights are determined using training data, and the ANN is then used to segment new data. ANNs can also be used in an unsupervised way as a clustering method. Different neural models have been developed for image segmentation, particularly Hopfield, Kohonen, radial basis functions and Pulse-Coupled Neural Networks (PCNNs), and MultiLayer Perceptrons (MLPs).[75-77,79-81] These methods work well in a noisy environment and consequently are interesting for processing nuclear medicine images, which are inherently noisy. For example, Keller and Mckinnon[79] used PCNNs for segmentation of nuclear ventilation/perfusion images of the lungs. A PCNN is physiologically motivated information processing model based on the mammalian visual cortex.[72] In this model, each neuron in the processing layer is directly tied to an image pixel or set of neighbouring input image pixels. Each neuron iteratively processes signals feeding from these nearby image pixels (i.e., feeding inputs) and linking from nearby neurons (i.e. linking inputs) to produce a pulse train.[79] There is no training involved for the standard PCNN. Ghosh *et al.*[77] presented an original neuro-fuzzy approach to extract objects from noisy images. This method takes into account the neighbourhood of the pixels and ensures the compactness of the extracted objects by extracting spatially compact regions through the process of self-organization using only one noise corrupted image. A MLP-type of neural network is used and trained using the well-known back-propagation (BP) algorithm. The advantage of this approach is that there is no supervised learning. The error of the proposed system is computed using concepts of fuzzy sets. Based on this approach, Behloul *et al.*[7] proposed a neural system to myocardium extraction in PET images. The noisy background of these images makes myocardium extraction and tracer uptake quantification a very difficult task.

Myocardium extraction is required for a good visualization and interpretation. Indeed, analysis of ^{18}F-FDG PET heart images performed by simple visual evaluation slice by slice is tedious and suffers from observer bias. As an alternative to this visual evaluation, information can be represented using a polar-map technique. This polar map can be thought of as the image that would be obtained if one looks at the LV as a 3D cone shaped object and projected onto a single plane. This form of display has the advantage that all areas of myocardium are represented in a single image, but it is less easy to assess the relationship between the defected region and coronary arterial territories.[80] Both slice by slice visualization and polar map do not allow an easy assessment of the extent of tracer uptake defect on the LV.[7] Thus, a 3D visualization method presenting accurate localization and extent of the disease is desirable.

Cardiac ^{18}F-FDG PET images are generally involved in viable myocardium assessment studies. The quantification of the viable tissue for a patient presenting a CAD can help to decide whether or not the patient will benefit

from a revascularisation procedure. In general, 3 levels of ^{18}F-FDG uptake are distinguished by the clinicians: low, intermediate and high. Normal or remote myocardium is expected to have a high uptake rate that is more than 70% of the maximal value in the image (assuming that the maximum value corresponds to a normal tissue). Seriously damaged or infracted tissues have a low uptake rate (less than 50%).[80] However, it is very difficult to diagnose ischemic processes involved in the studied dysfunctional myocardium based only on ^{18}F-FDG PET images and especially when ^{18}F-FDG uptake is intermediate (50 to 70 %). A medium (intermediate) uptake region may correspond to a partially infracted tissue (heterogeneous) or a homogeneous reduced metabolism function. Because of the doubt, some clinicians have rather classified these regions into the viable category. Ideally, a threshold value equal to 50% (or in the range of 50 to 60%) of the max value should allow the visualization of viable myocardium.[80] However, due to the noisy nature of PET images, simple thresholding techniques, generally provided in commercial software supplied to end-users, does not seem to be suitable for a 3D extraction and visualization of the viable myocardium.

Indeed, in some images, some pathological regions were not visible and thus a higher threshold value was needed. However, an increased threshold value tends to overestimate the extent of already detected defects and thins down the myocardium. There is obviously a need for more sophisticated methods to extract accurately viable myocardium from the background. To this end, Behloul *et al.*[7] proposed a Radial Basis Function Network (RBFN) to extract myocardium through a self-organization process using only one PET image. The error of the RBFN architecture is computed using the concept of fuzzy sets and fuzziness measures. RBFNs are mainly much easier to train than MLP since they establish the Radial Basis Function (RBF) parameters directly from the data, and training is primarily on the output layer.[80] The RBFN is a feed forward neural network which accomplishes an input-output non-linear mapping by linear combination of nonlinearly transformed inputs according to:

$$o_j = \sum_{i=1}^{M_f} w_{ij}\varphi_i(x) \tag{30}$$

where x is the input vector, o_j the output of the j^{th} output node and w_{ij} are the output linear combining weights. The $\varphi_i(x)$ are RBFs and M_f is the number of RBFs. Each component of the input vector x feeds forward to M_f RBF node whose output are linearly combined with weights into the network output node. An output of the network is a simple linear combination of the hidden neuron outputs. The most important feature of RBFs is that they are local i.e. they give a significant response only in a neighbourhood near a central point. Their response decreases monotonically with distance from a central point. RBF parameters are its center, shape, and width.

A typical local RBF is a Gaussian function centred at c_x and of width (or radius) r, having the form:

$$G(x, c_x, r) = \exp\left(\frac{\|x - c_x\|_R^2}{2r^2}\right) \tag{31}$$

where R is a positive definite matrix. RBFNs have traditionally been associated with radial function networks in a single hidden layer. The hidden and output layers are generally trained sequentially: the RBF parameters are first fixed and the optimal linear combining weights are then computed.[82,83] Once the number and parameters of the RBFs defined, the hidden layer performs a fixed nonlinear transformation; it maps the input space into a new space. The output layer then implements a linear combiner on this new space. The only output layer parameters to adjust are the weights of this linear combiner. In general, the output weights of a RBFN can be determined by a pseudo-inverse matrix. When applied to supervised learning with linear models, the Least Square (LS) principle leads to a particularly easy optimization problem. However, this approach can be computationally demanding when the training set is large. In this case, delta-rule type of BP is preferred since it is a less demanding technique. Behloul *et al.*[7] used the BP algorithm to train the output layer. The procedure for learning the correct set of weights consists in varying the weights so that the error (E) is reduced. In supervised training, a set of input-output couples is supplied. The error of the network is generally computed by:

$$E = 0.5 \times \sum_{j=1}^{N} \left(t_j - o_j\right)^2 \tag{32}$$

where o_j is the spontaneous output of node j and t_j is the expected output of node j, N is the total number of output units. The network is supposed to extract the heart from the noisy image based only on the contrast between myocardium and background. As human eyes would do, the network will have to focus on the object in the noisy image until it becomes "clear". Thus, a measure of fuzziness of the image can be considered as the error of the network, to be reduced.[81]

A fuzzy set F is characterized by a membership function $\mu_F(x)$ that associates to each point x in a space of points X a real number in the interval [0, 1]. The value $\mu_F(x)$ represents the grade of membership of x in F. $\mu_F(x) = 1$ indicates a strict containment of x in F and $\mu_F(x) = 0$ means that x does not belong to F. Any intermediate value would indicate the degree to which x is an element of F. Formally, a fuzzy set F that has a finite number of elements $x_1, x_2, \cdots, x_n$ is defined as a collection of pairs

$$F = \{(\mu_F(x_i), x_i), i = 1, 2, \cdots, n\} \tag{33}$$

A measure of fuzziness estimates the average ambiguity in a fuzzy set. Intuitively, the fuzziness of a crisp set using any measure should be zero (or minimum), as there is no ambiguity about whether an element belongs to the set or not. If the set is maximally ambiguous, then the fuzziness measure should be maximum. When the membership value approaches either 0 or 1, the ambiguity of the set decreases. Thus, a fuzzy set is most ambiguous when $\mu_F(x) = 0.5 \forall x$. The degree of fuzziness, I(F), of a fuzzy set F represents the amount of ambiguity in making the decision whether a point belongs to F or not. Such measures have been proposed by several authors.[84–88] Two mathematical models of fuzziness measures are used: the index of fuzziness and the fuzzy entropy.[77] Both measures presented below lie in the interval [0, 1]. The index of fuzziness of F is defined by $v(F) = \frac{2}{n^k} d(F, \underline{F})$, where $\underline{F}$ is the nearest ordinary set to F given by

$$\mu_{\underline{F}}(x) = \begin{cases} 0 & \text{if } \mu_F(x) \leq 0.5 \\ 1 & \text{if } \mu_F(x) > 0.5 \end{cases} \tag{34}$$

$d(F, \underline{F})$ is the distance between the two sets F and $\underline{F}$. The value of k depends on the type of the distance used. $k = 1$ corresponds to the generalized hamming distance and $k = 0.5$ to the Euclidean distance. The corresponding index of fuzziness is called the quadratic index of fuzziness $v_q^2(F)$.

$$v_q(F) = \frac{2}{\sqrt{n}} \sum_{i=1}^{n} \sqrt{\left(\mu_F(x_i) - \mu_{\underline{F}}(x_i)\right)^2} \tag{35}$$

The index of fuzziness reflects the ambiguity of an image by measuring the distance between its fuzzy property plane and the nearest ordinary plane. Pal and Pal[86] proposed an exponential entropy given by

$$H_z(F) = \frac{1}{n(\sqrt{e} - 1)} \sum_{i=1}^{n} \left(S_n(\mu_F(x_i)) - 1\right) \tag{36}$$

with

$$S_n(\mu_F(x_i)) = \mu_F(x_i) e^{1-\mu_F(x_i)} + (1 - \mu_F(x_i)) \mu_F(x_i) e^{\mu_F(x_i)}$$

$H_z(F)$ gives a measure of the average amount of difficulty in taking a decision on any pixel. Both terms (index and entropy) give an idea of "indefiniteness" of an image. They may be considered as measures of average intrinsic information that is received when one has to make a decision in order to classify the pixels described by a fuzzy set. Thus, these measures of fuzziness are used to compute the error of the SRBFN.

The architecture of SRBFN, derived form the approach of Ghosh *et al.*[77], presents three layers: the input layer, one hidden (RBF) layer and the output layer (Fig. 6). In every layer, there are $L_x \times L_y$ neurons (for an $L_x \times L_y$ image) each neuron corresponds to a single pixel. Neurons in the same layer are not connected to each other (Fig. 6b). Each neuron is connected to the corresponding neuron and its neighbours of the next layer. The

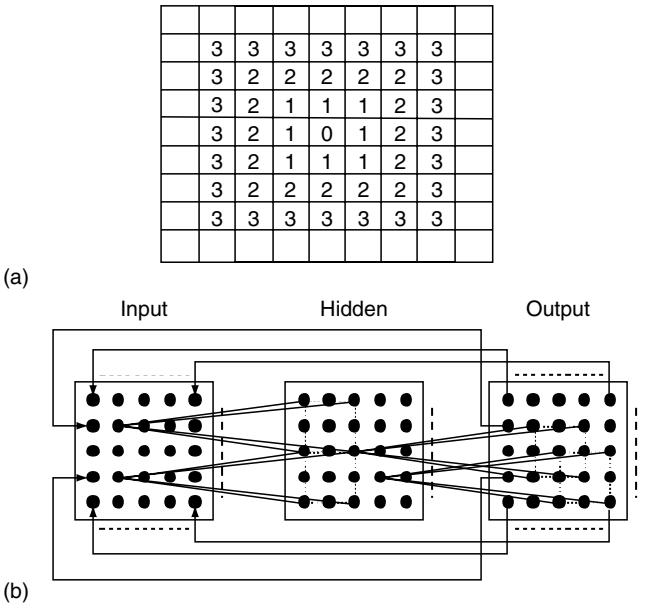


FIGURE 6. Illustration of the architecture of self organized RBFN. (a) Neighbourhood system. (b) RBFN topology. (Reprinted with permission from ref.[7]).

neighbourhood considered here corresponds to the $8(3 \times 3 - 1)$ nearest neighbours (N) or the $24(5 \times 5 - 1)$ nearest neighbours $\left(N^2\right)$ and so on (Fig. 6a).

In the SRBFN, the layers are partially connected and the neurons have spatial positions corresponding to the pixels in the image. The feedback connections from the output layer to the inputs make the network recurrent. The output is considered as input for the next iteration. Behloul[80] proposed a sigmoid function f for the activation of the output nodes $o_j = f\left(I_j\right)$ where $I_j = \sum_{i=1}^{M_f} w_{ij}\varphi_i(x)$ and

$$f(x) = 1/\left(1 + e^{-(x-\theta)}\right) \tag{37}$$

where θ is a bias value. The output of each neuron of the hidden and output layers lies in [0, 1]. It represents the degree of brightness of the corresponding pixel in the image. The measure of fuzziness of the fuzzy set bright is viewed as the error of the network. Once the weights have been adjusted according to the error measure, the neurons in the output layer are feed back to the corresponding neurons in the input layer. The output values of the output nodes which lie in [0, 1] are considered to be the brightness degrees of the corresponding pixels (0 for black, 1 for white, and any intermediate value is a grey level). The grey levels corresponding to the output values are considered as the input of the next iteration. The BP algorithm is used to train the output layer. The change in weights, such that the error E is reduced, is given

by $\Delta w_{ij} = \eta(-\partial E/\partial o_j) f'(I_j) o_i$ where η is a learning rate. The sigmoid function is used because of its simple derivative function: $f'(I_j) = \partial o_j/\partial I_j = o_j(1 - o_j)$. The mathematical derivations of the BP for quadratic index of fuzziness v_q^2 and fuzzy entropy H_z are defined as follows:

- *Quadratic index of fuzziness*
The error is defined by $E = v_q^2$ where

$$v_q^2 = (4/n)\left[\sum_j \{\min(o_j, (1 - o_j))\}^2\right] \tag{38}$$

and n is the number of output units. Thus,

$$-\partial E/\partial o_j = \begin{cases} -8o_j/n & \text{if } 0 \le oj \le 0.5 \\ 8(1 - o_j)/n & \text{if } 0.5 < oj \le 1 \end{cases} \tag{39}$$

$$\Delta w_{ij} = \begin{cases} \eta_1(-o_j) f'(I_j) o_j & \text{if } o \le oj \le 0.5 \\ \eta_1(1 - o_j) f'(I_j) o_j & \text{if } 0.5 < oj \le 1 \end{cases} \tag{40}$$

where $\eta = \eta_1 \times (4/n) \times 2$

- *Fuzzy entropy*
The error is defined by $E = H_z$ where

$$H_z = \frac{1}{n(\sqrt{e} - 1)} \sum_{j=1}^{n} \left(o_j e^{1-o_j} + (1 - o_j) e^{o_j}\right) \tag{41}$$

$$\frac{\partial H_z}{\partial o_j} = \frac{1}{n(\sqrt{e} - 1)} \left((1 - o_j) e^{1-o_j} - o_j e^{o_j}\right) \tag{42}$$

To make weight correction value minimum when the membership values of the elements are 0 or 1 and maximum when they are all 0.5, we take

$$\Delta w_{ij} = \eta_2 \frac{1}{\frac{\partial E}{\partial o_j}} f'(I_j) o_i \tag{43}$$

$$\Delta w_{ij} = -\eta \frac{1}{(1 - o_j) e^{1-o_j} - o_j e^{o_j}} f'(I_j) o_i \tag{44}$$

where $\eta = \eta_2 \times n(\sqrt{e} - 1)$

$$\Delta w_{ij} = \begin{cases} -\eta \frac{1}{(1-o_j) e^{1-o_j} - o_j e^{o_j}} f'(I_j) o_i & \text{if } 0 \le oj \le 0.5 \\ \eta \frac{1}{o_j e^{o_j} - (1-o_j) e^{1-o_j}} f'(I_j) o_i & \text{if } 0.5 \le oj \le 1 \end{cases} \tag{45}$$

At the first iteration, the input nodes receive the grey levels of the corresponding pixel. For the hidden layer, each RBF centre is localized at the

value of the corresponding pixel in the image. The total input to any node of the output layer lies in $[0, N_1]$ where N_1 is the number of links that a neuron has. The parameter θ of the sigmoids in the output nodes is set to $N_1/2$, since this value corresponds to the middle of the total input range.[7] All initial weights (between the hidden and output layers) are set to 1. The SRBFN presented by Behloul[80] aimed to extract a compact object from one noisy 2D image. It is easy to extend the architecture to process 3D images; however the computation complexity grows dramatically. It turned out that the result obtained when considering the 3D image as a set of 2D slices and applying the SRBFN to each slice was comparable (in terms of size and shape of the extracted object) to that when considering the total 3D image. The 2D images are processed independently in a sequential manner. Behloul considered the 3D PET volume as a set of 2D-images where two orthogonal views are presented: 63 (128×128)-slices and 128 (63×128)-slices; the intersection of the two independently extracted objects constitutes the final result of the myocardium extraction process.

Different sizes of neighbourhood have been tested and N^2 turned out to be the best compromise between processing time and accuracy of shape and width of the extracted myocardium. As pointed out by Ghosh *et al.*[77] and Behloul *et al.*[7], the index of fuzziness is better than the entropy measure for maintaining the compactness of the extracted objects. However, the shapes of objects are better preserved by the entropy measure. For the index of fuzziness, the learning rate is lower than that of the entropy measure. Low learning rate smoothes out noise and creates compact regions, while entropy measure enables the network to preserve object boundaries as learning rate is very high around the most ambiguous region ($o_j \approx 0.5$).[77] Figure 7 shows the results obtained using the index of fuzziness versus those obtained using the entropy measure. The diseased area appears as a hole in the myocardium (Fig. 8).

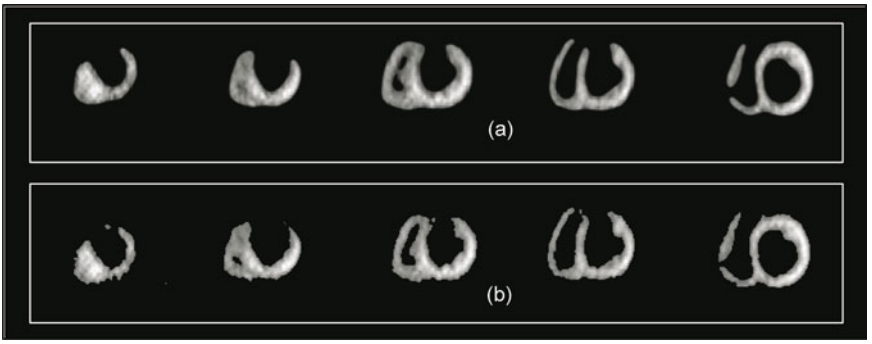


FIGURE 7. Object extraction results. (a) Index of fuzziness. (b) Fuzzy entropy. (Reprinted with permission from ref.[7]).

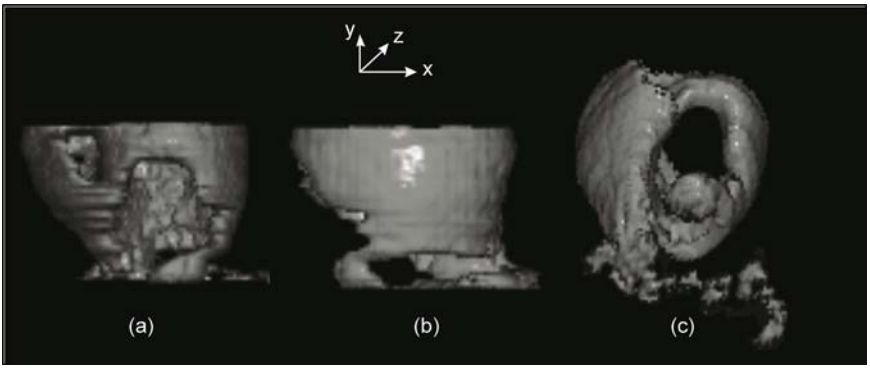


FIGURE 8. Volume rendered visualization of the entropy extracted myocardium: the diseased region appears as a hole in the myocardium. (a) Orthogonal view of the heart showing a defect in almost all the posterior wall of the LV. (b) 90° y-axis rotation of (a) showing the extent of the disease in the apical anterior wall of the LV. (c) 45° x-axis rotation of (a) showing both defects (posterior and apical anterior) and the RV. (Reprinted with permission from ref.[7]).

For the index of fuzziness (Fig. 7a), the diseased region and the myocardium width are overestimated (over-smoothed), while for entropy (Fig. 7b), the results are very much closer to manually delineated regions. Moreover, using the index of fuzziness, the network requires more time to stabilize. To take advantage of the main positive features of both measures in the same time: "smoothing and compactness" for index of fuzziness, and "good shape estimation" for the fuzzy entropy, both measures were combined. For the first iterations of the self-organizing process (2 to 3 iterations), index of fuzziness is used. Because of its low learning rate, this measure produces a contrast enhancement on the original images (Fig. 9a). The entropy measure is then used; it ensures a good shape extraction (Fig. 9b).

The sequential application of the two fuzziness measures presented the best results. Indeed, when using separately the fuzziness measures, the network was not able to detect some regions of the myocardium where there is a low ^{18}F-FDG uptake (the human eye was still able to allow delineation of such regions). Only the combination of both measures made the network successful in retaining these regions in the extracted myocardium (see arrows in Fig. 9b). Thus, it seems that the two error models complement each other perfectly for an accurate delineation of the myocardium in PET images.

3.8 Deformable Models

Deformable models are curves, surfaces or solids defined with an image or volume domain. These models deform under the influence of external and internal forces. The original deformable model, called *snake* model, was

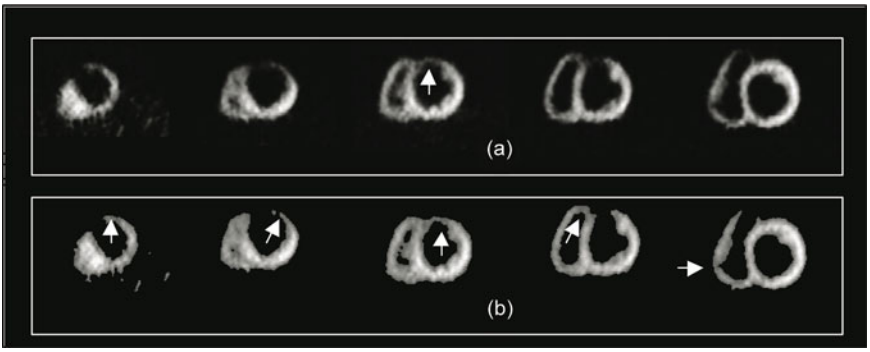


FIGURE 9. Combination of both fuzzy measures. (a) Images obtained after 2 iterations using only the index of fuzziness measure. (b) Final result using the entropy measure on the images shown in (a). (Reprinted with permission from ref.[7]).

introduced by Kass *et al.*[89] as an active spline reacting with image features. Basically, the snake model is an energy-minimizing spline controlled by external image forces such as lines and edges, and internal spline forces which impose a piecewise smoothness constraint. In practice, the user initializes the deformable model near the RoI and allows it to deform into place. User could manually fine-tune the filling by using interactive capabilities of the model. The snake deforms itself into conformity with the nearest salient contour by minimizing the snake energy, which pushes the snake toward salient image features and internal forces, which are responsible for smoothness of the curve. As the algorithm advances, the terms of snake functional energy can be adjusted to obtain a local minimum. The main advantages of deformable models are their ability to directly generate close parametric curves or surfaces from the image and their incorporation of a smoothness constraint that provides robustness to noise and spurious edges. A disadvantage is that they require manual interaction to place an initial model and choose appropriate parameters.[3]

The original formulation of the snake is for a continuous spline and image. The snake geometry is represented by a curve V (s, t) with parameters s (spatial index) and t (time index) defined on given open intervals Ω and T_m, respectively. By permitting the snake to have two deformational degrees of freedom in the plane, that is, the x and y coordinates, the active model, embedded in the plane image $(x, y) \in \mathcal{R}^2$, is represented as a time varying parametric contour by

$$\begin{aligned} V(s, t) = (x(s, t), y(s, t)) : s \in \Omega,\ t \in T_m \\ V : \Omega = [0, 1] \rightarrow \Re^2 \end{aligned} \tag{46}$$

Thus, at time $t = t_0$ the snake or the closed contour C is represented by its pair of coordinate functions $\{x(s, t_0), y(s, t_0)\}$, where $s \in [0, 1]$ is the

normalized arc length along C. Let A_d be a space of admissible deformations. The shape of the contour subject to an image I(x, y) can be dictated by its energy function E_{snake} as follows

$$\begin{aligned} &E_{snake} : A_d \rightarrow \Re \\ &E_{snake} = \int_{\Omega} \left[E_{int}(V(s,\, t)) + E_{image}(V(s,t)) + E_{con}(V(s,\, t)) \right] ds \end{aligned} \tag{47}$$

where E_{int}, E_{image}, and E_{con} refer to the internal, image and external constraint forces, respectively. The energy terms are defined cleverly in a way such that the final position of the contour will have a minimum energy (E_{min}).E_{int} represents the internal energy term of the snake which imposes the regularity on the curve by bending and stretching, and is given by

$$E_{int} = \frac{1}{2} \left\{ \alpha(s) \underbrace{\left\| \frac{\partial [V(s,\, t)]}{\partial s} \right\|^2}_{E_1} + \beta(s) \underbrace{\left\| \frac{\partial^2 [V(s,\, t)]}{\partial s^2} \right\|^2}_{E_2} \right\} \tag{48}$$

E_1 and E_2 represent the continuity and smoothness (curvature) terms, respectively. The contour is treated as an elastic rubber band possessing elastic energy. It discourages stretching by introducing a tension:

$$E_{elastic} = \frac{1}{2} \int_{\Omega} \alpha(s) \left\| \frac{\partial [V(s,\, t)]}{\partial s} \right\|^2 ds \tag{49}$$

$E_{elastic}$ has larger values where there is a gap in the curve. The weight $\alpha(s)$ allows to control elastic energy along different parts of the contour. The $\alpha(s)$ value determine the extent to witch the contour is allowed to stretch. If $\alpha(s) = 0$, a discontinuity can occur. For most applications, $\alpha(s)$ is assumed to be constant throughout the curve. In addition to being considered as an elastic band, the snake is also considered to behave like a thin metal strip giving rise to bending energy. This second term discourages bending. This energy is defined as sum of squared curvature of the contour:

$$E_{bending} = \frac{1}{2} \int_{\Omega} \beta(s) \left\| \frac{\partial^2 [V(s,\, t)]}{\partial s^2} \right\|^2 ds \tag{50}$$

Sharp corners or points of high curvature are characterized as high frequencies and bending energy is more sensitive for contours having such corners because the second derivative will be very high for such contours. $E_{elastic}$ has larger values when the curve is bending rapidly. The bending energy is minimum for a circle. The weight $\beta(s)$ value determine the extent to which the contour is allowed to bend at point s. Setting $\beta(s) = 0$ at point s means that we are relaxing our condition and allowing that *s* develops a corner. The total internal energy of the snake is the sum of elastic and

bending energies. Adjusting the weights $\alpha(s)$ and $\beta(s)$ controls the relative importance of the elastic and bending terms and therefore the internal energy of the contours.

E_{image} is derived from the image data over which the snake lies (object of interest). The three important features the snake can be attracted to are line, edge and termination functions. The total image energy can be expressed as a weighted combination of the three:

$$E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term} \tag{51}$$

The simplest useful image functional is the image intensity $E_{line} = I(x, y)$. Depending on w_{line} the snake is attracted to dark or light lines. The edge-based functional, E_{edge}, attracts the snake to contours with image gradients, that is, to locations of string edges. A way to define E_{edge} is as follows:

$$\begin{aligned} E_{edge}(V(s)) &= -\gamma(s)E_3(V(s)) \\ E_3(V(s)) &= \|\nabla I(V(s))\|^2 \\ E_3(V(s)) &= \|\nabla(G_\sigma(V(s)) \otimes I(V(s)))\|^2 \end{aligned} \tag{52}$$

$\gamma(s)$ is a parameter that controls the weight of image attraction. $\nabla I(V(s))$ denotes the gradient of the image I and $G_\sigma(V(s))$ is a 2D Gaussian function with standard derivation σ (the image is pre-processed with a low pass filter). It can be seen that larger σ will increase the blurring of the edges thereby increasing the capture range of the active contour. The termination functional, E_{term}, can be obtained by a function checking the curvature of level lines in a slightly smoothed image. Let $C(x, y) = G_\sigma(x, y)^* I(x, y)$ be the smoothed version of I(x,y) and $\theta = \tan^{-1}\left(\frac{C_y}{C_x}\right)$ the gradient angle. Curvature of contour in $C(x, y)$ can be written as:

$$\begin{aligned} E_{term} &= \frac{\partial\theta}{\partial R_\perp} \\ E_{term} &= \frac{\partial^2 C/\partial R_\perp^2}{\partial C/\partial R} \\ E_{term} &= \frac{C_{yy}C_x^2 - 2C_{xy}C_xC_y + C_{xx}C_y^2}{\left(C_x^2 + C_y^2\right)^{3/2}} \\ E_{term} &= \frac{\left(\partial^2 C/\partial y^2\right)(\partial C/\partial x)^2 - 2\left(\partial^2 C/\partial x\partial y\right)(\partial C/\partial x)(\partial C/\partial y) + \left(\partial^2 C/\partial x^2\right)(\partial C/\partial y)^2}{\left((\partial C/\partial x)^2 + (\partial C/\partial y)^2\right)^{3/2}} \end{aligned} \tag{53}$$

where $R = (\cos\theta,\ \sin\theta)$ is a unit vector in direction of the gradient and $R_\perp = (-\sin\theta,\ \cos\theta)$ is a vector perpendicular to gradient direction.

E_{con} represents the external constraint energy term that is responsible for guiding the snake towards the desired local minimum. It can be defined by the user, constructed from physical phenomena,[90] or derived form high-level interpretation.[89] For example, the user connects a spring between any point

p_1 in the spatial domain and a point p_2 on the snake. A snake trying to reduce its constraint energy is pulled towards the point p_1 defined by the user. Constraint energy can also be interpreted as a kind of prior for the system.

A contour is defined to lie in the position in which the snake reaches a local energy minimum. Before the minimization process, the snake formulation is discretized both in time and space domains. The spatial discretization is done by sampling the active contour V into N points $V(v_i, i = 1, 2, \ldots, N)$. For an initial contour (i.e., at $t = 0$), the N points are chosen to be evenly spaced on the contour. The time discretization is defined to start at time $t = 0$ with a constant time step Δt. The discretization of E_{int} is done by approximating the first and second derivatives by finite differences as follows:

$$\begin{aligned} E_1(i) &= \|v_i - v_{i-1}\|^2 = (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 \\ E_2(i) &= \|v_{i-1} - 2v_i - v_{i+1}\|^2 = (x_{i-1} - 2x_i - x_{i+1})^2 + (y_{i-1} - 2y_i - y_{i+1})^2 \end{aligned} \tag{54}$$

where x_i and y_i are the x and y coordinates of the v_i^{th} contour-point, respectively. Equation 54 is formulated to minimize the distance between neighbouring.[91] For simplicity we suppose that $w_{line} = w_{term} = 0$. Thus, the boundary is detected by iteratively deforming an initial approximation of the boundary by minimizing the reduced snake energy function:

$$E_{snake} = \sum_{i=1}^{N} \{\alpha(i)E_1(i) + \beta(i)E_2(i) + \gamma(i)E_3(i)\} \tag{55}$$

$$E_{min} = \min\{E_{snake}\} \tag{56}$$

Equation 56 was originally solved by using techniques of variational calculus and applying a finite difference method.[89] It was improved later on by Cohen and Cohen[92] using a finite element method. Another solution was suggested by Amini *et al.*[93] using dynamic programming that promotes stability of snake and hard constraints to be enclosed at cost of expensive computation. Williams and Shah[91] proposed another approach based on the Greedy algorithm, which is faster than the method of Amini *et al.*[93] Suppose that the contour has N points (Fig. 10), which are allowed to move to any point in a neighbourhood of size M at each iteration. The Greedy algorithm is described as follows:

Initialize α_i, β_i, and γ_i to 1 for all i
Do
/* Loop to move points to new locations */
For i = 1 to N /*N: number of points in contour*/

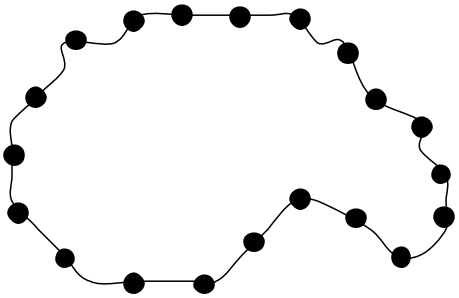


FIGURE 10. Curve representation with a set of N points $(v_i(x_i,y_i)$, $i = 1, \cdots, N)$.

$E_{min} \leftarrow \infty$

For $j = 1$ to N

$E(j) \leftarrow \alpha(i)E_1(j) + \beta(i)E_2(j) + \gamma(i)E_3(j)$

If $E_j < E_{min}$

$E_{min} \leftarrow E(j)$

$j_{min} \leftarrow j$

Move point v_i to location j_{min}

If j_{min} is not current location, ptsmoved ++

Until ptsmoved < Th_Mov

where ptsmoved is the number of points which have been moved to the new locations. Th_Mov is a threshold to decide whether we have converged according to the number of points that moved in each iteration. Figure 11 shows the result of the snake applied to a gated blood pool image. Fig. 11(a) shows a cardiac image acquired in the LAO projection providing the septal separation of the two ventricles. The aim is to delineate correctly the LV contour in order to estimate the EF. The initial contour drawn by the user is

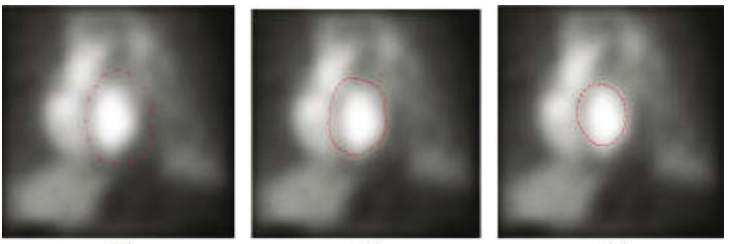


FIGURE 11. Snake applied to blood pool cardiac image. (a) Initial contour. (b) Contour obtained after 20 iterations. (c) The LV enclosed by the snake.

shown in Fig. 11(a). Contour obtained after 20 iterations with Th_Mov set to 6 is represented in Fig. 11(b). As shown in Fig. 11(c), the LV is well delimited. Results of delineation of the cortical structure from PET brain images are reported in ref.[94] These results show the interest to use deformable models for repeatable brain regional extraction applicable for regional quantification of tracer uptake.

3.9 Atlas-guided Approaches

Atlas-guided medical image segmentation could be a robust and efficient approach when a standard atlas or "template" for the particular organ and imaging modality being investigated is available. The atlas is produced by gathering detailed information on the anatomy of the structure to be segmented for a particular patient population. The template is then used as a reference frame for segmenting other images.[3] It is worth emphasizing that atlas-guided approaches are conceptually similar to classifiers described in section 3.3 keeping in mind that they are implemented in the spatial domain rather than in a feature space.

The standard atlas-guided approach considers segmentation as a registration problem (see chapter 9 for a detailed survey of registration techniques). The algorithm starts with a process called *atlas warping*, that is an estimate is made of the transformation matrix that matches a pre-segmented atlas image to the target image to be segmented. This transformation can be linear,[95] non-linear[96,97] or a combination of both[98] to handle efficiently anatomical variability. It is worth emphasizing that this issue is rather difficult to solve and might result in inaccurate segmentation of most dissimilar complexly-shaped structures exhibiting large variability in size and shape among subjects. Another important conceptual limitation for this technique, namely the existence of patient-specific anomalies that are obviously not modelled in an atlas obtained from a single or even an average representation of the population. It is therefore well accepted that atlas-guided approaches are more appropriate for segmentation of structures that do not vary significantly among subjects. The use of probabilistic atlases has proven to be useful in modelling anatomical variability at the expense of additional complexity and increase in computation time.[99] Atlas-guided approaches have been developed and applied mainly in MR brain imaging. Some of them are, however, suitable for an extension towards segmentation of nuclear medical images.

3.10 Analysis of Time Activity Curves

Gated blood pool scintigraphy (SPECT and PET) can measure the functional changes in tissues.[11,100] Theses modalities generate images demonstrating temporal changes in radioactive tracer distribution. For example, gated blood pool scintigraphy is a reliable non-invasive method for detecting

abnormalities of contraction and conduction of the heart. Dynamic imaging provides a time sequence of images taken throughout the whole cardiac cycle. The quantitative information contained in the images describes the physiological behaviour of the imaged structures such as the LV and RV.

In routine dynamic studies, RoI delineation is required for quantitative analysis. A common approach is to identify the anatomic structures by placing RoIs directly on the images, and the underlying tissue time-activity curves (TACs) are then extracted for subsequent analysis. This RoI analysis approach, although widely used in clinical and research settings, is operator-dependent and prone to reproducibility errors. A semi-or fully-automated RoI delineation (or segmentation) method offers advantages by reducing operator error/subjectivity, thereby improving reproducibility.[100] Automatic segmentation may be an alternative to manual RoI delineation. Cluster analysis has been investigated to partition sets of tissue TACs into a set of homogeneous TACs (clusters). It was found that a linear combination of those homogeneous TACs using a mixture model can accurately represent the original data while providing quantitative parametric images.[101,102] Wong *et al.*[100] used a non parametric C-means clustering technique to segment dynamic PET data into C characteristic curves automatically. However, large-scale and multi-centric studies for different applications are necessary to demonstrate the potential of automatic segmentation for replacing manual RoI delineation.

Conventional analysis of image sequences is usually performed by visual evaluation of differences from image to image in order to obtain qualitative information for example about the cardiac contraction kinetics or tissue perfusion. This is not the optimal way to interpret the changes appearing in the images along the series. This operation is tedious and suffers from observer bias. Furthermore, a sequence of images contains also spatially differentiated quantitative information describing the behaviour of the imaged structure, which is difficult to extract by visual evaluation. As an alternative to visual evaluation, the required information can be obtained from an image sequence using parametric analysis methods. Based on the first harmonic fit to the pixel time activity curve, Fourier analysis[103] has also been used to detect and describe wall motion abnormalities in gated blood pool scintigraphy.[104,105] The primary assumption of Fourier analysis is that the data are periodic and the second is that the transition between the first and the last frame of the study must be smooth. This analysis generates two parametric images, the amplitude image corresponding to the maximal change in counts within the image series regardless of the time these changes occur and the phase image corresponding to the time of maximal contraction (ED). In a first approximation, the amplitude image is proportional to the stroke volume. The phase image allows good separation of the ventricular regions (left or right) from the atria and this processing is useful to delineate automatically the LV contours in order to estimate the EF value from ventricular regions of interest.[4]

Let Vol(i, j, k) be the value of the pixel (i, j) of the k^{th} cardiac image. Two images, respectively, I_{cos} and I_{sin} are calculated using the following equations:

$$\begin{aligned} I_{cos}(i,j) &= \sum_{k=1}^{T_C} \cos\left[\frac{2\pi}{T_C}(k-1)\right] \times Vol(i,j,k) \\ I_{sin}(i,j) &= \sum_{k=1}^{T_C} \sin\left[\frac{2\pi}{T_C}(k-1)\right] \times Vol(i,j,k) \end{aligned} \tag{57}$$

where T_C is the number of images in the series corresponding to the cardiac cycle duration. The phase image is given by

$$I_{phase}(i,j) = \arctan\left(\frac{I_{sin}(i,j)}{I_{cos}(i,j)}\right) \tag{58}$$

and the amplitude by:

$$I_{amp}(i,j) = \sqrt{(I_{cos}(i,j))^2 + (I_{sin}(i,j))^2} \tag{59}$$

For cardiac studies, the amplitude image is a picture of pixel-by-pixel stroke volume. The rationale for using the phase image rests on the observation that ischemia reduces the velocity of myocardial contraction and thus local abnormalities of wall motion are frequently associated with changes in the time of wall movement.

In spite of the apparent mathematical difficulties, Factor Analysis of Dynamic Structures (FADS) has gained clinical acceptance for cardiac studies at equilibrium with abnormalities of both contraction and conduction.[106–108] This analysis assumes that any time activity curve is a weighted sum of a limited number, q, of pure time activity evolutions, called physiological components.[107] These components correspond to regions of similar temporal behaviour.

Assuming that image sequence has m frames each containing n pixels, this can be expressed as a *q* -factor problem:[108]

$$X = \hat{X} + E;\ \hat{X} = FA \tag{60}$$

where $X, \hat{X}$, E, F and A are m × n,m × n, m × n, m × q and q × n matrices, respectively. X is the data matrix (noise corrupted observation) whose columns contain TACs of each pixel; $\hat{X}$ is the relevant part of the observation; E is an error matrix; F contains the q temporal factors in its columns and A contains the q factor images in its rows. The FADS solution involves choosing the number of factors q, minimizing the size of E in a suitable sense and determining the composition of $\hat{X}$ into FA by appropriate constraints.[108] For dynamic studies in nuclear medicine a Principal Component Analysis (PCA) may be performed on the population of pixels (TAC for each pixel) in the image to extract the first p components. A Singular Value

Decomposition (SVD) of the row space of $X(m < n)$ gives $X = PL^{0.5}Q^T$. P and Q are $m \times m$ and $n \times m$ column orthogonal matrices and L is the $m \times m$ diagonal matrix with non-negative eigenvalues in decreasing order. Let P_p and Q_p represent the first p columns of P and Q; L_P the upper $p \times p$ submatrix of L; and $D_p = P_p L_p^{0.5} Q_p^T$ the $m \times n$ matrix. Then $X = X_p + \varepsilon_p = C_e W_e + \varepsilon_p$ where $C_e = P_p$; $W = L_p^{0.5} Q_p^T$. C_e is an $m \times p$ matrix containing the orthogonal components in its columns; W_e is a $p \times n$ matrix containing the pixel coefficients for each component in its corresponding rows; and ε_p is the residual error matrix. It is now assumed that $p = q$. Since PCA minimizes residual variance for a given number of components, it is natural to estimate E by ε_p and thus estimate $\hat{\mathbf{X}} = C_e W_e$. Transforming a PC solution to p-factor solution involves finding a non-singular $p \times p$ transformation matrix V such that $F = C_e V$; $A = V^{-1} W_e$. The columns of W (one per pixel) will plot as points q-dimensional feature space, as will the columns of V (one per structure). The problem reduces to locating the points in feature space corresponding to the columns of V. Since nuclear medicine data are in principle non-negative, the elements of F and A should be non-negative.[108]

The estimated physiological components given by FASD are used to reconstruct functional images (factor images). This method has an obvious similarity with Fourier analysis: while Fourier analysis is based on sine and cosine functions, no such restriction is placed on the shape of the PCs. Although factor analysis provides a valid representation of wall motion and conduction abnormalities, some limitations must be considered. The quality of the results and the number of significant factors depend on the signal to noise ratio of the dynamic series.[107]

Recently, a new technique designed to analyze nuclear cardiac image series was proposed.[11] Compared to Fourier analysis, this method does not assume that the data are periodic and no restriction is placed on the shape of the pixel TAC. This method segments the time series of the cardiac images in regions of similar temporal behaviour (components) but, contrary to PCA, knowledge of their number is not needed to run the algorithm.[11] The aim of this technique is to capture the essential of the sequence while reducing the amount of image data presented to the clinician for diagnostic interpretation. This method is based on the measure of similarity between the temporal response of pixels in a reference RoI and the temporal response of the other pixels in the image series. The calculated similarity provides quantitative information about the degree of local similarity in comparison with the RoI taken as a reference. Template matching is perhaps the most straightforward way of locating the presence of an object of interest within an image field. The template is a duplicate of the sought object. If the template matches an unknown object and if it is sufficiently close, the unknown object is labelled as the template object. In nuclear cardiac studies, the template is a fixed reference TAC describing the temporal evolution of a

cardiac RoI during the cardiac cycle. The image field to be searched is the time series of cardiac images. For dynamic analysis, template matching may be viewed as a temporal one. For example, the template may therefore be a TAC of a subregion of the LV or the whole ventricle. Similarity measures provide quantitative means of determining the degree of temporal match between the time sequence of images and the template. The similarity used is the covariance function. The covariance value of the $(i, j)^{th}$ pixel is given by:

$$\mathrm{Cov}(i, j) = \frac{1}{T_C} \sum_{t=1}^{T_C} (\mathrm{Vol}(i, j, t) - \mu_{AC}(i, j))(R(t) - \mu_R) \tag{61}$$

where

$$\begin{aligned} R(t) &= \frac{1}{M_C} \sum_{(i,j)\in \mathrm{RoI}} \mathrm{Vol}(i, j, t) \\ \mu_R &= \frac{1}{T_C} \sum_{t=1}^{T_C} R(t),\ \mu_{AC}(i, j) = \frac{1}{T_C} \sum_{t=1}^{T_C} \mathrm{Vol}(i, j, t) \end{aligned} \tag{62}$$

$R(t)$ is the reference series, μ_R the mean value of the reference series and $\mu_{AC}(i, j)$ the mean value of the time activity curve of the $(i, j)^{th}$ pixel. M_C is the total number of pixels in the RoI.

There is no major problem of spatial registration for this kind of images because the images of the series correspond to the same morphologic projection even if its shape changes with time. The method is based on the computation, pixel by pixel, of the covariance (Eq. 61) between two time series representing the TAC of any pixel and a reference time series. The generated similarity map is an image where the value of each pixel represents the degree of temporal similarity to the reference. This similarity measure is less than the absolute value of $\sigma_{AC}\sigma_R$, where σ_{AC} and σ_R are the standard deviations of the TAC and the reference, respectively. A positive value of the covariance indicates that the TAC and the reference variables vary in the same sense. While, a negative value indicates that these two variables vary in opposite sense. It is important to keep in mind that all the pixels that match or mismatch the reference are equally important to describe or to interpret the information contained in the image series.

Gated cardiac blood pool images are analyzed by the covariance technique. The imaging is performed in LAO projection. Figure 12 shows the placement of four RoI references. For covariance analysis, only one reference is used. Each reference time series is calculated in, for example, a 5×5 pixels window size of the corresponding RoI. Once a covariance map is calculated for all pixels, with respect to the chosen reference RoI, its values are mapped into 0 through 255. The map is displayed with 256 colours look-up table and a colour coded image represents the distribution of the temporal degree of similarity to be assessed. The red colour corresponds to a maximum covariance value while the blue colour corresponds to minimum

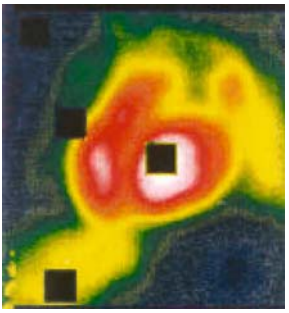


FIGURE 12. End-diastolic image with the reference RoIs of the background (top left), the liver (bottom left), the LV (middle right) and the atria (middle left). (Reprinted with permission from ref.[11]).

covariance value. Like in the phase analysis, the strength of the covariance image (similarity map) lies in its relative and not in its absolute values. For a ventricular reference, the similarity maps of the ventricles are displayed in red and the atria in blue. This is expected since the atrial pixels are out of phase with the ventricular ones. Figure 13 shows the results obtained in four pathological patients (from top to bottom: patients 1–4) with RoI reference placed in the LV region. The left and the middle columns of figure 13 represent the end diastolic and systolic images of the four patients, respectively. The corresponding similarity maps are shown in the right column. In patient 1, the similarity map shows the septal portion of the LV with a delayed temporal response compared to the remaining segments of the ventricle. Although the LV has a normal morphology, it presents an anteroseptal and hypokinesia. Patient 2 has a largely dilated LV. Note the opposition of temporal responses of the ventricular and the atria are well demonstrated. As the LV, the RV is dilated but its anterobasal portion is hypokinetic. Patient 3 also has a dilated LV. The covariance map reveals, in the LV, an akinesis of the inferoapical segment and dyskinesis of the distal septum. In patient 4, the map shows that there is an akinesis of the inferior part of the RV with a limited dyskinetic region. One may also note a hypokinesis of the septal segment of the LV.

4. Clinical and Research Applications of Image Segmentation

Image segmentation has found numerous clinical and research applications both in diagnostic and therapeutic nuclear medicine. This includes estimation of organ (e.g. liver, spleen, heart, thyroid . . .) or tumour volumes as well as target treatment volumes definition in radiation therapy, extraction of parameters of clinical relevance such as left ventricular region in nuclear

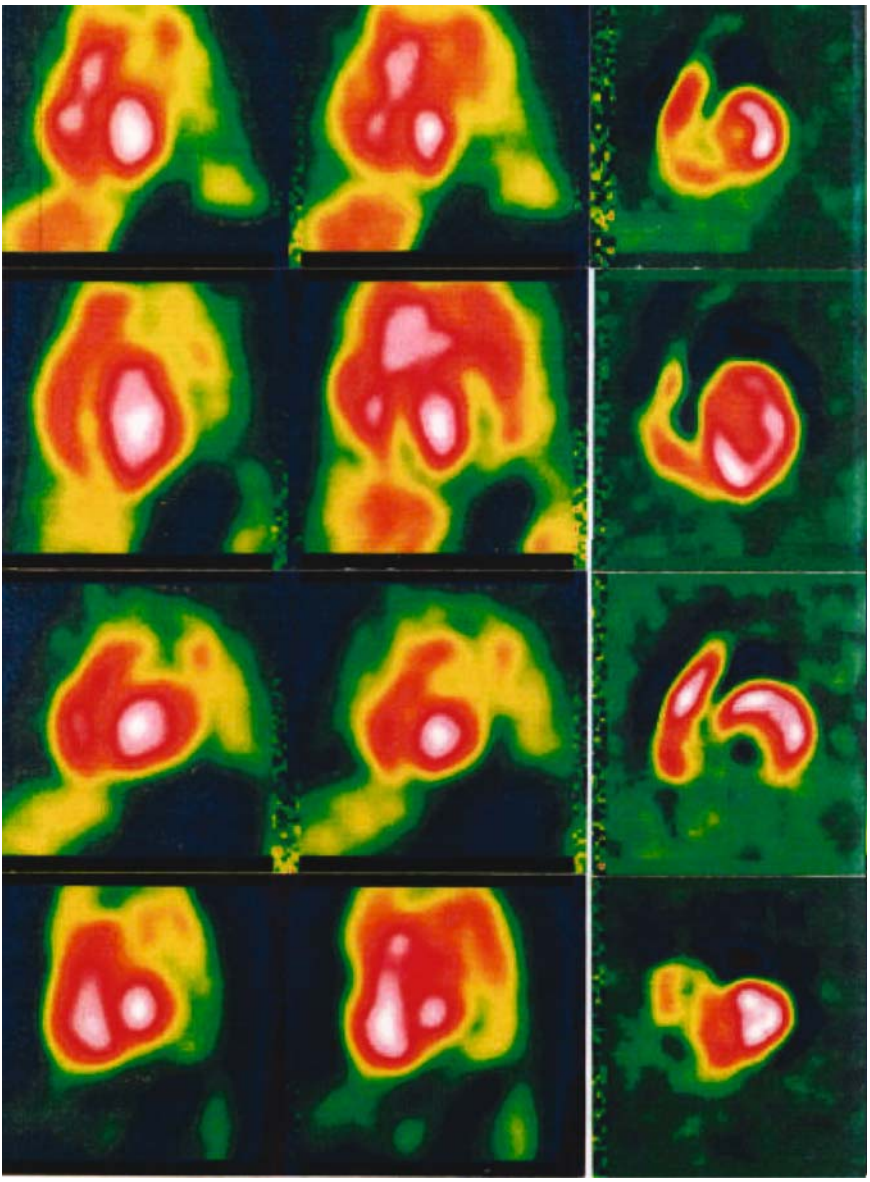


FIGURE 13. Study of four patients showing end-diastolic images (left column), end-systolic images (middle column) and the covariance map of the LV RoI reference (right column). (Reprinted with permission from ref.[11]).

cardiology, automated RoI delineation for structures of interest in dynamic functional imaging, determination of the attenuation map in emission tomography and construction of voxel-based anthropomorphic phantoms based on high resolution anatomical images. For the latter, the interested reader is refereed to a recent book describing the development of such computational models in connection with Monte Carlo modelling tools in nuclear imaging.[109]

In diagnostic and oncological nuclear medicine imaging, if the intention is to measure the volume and uptake in a specific organ or lesion, it may be convenient to make preliminary phantom studies on an object of similar shape and size with the assumption that the size, shape, location, and density of the simulated object and surrounding tissues are comparable to that of the patient to be measured. Absolute quantitation by SPECT and PET is feasible now and has been shown to be a clinically reliable and useful technique. For example, the accurate determination of thyroid volume may be an important factor in the calculation of the most appropriate dose of therapeutic radioiodine.[9,110] The volume measurement of different organs and lesions using SPECT and PET has been widely applied in the clinical field.[10,29,111–113] The major problems encountered in functional volume quantitation are image segmentation and imperfect system transfer function. The difficulty in image segmentation is compounded by the low spatial resolution and high noise characteristics of nuclear images.[114] Long *et al.*[28] have evaluated the relative performance of a number of image segmentation methods for volume quantitation on simulated SPECT images. This included: (i) manual methods in which the operator defines the surface boundary of an image structure by delineating its cross-sectional edges in the contiguous slices incorporating it, (ii) algorithms based on simple thresholding techniques to define the edge of the organ where voxels are flagged as part of the volume of interest if they contain counts greater than some pre-defined threshold value, (iii) the adaptive thresholding method which uses a nonparametric procedure of automatic threshold selection formulated by Otsu,[20] and (iv) two-dimensional (2-D) and three-dimensional (3-D) implementations of a multifeature determination of edge location based on calculation of the gradient (magnitude and direction of the greatest rate of change in counts). They conclude that the 3-D gradient-based method of image segmentation is the one requiring minimal operator intervention while providing the most accurate and consistent estimates of object volume across changes in object size and contrast. A survey of different methods proposed for evaluation of image segmentation algorithms including the criteria for choosing figures of merit is given elsewhere.[115]

Segmentation of radionuclide attenuation images for the purpose of reducing noise propagation in short transmission scanning or X-ray CT and MRI for identifying tissues with different attenuation properties is also an appealing research area.[116] In particular, Pan *et al.*[117] reported on an accurate segmentation method to determine the attenuation map from Compton

scatter and photopeak window SPECT projection data. During the last decade, techniques using transmission image segmentation and tissue classification have been proposed to minimize the acquisition time and increase the accuracy of the attenuation correction process, while preserving or even reducing the noise level. The reconstructed transmission image pixels are segmented into populations of uniform attenuation. In a clinical setting, segmentation algorithms must be designed to balance image quality and computational time of the emission tomograms. The majority of segmentation methods used for attenuation correction fall into one of the following two classes (see chapter 6): histogram-based thresholding techniques[118–120] and fuzzy-clustering based segmentation techniques.[12,13] Other interesting approaches to segment noisy transmission data include the use of active contour models,[121] neural networks[122] morphological segmentation[123] and hidden Markov modelling.[124] An alternative to segmentation of transmission images with the goal of reducing noise in PET transmission measurements is the use of coregistered segmented MRI data in functional brain imaging.[125] The T_1-weighted brain MR images were segmented by means of a fuzzy clustering technique which identifies tissues of significantly different density and composition. A promising technique for segmentation of skull in human T_1-weighted MR images, which could be useful for more robust implementation of the method mentioned above, consists in performing skull segmentation using a sequence of mathematical morphological operations.[126]

5. Summary

It is gratifying to see in overview the progress that image segmentation has made in the last ten years, from operator-dependent manual delineation of structures, through simple thresholding, the use of classifiers and fuzzy clustering, and more recently atlas-guided approaches incorporating prior information. Recent developments have been enormous particularly in the last ten years, the main opportunities striving towards improving the accuracy, precision, and computational speed through efficient implementation in conjunction with decreasing the amount of operator interaction. The application of medical image segmentation is well established in research environments and is still limited in clinical settings to institutions with advanced physics and extensive computing support. As the above mentioned challenges are met, and experience is gained, implementation of validated techniques in commercial software packages will be useful to attract the interest of the clinical community and increase the popularity of these tools. It is expected that with the availability of computing power in the near future, more complex and ambitious computer intensive segmentation algorithms will become clinically feasible.

Acknowledgments. HZ was supported in part by the Swiss National Science Foundation under grant SNSF 3152A0-102143 and Geneva University Hospital R&D funds under grant PRD-04-1-08. AB wishes to express his sincere thanks to Dr. F. Behloul, Leiden University Medical Center (Department of Radiology) for her helpful discussions and suggestions. Illustrations presented in *Artificial Neural Networks* section are essentially inspired from her PhD dissertation.

References

1. Suetens P., Bellon, E., Vandermeulen, D. *et al.*, Image segmentation: methods and applications in diagnostic radiology and nuclear medicine. *Eur J Radiol* **17:** 14-21 (1993).
2. Deklerck R., Cornelis, J. and Bister, M., Segmentation of medical images. *Image Vision Comput* **11:** 1101-1113 (1993).
3. Pham D. L., Xu, C. and Prince, J. L., A survey of currents methods in medical image segmentation. *Ann Rev Biomed Eng* **2:** 315-337 (2000).
4. Boudraa A. O., Mallet, J.-J., Besson, J.-E. *et al.*, Left ventricle automated detection method in gated isotopic ventriculography using fuzzy clustering. *IEEE Trans Med Imaging* **12:** 451-465 (1993).
5. Boudraa A. O., Champier, J., Cinotti, L. *et al.*, Delineation and quantitation of brain lesions by fuzzy clustering in positron emission tomography. *Comput Med Imaging Graph* **20:** 31-41 (1996).
6. Boudraa A. O., Arzi, M., Sau, J. *et al.*, Automated detection of the left ventricular region in gated nuclear cardiac imaging. *IEEE Trans Biomed Eng* **43:** 430-437 (1996).
7. Behloul F., Boudraa, A.O., Lelieveldt, B. P. *et al.*, Myocardium extraction in positron emission tomography based on soft computing. *Comput Med Imaging Graph* **25:** 277-286 (2001).
8. Bezdek J. C., *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York, (1981).
9. Zaidi H., Organ volume estimation using SPECT. *IEEE Trans Nucl Sci* **43:** 2174-2182 (1996).
10. Erdi Y. E., Mawlawi, O., Larson, S. M. *et al.*, Segmentation of lung lesion volume by adaptive positron emission tomography image thresholding. *Cancer* **80:** 2505-2509 (1997).
11. Boudraa A. O., Champier, J., Djebali, M. *et al.*, Analysis of dynamic nuclear cardiac images by covariance function. *Comput Med Imaging Graph* **23:** 181-191 (1999).
12. Bettinardi V., Pagani, E., Gilardi, M. *et al.*, An automatic classification technique for attenuation correction in positron emission tomography. *Eur J Nucl Med* **26:** 447-458 (1999).
13. Zaidi H., Diaz-Gomez, M., Boudraa, A. O. *et al.*, Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging. *Phys Med Biol* **47:** 1143-1160 (2002).
14. Peter J., Freyer, R., Smith, M. *et al.*, Nuclear medicine image segmentation using a connective network. *IEEE Trans Nucl Sci* **44:** 1583-1590 (1997).

15. Clarke L. P., Velthuizen, R. P., Camacho, M. A. *et al*., MRI segmentation: methods and applications. *Magn Reson Imaging* **13:** 343-368 (1995).
16. Pal N. R. and Pal, S. K., A review on image segmentation techniques. *Pattern Recognition* **26:** 1277-1294 (1993).
17. Gonzallez R. C. and Woods, R. E., *Digital image processing*., Pearson Education International, 2nd edition, Upper Saddle River, NJ, (2002).
18. Canny J. F., A computational approach to edge detection. *IEEE Trans Pattern Anal Machine Intell* **8:** 679-698 (1986).
19. Albregtsen F., ''Non-parametric histogram thresholding methods. Error versus relative object area'' Proc. 8th Scandinavian Conf. Image Processing, *pp* 273-280 (1993).
20. Otsu N., A thresholding selection method from gray-level histograms. *IEEE Trans Sys Man Cyber* **9:** 62-66 (1979).
21. Reddi S. S., Rudin, S. F. and Keshavan, H. R., An optimal multiple threshold scheme for image segmentation. *IEEE Trans Syst Man and Cybern* **14:** 661-665 (1984).
22. Kittler J. and Illingworth, J., Minimum error thresholding. *Patt Recogn* **19:** 41-47 (1986).
23. Pun T., A new method for gray-level picture thresholding using the entropy of the histogram. *Signal Process* **2:** 223-237 (1980).
24. Pun T., Entropic thresholding: a new approach. *Comput Graph Image Process* **16:** 210-239 (1981).
25. Kapur J. N., Sahoo, P. K. and Wong, A., A new method for gray-level picture thresholding using the entropy of the histogram. *Graph Models Imag Proc* **29:** 273-285 (1985).
26. Mortelmans L., Nuyts, J., Van Pamel, G. *et al*., A new thresholding method for volume determination by SPECT. *Eur J Nucl Med* **12:** 284-290 (1986).
27. Groshar D., Frankel, A., Iosilevsky, G. *et al*., Quantitation of renal uptake of technetium-99m DMSA using SPECT. *J Nucl Med* **30:** 246-250 (1989).
28. Long D. T., King, M. A. and Sheehan, J., Comparative evaluation of image segmentation methods for volume quantitation in SPECT. *Med Phys* **19:** 483-489 (1992).
29. Alaamer A. S., Fleming, J. S. and Perring, S., Evaluation of the factors affecting the accuracy and precision of a technique for quantification of volume and activity in SPECT. *Nucl Med Commun* **15:** 758-771 (1994).
30. Savolainen S., Pohjonen, H., Sipila, O. *et al*., Segmentation methods for volume determination with 111In/99mTc SPET. *Nucl Med Commun* **16:** 370-377 (1995).
31. Fleming J. S. and Alaamer, A. S., A rule based method for context sensitive threshold segmentation in SPECT using simulation. *Phys Med Biol* **43:** 2309-2323 (1998).
32. Wang S. and Haralick, R. M., Automatic multithreshold selection. *Comput Vis Graph Imag Process* **25:** 46-67 (1984).
33. Carlotto J. M., Histogram analysis using scale-space approach. *IEEE Trans Pattern Anal Mach Intell* **9:** 121-129 (1987).
34. Papamarkos N. and Gatos, B., A new approach for multithreshold selection. *Graph Models Imag Proc* **56:** 357-370 (1994).
35. Van Laere K. J., Warwick, J., Versijpt, J. *et al*., Analysis of clinical brain SPECT data based on anatomic standardization and reference to normal data: an ROC-

based comparison of visual, semiquantitative, and voxel-based methods. *J Nucl Med* **43:** 458-469 (2002).
36. Slomka P. J., Hurwitz, G. A., St. Clement, G. *et al.*, Three-dimensional demarcation of perfusion zones corresponding to specific coronary arteries: application for automated interpretation of myocardial SPECT. *J Nucl Med* **36:** 2120-2126 (1995).
37. Skurichina M. and Duin, R. P. W., "Stabilizing classifiers for very small sample sizes" Proc. 13th Int. Conf. Patt. Recogn. Vol. 12; *pp* 891-896 (1996).
38. Raudys S. J. and Duin, R. P. W., Expected classification error of the Fisher linear classifier with pseudo-inverse covariance matrix. *Patt Recogn Letters* **19:** 385-392 (1998).
39. Hamamoto Y., Fujimoto, Y. and Tomita, S., On the estimation of a covariance matrix in designing Parzen classifiers. *Patt Recogn* **29:** 1751-1759 (1996).
40. Vapnik V. N., *Statistical learning theory*, John Wiley & Sons Inc, New York, (1998).
41. Stoeckel J., Malandain, G., Migneco, O. *et al.*, "Classification of SPECT images of normal subjects versus images of Alzheimer's disease patients" in Conf. Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI'01), Utrecht, The Netherlands, Vol. 2208; *pp* 666-674 (2001).
42. Raudys S. J. and Jain, A. K., Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Trans Patt Anal Mach Intell* **13:** 252-264 (1991).
43. Xie X. L. and Beni, G., A validity measure for fuzzy clustering. *IEEE Trans Pattern Anal Machine Intell* **13:** 841-847 (1991).
44. Pal N. R. and Bezdek, J. C., On cluster validity for the fuzzy c-means model. *IEEE Trans Fuzzy Syst* **3:** 370-379 (1995).
45. Boudraa A. O., Dynamic estimation of number of clusters in data sets. *IEEE Electronics Letters* **35:** 1606-1608 (1999).
46. Coleman G. B. and Andrews, H. C., Image segmentation by clustering. *Proc. IEEE* **5:** 773-785 (1979).
47. Dempster A. P., Laird, N. M. and Rubin, D. B., Maximum likelihood estimation from incomplete data via the EM algorithm. *J Royal Stat Soc B* **39:** 1-38 (1977).
48. Acton P. D., Pilowsky, L. S., Kung, H. F. *et al.*, Automatic segmentation of dynamic neuroreceptor single-photon emission tomography images using fuzzy clustering. *Eur J Nucl Med* **26:** 581-590 (1999).
49. Boudraa A. O. and Clarysse, P., "Fast fuzzy grey level image segmentation method" World Congress on Medical Physics and Biomedical Engineering, Nice, France, Vol. 35; *pp* 686 (1997).
50. Krishnapuram R. and Keller, J. M., A possibilistic approach to clustering. *IEEE Trans Fuzzy Syst* **1:** 98-110 (1993).
51. Beni G. and Liu, X., A least biased fuzzy clustering method. *IEEE Trans Pattern Anal Machine Intell* **16:** 954-960 (1994).
52. Barni M., Cappellini, V. and Mecocci, A., Comments on "A possibilistic approach to clustering". *IEEE Trans Fuzzy System* **4:** 393-396 (1996).
53. Marr D. and Hildreth, E., Theory of edge detection. *Proc Roy Soc London* **207:** 187-217 (1980).
54. Huertas A. and Medioni, G., Detection of intensity changes with subpixel accuracy using Laplacian-Gaussian masks. *IEEE Trans Patt Anal Mach Intell* **8:** 651-664 (1986).

55. Wang C. X., Small, L., Synder, W. E. *et al.*, ''Edge gated cardiac nuclear medicine images'' 7th IEEE Symposium on computer-based medical systems, Winston-Salem, NC, *pp* 28-33 (1994).
56. Dai X., Snyder, W. E., Bilbro, G. L. *et al.*, Left-ventricle boundary detection from nuclear medicine images. *J Digit Imaging* **11:** 10-20 (1998).
57. Long D. T., King, M. A. and Gennert, M. A., Development of a 3D gradient-based method for volume quantitation in SPECT. *IEEE Trans Nucl Sci* **38:** 748-754 (1991).
58. Hillel P. G. and Hastings, D. L., A three-dimensional second-derivative surface-detection algorithm for volume determination on SPECT images. *Phys Med Biol* **38:** 583-600 (1993).
59. Nihsimura Y., Katafuchi, T., Hirase, Y. *et al.*, Measurement of left ventricular chamber and myocardial volume in hypertrophic cardiomyopath patients by ECG-gated myocardial perfusion SPECT: Application of newly developed edge-detection algorithm. *Jpn J Radiol Technol* **58:** 1586-1591 (2002).
60. Langleben D. D., Austin, G., Krikorian, G. *et al.*, Interhemispheric asymmetry of regional cerebral blood flow in prepubescent boys with attention deficit hyper-activity disorder. *Nucl Med Comm* **22:** 1333-1340 (2001).
61. Li S. Z., *Markov Random field modeling in computer vision*, Springer-Verlag, Tokyo, (1995).
62. Geman S. and Geman, D., Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Trans Patt Anal Mach Intell* **6:** 721-741 (1984).
63. Besag J., On the statistical analysis of dirty pictures. *J Royal Stat Soc B* **48:** 259-302 (1986).
64. Derin H. and Elliott, H., Modeling and segmentation of noisy and textured images using Gibbs random fields. *IEEE Trans Patt Anal Mach Intell* **9:** 39-55 (1987).
65. Chen J. L., Gunn, S. R., Nixon, M. S. *et al.*, ''Markov random field models for segmentation of PET images'' in: *Proc. Information Processing in Medical Imaging*, edited by and R.M. Leahy M.F. Insana, (2001), *pp* 468-474.
66. Lin K. P., Chen, T. S., Yao, W. J. *et al.*, ''Dynamic PET-FDG image segmentation using Markov Random field method'' Proc SPIE Int Soc Opt Eng, Vol. 3338; *pp* 1198-1204 (1998).
67. Mignotte M. and Meunier, J., ''Unsupervised restoration of brain SPECT volumes'' Vision Interfaces (VI 2000), Montreal, Canada, *pp* 55-60 (2000).
68. Mignotte M., Meunier, J., Soucy, J. P. *et al.*, ''Segmentation and classification of brain SPECT images using 3D MRF and density mixture estimations'' 5th World Multi-Conference on Systemics, Cybernetics and Informatics, Concepts and Applications of Systemics and Informatics, Orlando, FL, USA, Vol. 10; *pp* 239-244 (2001).
69. Lin H. D., Lin, K. P., Chung, B. T. *et al.*, ''Using kinetic parameter analysis of dynamic FDOPA-PET for brain tissue classification'' Proc SPIE Int Soc Opt Eng, Vol. 4683; *pp* 436-449 (2002).
70. Kohonen T., *Self-organization and associative memory*, Springer-Verlag, Berlin, (1989).
71. Pao Y. H., *Adaptive pattern recognition and neural networks*, Addison-Wesley, New York, (1989).

72. Eckhorn R., Reitboeck, H. J., Arendt, M. *et al*., Feature linking via synchronization among distributed assemblies: simulation of results from cat visual cortex. *Neural Computation* **2:** 293-307 (1990).
73. Hertz J., Krogh, A. and Palmer, R. G., *Introduction to the theory of neural computation*, Addison-Wesley, New York, (1991).
74. Kosko B., *Neural networks and fuzzy systems: A dynamical systems approach to machine intelligence*, Prentice-Hall, Englewood Cliffs, New Jersey, (1991).
75. Blanz W. E. and Pal, S. L., Image segmentation using a neural network. *Biol. Cybern*. **66:** 151-158 (1991).
76. Ghosh A., Pal, N. R. and Pal, S. K., Object background classification using Hopfield type neural network. *Int J Pattern Recogn Mach Intell* **6:** 989-1008 (1992).
77. Ghosh A., Pal, N. R. and Pal, S. K., Self-organization for object extraction using multilayer neural networks and fuzziness measures. *IEEE Trans Fuzzy Syst* **1:** 54-68 (1993).
78. Haykin S., *Neural networks. A comprehensive foundation*, Macmillan, New York, (1994).
79. Keller P. E. and McKinnon, A. D., "Segmentation of medical imagery with pulse-coupled neural networks." Int Joint Conf Neural Networks, Washington, DC, USA, Paper #130, available on CDROM (1999).
80. Behloul F., "Fusion de données TEP-IRM par méthodes neuro-floues pour l'étude de la viabilité du myocarde" Ph.D Thesis, INSA de Lyon, 1999.
81. Behloul F., Lelieveldt, B. P., Boudraa, A.O. *et al*., Neuro-fuzzy systems for computer-aided myocardial viability assessment. *IEEE Trans Med Imaging* **20:** 1302-1313 (2001).
82. Musavi M. T., Ahmed, W., Chan, K. H. *et al*., On the training of radial basis function classifiers. *Neural Networks* **5:** 595-603 (1992).
83. Hwang Y. S. and Bang, S. Y., An Efficient method to construct a radial basis function neural network classifier. *Neural Networks* **8:** 1495-1503 (1997).
84. Kang H. B. and Walker, E. L., Characterising and controlling approximation in hierarchical perceptual grouping. *Fuzzy Sets Syst* **65:** 187-223 (1994).
85. Kosko B., Fuzzy entropy and condition. *Inform Sci* **40:** 165-174 (1989).
86. Pal N. R. and Pal, S. K., Higher order of fuzzy entropy and hybrid entropy of a set. *Inform Sci* **61:** 211-231 (1992).
87. Deluca A. and Termini, S., A definition of non probabilistic entropy in the setting of fuzzy set theory. *Inform and Control* **20:** 301-312 (1972).
88. Kandel A., *Fuzzy mathematical technique with applications*, Addison-Wesley, New York, (1986).
89. Kass M., Witkin, A. and Terzopoulos, D., Snakes: active contour models. *Int J Comput Vision* **1:** 321-331 (1988).
90. Leymarie F. and Levine, M. D., Tracking deformable objects in the plane using an active contour model. *IEEE Trans Patt Anal Mach Intell* **15:** 617-634 (1993).
91. William D. J. and Shash, M., A fast algorithm for active contours and curvature estimation. *Image Understanding* **55:** 14-26 (1992).
92. Cohen L. D. and Cohen, I., "A finite element method applied to new active contour models and 3D reconstruction from cross-section" Proc. IEEE, 3rd Int. Conf. Computer Vision, Osaka, Japan, *pp* 587-591 (1990).

93. Amini A. A., Tehrani, S. and Weymouth, T. E., "Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints" Proc. IEEE 2rd Int. Conf. Computer Vision, *pp* 855-867 (1988).
94. Mykkanen J. M., Juhola, M. and Ruotsalainen, U., Extracting VOIs from brain PET images. *Int J Med Inf* **58-59:** 51-57 (2000).
95. Lancaster J., Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Hum Brain Mapp* **5:** 238-242 (1997).
96. Ashburner J. and Friston, K. J., Nonlinear spatial normalization using basis functions. *Hum Brain Mapp* **7:** 254-266 (1999).
97. Minoshima S., Koeppe, R. A., Frey, K. A. *et al*., Anatomic standardization: linear scaling and nonlinear warping of functional brain images. *J Nucl Med* **35:** 1528-1537 (1994).
98. Sandor S. and Leahy, R., Surface-based labeling of cortical anatomy using a deformable atlas. *IEEE Trans Med Imaging* **16:** (1997).
99. Thompson P. and Toga, A. W., Detection, visualization and animation of abnormal anatomic structure with a probabilistic brain atlas based on random vector field transformations. *Med Image Analysis* **1:** 271-294 (1997).
100. Wong K.-P., Dagan, F., Meikle, S. R. *et al*., Segmentation of dynamic PET images using cluster analysis. *IEEE Trans Nucl Sci* **49:** 200-207 (2002).
101. O' Sullivan F., Imaging radiotracer model parameters in PET: A mixture analysis approach. *IEEE Trans Med Imaging* **12:** 399-412 (1993).
102. Kimura Y., Hsu, H., Toyama, H. *et al*., Improved signal-to-noise ratio in parametric images by cluster analysis. *Neuroimage* **9:** 554-561 (1999).
103. Bossuyt A., Deconinck, F., Lepoudre, R. *et al*., "The temporal Fourier transform applied to functional isotopic imaging" in: *Information processing in medical imaging*, edited by R Di Paola and E Kahn INSERM 88, Paris, (1979), *pp* 397-408.
104. Botvinick E., Dunn, R., Frais, M. *et al*., The phase image: its relationship to patterns of contraction and conduction. *Circulation* **65:** 551-560 (1982).
105. Pavel D., Byrom, E., Lam, W. *et al*., Detection and quantification of regional wall motion abnormalities using phase analysis of equilibrium gated cardiac studies. *Clin Nucl Med* **8:** 315-321 (1983).
106. Barber D. C., The use of principal components in the quantitative analysis of gamma camera dynamic studies. *Phys Med Biol* **25:** 283-292 (1980).
107. Cavailloles F., Bazin, J. P. and Di Paola, R., Factor analysis in gated cardiac studies. *J Nucl Med* **25:** 1067-1079 (1984).
108. Houston A. S. and Sampson, W. F., A quantitative comparison of some FADS methods in renal dynamic studies using simulated and phantom data. *Phys Med Biol* **42:** 199-217 (1997).
109. Zaidi H. and Sgouros, G., *Therapeutic applications of Monte Carlo calculations in nuclear medicine*, Institute of Physics Publishing, Bristol, (2002).
110. Zaidi H., Comparative methods for quantifying thyroid volume using planar imaging and SPECT *J Nucl Med* **37:** 1421-1426 (1996).
111. Erdi Y. E., Wessels, B. W., Loew, M. H. *et al*., Threshold estimation in single photon emission computed tomography and planar imaging for clinical radioimmunotherapy. *Cancer Res* **55:** 5823s-5826s (1995).
112. Erdi Y. E., Humm, J. L., Imbriaco, M. *et al*., Quantitative bone metastases analysis based on image segmentation. *J Nucl Med* **38:** 1401-1406 (1997).

113. Sjorgreen K., Ljungberg, M. and Strand, S. E., Parameters influencing volume and activity quantitation in SPECT. *Acta Oncol* **35:** 323-330 (1996).
114. King M. A., Long, D. T. and Brill, A. B., SPECT volume quantitation: influence of spatial resolution, source size and shape, and voxel size. *Med Phys* **18:** 1016-1024 (1991).
115. Zhang Y. J., A survey on evaluation methods for image segmentation. *Patt Recogn Letters* **29:** 1335-1346 (1996).
116. Zaidi H. and Hasegawa, B. H., Determination of the attenuation map in emission tomography. *J Nucl Med* **44:** 291-315 (2003).
117. Pan T.-S., King, M. A., de Vries, D. J. *et al*., Segmentation of the body and lungs from Compton scatter and photopeak window data in SPECT: a Monte-Carlo investigation. *IEEE Trans Med Imaging* **15:** 13-24 (1996).
118. Meikle S. R., Dahlbom, M. and Cherry, S. R., Attenuation correction using count-limited transmission data in positron emission tomography. *J Nucl Med* **34:** 143-150 (1993).
119. Xu M., Cutler, P. and Luk, W., An adaptive local threshold segmented attenuation correction method for whole-body PET imaging. *IEEE Trans Nucl Sci* **43:** 331-336 (1996).
120. Bilger K., Kupferschlager, J., Muller-Schauenburg, W. *et al*., Threshold calculation for segmented attenuation correction in PET with histogram fitting. *IEEE Trans Nucl Sci* **48:** 43-50 (2001).
121. Tai Y.-C., Lin, K.-P., Dahlbom, M. *et al*., A hybrid attenuation correction technique to compensate for lung density in 3-D total body PET. *IEEE Trans Nucl Sci* **43:** 323-330 (1996).
122. Yu S. K. and Nahmias, C., Segmented attenuation correction using artificial neural networks in positron tomography. *Phys Med Biol* **41:** 2189-2206 (1996).
123. Riddell C., Brigger, P., Carson, R. E. *et al*., The watershed algorithm: a method to segment noisy PET transmission images. *IEEE Trans Nucl Sci* **46:** 713-719 (1999).
124. Anderson J. M. M., Srinivasan, R., Mair, B. A. *et al*., Hidden Markov model based attenuation correction for positron emission tomography. *IEEE Trans Nucl Sci* **49:** 2103-2111 (2002).
125. Zaidi H., Montandon, M.-L. and Slosman, D. O., Magnetic resonance imaging-guided attenuation and scatter corrections in three-dimensional brain positron emission tomography. *Med Phys* **30:** 937-948 (2003).
126. Dogdas B., Shattuck, D. W. and Leahy, R. M., "Segmentation of the skull in 3D human MR images using mathematical morphology" Proceedings of SPIE Medical Imaging Conference, San Diego, USA, Vol. 4684; *pp* 1553-1562 (2002).

11
Monte Carlo Modeling in Nuclear Medicine Imaging

H. Zaidi*

1. Conceptual Role of Monte Carlo Simulations

The Monte Carlo method describes a very broad area of science, in which many processes, physical systems and phenomena are simulated by statistical methods employing random numbers. The general idea of Monte Carlo analysis is to create a model, which is as similar as possible to the real physical system of interest, and to create interactions within that system based on known probabilities of occurrence, with random sampling of the probability density functions (PDFs). As the number of individual events (called histories) is increased, the quality of the reported average behaviour of the system improves, meaning that the statistical uncertainty decreases. Virtually, any complex system can in principle be modelled; perhaps there is a desire to model the number of cars passing a particular intersection during certain times of the day, to optimize traffic management, or to model the number of people that will make transactions in a bank, to evaluate the advantages of different queuing systems. If the distribution of events that occur in a system is know from experience, a PDF can be generated and sampled randomly to simulate the real system. A detailed description of the general principles of the Monte Carlo method is given elsewhere.[1-4]

In the specific application of interest in this volume (simulation of imaging systems), radiation transport is simulated by the creation of charged particles or photons from a defined source region, generally with a random initial orientation in space, with tracking of the particles as they travel through the system, sampling the probability PDFs for their interactions to evaluate their trajectories and energy deposition at different points in the system. The interactions determine the penetration and motion of particles, but, more importantly, the energy deposited during each interaction gives the radiation absorbed dose, when divided by the appropriate values of mass. With sufficient numbers of interactions, the mean absorbed dose at points of interest will

*PD Dr H. Zaidi, Geneva University Hospital, Division of Nuclear Medicine, CH-1211 Geneva, Switzerland

be given with acceptable uncertainties. The central issues include how well the real system of interest can be simulated by a geometrical model, how many histories (i.e. how much computer time) are needed to obtain acceptable uncertainties (usually around 5%, no more than 10%) and how can measured data be used to validate the theoretical calculations.

Monte Carlo techniques have become one of the most popular tools in different areas of medical physics in general and medical imaging in particular following the development and subsequent implementation of powerful computing systems for clinical use.[1] In particular, they have been extensively applied to simulate processes involving random behaviour and to quantify physical parameters that are difficult or even impossible to calculate analytically or to determine by experimental measurements. The applications of the Monte Carlo method in medical physics cover almost all topics, including radiation protection, diagnostic radiology, radiotherapy and nuclear medicine, with an increasing interest in exotic and new applications, such as intravascular radiation therapy, boron neutron capture therapy and synovectomy. With the rapid development of computer technology, Monte Carlo-based reconstruction in emission tomography and treatment planning for radiation therapy is becoming practicable.

This chapter briefly reviews Monte Carlo simulation packages used worldwide and summarizes their application in diagnostic and therapeutic nuclear medicine. Emphasis is given to applications involving accurate modelling of radiation transport in matter. Object modelling through mathematical and voxel-based phantoms is also summarised highlighting potential advantages and drawbacks of each approach in modelling simulated objects. The historical developments and computational aspects of the Monte Carlo method mainly related to random number generation, radiation transport, sampling and variance reduction fall outside the scope of this chapter and are discussed in the above mentioned reviews and books. It should be emphasized that due to limited space, the references contained herein are for illustrative purposes and are not inclusive; no implication that those chosen are better than others not mentioned is intended.[5]

2. Monte Carlo Software Packages

Image modelling was employed in the 1970s by Schulz[6] who devised a computer program simulating a rectilinear scanner to study the influence of different imaging protocols on the detectability of lesions. Simulation of gamma camera imaging to assess qualitatively and quantitatively the image formation process and interpretation[7] and to assist development of collimators[8] using deterministic methods and simplifying approximations have been developed mainly to improve speed of operation. For example, Beekman[9] developed a fast analytic simulator of tomographic projection data taking into account attenuation, distance dependent detector response, and scatter

based on an analytical point spread function (PSF) model. Several simplifying approximations were also adopted to improve speed of operation: restriction of the extent of the primary and scatter PSFs, coarse sampling of the PSFs in the direction perpendicular to the camera face and use of a circularly symmetric scatter function.[10]

Many Monte Carlo programs have been in use in the field of nuclear imaging and internal dosimetry with many of them available as open source codes in the public domain. Basically there are two categories of software packages: general purposes Monte Carlo codes developed for high-energy or general medical physics applications and dedicated software packages developed mainly and optimized for nuclear medicine imaging applications. These are highly sophisticated tools requiring advanced computer skills and familiarity with radiation transport theory. Each category has its own advantages and drawbacks; the motivations for the choice of a particular code being mainly dictated by availability of code and documentation, the user's needs, experience and computer skills.

2.1 General-purpose Monte Carlo Programs

Table 1 lists widely used public domain Monte Carlo codes together with a short description of their key features. Most of the packages mentioned below run virtually on different platforms and operating systems and are available either directly from the developers or through the official channels (RSIC or NEA). One of the most popular computer codes developed specifically for medical physics applications is the electron gamma shower (EGS) system, which is a general-purpose package for Monte Carlo simulation of the coupled transport of electrons and photons in an arbitrary geometry for particles with energies from a few keV up to several TeV.[11] The code represents the state-of-the-art of radiation transport simulation because it is very flexible, well documented and extensively tested. Some have referred to the EGS code as the de facto gold standard for clinical radiation dosimetry. EGS is written in MORTRAN, a FORTRAN pre-processor with powerful macro capabilities. EGS is a 'class II' code that treats knock-on electrons and bremsstrahlung photons individually. Such events require predefined energy thresholds and pre-calculated data for each threshold, determined with the cross-section generator PEGS. A new version called EGSnrc reflecting substantial changes and improvements made to the original EGS4 code has been recently reported.[12] Many other ambitious fully integrated all-particle radiation transport codes are now available for users (table 1).

Some research groups devised their own Monte Carlo computer codes built on top of one of the general-purpose codes mentioned above, which serves as core layer giving the opportunity to the developer to construct application-specific modules in a hierarchical layer architecture. This includes the SimSPECT code based on MCNP,[26,27] PET-EGS based on EGS4,[15] and more recently GATE based on GEANT4 code developed at CERN.[34] It should be

TABLE 1. Key features of general-purpose public domain Monte Carlo codes used in modelling nuclear medicine imaging systems.

MC code	General description	Example references
EGS4[11]/EGSnrc[12]	Coupled photons/electrons transport in any material through user specified geometries. Simulation of nuclear imaging systems not specifically included and requires an extensive amount of user programming in Mortran.	13-22
ITS including TIGER CYLTRAN and ACCEPT[23]	Coupled photons/electrons transport in any material through slabs, cylinders or combinatorial. Simulation of nuclear imaging systems not specifically included and requires an extensive amount of user programming in Fortran.	
MCNP[24]/ MCNPX[25]	Coupled neutrons/photons/electrons transport in any material through user-generalised geometry. Simulation of nuclear imaging systems not specifically included and requires an extensive amount of user manipulation of input date files to model complex geometries.	26-29
GEANT[30,31]	Coupled photons/electrons transport in any material through combinatorial geometry. Simulation of nuclear imaging systems not specifically included and requires an extensive amount of user programming in C/C++.	32-34
PENELOPE[35,36]	Coupled photons/electrons transport in any material through combinatorial geometry. Simulation of nuclear imaging systems not specifically included and requires an extensive amount of user programming in C/C++.	37-39
FLUKA[40]	Coupled photons/electrons transport in any material through combinatorial geometry. Simulation of nuclear imaging systems not specifically included and requires an extensive amount of user programming in C/C++.	41

emphasized that as opposed to radiation dosimetry where many studies investigated potential differences between the results obtained using different computational Monte Carlo codes,[42] very few studies addressed this issue in simulation of nuclear medicine imaging. In particular, the various Monte Carlo codes developed to simulate nuclear medicine imaging systems use different photon cross section libraries, which might impact the outcome of the generated data sets. It has been shown that the effect of using different photon cross-section libraries is noticeable on both generated PET projection data and reconstructed images.[43] Standardization of interaction libraries and other common computational models will help to eliminate potential differences between the results obtained with different codes.

2.2 Dedicated Monte Carlo Simulation Packages

Table 2 lists popular and widely used dedicated Monte Carlo codes together with a short description of their key features. The computer codes developed by Dresser[44] and Beck,[45] are among the oldest Monte Carlo programs

TABLE 2. Key features of dedicated Monte Carlo codes used to simulate nuclear medical imaging systems.

MC code	General description
SIMSET[51,52]	Photons transport in any material through voxel-based phantoms. Simulation of SPECT and PET imaging systems included. User modules written in C could be linked.
SIMIND[53,54]	Photons transport in any material through shape- or voxel-based phantoms. Simulation of SPECT imaging systems included. User modules written in Fortran could be linked.
Unnamed[55,56]	Photons transport in any material through shape-based phantoms. Simulation of SPECT imaging systems included. User modules written in Fortran/C could be linked.
MCMATV[57,58]	Photons transport in any material through voxel-based phantoms. Simulation of SPECT imaging systems included. User modules written in Fortran could be linked.
PETSIM[59,60]	Photons transport in any material through shape-based phantoms. Simulation of PET imaging systems included. User modules written in Fortran could be linked.
EIDOLON[49,50]	Photons transport in any material through shape-or voxel-based phantoms. Simulation of 3D PET imaging systems included. User modules written in C/ Objective-C could be linked.
PET-SORTEO[61]	Photons transport in any material through shape-or voxel-based phantoms. Simulation of 3D PET and transmission imaging systems included.

developed specifically for modelling scintillation cameras. The latter was extensively used by the group of Duke University for scatter modelling, correction and image reconstruction through inverse Monte Carlo in SPECT.[46] Likewise, Keller and Lupton pioneered the modelling of cylindrical and multi-ring PET systems using the Monte Carlo method.[47,48]

An example of dedicated codes is the *Eidolon* Monte Carlo simulator, which was developed using modern software engineering techniques mainly for fully 3D multi-ring PET imaging.[49] The code was written in Objective-C, an object-oriented programming language based on ANSI C. The first version of the program was developed using the NextStep development environment. A modular design featuring dynamically loadable program elements or bundles was adopted for software design. The basic building block is a *model element* object class which allows elements to be browsed, inspected, adjusted, created and destroyed through a graphical inspector. The graphical interface allows the user to select scanner parameters such as the number of detector rings, detector material and sizes, energy discrimination thresholds and detector energy resolution. It also allows to choose either a complex anthropomorphic phantom or a set of simple 3D shapes, such as parallelepiped, ellipsoid or cylindroid for both the annihilation sources and the scattering media, as well as their respective activity concentrations and chemical compositions. The user has the possibility to view the reference source image and sinogram data sets as they are generated and are

periodically updated. An implementation of the software on a high-performance parallel platform was also reported.[50] The current version runs on most of the current platforms and operating systems supporting the GNU C compiler, which should allow subdividing time-consuming simulations on geographically distributed platforms taking advantage of the latest developments in Grid technology.

3. Application of Monte Carlo Techniques in Nuclear Medicine Imaging

There has been an enormous increase and interest in the use of Monte Carlo techniques in all aspects of nuclear imaging instrumentation design and quantification, including planar imaging, SPECT, PET and multimodality imaging devices.[2,3,62] However, due to computer limitations, the method has not yet fully lived up to its potential. With the advent of high speed supercomputers, the field has received increased attention, particularly with parallel algorithms, which have much higher execution rates. Figure 1 illustrates the principles and main components of Monte Carlo or statistical simulation as applied to a cylindrical multi-ring PET imaging system.[49] Assuming that the behaviour of the imaging system can be described by PDFs, then the Monte Carlo simulation can proceed by sampling from these PDFs, which

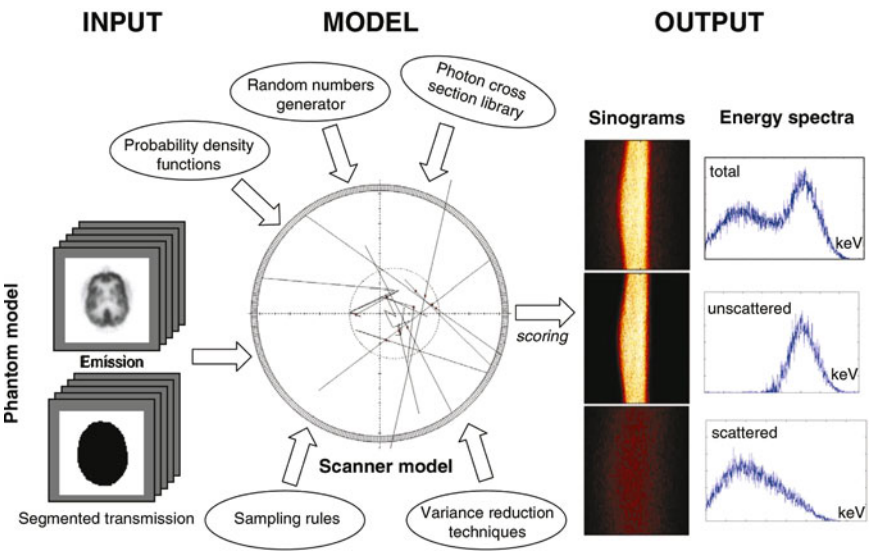


FIGURE 1. Principles and main components of a Monte Carlo program dedicated for simulation of cylindrical multi-ring PET imaging systems.

necessitates a fast and effective way to generate uniformly distributed random numbers. Photon emissions are generated within the phantom and are transported by sampling from PDFs through the scattering medium (transmission image) and detection system until they are absorbed or escape the volume of interest without hitting the crystal matrices or with insufficient energy deposit. The outcomes of these random samplings, or trials, must be accumulated or tallied in an appropriate manner to produce the desired result, but the essential characteristic of the Monte Carlo method is the use of random sampling techniques to arrive at a solution of the physical problem.

Following the review by Zaidi,[2] the applications of the Monte Carlo method in nuclear medical imaging cover almost all topics including detector modelling and systems design, image correction and reconstruction techniques, internal dosimetry and pharmacokinetic modelling, with an increasing enthusiastic interest in exotic and exciting new applications such as online PET monitoring of radiation therapy beams.[63] The use of Monte Carlo techniques in the last two areas falls outside the scope of this survey whereas applications in other fields are only briefly summarized in the following sections given the huge amount of literature available today. With the rapid development of computer technology, Monte Carlo-based modelling in a clinical setting is becoming practical.

3.1 Detector Modelling

Since many important questions can be addressed through detailed Monte Carlo analysis, simulation of detector responses and efficiencies is one of the areas which has received considerable attention during the last 4 decades with an early contribution to the field due to Zerby.[64] The hypothesis is that the performance of a single detector element must be extrapolated using Monte Carlo simulations to predict the performance of a multi-element module or a complete camera system.

Many detector modelling applications were developed in the PET field including the pioneering work of Derenzo[65] who simulated arrays of detectors of different materials and sizes to study the effect of the inter-crytal septa and later on to optimize the optical coupling between BGO crystals and PMTs[66] by taking into account the reflection and scattering along the detection system. The search for an appropriate detector for this imaging modality was conducted in a comparative study of several crystals including BGO, CsF and NaI(Tl),[67] BaF_2[68] used in time-of-flight PET, and liquid Xenon.[69] Binkley[70] modelled the impulse response of a PMT, front-end amplifier, and constant fraction discriminator to evaluate the effects of front-end bandwidth and constant fraction delay and fraction for timing-system optimizations of BGO scintillation detectors.

The position-dependent light distribution has been used to measure the 511 keV photon interaction position in the crystal on an event-by-event

basis to reduce radial elongation.[71] Comanor[72] investigated algorithms to identify and correct for detector Compton scatter in hypothetical PET modules with $3 \times 3 \times 30$ mm BGO crystals coupled to individual photosensors. The true crystal of first interaction was determined by the simulation for eventual comparison with the crystal identified by a given algorithm. They reported a misidentification fraction of 12% if the detector has good energy and position resolution when using the position of interaction to identify forward scatter. A simulation tool modelling position encoding multicrystal detectors for PET that treats the interactions of annihilation photons in a scintillator, the geometry of the multicrystal array, as well as the propagation and detection of individual scintillation photons has been designed[73] to assist design studies of a whole-body PET tomograph with the capacity to correct for the parallax error induced by the DOI of annihilation photons.[74] The experimental energy, depth, and transverse position resolutions of BGO block detectors were used as main inputs to the simulations to avoid extensive light transport in position encoding blocks, which can otherwise be accomplished very efficiently using DETECT2000 Monte Carlo simulation of the light transport in scintillation crystals.[75]

The quest for the ideal scintillation crystal to optimize design features of future fully 3D, large axial field-of-view fifth generation PET scanners technology led to comparative assessment studies evaluating the relative performance of different scintillators currently in use or under development for PET applications for the same geometric design (see their physical characteristics in chapter 1). Among the many figures of merit that can be assessed, Figure 2 shows the Noise Equivalent Count Rate (NECR) metric

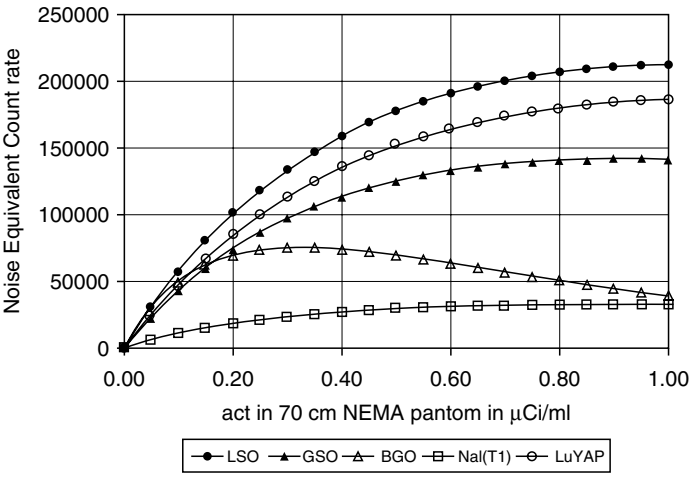


FIGURE 2. Monte Carlo simulation-based NEC comparison for the ECAT EXACT PET scanner for different scintillation crystals with a 450 keV LLD setting. Only detector dead time has been modelled. Reprinted with permission from ref.[76]

of count rate performance measured using the standard NEMA NU-2 cylindrical phantom, clearly demonstrating the advantage of LSO as the detector material for PET.[76]

3.2 Imaging Systems and Collimators Design

Monte Carlo techniques were extensively used to analyze the performance of new collimators design for planar scintillation cameras,[56,77,78] SPECT[79] and PET imaging.[80-82] Practical guidance could be offered for understanding trade-offs that must be considered for clinical imaging. Selective comparisons among different collimators could also be presented for illustrative and teaching purposes. Approaches to the collimator optimization problem, as well as more sophisticated 'task-dependent' treatments and important considerations for collimators design have been performed.[83] The well-known imaging performance parameters of parallel-hole collimators could be compared with those of fan- and cone-beam collimators which have enjoyed considerable success in recent years, particularly for brain SPECT. Monte Carlo calculations were also used to aid in the development of a method for imaging administered therapeutic doses of ^{131}I by using thick lead sheets to the front face of a high-energy parallel-hole collimator.[84]

There has been renewed interest in pinhole collimation for high resolution imaging of small organs and small animals owing to its improved spatial resolution and increase in sensitivity as the distance between the source and the pinhole aperture decreases. The effect of pinhole aperture design parameters on spatial resolution, angle-dependent sensitivity, edge penetration, and scatter fraction for high resolution pinhole imaging was also investigated using Monte Carlo modelling.[29,85,86] Acceleration of simulation time and improvement of simulation efficiency were also considered through the use of forced detection[87] and kernel-based forced detection.[88]

Likewise, Monte Carlo simulations were used in PET to determine the effects of crystals with straight and pointed tips and septa on spatial resolution and efficiency,[89] to compare the singles to true coincident events ratios in well collimated single-slice, multi-slice and open geometry 3D configurations,[80] to evaluate tungsten inter-plane septa of different thicknesses and geometries[81] and to assess the effect of collimation on the scatter fraction.[90]

The Monte Carlo method has also been widely used to optimize the design of SPECT cameras[91,92] and PET tomographs operating in 2D[47,48,93] and fully 3D[13,32,94] acquisition modes. In addition, several researchers used Monte Carlo simulations to study design aspects of dedicated units for brain imaging,[95,96] positron emission mammography[21] and to investigate potential novel design geometries of dedicated small animal imaging systems for both SPECT[97,98] and PET.[99-105]

3.3 *Development and Evaluation of Image Correction and Reconstruction Techniques*

Monte Carlo simulations have been shown to be very useful for the development, validation and comparative evaluation of image correction and reconstruction techniques since it is possible to obtain a reference image to which corrected/reconstructed images should be compared. The development of the inverse Monte Carlo reconstruction technique was a logical consequence of the search for a unified reconstruction algorithm.[106] This method converts the inverse problem into a system of algebraic equations that can be solved by standard analytical or numerical techniques. The principal merits of the method are that, like direct Monte Carlo, it can be applied to complex and multivariable problems, and variance reduction procedures can be applied. The same group used inverse Monte Carlo to perform tomographic reconstruction for SPECT with simultaneous compensation for attenuation, scatter, and distance-dependent collimator resolution.[46] A detection probability matrix is formed by Monte Carlo solution to the photon transport equation for SPECT acquisition from a unit source activity in each reconstruction source voxel. The off-line computation of the system probability matrix by means of Monte Carlo simulations combined with iterative reconstruction proved to be adequate for improvement of image quality and spatial resolution recovery.[107] Given the sparsity and size of the system matrix, some solutions were proposed to tackle this problem through exploitation of symmetries and use of database management system.[108] The interest in fully 3D Monte Carlo-based statistical reconstruction approaches spurred the development of computationally efficient algorithms capable of obtaining highly accurate quantitative data in clinically acceptable computation times.[109]

On the other hand, the capability to theoretically model the propagation of photon noise through emission computed tomography reconstruction algorithms is fundamental in evaluating both the quality and quantitative accuracy of reconstructed images as a function of parameters of the algorithm. Monte Carlo methods can be used to check the validity of the predictions of the theoretical formulations through computation of the sample statistical properties of algorithms under evaluation.[110-112]

Monte Carlo calculations have also been found to be powerful tools to quantify and correct for partial volume effect, photon attenuation and scattering in nuclear medicine imaging since the user has the ability to separate the detected photons into their components: primary events, scatter events, contribution of down-scatter events, etc. Monte Carlo modelling thus allows a detailed investigation of the spatial and energy distribution of Compton scatter which would be difficult to perform using present experimental techniques, even with very good energy resolution semiconductor detectors. Figure 3 shows the energy pulse-height distribution obtained by simulation of a gamma emitting ^{99m}Tc line source in the centre

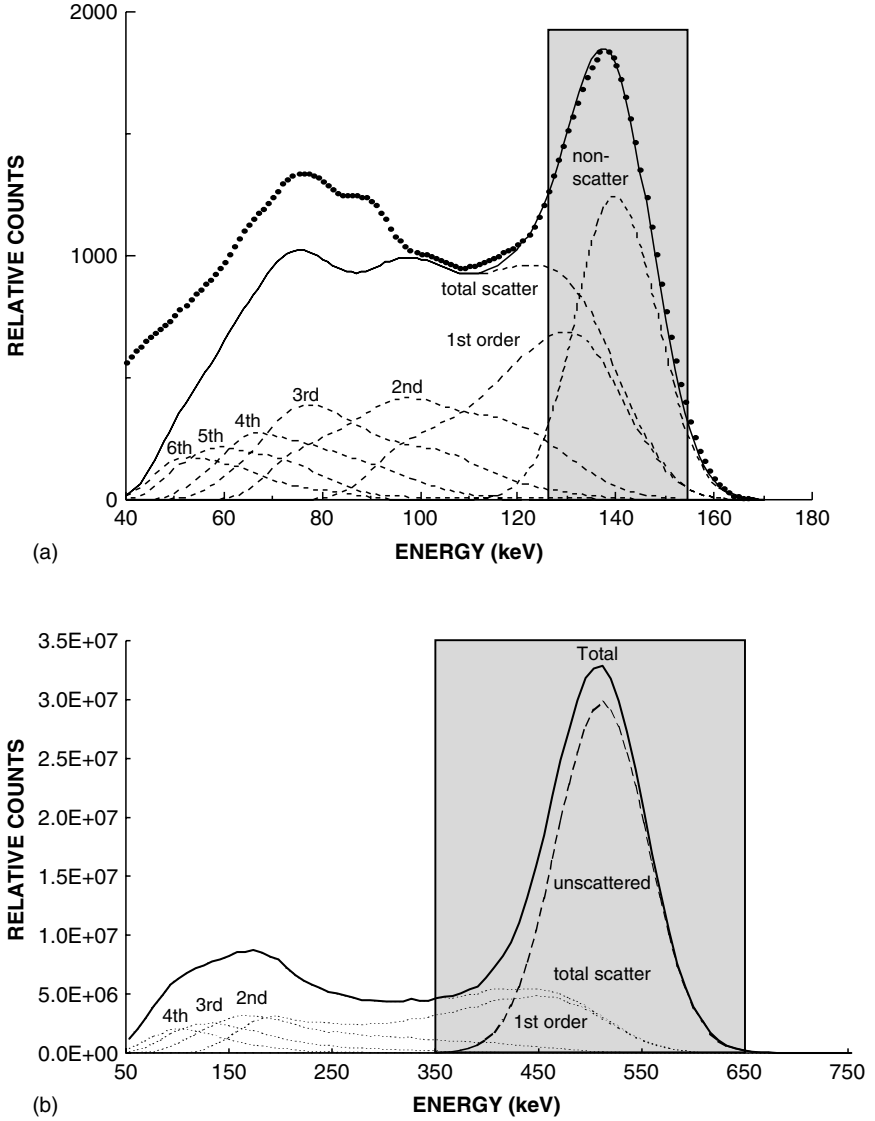


FIGURE 3. A. An energy spectra for a gamma emitting ^{99m}Tc line source on the axis of a water-filled cylinder simulated using the Monte Carlo method. The spectra due to primary and scattered photons (solid line) is separated into different contributions (total scattering or different orders of photon scattering). The distributions of the various orders of scattered and unscattered photons are shown by broken lines. The experimentally measured spectrum is also shown (dots). **B.** Illustration of the energy distribution due to unscattered and scattered photons resulting from the simulation of a 20 cm diameter cylinder filled with a uniform positron-emitting ^{18}F source separated into different contributions (total scattering or different orders of photon scattering). Typical energy acquisition windows for both cases are also shown. (Adapted from refs.[124,125] and reprinted with permission from ref.[126]).

of a water-filled cylindrical phantom and a positron-emitting ^{18}F uniform cylindrical source. The scattered events in the energy pulse-height distribution have been separated according to the order of scattering. It is clear from viewing Figure 3 that events from some scattered photons will not be rejected by the usual [126-154 keV] and [350-650 keV] energy discrimination windows used in SPECT and PET, respectively, due to the limited energy resolution of the detectors. Scattered photons which fall within the photopeak window consist mainly of photons which have only scattered once (1^{st} order). The lower level threshold can thus be easily changed and its effect on the scatter component studied in an effective way.[113,114]

Modelling the scatter response function and development and assessment of scatter correction methods in both SPECT and PET using Monte Carlo simulations is described extensively in chapter 7 of this volume and will not be repeated here. Monte Carlo simulation of the spatial characteristics of scatter originating from activity outside the field-of-view or scatter from matter outside the FOV in 3D PET proved that the spatial distribution of multiple scatter is quite different from the single scatter component which might preclude the rescaling of the latter to take into account the effect of the former for scatter correction purposes.[17] In addition, the spatial distributions of scatter from external matter and external activity were found to be different.[115]

Several studies addressed the issue of the impact of scatter in radionuclide transmission[33,116] and x-ray CT scanning[117,118] as well as the influence of down-scatter from emission (e.g ^{99m}Tc to transmission (e.g. ^{153}Gd) data on the accuracy of the derived attenuation map in emission tomography.[119–121] The comparative assessment of different attenuation correction strategies in lung [122] and brain[123] SPECT were also conducted using Monte Carlo simulations of a digital thorax and brain phantom, respectively.

Monte Carlo simulations are also powerful tools to investigate the limits of algorithms developed for correction of partial volume effect by allowing to replicate realistic conditions in emission tomography for a wide range of practical situations and to study the effect of object shape and size on recovery performance.[127] In addition, the assessment of the impact of inherent assumptions such as accurate characterisation of system response function, perfect registration between MRI/CT and PET/SPECT and anatomical image segmentation as well as other hypotheses regarding tracer distribution on quantification bias is more straightforward compared to experimental approaches.[128,129]

3.4 Dosimetry and Treatment Planning

3.4.1 Calculation of Absorbed Fractions

There is broad consensus in accepting that the earliest Monte Carlo calculations in medical physics were made in the area of nuclear medicine, where the technique was used for internal dosimetry modelling and computations.

Formalism and data based on Monte Carlo calculations, developed by the Medical Internal Radiation Dose (MIRD) Committee of the Society of Nuclear Medicine, have been published as pamphlets in a series of supplements to the Journal of Nuclear Medicine, the first one being published in 1968.[130] Some of these pamphlets made extensive use of Monte Carlo calculations to derive specific absorbed fractions (AFs) for photon sources uniformly distributed in organs of mathematical phantoms. This was extended later to electrons,[131] beta particles and positron emitters.[132]

Monte Carlo calculations for photons were performed using a computer code called ALGAM, which created photons at random positions within any source region (organ or tissue), gave these photons a random orientation in 4π space and then followed them through various Compton and photoelectric interactions (coherent scattering was neglected because of its low contribution to the total cross-section and pair productions events were quite rare, as starting energies did not exceed 4 MeV) until the photon reached a certain critical low cut-off energy and was assumed to be locally absorbed, or until it escaped the surface of the body (at which point the probability of scatter from an air molecule and redirection towards the body was assumed to be negligibly low). With repeated sampling of the source, which at this time generally involved only tens of thousands of trials (histories), a statistical average behaviour of particles originating in this source could be obtained for other regions of the body of interest to radiation dose assessment (target regions). This behaviour was reported as the fraction of energy emitted in the source that was absorbed in a target (AF), with an associated uncertainty (reported as the coefficient of variation). These AFs were thus a considerable improvement over the values given in ICRP Publication 2, as the organ geometries were more realistic, and more importantly, the organs could irradiate each other, whereas in the ICRP 2 model an organ could irradiate only itself. These AFs were used later by the ICRP in updated assessments for workers. Of more interest for this volume is that they found a more immediate application in dose assessments for nuclear medicine patients, owing to the monumental efforts of the newly formed MIRD Committee. In a flurry of publications in its early years, this committee published decay data, methods for kinetic analyses, the AFs from the ALGAM code calculations, dose conversion factors for over 100 nuclides of interest to nuclear medicine, dose calculations for various radiopharmaceuticals, methods for small scale dose calculations with electrons, and other interesting practical scientific documents. AFs for these phantoms were developed using the ALGAMP code (the P signifying a parameterized version of the code, allowing the substitution of parameters giving the radii and positions of the various organs at different ages). These values were published in an ORNL document,[133] but never officially adopted in the MIRD or other peer reviewed literature. Nonetheless, they were widely accepted and used for dose calculations for individuals of different ages.

The application of the Monte Carlo method to internal radiation dosimetry is further emphasized in two recent MIRD pamphlets. In MIRD Pamphlet No. 15,[134] the EGS4 Monte Carlo radiation transport code was used to revise substantially the dosimetric model of the adult head and brain originally published in MIRD Pamphlet No. 5.[135] Pamphlet No. 17[136] demonstrates the utility of the MIRD formalism for the calculation of the non-uniform distribution of radiation absorbed dose in different organs through the use of radionuclide specific S-values defined at the voxel level. Previously calculated AFs for unit density spheres in an infinite unit density medium for photon and electron emitters have been recently re-evaluated using both the EGS4 and MCNP-4B Monte Carlo codes. Moreover, Stabin and Yoriyaz[137] used the MCNP-4B code to calculate AFs for photons in the VoxelMan phantom of Zubal *et al.*,[138] and the results were compared with reference values from traditional MIRD and ORNL phantoms, whereas Chao and Xu used the EGS4 code to estimate specific AFs from internal electron emitters for the VIP-Man model with energies from 100 keV to 4 MeV.[139] More recently, they used the same code and a voxel-based head/brain model developed from the Visible Human images to calculate S-values for typical nuclear brain imaging procedures[140] and compared the results with those calculated from the revised MIRD stylized head/brain model.[134]

3.4.2 Derivation of Dose Point Kernels

In most cases Monte Carlo calculations are used to simulate the random distribution of sources or targets, whereas the actual dosimetric calculation is performed using so called dose point kernels. Such kernels, usually spherical and calculated for monoenergetic sources, describe the pattern of energy deposited at various radial distances from photon and electron or beta point sources. Dose point kernels can be calculated using analytical or Monte Carlo methods. Hybrid approaches (analytical calculations using Monte Carlo data) have also been considered to decrease the computation time.[141] Three Monte Carlo codes have mainly been used for this purpose, namely ETRAN, the ACCEPT code of the ITS system and EGS4. Limitations and constraints of some of these codes have been reported in the literature, the impact of which on the calculated kernels is difficult to evaluate. ETRAN, for instance, had an incorrect sampling of the electron energy loss straggling, which has been corrected for in the ITS3 system (based on ETRAN). EGS4 did not include the accurate electron physics and transport algorithms, which have been incorporated in the recent EGSnrc system. Furhang[142] generated photon point dose kernels and absorbed fractions in water for the full photon emission spectrum of the radionuclides of interest in nuclear medicine by simulating the transport of particles using Monte Carlo techniques. The kernels were then fitted to a mathematical expression.

A unified approach for photon and beta particle dosimetry has been proposed by Leichner[143] by fitting Berger's tables for photons and electrons

to generate an empirical function that is valid for both photons and beta particles. Both point kernel and Monte Carlo techniques can therefore be effectively employed to calculate absorbed dose to tissue from radionuclides that emit photons or electrons. The latter are computationally much more intensive; however, point kernel methods are restricted to homogeneous tissue regions that can be mathematically described by analytical geometries, whereas Monte Carlo methods have the advantage of being able to accommodate heterogeneous tissue regions with complex geometric shapes.

3.4.3 Patient-specific Dosimetry and Treatment Planning

To perform real patient-specific dose calculations, a patient-specific physical model to be used with patient-specific biokinetic data is required. Individual patients not only have significantly different uptake and retention half-lives of activity of the radioactive agent, but also have significantly different physical characteristics and radiosensitivities. If our goal is to optimize patient therapies, their individual parameters should be accounted for as much as possible during treatment planning. Currently, the preferred strategy with radiolabelled antibodies is to use personalized patient dosimetry, and this approach may become routinely employed clinically. The dose distribution pattern is often calculated by generalizing a point source dose distribution,[144] but direct calculation by Monte Carlo techniques is also frequently reported, since it allows media of inhomogeneous density to be considered.[145,146]

The development of a nuclear imaging-based 3-D treatment planner is an area of considerable research interest and several dose calculation algorithms have been developed.[4] The essential steps required in developing a 3-D treatment planning program for radioimmunotherapy are discussed in chapter 17 of this volume. Basically, projection data acquired from an emission tomographic imaging system are processed to reconstruct transverse section images, which yield a count density map of source regions in the body. This count density is converted to an activity map using the sensitivity derived from a calibration phantom. In the final step this activity distribution is converted to a dose rate or dose map, either by convolving the activity distribution with dose point kernels or by direct Monte Carlo calculations. To elaborate a treatment plan for an individual patient, prospective dose estimates can be made by using a tracer activity of radiolabelled antibody to obtain biodistribution information prior to the administration of a larger therapeutic activity. The clinical implementability of treatment planning algorithms will depend to a significant extent on the time required to generate absorbed dose estimates for a particular patient.

Many specially designed software packages have been developed for patient-specific dosimetry and treatment planning.[4] The MABDOSE and DOSE3D computer codes adapt the standard geometrical phantoms, allowing the placement of a single or multiple tumours in various locations to

estimate dose contributions from these tumours to normal organs, but do not at present use patient images. These codes work with stylized representations of average individuals, and give the average dose to whole organs. The RTDS code employs either the standard MIRDOSE phantom set or its own transport algorithms in a limited body space, based on voxel source kernels to produce average organ doses or dose distributions within specified organs or tissues of the body. More sophisticated approaches combine anatomical (CT or MRI) and functional radionuclide (single photon emission computed tomography (SPECT) and positron emission tomography (PET)) images to compute patient-specific absorbed dose distributions and dose–volume histograms similar to the treatment planning programs used in external beam radiotherapy. Several software packages have been devised and validated by different research groups, including the 3D-ID code, SIMDOS and its more recent version based on the EGS4 system, the RMDP package and the SCMS code. A detailed description of some of these tools is provided in chapter 17 and ref.[4] It is worth emphasizing that, with some exceptions; very few have been used in clinical environments.

4. Development of Anthropomorphic Phantoms

Conceptually, the purpose of a physical or computerized phantom is to represent an organ or body region of interest, to allow modelling the biodistribution of a particular radiotracer and the chemical composition of the scattering medium which absorbs and scatters the emitted radiation in a manner similar to biological tissues. In other terms, a phantom is a mathematical model designed to represent an organ or tissue of the body, an organ system, or the whole-body.[147]

Computerized anthropomorphic phantoms can either be defined by mathematical (analytical) functions or digital (voxel-based) volume arrays.[2] Analytic phantoms consist of regularly shaped continuous objects defined by combinations of simple mathematical geometries (e.g. right circular cylinders, spheres, or disks), whereas voxel-based phantoms are mainly derived from segmented tomographic images of the human anatomy obtained by either x-ray computed tomography (CT) or magnetic resonance imaging (MRI). Any complex activity and corresponding attenuation distributions can therefore be modelled without being confined to geometrical shapes. Analytical phantoms, however, have the advantage of being able to model anatomical variability and dynamic organs easily. In addition, the disadvantage of the voxelized approach is that inherent errors are introduced due to the phantom voxelization. The discretization errors inherent in the voxelized representation may be reduced by finer sampling of the discretized phantoms. More refined analysis of advantages and drawbacks of analytical versus voxelized phantom representation for simulation of imaging systems is described elsewhere.[148,149]

4.1 Stylized Mathematical Models

The mathematical specifications for phantoms that are available assume a specific age, height and weight. People, however, exhibit a variety of shapes and sizes. The first breakthrough in the use of Monte Carlo techniques was the development of the Fisher–Snyder heterogeneous, hermaphrodite, anthropomorphic model of the human body in the 1970s.[135] This phantom consisted of spheres, ellipsoids, cones, tori and subsections of such objects, combined to approximate the geometry of the body and its internal structures. The representation of internal organs with this mathematical phantom is very crude, since the simple equations can only capture the most general description of an organ's position and geometry. The original phantom developed was intended mainly to represent a healthy average adult male, which well characterized the working population of its time. The phantom did have both male and female organs, but most structures represented the organs of "Reference Man", as defined by the International Commission on Radiological Protection (ICRP) from an extensive review of medical and other literature, restricted primarily to European and North American populations. Both due to the makeup of the nuclear medicine population and the diversifying worker population, the need for other phantoms arose. In 1987, Cristy and Eckerman[133] of Oak Ridge National Laboratory (ORNL) developed a series of phantoms representing children of different ages, one of which (the 15 year old) also served as a model for the adult female. Bouchet and Bolch[150] developed a series of five dosimetric head and brain models to allow more precise dosimetry in paediatric neuroimaging procedures. More recently, a new rectal model and a dynamic urinary bladder model have also been proposed. To develop more patient-specific dosimetry, new mathematical models for adults of different height have been developed using anthropometric data.[151,152] Mathematical anthropomorphic phantoms are continuously being improved. Recent three- and four-dimensional computer phantoms seek a compromise between ease of use, flexibility and the accurate modelling of populations of patient anatomies, and attenuation and scatter properties as well as biodistributions of radiopharmaceuticals in patients. Current developments are aimed at computer phantoms that are flexible while providing the accurate modelling of patient populations. The use of dynamic anthropomorphic phantoms in Monte Carlo simulations is becoming possible, owing to the increasing availability of computing power. This includes the development of appropriate primitives that allow the accurate modelling of anatomical variations and patient motion, such as superquadrics[153] and non-uniform rational B-spline surfaces (NURBS).[154]

While much effort has been devoted towards the creation of human models, few research studies have been initiated to address the need for models supporting small animal imaging and dosimetry research. The recent interest in small animal imaging instrumentation for molecular imaging research has spurred the development of realistic computerized phantoms

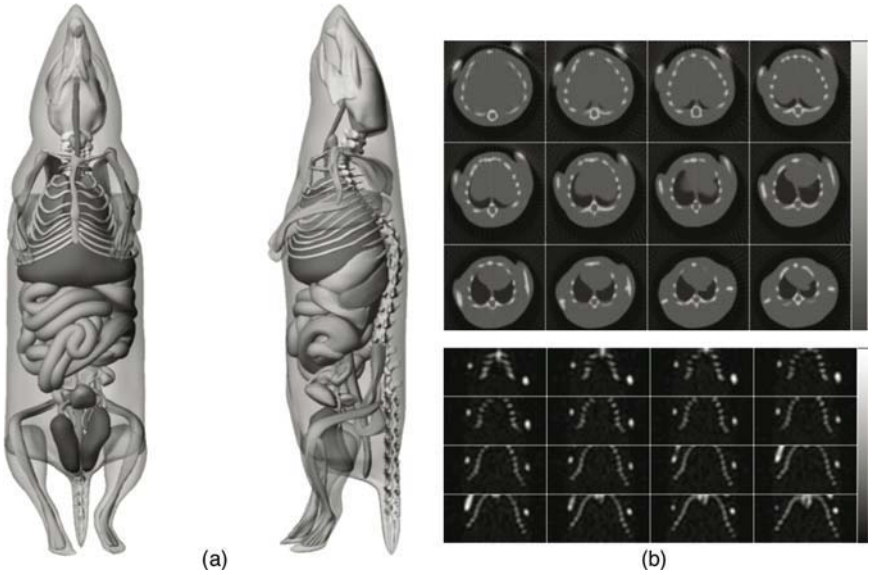


FIGURE 4. **A.** 3-D surface renderings showing anterior (left) and lateral (right) views of the digital mouse phantom. **B.** Representative coronal slices of reconstructed SPECT images simulating the uptake of Tc-99m MDP (top) and cone-beam X-ray CT images generated from the mouse phantom. Reprinted with permission from ref.[155]

modelling their anatomy and physiological functions. One such example is the 4-D (space and time) digital mouse phantom based on high-resolution 3-D magnetic resonance microscopy acquired data and NURBS formalism mentioned above for modelling the organ shapes.[155] 3-D surface renderings illustrating anterior and lateral views of the digital mouse phantom are shown in Figure 4. The NURBS primitives can elegantly model the complex organ shapes and structures, providing the foundation for a realistic model of the 3-D mouse anatomy.

4.2 Tomographic Voxel-Based Phantoms

Modelling for imaging and dosimetry applications is best done with phantom models that match the gross parameters of an individual patient. Anthropomorphic phantoms with internally segmented structures make clinically realistic Monte Carlo simulations possible. Zubal[138] developed a typical anthropomorphic voxel-based adult phantom by the manual segmentation of CT transverse slices of a living human male performed by medical experts. A computerized 3-D volume array modelling all the major internal structures of the body was then created. Each voxel of the volume contains an index number designating it as belonging to a given organ or internal structure. These indexes can then be used to assign a value, corresponding to, for

example, density or activity. The phantom data are available as a $128 \times 128 \times 246$ matrix with a cubic resolution of 4 mm. The same group has also developed a high resolution brain phantom based on an MRI scan of a human volunteer, which can be used for detailed investigations in the head. The torso phantom was further improved by copying the arms and legs from the visible human (VH) and attaching them to the original torso phantom.[156] However, the arms of the VH cadaver were positioned over the abdominal part, which limited the usefulness of the phantom for simulations of whole body scanning. This problem was tackled by mathematically straightening the arms out along the phantom's side.[157] More recently, a new voxel-based whole body model, called VIP-Man,[158] and a head/brain model[140] have been developed using high resolution transversal colour photographic images obtained from the National Library of Medicine's VH Project.[159] A group at the National Research Center for Environment and Health (GSF) in Germany has also been developing some voxel-based phantoms. The GSF voxel phantoms family tends to cover persons of individual anatomy and includes at the moment two paediatric and five adult phantoms of both sexes, different ages and stature, and several others are under construction.[160]

Despite the difficulties and ethical issues, the efforts devoted towards the development of voxel-based paediatric phantoms and pregnant female are further emphasized in the construction of the 2 UF phantoms (female newborn and 2-month male)[161] and the 30-week pregnant female tomographic model[162] opening the door to prospective imaging and dosimetry studies that would otherwise be difficult to perform. Several national and international initiatives were also undertaken to develop Korean- and Chinese-type[163] equivalents to the above referenced VH project. Table 3 summarizes state-of-the-art voxel-based computational models available today for radiological research.[147,164]

5. Concluding Remarks and Future Directions

The use of the Monte Carlo method to simulate radiation transport has become the most accurate means of simulating nuclear medical imaging systems with the aim of optimizing instrumentation design or improving quantitation and predicting absorbed dose distributions and other quantities of interest in radiation treatments of cancer patients using either external or radionuclide radiotherapy. This trend has continued for the estimation of the absorbed dose in diagnostic procedures using radionuclides as well as the assessment of image quality and the quantitative accuracy of radionuclide imaging. As a consequence of this generalized use, many questions are being raised, primarily about the need and potential of Monte Carlo techniques, but also about how accurate they really are, what would it take to apply them clinically and make them available widely to the nuclear medicine community at large. Many of these questions will be answered when

TABLE 3. Summary of available voxel-based tomographic phantoms with their main characteristics.

Model	Age/ Gender	Race	Height (cm)	Weight (Kg)	Body region	Modality	Voxel (mm^3)	Reference
Baby	8-w/f	Caucasian	57	4.2	WB[a]	CT	2.9	[165]
Child	7-y/m	Caucasian	115	21.7	WB	CT	19	[165]
VoxelMan[b]	35-y/m	Caucasian	175	70	head to thigh	CT	56.3	[138,166]
MANTISSUE 3-6/VOXTISS8	modified oxelMan[c]	Caucasian	175.3	70	WB	CT	46.7	[156,157]
NORMAN	adult/m	Caucasian	176	73	WB	MRI	8	[167-169]
MAX	modified VoxelMan[d]	Caucasian	175	75	WB	CT	46.7	[170]
FAX	37-y/f	Caucasian	165	63.4	WB	CT	2.7	[171]
Golem	38-y/m	Caucasian	176	69	WB	CT	34.6	[172,173]
ADELAIDE	14-y/f	Caucasian	157	48	torso	CT		[174]
VIP-man	38-y/m	Caucasian	186	103	WB	colour photos	0.1	[139,158,159]
Otoko	adult/m	Japanese	170	65	WB	CT	9.6	[175]
Onago	adult/f	Japanese	162	57	WB	CT	9.6	[176,177]
UF newborn	6-d/f	Caucasian		3.8	WB	CT	0.35	[161]
UF2-month	~2-m/m	Caucasian		5.4	WB	CT	0.3	[161]
Visible-human	38-y/m	Caucasian	186	103	head to knee	CT	4.3	[178]
Frank	48-y/m	Caucasian	174	95	head to trunk	CT	2.8	[160,178]
Donna	40-y/f	Caucasian	170	79	WB	CT	35.2	[160,179]
Helga	26-y/f	Caucasian	170	81	head to thigh	CT	9.6	[178,179]
Irene	32-y/f	Caucasian	163	51	WB	CT	17.6	[178,179]
Nagaoka-man	22-y/m	Japanese	173	65	WB	MRI	8	[180]
Nagaoka-woman	22-y/m	Japanese	160	53	WB	MRI	8	[180]
KR-man	28-y/m	Korean				MRI		[181]
KORMAN	30-y/m	Korean	170	68	WB	MRI	40	[182]
Un-named	9-m/m	Caucasian			head to thigh	CT		[183]
Pregnant Woman	31-y/f (30-w pregnant)	Caucasian		91		CT	6.2	[162,184]
Visible Korean Human	31-y/m	Korean	164	55	WB	Colour photos	0.01	[163]

[a]Whole-body

[b]Voxelman represents the combination of the segmented head and body CT images without legs and arms

[c]MANTISSUE3-6 is Voxelman with arms and legs added (crossed in front of body) from segmented Visible Man's colour cross-sections whereas VOXTISS8 is MANTISSUE3-6 with arms mathematically straightened out along the phantom's side

Monte Carlo techniques are implemented and used for more routine calculations and for in-depth investigations.

The widespread availability of high performance computing and popularity of simulations stimulated the development of dedicated simulation tools for other imaging modalities including magnetic resonance imaging[185-187] and x-ray CT.[117,118,188,189] For example, an x-ray CT Monte Carlo simulator based on the MCNP4C radiation transport computer code was developed for simulation of fan- and cone-beam CT scanners.[118] A user friendly interface running under Matlab 6.5.1 creates the scanner geometry at different views as MCNP4C's input file. The full simulation of x-ray tube, phantom and detector with single-slice, multi-slice and flat panel detector configurations was considered. The simulator was validated using experimental measurements from different phantoms containing steel, teflon, water, polyethylene and air cylinders with different sizes on the GE HiSpeed X/iF (General Electric Healthcare Technologies, Waukesha, WI) fan-beam scanner and the SkyScan 1076 (SkyScan, Aartselaar, Belgium) small animal cone- beam CT scanner. Figure 5 illustrates typical reconstructed images of

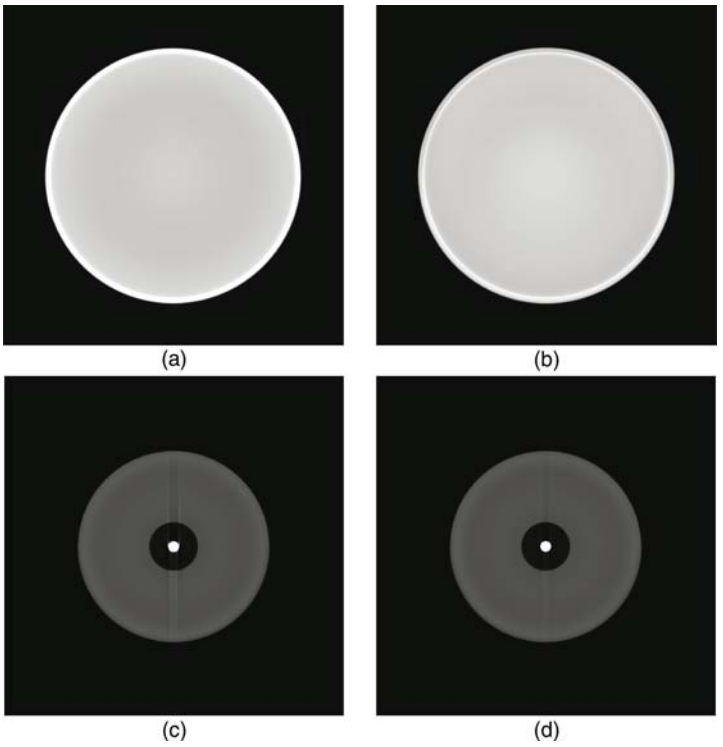


FIGURE 5. Filtered backprojection reconstructions of (**A**) measured (**B**) Monte Carlo simulated x-ray projection data corresponding to a homogeneous cylindrical water phantom. (**C-D**) Same as above for a 20 cm diameter polyethylene phantom containing two centred air and steel cylinders with diameters 5 cm and 1 cm, respectively.

measured (left column) and Monte Carlo simulated (right column) data sets corresponding to a homogeneous cylindrical water phantom and a 20 cm diameter polyethylene phantom containing two centred air and steel cylinders with diameters 5 cm and 1 cm, respectively. The projection data were reconstructed without correction for scatter and beam hardening effects, generally incorporated in commercial software. Note the good agreement between the measured and simulated images.

It is clear that a new generation of dose modelling tools needs to be developed for use with internal emitter therapy in nuclear medicine. It is unfortunate to use in the 21st century standardized, geometrical phantoms to perform dose calculations for individual patients if the physician is asking for meaningful information to be used in planning patient therapy. The evolution of the methodology that was followed for external beam radiotherapy treatment planning must be followed for internal emitter therapy. The technology now exists to develop patient-specific 3-D dose distributions, based on fusion of anatomical (CT or MRI) and functional (SPECT or PET) data, with individualized Monte Carlo calculations done in a reasonable amount of time using high powered computing workstations or distributed computing networks (see chapter 17). The combination of realistic computer phantoms and accurate models of the imaging process allows the simulation of nuclear medicine data that are ever closer to actual patient data. Simulation techniques will find an increasingly important role in the future of nuclear medicine in light of the further development of realistic computer phantoms, the accurate modelling of projection data and computer hardware. However, caution must be taken to avoid errors in the simulation process, and verification via comparison with experimental and patient data is essential.

References

1. Andreo A., Monte Carlo techniques in medical radiation physics. *Phys Med Biol* **36:** 861-920 (1991).
2. Zaidi H., Relevance of accurate Monte Carlo modeling in nuclear medical imaging. *Med Phys* **26:** 574-608 (1999).
3. Ljungberg M., Strand, S.-E. and King, M. A., *Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging,* IOP Publishing, Bristol, (1998).
4. Zaidi H. and Sgouros, G., *Therapeutic applications of Monte Carlo calculations in nuclear medicine,* Institute of Physics Publishing, Bristol, (2002).
5. Zaidi H., ''Monte Carlo techniques in diagnostic and therapeutic nuclear medicine'' *Proc*. of the International Symposium on Standards and Codes of Practice in Medical Radiation Dosimetry. ISBN 92-0-111403-6, Vienna (Austria), 25-28 November 2002, Vol. 868; *pp* 29-44 (2002).
6. Schulz A. G., Knowles, L. G., Kohlenstein, L. C. *et al*., Quantitative assessment of scanning-system parameters. *J Nucl Med* **11:** 61-68 (1970).
7. Dye R. E. A., Simulation of clinical scintigrams for nuclear medicine imaging devices. *Phys Med Biol* **33:** 1329-1334 (1988).

8. Gantet P., Esquerre, J. P., Danet, B. *et al.*, A simulation method for studying scintillation camera collimators. *Phys Med Biol* **35:** 659-669 (1990).
9. Beekman F. J. and Viergever, M. A., Fast SPECT simulation including object shape dependent scatter. *IEEE Trans Med Imaging* **14:** 271-282 (1995).
10. Fleming J. S., Evaluation of a technique for simulation of gamma camera images. *Phys Med Biol* **41:** 1855-1861 (1996).
11. Nelson W. R., Hirayama, H. and Rogers, D. W. O., The EGS4 code system. Stanford Linear Accelerator Center; Report SLAC-256, 1985.
12. Kawrakow I., Mainegra-Hing, E. and Rogers, D. W. O., The EGSnrc Code System: Monte Carlo Simulation of Electron and Photon Transport Ionizing Radiation Standards, National Research Council of Canada; 2003.
13. Del Guerra A. and Nelson, W. R., "Positron emission tomography applications of EGS" in: *Monte Carlo transport of electrons and photons*, edited by Nelson W. R. and Rindi A. Jenkins T M Plenum Publishing Corporation, New York, (1988), *pp* 469-484.
14. Bollini D., Del Guerra, A., Di Domenico, G. *et al.*, Sub-millimeter planar imaging with positron emitters: EGS4 code simulation and experimental results. *IEEE Trans Nucl Sci* **44:** 1499-1502 (1997).
15. Castiglioni I., Cremonesi, O., Gilardi, M. C. *et al.*, Scatter correction techniques in 3D PET: a Monte Carlo evaluation. *IEEE Trans Nucl Sci* **46:** 2053-2058 (1999).
16. Castiglioni I., Cremonesi, O., Gilardi, M.-C. *et al.*, A Monte Carlo model of noise components in 3D PET. *IEEE Trans Nucl Sci* **49:** 2297-2303 (2002).
17. Adam L. E., Karp, J. S. and Brix, G., Investigation of scattered radiation in 3D whole-body positron emission tomography using Monte Carlo simulations. *Phys Med Biol* **44:** 2879-2895 (1999).
18. Adam L. E., Karp, J. S. and Freifelder, R., Energy-based scatter correction for 3-D PET scanners using NaI(T1) detectors. *IEEE Trans Med Imaging* **19:** 513-521 (2000).
19. Narita Y., Eberl, S., Iida, H. *et al.*, Monte Carlo and experimental evaluation of accuracy and noise properties of two scatter correction methods for SPECT. *Phys Med Biol* **41:** 2481-2496 (1996).
20. Narita Y., Iida, H., Eberl, S. *et al.*, Monte Carlo evaluation of accuracy and noise properties of two scatter correction methods for 201Tl cardiac SPECT. *IEEE Trans Nucl Sci* **44:** 2465-2472 (1997).
21. Motta A., Righi, S., Guerra, A. D. *et al.*, A full Monte Carlo simulation of the YAP-PEM prototype for breast tumor detection. *Nucl Instr Meth A* **527:** 201-205 (2004).
22. Conti M., Casey, M. E., Eriksson, L. *et al.*, Benchmarking a Monte Carlo simulation code on a prototype LSO scanner. *IEEE Trans Nucl Sci* **49:** 644-648 (2002).
23. Halbleib J. A., Kensek, R. P., Valdez, G. D. *et al.*, ITS: The Integrated TIGER Series of electron/photon transport codes - version 3.0. *IEEE Trans Nucl Sci* **39:** 1025-1030 (1992).
24. Briesmeister J. F., MCNP - A general Monte Carlo N-particle transport code. version 4C. Los Alamos National Laboratory, NM; Report LA-13709-M, 2000.
25. Hendricks J. S., McKinney, G. W., Waters, L. S. *et al.*, MCNPX, VERSION 2.5.e. Los Alamos National Laboratory, NM; Report LA-UR-04-0569, 2004.

26. Yanch J. C., Dobrzeniecki, A. B., Ramanathan, C. *et al*., Physically realistic Monte Carlo simulation of source, collimator and tomographic data acquisition for emission computed tomography. *Phys Med Biol* **37:** 853-870 (1992).
27. Yanch J. C. and Dobrzeniecki, A. B., Monte Carlo simulation in SPECT: Complete 3-D modeling of source, collimator and tomographic data acquisition. *IEEE Trans Nucl Sci* **40:** 198-203 (1993).
28. Du Y., Frey, E. C., Wang, W. T. *et al*., Combination of MCNP and SimSET for Monte Carlo simulation of SPECT with medium- and high-energy photons. *IEEE Trans Nucl Sci* **49:** 668-674 (2002).
29. Song T. Y., Choi, Y., Chung, Y. H. *et al*., Optimization of pinhole collimator for small animal SPECT using Monte Carlo simulation. *IEEE Trans Nucl Sci* **50:** 327-332 (2003).
30. Brun R., Bruyant, F., Maire, M. *et al*., GEANT detector description and simulation tool. CERN; Report DD/EE/84-1, 1994.
31. Agostinelli S., Allison, J., Amako, K. *et al*., GEANT4- a simulation toolkit. *Nucl Instrum Methods A* **506:** 250-303 (2003).
32. Michel C., Bol, A., Spinks, T. *et al*., Assessment of response function in two PET scanners with and without interplane septa. *IEEE Trans Med Imaging* **10:** 240-248 (1991).
33. Wegmann K., Adam, L.-E., Livieratos, L. *et al*., Investigation of the scatter contribution in single photon transmission measurements by means of Monte Carlo simulations. *IEEE Trans Nucl Sci* **46:** 1184-1190 (1999).
34. Jan S., Santin, G., Strul, D. *et al*., GATE: a simulation toolkit for PET and SPECT. *Phys Med Biol* **49:** 4543-4561 (2004).
35. Sempau J., Acosta, E., Baro, J. *et al*., An algorithm for Monte Carlo simulation of the coupled electron-photon transport. *Nucl Instr Meth B* **132:** 377-390 (1997).
36. Salvat F., Fernandez-Varea, J., Costa, E. *et al*., *PENELOPE - A Code System for Monte Carlo Simulation of Electron and Photon Transport*, Workshop Proceedings, Issy-les-Moulineaux, France, 5-7 November 2001.
37. Cot A., Sempau, J., Pareto, D. *et al*., Evaluation of the geometric, scatter, and septal penetration components in fan-beam collimators using Monte Carlo simulation. *IEEE Trans Nucl Sci* **49:** 12-16 (2002).
38. Cot A., Sempau, J., Pareto, D. *et al*., Study of the point spread function (PSF) for ^{123}I SPECT imaging using Monte Carlo simulation. *Phys Med Biol* **49:** 3125-3136 (2004).
39. Sanchez-Crespo A., Andreo, P. and Larsson, S. A., Positron flight in human tissues and its influence on PET image spatial resolution. *Eur J Nucl Med Mol Imaging* **31:** 44-51 (2004).
40. Fasso A., Ferrari, A. and Sala, P. R., "Electron-photon transport in FLUKA: status" in: *Proceedings of the Monte Carlo 2000 Conference*, edited by A Kling, Barao, F, Nakagawa, M, Tavora, L, Vaz, P Springer-Verlag, Berlin, (2001), *pp* 159-164.
41. Parodi K. and Enghardt, W., Potential application of PET in quality assurance of proton therapy. *Phys Med Biol* **45:** N151-56 (2000).
42. Tanner R. J., Chartier, J.-L., Siebert, B. R. L. *et al*., Intercomparison on the usage of computational codes in radiation dosimetry. *Radiat Prot Dosimetry* **110:** 769-780 (2004).

43. Zaidi H., Comparative evaluation of photon cross section libraries for materials of interest in PET Monte Carlo simulations. *IEEE Trans Nucl Sci* **47:** 2722-2735 (2000).
44. Dresser M. M. and Knoll, G. F., Results of scattering in radioisotope imaging. *IEEE Trans Nucl Sci* **20:** 266-272 (1973).
45. Beck J. W., Jaszczak, R. J., Coleman, R. E. *et al*., Analyzing SPECT including scatter and attenuation using sophisticated Monte Carlo modeling methods. *IEEE Trans Nucl Sci* **29:** (1982).
46. Floyd C. E., Jaszczak, R. J., Greer, K. L. *et al*., Inverse Monte Carlo as a unified reconstruction algorithm for ECT. *J Nucl Med* **27:** 1577-1585 (1986).
47. Keller N. A. and Lupton, J. R., PET detector ring aperture function calculations using Monte Carlo techniques. *IEEE Trans Nucl Sci* **30:** 676-680 (1983).
48. Lupton L. R. and Keller, N. A., Performance study of single-slice positron emission tomography scanners by Monte Carlo techniques. *IEEE Trans Med Imaging* **2:** 154-168 (1983).
49. Zaidi H., Scheurer, A. H. and Morel, C., An object-oriented Monte Carlo simulator for 3D cylindrical positron tomographs. *Comput Methods Programs Biomed* **58:** 133-145 (1999).
50. Zaidi H., Labbé, C. and Morel, C., Implementation of an environment for Monte Carlo simulation of fully 3D positron tomography on a high-performance parallel platform. *Parallel Comput*. **24:** 1523-1536 (1998).
51. Harrison R. L., Vannoy, S. D., Haynor, D. R. *et al*., "Preliminary experience with the photon history generator module for a public-domain simulation system for emission tomography" 1999 Records of IEEE Nuclear Science Symposium and Medical Imaging Conference, San Francisco, *pp* 1154-1158 (1993).
52. Lewellen T., Harrison, R. L. and Vannoy, S., "The SIMSET program" in: *Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging*, edited by M Ljungberg, S-E Strand, and M A King Institute of Physics Publishing, Bristol, (1998), *pp* 77-92.
53. Ljungberg M. and Strand, S.-E., A Monte Carlo program for the simulation of scintillation camera characteristics. *Comput Methods Programs Biomed* **29:** 257-272 (1989).
54. Ljungberg M., "The SIMIND Monte Carlo program" in: *Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging*, edited by M Ljungberg, S-E Strand, and M A King Institute of Physics Publishing, Bristol, (1998), *pp* 145-163.
55. De Vries D. J., Moore, S. C., Zimmerman, R. E. *et al*., Development and validation of a Monte Carlo simulation of photon transport in an Anger camera. *EEE Trans Med Imaging* **9:** 430-438 (1990).
56. de Vries D. and Moore, S., "Monte Carlo simulation of photon transport in gamma camera collimators" in: *Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging*, edited by M Ljungberg, S-E Strand, and M A King Institute of Physics Publishing, Bristol, (1998), *pp* 125-144.
57. Smith M. F., Floyd, C. E. and Jaszczak, R. J., A vectorized Monte Carlo code for modeling photon transport in SPECT. *Med Phys* **20:** 1121-1127 (1993).
58. Smith M. F., "Vectorized Monte Carlo code for modelling photon transport in nuclear medicine" in: *Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging*, edited by M Ljungberg, S-E Strand, and M A King Institute of Physics Publishing, Bristol, (1998), *pp* 93-109.

59. Thompson C. J., Cantu, J.-M. and Picard, Y., PETSIM: Monte Carlo program simulation of all sensitivity and resolution parameters of cylindrical positron imaging systems. *Phys Med Biol* **37:** 731-749 (1992).
60. Thompson C. and Picard, Y., "PETSIM: Monte Carlo simulation of positron imaging systems" in: *Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging*, edited by M Ljungberg, S-E Strand, and M A King Institute of Physics Publishing, Bristol, (1998), *pp* 233-248.
61. Reilhac A., Lartizien, C., Costes, N. *et al.*, PET-SORTEO: a Monte Carlo-based simulator with high count rate capabilities. *IEEE Trans Nucl Sci* **51:** 46-52 (2004).
62. Buvat I. and Castiglioni, I., Monte Carlo simulations in SPET and PET. *Q J Nucl Med* **46:** 48-61 (2002).
63. Ponisch F., Parodi, K., Hasch, B. G. *et al.*, The modelling of positron emitter production and PET imaging during carbon ion therapy. *Phys Med Biol* **49:** 5217-5232 (2004).
64. Zerby C. D., "A Monte Carlo calculation of the response of gamma-ray scintillation counters" in: *Methods in Computational Physics*, edited by Fermbach S Alder B, Rotenberg M New York Academic, (1963), *pp* 89-134.
65. Derenzo S. E., Monte Carlo calculations of the detection efficiency of NaI(Tl), BGO, CsF, Ge and plastic detectors for 511 keV photons. *IEEE Trans Nucl Sci* **28:** 131-136 (1981).
66. Derenzo S. E. and Riles, J. K., Monte Carlo calculations of the optical coupling between bismuth germanate crystals and photomultiplier tubes. *IEEE Trans Nucl Sci* **29:** 191-195 (1982).
67. Bottigli U., Guzzardi , R., Mey, M. *et al.*, Monte Carlo simulation and experimental tests on BGO, CsF and NaI(Tl) crystals for positron emission tomography. *J Nucl Med Allied Sci* **29:** 221-227 (1985).
68. Bice A. N., Lewellen, T. K., Miyaoka, R. S. *et al.*, Monte Carlo simulation of BaF2 detectors used in time-of-flight positron emission tomography. *IEEE Trans Nucl Sci* **37:** 696-701 (1990).
69. Lopes M. I., Chepel , V., Carvalho, J. C. *et al.*, Performance analysis based on a Monte Carlo simulation of a liquid Xenon PET detector. *IEEE Trans Nucl Sci* **42:** 2298-2302 (1995).
70. Binkley P. F., Optimization of scintillation detector timing systems using Monte Carlo analysis. *IEEE Trans Nucl Sci* **41:** 386-393 (1994).
71. DeVol T. A., Moses, W. W. and Derenzo, S. E., Monte Carlo optimization of depth-of-interaction resolution in PET crystals. *IEEE Trans Nucl Sci* **40:** 170-174 (1993).
72. Comanor K. A., Virador, P. R. G. and Moses, W. W., Algorithms to identify detector Compton scatter in PET modules. *IEEE Trans Nucl Sci* **43:** 2213-2218 (1996).
73. Tsang G., Moisan, C. and Rogers, J. G., A simulation to model position encoding multicrystal PET detectors. *IEEE Trans Nucl Sci* **42:** 2236-2243 (1995).
74. Moisan C., Rogers, J. G., Buckley, K. R. *et al.*, Design studies of a depth-encoding large aperture PET camera. *IEEE Trans Nucl Sci* **42:** 1041-1050 (1995).
75. Cayouette F., Zhang, N. and Thompson, C. J., Monte Carlo simulation using DETECT2000 of a multilayered scintillation block and fit to experimental data. *IEEE Trans Nucl Sci* **50:** 339-343 (2003).

76. Eriksson L., Townsend, D., Eriksson, M. *et al*., Experience with scintillators for PET: towards the fifth generation of PET scanners. *Nucl Instr Meth A* **525:** 242-248 (2004).
77. Webb S., Binnie, D. M., Flower, M. A. *et al*., Monte Carlo modelling of the performance of a rotating slit-collimator for improved planar gamma-camera imaging. *Phys Med Biol* **37:** 1095-1108 (1992).
78. Moore S., de Vries, D., Penney, B. *et al*., ''Design of a collimator for imaging In-111'' in: *Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging*, edited by M Ljungberg, S-E Strand, and M A King Institute of Physics Publishing, Bristol, (1998), *pp* 183-193.
79. Kimiaei S., Ljungberg, M. and Larsson, S. A., Evaluation of optimally designed planar-concave collimators in single-photon emission tomography. *Eur J Nucl Med* **24:** 1398-1404 (1997).
80. Thompson C. J., The effect of collimation on single rates in multi-slice PET. *IEEE Trans Nucl Sci* **36:** 1072-1077 (1989).
81. Digby W. M., Dahlbom, M. and Hoffman, E. J., Detector, shielding and geometric design factors for a high-resolution PET system. *IEEE Trans Nucl Sci* **37:** 664-670 (1990).
82. Groiselle C. J., D'Asseler, Y., Kolthammer, J. A. *et al*., A Monte Carlo simulation study to evaluate septal spacing using triple-head hybrid PET imaging. *IEEE Trans Nucl Sci* **50:** 1339-1346 (2003).
83. Moore S. C., deVries, D. J., Nandram, B. *et al*., Collimator optimization for lesion detection incorporating prior information about lesion size. *Med Phys* **22:** 703-713 (1995).
84. Pollard K. R., Bice, A. N., Eary, J. F. *et al*., A method for imaging therapeutic doses of iodine-131 with a clinical gamma camera. *J Nucl Med* **33:** 771-776 (1992).
85. Smith M. F. and Jaszczak, R. J., The effect of gamma ray penetration on angle-dependent sensitivity for pinhole collimation in nuclear medicine. *Med Phys* **24:** 1701-1709 (1997).
86. van der Have F. and Beekman, F. J., Photon penetration and scatter in micro-pinhole imaging: a Monte Carlo investigation. *Phys Med Biol* **49:** 1369-1386 (2004).
87. Wang H., Jaszczak, R. J. and Coleman, R. E., Monte Carlo modeling of penetration effect for iodine-131 pinhole imaging. *IEEE Trans Nucl Sci* **43:** 3272-3277 (1996).
88. Gieles M., de Jong, H. W. and Beekman, F. J., Monte Carlo simulations of pinhole imaging accelerated by kernel-based forced detection. *Phys Med Biol* **47:** 1853-1867 (2002).
89. Thompson C. J., Roney, J. M., Lecomte, R. *et al*., Dependence of the coincidence aperture function of narrow BGO crystals on crystal shape and light encoding schemes. *Phys Med Biol* **31:** 491-506 (1986).
90. Thompson C. J., The effect of collimation on scatter fraction in multi-slice PET *IEEE Trans Nucl Sci* **35:** 598-602 (1988).
91. Bradshaw J., Burnham, C. and Correia, J., Application of Monte Carlo methods to the design of SPECT detector systems. *IEEE Trans Nucl Sci* **32:** 753-757 (1985).
92. Heanue J. A., Brown, J. K., Tang, H. R. *et al*., A bound on the energy resolution required for quantitative SPECT. *Med Phys* **23:** 169-173 (1996).

93. Thompson C. J., The effects of detector material and structure on PET spatial resolution and efficiency. *IEEE Trans Nucl Sci* **37:** 718-724 (1990).
94. Dahlbom M., Rosenquist, G., Eriksson, L. *et al*., A study of the possibility of using multi-slice PET systems for 3D imaging. *IEEE Trans Nucl Sci* **36:** 1066-1071 (1989).
95. Moses W. W., Virador, P. R. G., Derenzo, S. E. *et al*., Design of a high-resolution, high-sensitivity PET camera for human brains and small animals. *IEEE Trans Nucl Sci* **41:** 1487-1491 (1997).
96. Braem A., Chamizo Llatas, M., Chesi, E. *et al*., Novel design of a parallax free Compton enhanced PET scanner. *Nucl Instr Meth A* **525:** 268-274 (2004).
97. Schramm N. U., Ebel, G., Engeland, U. *et al*., High-resolution SPECT using multipinhole collimation. *IEEE Trans Nucl Sci* **50:** 315-320 (2003).
98. Lazaro D., Buvat, I., Loudos, G. *et al*., Validation of the GATE Monte Carlo simulation platform for modelling a CsI(Tl) scintillation camera dedicated to small-animal imaging. *Phys Med Biol* **49:** 271-285 (2004).
99. Miyaoka R. S., Dynamic high resolution positron emission imaging of rats. *Biomed Sci Instrum* **27:** 35-42 (1991).
100. Pavlopoulos S. and Tzanakos, G., Design and performance evaluation of a high-resolution small animal positron tomograph. *IEEE Trans Nucl Sci* **43:** 3249-3255 (1996).
101. Bevilacqua A., Bollini, D., Del Guerra, A. *et al*., A 3-D Monte Carlo simulation of a small animal positron emission tomograph with millimeter spatial resolution. *IEEE Trans Nucl Sci* **46:** 697-701 (1999).
102. Lartizien C., Reilhac, A., Costes, N. *et al*., Monte Carlo simulation-based design study of a LSO-LuAP small animal PET system. *IEEE Trans Nucl Sci* **50:** 1433-1438 (2003).
103. Seidel J., Vaquero, J. J. and Green, M. V., Resolution uniformity and sensitivity of the NIH ATLAS small animal PET scanner: Comparison to simulated LSO scanners without depth-of-interaction capability. *IEEE Trans Nucl Sci* **50:** 1347-1350 (2003).
104. Rannou F. R., Kohli, V., Prout, D. L. *et al*., Investigation of OPET performance using GATE, a Geant4-based simulation software. *IEEE Trans Nucl Sci* **51:** 2713-2717 (2004).
105. Vaska P., Woody, C. L., Schlyer, D. J. *et al*., RatCAP: miniaturized head-mounted PET for conscious rodent brain imaging. *IEEE Trans Nucl Sci* **51:** 2718-2722 (2004).
106. Floyd C. E., Jaszczak, R. J. and Coleman, R. E., Inverse Monte Carlo: a unified reconstruction algorithm for SPECT. *IEEE Trans Nucl Sci* **32:** 779-785 (1985).
107. Boning G., Pichler, B. J., Rafecas, M. *et al*., Implementation of Monte Carlo coincident aperture functions in image generation of a high-resolution animal positron tomograph. *IEEE Trans Nucl Sci* **48:** 805-810 (2001).
108. Rafecas M., Mosler, B., Dietz, M. *et al*., Use of a Monte Carlo-based probability matrix for 3-D iterative reconstruction of MADPET-II data. *IEEE Trans Nucl Sci* **51:** 2597-2605 (2004).
109. Beekman F. J., de Jong, H. W. A. M. and van Geloven, S., Efficient fully 3-D iterative SPECT reconstruction with monte carlo-based scatter compensation. *IEEE Trans Med Imaging* **21:** 867-877 (2002).
110. Wilson D. W., Tsui, B. M. W. and Barrett, H., Noise properties of the EM algorithm: II. Monte Carlo simulations. *Phys Med Biol* **39:** 847-871 (1994).

111. Wang W. and Gindi, G., Noise analysis of MAP-EM algorithms for emission tomography. *Phys Med Biol* **42:** 2215-2232 (1997).
112. Soares E. J., Glick, S. J. and Hoppin, J. W., Noise characterization of block-iterative reconstruction algorithms: II. Monte Carlo simulations. *IEEE Trans Med Imaging* **24:** 112-121 (2005).
113. Lewellen T. K., Harrison, R. L. and Kohlmyer, S. G., Effect of lower energy threshold on single and multiple scatter distributions in positron volume imaging. *IEEE Trans Nucl Sci* **46:** 1129-1135 (1999).
114. Zaidi H., Reconstruction-based estimation of the scatter component in positron emission tomography *Ann Nucl Med Sci* **14:** 161-171 (2001).
115. Laymon C. M., Harrison, R. L., Kohlmyer, S. G. *et al*., Characterization of single and multiple scatter from matter and activity distributions outside the FOV in 3-D PET. *IEEE Trans Nucl Sci* **51:** 10-15 (2004).
116. Ogawa K., Kawamura, Y., Kubo, A. *et al*., Estimation of scattered photons in gamma ray transmission CT using Monte Carlo simulations. *IEEE Trans Nucl Sci* **44:** 1225-1230 (1997).
117. Colijn A. P., Zbijewski, W., Sasov, A. *et al*., Experimental validation of a rapid Monte Carlo based micro-CT simulator. *Phys Med Biol* **49:** 4321-4333 (2004).
118. Ay M. and Zaidi, H., Development and validation of MCNP4C-based Monte Carlo simulator for fan- and cone-beam x-ray CT. (2005) *submitted*
119. Ljungberg M., Strand, S.-E., Rajeevan, N. *et al*., Monte Carlo simulation of transmission studies using a planar source with a parallel collimator and a line source with a fan-beam collimator. *IEEE Trans Nucl Sci* **41:** 1577-1584 (1994).
120. de Jong H. W., Wang, W. T., Frey, E. C. *et al*., Efficient simulation of SPECT down-scatter including photon interactions with crystal and lead. *Med Phys* **29:** 550-560 (2002).
121. Bokulic T., Vastenhouw, B., De Jong, H. *et al*., Monte Carlo-based down-scatter correction of SPECT attenuation maps. *Eur J Nucl Med Mol Imaging* **31:** 1173-1181 (2004).
122. Gustafsson A., Bake, B., Jacobsson, L. *et al*., Evaluation of attenuation corrections using Monte Carlo simulated lung SPECT. *Phys Med Biol* **43:** 2325-2336 (1998).
123. Arlig A., Gustafsson, A., Jacobsson, L. *et al*., Attenuation correction in quantitative SPECT of cerebral blood flow: a Monte Carlo study. *Phys Med Biol* **45:** 3847-3859 (2000).
124. Kojima A., Matsumoto, M., Takahashi, M. *et al*., Effect of energy resolution on scatter fraction in scintigraphic imaging: Monte Carlo study. *Medical Physics* **20:** 1107-13 (1993).
125. Meikle S. R. and Badawi, R. D., "Quantitative Techniques in Positron Emission Tomography" in: *Positron Emission Tomography: Basic Science and Clinical Practice*, edited by P E Valk, D L Bailey, D W Townsend *et al*. Springer, London, (2003), *pp* 115-146.
126. Zaidi H. and Koral, K. F., Scatter modelling and compensation in emission tomography. *Eur J Nucl Med Mol Imaging* **31:** 761-782 (2004).
127. Dewaraja Y. K., Ljungberg, M. and Koral, K. F., Monte Carlo evaluation of object shape effects in iodine-131 SPET tumor activity quantification. *Eur J Nucl Med* **28:** 900-906 (2001).

128. Frouin V., Comtat, C., Reilhac, A. *et al*., Correction of partial volume effect for PET striatal imaging: fast implementation and study of robustness. *J Nucl Med* **43:** 1715-1726 (2002).
129. Soret M., Koulibaly, P. M., Darcourt, J. *et al*., Quantitative accuracy of dopaminergic neurotransmission imaging with ^{123}I SPECT. *J Nucl Med* **44:** 1184-1193 (2003).
130. Loevinger R. and Berman, M., MIRD Pamphlet No. 1: A schema for absorbed-dose calculation for biologically distributed radionuclides. *J Nucl Med* **9:** 7-14 (1968).
131. Stabin M. G. and Konijnenberg, M. W., Re-evaluation of absorbed fractions for photons and electrons in spheres of various sizes. *J Nucl Med* **41:** 149-160 (2000).
132. Bice A. N., Links, J. M., Wong, D. F. *et al*., Absorbed fractions for dose calculations of neuroreceptor PET studies. *Eur J Nucl Med* **11:** 127-131 (1985).
133. Cristy M. and Eckerman, K. F., Specific absorbed fractions of energy at various ages from internal photon sources. I Methods, II one year old, III five year old, IV ten year old, V fifteen year old male and adult female, VI new-born and VII adult male. Oak Ridge National Laboratory; Report ORNL/TM 8381/V1–V7, 1987.
134. Bouchet L. G., Bolch, W. E., Weber, D. A. *et al*., MIRD Pamphlet No. 15: Radionuclide S values in a revised dosimetric model of the adult head and brain. Medical Internal Radiation Dose. *J Nucl Med* **40:** 62S-101S (1999).
135. Snyder W., Ford, M. R. and Warner, G., *Estimates of specific absorbed fractions for photon sources uniformly distributed in various organs of a heterogeneous phantom. Pamphlet No. 5, revised*., Society of Nuclear Medicine, New York, (1978).
136. Bolch W. E., Bouchet, L. G., Robertson, J. S. *et al*., MIRD pamphlet No. 17: the dosimetry of nonuniform activity distributions–radionuclide S values at the voxel level. Medical Internal Radiation Dose Committee. *J Nucl Med* **40:** 11S-36S (1999).
137. Stabin M. G. and Yoriyaz, H., Photon specific absorbed fractions calculated in the trunk of an adult male voxel-based phantom. *Health Phys* **82:** 21-44 (2002).
138. Zubal I. G., Harrell , C. R., Smith, E. O. *et al*., Computerized 3-dimensional segmented human anatomy. *Med Phys* **21:** 299-302 (1994).
139. Chao T. C. and Xu, X. G., Specific absorbed fractions from the image-based VIP-Man body model and EGS4-VLSI Monte Carlo code: internal electron emitters. *Phys Med Biol* **46:** 901-927 (2001).
140. Chao T.-c. and Xu, X. G., S-values calculated from a tomographic head/brain model for brain imaging. *Phys Med Biol* **49:** 4971-4984 (2004).
141. Bardies M. and Myers, M. J., Computational methods in radionuclide dosimetry. *Phys Med Biol* **41:** 1941-55 (1996).
142. Furhang E. E., Sgouros, G. and Chui, C. S., Radionuclide photon dose kernels for internal emitter dosimetry. *Med Phys* **23:** 759-764 (1996).
143. Leichner P. K., A unified approach to photon and beta particle dosimetry. *J Nucl Med* **35:** 1721-1729 (1994).
144. Sgouros G., Chiu, S., Pentlow, K. S. *et al*., Three-dimensional dosimetry for radioimmunotherapy treatment planning. *J Nucl Med* **34:** 1595-1601 (1993).
145. Yoriyaz H., Stabin, M. G. and dos Santos, A., Monte Carlo MCNP-4B-based absorbed dose distribution estimates for patient-specific dosimetry. *J Nucl Med* **42:** 662-669 (2001).

146. Ljungberg M., Sjogreen, K., Liu, X. *et al*., A 3-dimensional absorbed dose calculation method based on quantitative SPECT for radionuclide therapy: evaluation for (131)I using monte carlo simulation. *J Nucl Med* **43:** 1101-1109 (2002).
147. Poston J. W., Bolch, W. and Bouchet, L., "Mathematical models of the human anatomy" in: *Therapeutic applications of Monte Carlo calculations in nuclear medicine*, edited by H Zaidi and G Sgouros Institute of Physics Publishing, London, (2002), *pp* 108-132.
148. Peter J., Tornai, M. P. and Jaszczek, R. J., Analytical versus voxelized phantom representation for Monte Carlo simulation in radiological imaging. *IEEE Trans Med Imaging* **19:** 556-564 (2000).
149. Goertzen A. L., Beekman, F. J. and Cherry, S. R., Effect of phantom voxelization in CT simulations. *Med Phys* **29:** 492-498 (2002).
150. Bouchet L. G. and Bolch, W. E., Five pediatric head and brain mathematical models for use in internal dosimetry. *J Nucl Med* **40:** 1327-1336 (1999).
151. Clairand I., Bouchet, L. G., Ricard, M. *et al*., Improvement of internal dose calculations using mathematical models of different adult heights. *Phys Med Biol* **45:** 2771-2785 (2000).
152. Huh C. and Bolch, W. E., A review of US anthropometric reference data (1971-2000) with comparisons to both stylized and tomographic anatomic models. *Phys Med Biol* **48:** 3411-3429 (2003).
153. Peter J., Gilland, D. R., Jaszczak, R. J. *et al*., Four-dimensional superquadric-based cardiac phantom for Monte Carlo simulation of radiological imaging systems. *IEEE Trans Nucl Sci* **46:** 2211-2217 (1999).
154. Segars W. P., Lalush, D. S. and Tsui, B. M. W., Modeling respiratory mechanics in the MCAT and spline-based MCAT phantoms. *IEEE Trans Nucl Sci* **48:** 89-97 (2001).
155. Segars W. P., Tsui, B. M., Frey, E. C. *et al*., Development of a 4-D digital mouse phantom for molecular imaging research. *Mol Imaging Biol* **6:** 149-159 (2004).
156. Dawson T. W., Caputa, K. and Stuchly, M. A., A comparison of 60 Hz uniform magnetic and electric induction in the human body. *Phys Med Biol* **42:** 2319-2329 (1997).
157. Sjogreen K., Ljungberg, M., Wingardh, K. *et al*., Registration of emission and transmission whole-body scintillation-camera images. *J Nucl Med* **42:** 1563-1570 (2001).
158. Xu X. G., Chao, T. C. and Bozkurt, A., VIP-Man: an image-based whole-body adult male model constructed from color photographs of the Visible Human Project for multi-particle Monte Carlo calculations. *Health Phys* **78:** 476-486 (2000).
159. Spitzer V. M. and Whitlock, D. G., The Visible Human Dataset: the anatomical platform for human simulation. *Anat Rec* **253:** 49-57 (1998).
160. Petoussi-Henss N., Zanki, M., Fill, U. *et al*., The GSF family of voxel phantoms. *Phys Med Biol* **47:** 89-106 (2002).
161. Nipper J. C., Williams, J. L. and Bolch, W. E., Creation of two tomographic voxel models of paediatric patients in the first year of life. *Phys Med Biol* **47:** 3143-3164 (2002).
162. Shi C., "The development and application of a tomographic model from CT images for internal dose calculation to pregnant woman," Ph.D dissertation, Rensselaer Polytechnic Institute, 2004.

163. Park J. S., Chung, M. S., Hwang, S. B. *et al.*, Visible Korean Human. Improved serially sectioned images of the entire body. *IEEE Trans Med Imaging* **24:** 352-360 (2005).
164. Caon M., Voxel-based computational models of real human anatomy: a review. *Radiat Environ Biophys* **42:** 229-235 (2004).
165. Zankl M., Veit, R., Williams, G. *et al.*, The construction of computer tomographic phantoms and their application in radiology and radiation protection. *Radiation and Environmental Biophysics* **27:** 153-64 (1988).
166. Zubal I. G. and Harrell, C. R., Voxel-based Monte Carlo calculations of nuclear medicine images and applied variance reduction techniques. *Image Vision Computing* **10:** 342-348 (1992).
167. Jones D. G., A realistic anthropomorphic phantom for calculating organ doses arising from external photon irradiation. *Radiat Prot Dosimetry* **72:** 21-29 (1997).
168. Jones D. G., A realistic anthropomorphic phantom for calculating specific absorbed fractions of energy deposited from internal gamma emitters. *Radiat Prot Dosimetry* **79:** 411-414 (1998).
169. Dimbylow P. J., FDTD calculations of the whole-body averaged SAR in an anatomically realistic voxel model of the human body from 1 MHz to 1 GHz. *Phys Med Biol* **42:** 479-490 (1997).
170. Kramer R., Vieira, J. W., Khoury, H. J. *et al.*, All about MAX: a male adult voxel phantom for Monte Carlo calculations in radiation protection dosimetry. *Phys Med Biol* **48:** 1239-1262 (2003).
171. Kramer R., Khoury, H. J., Vieira, J. W. *et al.*, All about FAX: a Female Adult voXel phantom for Monte Carlo calculation in radiation protection dosimetry. *Phys Med Biol* **23:** (2004).
172. Petoussi-Henss N. and Zankl, M., Voxel anthropomorphic models as a tool for internal dosimetry. *Radiat Prot Dosimetry* **79:** 415-418 (1998).
173. Zankl M. and Wittmann, A., The adult male voxel model "Golem" segmented from whole-body CT patient data. *Radiat Environ Biophys* **40:** 153-162 (2001).
174. Caon M., Bibbo, G. and Pattison, J., An EGS4-ready tomographic computational model of a 14-year-old female torso for calculating organ doses from CT examinations. *Phys Med Biol* **44:** 2213-2225 (1999).
175. Saito K., Wittmann, A., Koga, S. *et al.*, Construction of a computed tomographic phantom for a Japanese male adult and dose calculation system. *Radiat Environ Biophys* **40:** 69-75 (2001).
176. Kinase S., Zankl, M., Kuwabara, J. *et al.*, "Evaluation of specific absorbed factions in voxel phantoms using Monte Carlo simulation" Proceedings of Radiation Risk Assessment in the 21st Century, EPA/JAERI Workshop, Las Vegas, November 5-7 (2001), *pp* 118-127 (2001).
177. Sato K., Noguchi, H., Saito, K. *et al.*, "Development of CT voxel phantoms for Japanese" Proceedings of Radiation Risk Assessment in the 21st Century, EPA/JAERI Workshop, Las Vegas, November 5-7 (2001), *pp* 102-110 (2001).
178. Zankl M., Fill, U., Petoussi-Henss, N. *et al.*, Organ dose conversion coefficients for external photon irradiation of male and female voxel models. *Phys Med Biol* **47:** 2367-2385 (2002).
179. Fill U. A., Zankl, M., Petoussi-Henss, N. *et al.*, Adult female voxel models of different stature and photon conversion coefficients for radiation protection. *Health Phys* **86:** 253-272 (2004).

180. Nagaoka T., Watanabe, S., Sakurai, K. *et al.*, Development of realistic high-resolution whole-body voxel models of Japanese adult males and females of average height and weight, and application of models to radio-frequency electromagnetic-field dosimetry. *Phys Med Biol* **49:** 1-15 (2004).
181. Lee C. and Lee, J., The Korean reference adult male voxel model ''KRman'' segmented from whole-body MR data and dose conversion coefficients. [abstract] *Health Phys* **84 [Suppl]:** S163 (2003).
182. Lee C. and Lee, J., Korean adult male voxel model KORMAN segmented from magnetic resonance images. *Med Phys* **31:** 1017-1022 (2004).
183. Lee C. and Bolch, W., Construction of a tomographic computational model of a 9-mo-old and its Monte Carlo calculation time comparison between the MCNP4C and MCNPX codes. [abstract] *Health Phys* **84 [Suppl]:** S259 (2003).
184. Shi C. and Xu, X. G., Development of a 30-week-pregnant female tomographic model from computed tomography (CT) images for Monte Carlo organ dose calculations. *Med Phys* **31:** 2491-2497 (2004).
185. Olsson M. B., Wirestam, R. and Persson, B. R., A computer simulation program for MR imaging: application to RF and static magnetic field imperfections. *Magn Reson Med* **34:** 612-617 (1995).
186. Kwan R. K., Evans, A. C. and Pike, G. B., MRI simulation-based evaluation of image-processing and classification methods. *IEEE Trans Med Imaging* **18:** 1085-97 (1999).
187. Hacklander T., Mertens, H. and Cramer, B. M., [Computer simulation of a clinical magnet resonance tomography scanner for training purposes]. *Rofo* **176:** 1151-1156 (2004).
188. Colijn A. P. and Beekman, F. J., Accelerated simulation of cone beam X-ray scatter projections. *IEEE Trans Med Imaging* **23:** 584-590 (2004).
189. Ay M., Shahriari, M., Sarkar, S. *et al.*, Monte Carlo simulation of x-ray spectra in diagnostic radiology and mammography using MCNP4C. *Phys Med Biol* **49:** 4897-4917 (2004).

12
Tracer Kinetic Modeling in Nuclear Medicine: Theory and Applications

M. Bentourkia* and H. Zaidi†

1. Introduction

Over the last two decades, Positron Emission Tomography (PET) has advanced elegantly in the clinical and research arenas and is now providing a unique tool for investigating the biochemistry of living organs non-invasively. The wide range of radiolabelled molecules makes it feasible to explore a variety of biochemical, physiological and pharmacological processes. Since each radiotracer is characterised by its particular kinetic behaviour in the human body, the quantification of this behaviour is a critical component for the improvement of imaging protocols and rapid translation from research and development to clinical routine applications.

Suitable image correction and reconstruction algorithms combined with tracer kinetic modeling techniques are needed to assess parametric or quantitative biological images from the available 4D images (space and time). An appropriate mathematical model[1] is generally used to fit the time-activity curves (TAC) of a region or volume of interest (which can be as small as a voxel), thus allowing the assessment of biological parameters. It is worth emphasizing that even if production of parametric images at the voxel level is possible, the noisy TACs make it difficult to extract accurate parameters without robust analysis strategies.

The different quantitative tracer kinetic modeling strategies are generally divided in two broad classes of method:[2] model-driven methods[3-6] and data-driven methods.[7-10] The difference between the two approaches lies in the fact that no *a priori* decision about the most appropriate model is required in the latter, which in this case is estimated directly from the kinetic data.

For some practical reasons, especially the capability to provide functional and metabolic studies, we focus our effort in this chapter on the PET

*Dr M. Bentourkia, Dept. of Nuclear Medicine and Radiobiology, 3001 Sherbrooke (Qc) J1H 5N4, Canada
†PD Dr H. Zaidi, Division of Nuclear Medicine, Geneva University Hospital, CH-1211 Geneva, Switzerland

modality. However, this investigation is not exclusive to PET; in fact, the same analysis can be conducted by analogy using other imaging modalities.

The main advantages of PET are: 1) its ability to simultaneously measure many slices of the body; 2) to use radiolabelled biological molecules which provoke biochemical changes in tissue; 3) its accuracy in measuring the radioactivity concentration in the body by accurately correcting for physical degrading factors. The main disadvantages of PET can be resumed in three points: 1) because PET operates with isotopes of short half-life, it needs a cyclotron in the vicinity; 2) since any tissue or organ can be imaged with PET (even bones), to be fully exploited, this modality necessitates several radiolabelled tracers, imaging protocols, data analysis software and kinetic models etc..., which mobilize large resources and a big team of specialists in many fields; 3) the measured signal with PET depends on the limited amount of the radioactivity injected in the patient, hence the PET detectors have to be large enough to collect photons and to allow images of good statistics. Consequently, the size of these detectors is the cause of the reduced spatial resolution and the partial volume effect. Because of these three drawbacks, PET has been considered for a long time (more than a decade) as a research tool only. Now that PET has gained in recognition, its widespread over the world is considerable. Meanwhile, the clinicians are still using PET as image generator to visually observe abnormal tissues.

The main purpose of this chapter is to revisit kinetic modeling in PET and to present a framework for kinetic modeling for a wide range of problems in nuclear medicine. Emphasis is given to practical applications in order to understand the potential, and may be to take full advantage, of this imaging modality. The basic concepts are presented in numerous publications and will only be briefly discussed here. Interested readers are encouraged to consult the detailed and comprehensive reviews[6,11] and book chapters[12,13] published on the subject. It is believed that if specific imaging protocols are applied to allow extracting physiological parameters, PET imaging could provide an interesting database in clinical and research settings.

2. Radiolabelled Tracers Used in PET

The radiotracers used in PET imaging are based on biological molecules. The mostly used molecules are labelled with $^{11}C(T_{1/2}=20.4\text{min})$,$^{13}N(T_{1/2}=9.96\text{min})$,$^{15}O(T_{1/2}=2.07\text{min})$ *and* $^{18}F(T_{1/2}=109.7\text{min})$. Since there is no positron emitter of hydrogen, fluorine-18 is used as a hydrogen substitute. Positron emitters of Cu, Zn, K, Br, Rb, I, P, Fe, Ga and others are also used. From these considerations, it appears relatively simple to label a natural or a pharmacological molecule targeting a specific organ or tissue and to proceed with PET imaging. Consequently, nearly any apparent anomaly in the body or any task to be accomplished by an organ can be measured with PET.

Till now PET has been used in several fields in medicine such as neurology, cardiology and oncology. It is used to study metabolism and perfusion,[4,14,15] receptor density,[16,17] and cancer therapy.[18-20] However, PET is better exploited in research where several basic research studies are conducted on high resolution dedicated small animal scanners.[21,22] Now PET imaging has moved elegantly to include gene therapy[23,24] and molecular imaging.[25-27]

3. Tracer Kinetic Modeling: General Considerations

A tracer is developed in order to measure a physiological parameter of interest in a specific tissue. The biochemical process of the tracer in the tissue needs to be known in order to estimate the concentration of the reactants and products over time, and hence to obtain the reaction rates. An appropriate tracer kinetic model is designed to describe the biochemical process steps in the tissue. This model accounts for the initial state of the tracer as delivered by blood to the cells, and subsequently each of the form of the tracer in tissue. In other words, the PET scanner measures the concentrations of the isotope irrespective to which molecule it is bounded. Note that the labelled molecule can be split into two molecules or more, or it can split and its parts can bind to other molecules. Basically, each step of this process (i.e. each different labelled product molecule) can form a compartment. In some instances, when the dissociation of the labelled molecule is fast, these two steps (pre- and post-dissociation) can be assumed a single compartment. Hypothetically, a tracer which can transform in less than three products can be assumed as a potentially good tracer.

In fact there are other crucial characteristics of a good tracer. We cite here the most relevant: 1) the extraction of the tracer by the cells should be rapid; 2) its retention in the cells as initial tracer or as products of the tracer can provide a good signal in the tissue to be imaged; 3) the tissue to be imaged should utilize this labelled molecule independently of the other available molecules; 4) the tracer should not generate metabolites since these can be measured in blood while they don't penetrate the cells.

4. Compartmental Modeling

The biochemical transfer of the labelled molecule from one form to another is usually described by kinetic compartments. Hypothetically each compartment represents the history of the tracer as a physiological or biochemical process. Basically, the tracer is conveyed in blood till the site of interest where the blood supplies nutrients to the cells through capillaries. The next step is the presence of the tracer in the extracellular space, then it crosses the cell membrane to enter the cell where, normally, it's ingested. The

compartments interact by exchanging the tracer itself or its products. There may be also an exchange between a compartment and its surroundings such as an excretion. The whole process can be described by means of a mathematical model. This model should be able to estimate the amount of the radioactivity concentration in each compartment and the rate of exchange between these compartments. In PET imaging, these rates, usually called rate constants, directly provide the physiological parameters characterizing the behaviour of the tracer in the tissue of interest.

The PET scanner measures the whole concentration of the radioactivity in the field of view irrespective of the type of the emitting material: tissue, organs or blood. The reconstructed images must be corrected for the usual effects such as normalisation, randoms, scatter and attenuation, and decay of the radioactivity during the measurement (see other chapters of this book). Since the kinetic model is expressed in terms of time, the measurements should be performed in terms of frames. Depending on the duration of the administration of the tracer, the time framing should be sampled finely to allow assessment of the time course of the radiotracer in tissue. Also, to be able to assess the response of the tissue to a delivered radiotracer, the amount of the tracer in blood should be known as precisely as possible as a function of time. This function is called the input function, the blood curve or the plasma curve.

The kinetic analysis can be processed by either applying the model on a pixel by pixel basis, which produces parametric images, or by grouping pixels, representing homogeneous structures, in regions of interest (RoIs). The time course (or time-activity curve) of a single pixel is usually noisy and for some tracers of short half-life it is impossible to adjust to the model. The RoIs method appears as an appreciable solution since it averages the radioactivity of the pixels included in the RoI, thus allowing to process data with better statistical properties and to reduce computing time since only a few RoIs are analyzed instead of thousands of pixels. Typical RoIs drawn on an ^{18}F-fluorodeoxyglucose (FDG)-PET image of a slice acquired through the mid-brain of a normal human subject are shown in figure 1.

5. Input Function

The tracer plasma time activity curve is necessary for rate constants calculations. The input curve can be derived from blood sampling, reconstructed images or from a population database. Blood samples are withdrawn from a vein of the patient or the animal at a sufficient rate to describe the time course of the bolus. When the input curve is obtained by drawing RoIs on the left ventricle (LV) blood pool PET images, these images should be highly sampled to accurately reproduce in time the tracer delivery to tissue. If the input curve is extracted from a population sample, one or more blood samples are withdrawn from the subject to scale the averaged population input curve.

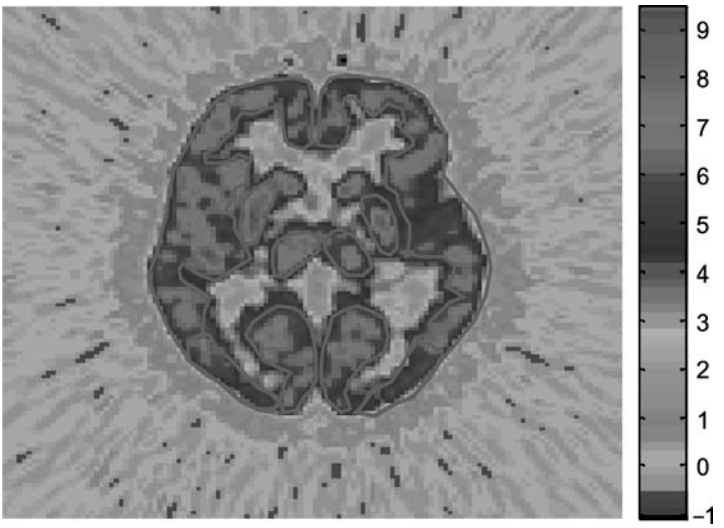


FIGURE 1. Representative RoIs defined for the quantitative evaluation of brain PET data acquired with ^{18}F-FDG in a normal brain. The color bar has relative units due to the image reconstruction filtering.

A rapid bolus injection is preferred to a slow bolus or an infusion because it allows a rapid delivery of the tracer to the cells, hence reducing the gradient between arterial plasma and interstitial fluid.[28] However, the difficulty resides in the rapid sampling during the injection. External blood samples withdrawn from the patient either automatically or manually should not contaminate each other, and should be sampled at 5 sec for the first minute for small animals and for the first 2 minutes for humans and large size animals. The sampling frequency can be progressively reduced for later times; this could vary from 30 sec till 10 min or more for lengthy scans. The high frequency sampling till shortly after the injection not only allows to accurately determine the peak of the input curve but also to generally estimate the rise of the metabolites in the blood. Two minutes after a bolus injection, the concentration of the radioactivity in blood reduces very slowly and can be described with a few blood samples. A last sample at some seconds after the end of the PET scan is mandatory. Note that for small animals such as rats and mice, because of the fragility of the veins it is difficult to install catheters and to withdraw blood manually or automatically. In addition, the blood quantity in these animals is very limited and could not allow enough samples to determine the input curve. In the case of external blood sampling, the samples need to be cross-calibrated with the PET scans. This is achieved by measuring a radioactive solution in both the PET scanner and the well counter where the blood samples are measured.[29] When estimating the input curve from RoIs such as the image of the LV blood pool or a vein or artery by means of any method, care must

be taken to correct for partial volume effect and spillover from neighbouring tissues (see chapter 8). A disadvantage of this method is the lack in getting enough statistics in the images acquired in reduced intervals (nearly 5 sec). The advantage is that no cross calibration is needed since the input curve has the same units as the PET data. For input curves extracted from both external sampling and LV blood pool, it is possible to interpolate the data for finer sampling.[30] The idea of interpolating provides two benefits: 1) because the kinetic model integrates data over the dynamic scan intervals (frames) and the noiseless analytical input function is suitable for this integration, and 2) it allows to obtain a complete input function even with a few samples after the peak sampling.

For some tracers, there are labelled metabolites which are contained in blood. These metabolites do not enter the cells while they are counted in the input curve time course. The contribution of the metabolites could be determined experimentally or it could be incorporated in the kinetic model.

6. Application of Compartmental Modeling

To better understand the concepts of compartmental modeling and to be able to relate the kinetic modeling parameters to physiological processes, the procedures described above are illustrated through real data acquired with PET. The authors preferred to concentrate their efforts in presenting a full and detailed analysis in two distinct applications, to allow the reader taking advantage of this chapter in designing her or his own compartmental models, instead of spreading their efforts in discussing diverse topics related to the subject matter. The two particular cases presented in this chapter are the glucose metabolism in the brain measured with ^{18}F-FDG and myocardial blood flow measured with ^{13}N-Ammonia.

6.1 ^{18}F-*FDG Model*

The ^{18}F-FDG model was derived from autoradiographic ^{14}C-deoxyglucose (^{14}C-DG) to measure glucose utilization in animals.[3] This method is developed below for ^{18}F-FDG. The advantage in using DG (or FDG) over labelled glucose is its potential to be trapped and to accumulate in brain tissue as a primary or secondary phosphorylation products.[4,31-33] DG (or FDG) and glucose in the plasma compete for transport through the blood-brain barrier (BBB) and for the enzyme hexokinase. Glucose is metabolized further while DG-6-PO4 (or FDG-6-PO4) is trapped in brain tissue, and released with a very low clearance.

The ^{18}F-FDG kinetic model depends on several assumptions mostly:[31,32]

1. The tissue compartment should be homogeneous with respect to blood flow, transport rates of glucose and DG (or FDG) between plasma and tissue, and rates of phosphorylation;
2. Glucose metabolism should be in a steady-state. The rates of glucose utilization, plasma glucose concentration and the concentrations of all the substrates and intermediates of the glycolytic pathway should be constant during the measurement.

The general three-compartment model for ^{18}F-FDG is shown in figure 2. In this diagram, C_a, C_f and C_m are compartments for arterial, free and metabolized tracer, respectively. K_1 is the perfusion constant and has units of ml of blood/g of tissue/min or ml of blood/ml of tissue/min (tissue density = 1.04 g/ml). k_2, k_3 and k_4 are rate constants and have units of min^{-1}. The dashed line corresponds to what the PET scanner measures (CPET). In fact it measures the concentration of the radioactivity in the capillaries and the amount of the tracer in the interstitial liquid and in the cells. In order to satisfy the assumptions above, the radiotracer should be delivered to tissue as a bolus. Usually, an intravenous injection is made in less than 20 sec followed with a flush.

The kinetic model for ^{18}F-FDG is described by the differential equations directly based on the compartmental model shown in figure 2:[4]

$$\begin{aligned}\frac{dC_f(t)}{dt} &= K_1 C_a(t) - (k_2 + k_3)C_f(t) + k_4 C_m(t)\\ \frac{dC_m(t)}{dt} &= k_3 C_f(t) - k_4 C_m(t)\end{aligned} \tag{1}$$

Solving Eq. 1 gives:

$$\begin{aligned}C_f(t) &= \frac{K_1}{\alpha_2 - \alpha_1}[(k_4 - \alpha_1)e^{-\alpha_1 t} + (\alpha_2 - k_4)e^{-\alpha_2 t}] \otimes C_a(t)\\ C_m(t) &= \frac{K_1 k_3}{\alpha_2 - \alpha_1}(e^{-\alpha_1 t} - e^{-\alpha_2 t}] \otimes C_a(t)\end{aligned} \tag{2}$$

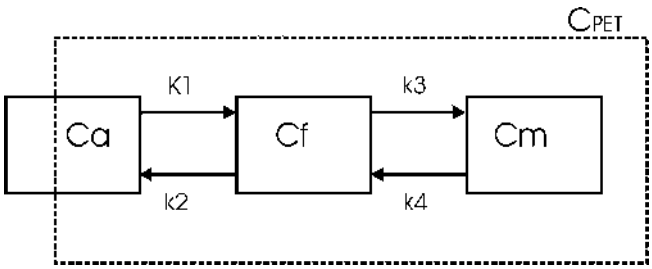


FIGURE 2. Diagram of the three-compartment model for ^{18}F-FDG illustrating arterial (C_a) free (C_f) and metabolized (C_m) tracer, the perfusion constant (K_1) and the rate constants ($k_2 - k_4$). The dashed line corresponds to what the PET scanner measures (C_{PET}).

where the symbol $\otimes$ stands for the operation of convolution and α_1 and α_2 are defined as:

$$\begin{aligned}\alpha_1 &= [k_2 + k_3 + k_4 - \sqrt{(k_2 + k_3 + k_4)^2 - 4k_2k_4}]/2 \\ \alpha_2 &= [k_2 + k_3 + k_4 + \sqrt{(k_2 + k_3 + k_4)^2 - 4k_2k_4}]/2\end{aligned} \tag{3}$$

The sum of the two compartments $C_f(t)$ and $C_m(t)$ plus a fraction (k_5) of the input curve should match the PET data:

$$C_{PET}(t)=\frac{K_1}{\alpha_2-\alpha_1}[(k_3+k_4-\alpha_1)e^{-\alpha_1 t}+(\alpha_2-k_3-k_4)e^{-\alpha_2 t}]\otimes C_a(t)+k_5 C_a(t) \tag{4}$$

Once the kinetic parameters are defined, the ultimate physiological parameter of interest is the metabolic rate for glucose (MRGlc). MRGlc is calculated by means of the rate constants but also as a function of the lumped constant (LC) and the glycemia (gl) as:

$$\begin{aligned}\text{MRGlc}(\mu\text{moles}/100\text{g}/\text{min}) = &\frac{\text{gl(mg of glucose}/100\text{ml of plasma)}}{\text{LC } 0.18(\text{mg}/\mu\text{moles})} \\ &\times \frac{\text{K}_1(\text{ml}/\text{g}/\text{min})\text{k}_3(\text{min}^{-1})}{k_2(\text{min}^{-1}) + k_3(\text{min}^{-1})}\end{aligned} \tag{5}$$

The lumped constant is defined as the ratio of the arterio-venous extraction fraction of ^{18}F-FDG to that of glucose under steady-state conditions and when k_4 is small. LC was calculated by Phelps *et al.*[4] to be 0.42 in human brain, while it was estimated by Reivich *et al.*[31] to be 0.52. Both values are now widely used for brain studies.

The glycemia is defined as the capillary plasma glucose concentration in steady-state. It is obtained by averaging glycemia values measured in two or three blood samples withdrawn through the PET scan. In order to get a stable glucose concentration in blood, the patient is usually asked to fast for at least 4 hours prior to the PET study.

The unit of MRGlc is μmoles of glucose /100g of tissue/min. Some workers prefer to express the tissue in unit of ml. The factor 0.18 in the denominator of Eq. 5 is deduced from the fact that 1 μmole of ^{18}F-FDG = 0.180 mg, and serves to convert MRGlc in μmoles.

6.1.1 Non-linear least squares fitting

In order to calculate MRGlc as expressed in Eq. 5, the rate constants K_1-k_4 must be known. These are estimated from the kinetic model and the PET data. The PET data should be acquired at predefined intervals of time or frames. The more they are finely sampled the more the time course of ^{18}F-FDG in brain can be accurately defined. The drawback of the fine sampling is that the data become noisy, and the late scans of the study would have a reduced radioactivity concentration in the body which makes

them more noisy. Another limitation is the storage (space) and analysis (time) of so many data. Although the half-life of ^{18}F-FDG (109.7 min) allows a study of 2 hours with good statistics in the data, 40 min is generally enough to accurately estimate the rate constants.[34]

The nonlinear least squares (NLLS) fitting is based on algorithms such as the one developed by Levenberg and Marquardt.[35,36] The algorithm is first supplied with the initial values for the rate constants and other parameters as described below (Figure 3). The brain images measured with PET are first used to draw regions of interest for each homogeneous structure. This drawing can be accomplished with the help of magnetic resonance images (MRI) coregistered to the PET images to better delineate the anatomical structures. These RoIs are defined in 4D, i.e. on each image slice and on each frame. The mean values of each RoI are further divided by the respective frame lengths to give values in units of counts/pixel/sec or MBq/pixel/sec. These mean values with the mid-frame times make the TACs. These TACs together with the input curve and the initial parameters are supplied to the algorithm to return the rate constants for each RoI.

The other parameters needed in ^{18}F-FDG brain modeling are the dispersion, the delay and the blood volume (bv). The input curve is measured by sampling which doesn't really reflect the supply of the tracer in blood. The mechanical injection of the tracer is further modified (modelled) by the organism, and this modification can be partly corrected by introducing the parameter for the dispersion. The time delay serves to adjust the time at which the tracer is supplied to the cells in the brain. This time is different from that at which the blood samples are withdrawn from another part of the body (arm of the patient). The delay can be thought as a shift of the input curve with respect to the brain TACs. The blood volume is always counted in any RoI since blood is present in the capillaries everywhere in the brain. The blood volume is accounted for by inserting a fraction of the input curve as shown in Eq. 4 where k_5 reflects the value of the blood volume ($bv = k_5$). In the normal brain, bv is generally between 3% and 8% of the input curve ($k_5 = 0.03 - 0.08$).

The NLLS fitting algorithm should constrain the fitting parameters to non-negative values only. It is also recommended to put upper constraints on these parameters to help the fast convergence of the algorithm. Another recommendation is to not modify the initial parameter values for a type of a study. These initial values are obtained from a database in normal subjects.

Figure 3 displays an example of fitting a TAC extracted from the right temporal region in a brain of a volunteer. The data were acquired with a CTI/Siemens ECAT HR scanner in 3D mode with the following time sequence: 10×60 sec and 5×600 sec. The blood samples were manually withdrawn at a rate of 5 sec during the first 2 min then at increased intervals of time, resulting in 30 samples from 5 sec till 3635 sec. A complete analysis can be found in Bentourkia *et al.*[37] where brain ^{18}F-FDG and ^{15}O-water images were coregistered with MRI images in a comparative study between cerebral blood flow and glucose metabolism as a function of age.

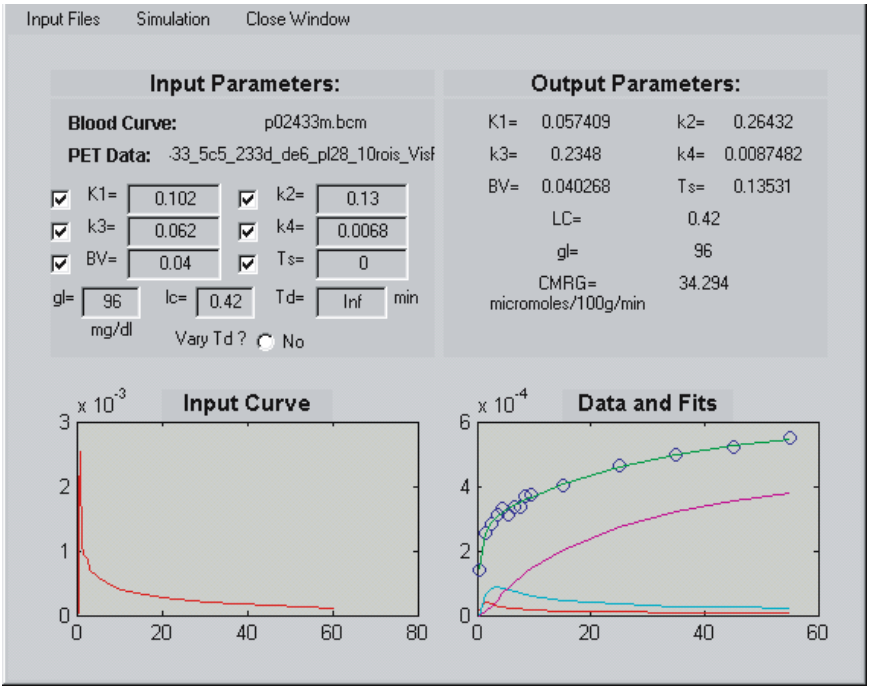


FIGURE 3. Example of PET ^{18}F-FDG analysis of brain data. The interface used at the Metabolic and Functional Imaging Centre, Sherbrooke, Canada, shows the input parameter values which can be interactively modified or selected to be fitted to the data by checking the squares beside the parameters. The values for the lumped constant, the glycemia and the time to be considered for analysis can also be entered. The input curve and a single TAC or a list of TACs can be entered at the submenus under the File menu. The software generates the output parameters in the same window and displays the plots of the input curve and the three compartment functions together with the measured TAC and the fit. These curves are also saved to disk for each TAC. Another file listing all the TACs fitted and their corresponding final parameters with MRGlc and goodness of fit (chi square) is also written to disk.

6.1.2 Graphical method

The graphical method was first developed by Patlak *et al.*[8] and is sometimes referred to as patlak plot. The method was applied selectively to ^{18}F-FDG in comparison to other methods.[38,39] It is based on the assumption that $k_4 = 0$ and that there is a steady-state condition later after the bolus injection where the function of plasma concentration can be considered constant in comparison to the exponential function. Making $k_4 = 0$ in Eq. 3 leads to:

$$\begin{aligned} \alpha_1 &= 0 \\ \alpha_2 &= k_2 + k_3 \end{aligned} \tag{6}$$

Replacing α_1 and α_2 by their values in Eq. 6, Eq. 4 becomes, without considering the last term:

$$C_{PET}(t) = \frac{K_1}{k_2 + k_3}[k_3 + k_2 e^{-(k_2+k_3)t}] \otimes C_a(t) \tag{7}$$

Using the definition of the convolution operation, Eq. 7 becomes:

$$C_{PET}(t) = \frac{K_1}{k_2 + k_3}\left[k_3 \int_0^t C_a(u)du + k_2 e^{-(k_2+k_3)t} \int_0^t C_a(u) e^{(k_2+k_3)u} du\right] \tag{8}$$

Now we consider the steady-state where $C_a(t)$ is a constant beside the exponential function, Eq. 8 becomes:

$$C_{PET}(t) = \frac{K_1}{k_2 + k_3}\left[k_3 \int_0^t C_a(u)du + \frac{k_2}{k_2 + k_3} C_a(t)\right] \tag{9}$$

By dividing both sides of Eq. 9 by $C_a(t)$ we obtain:

$$C_{PET}(t)/C_a(t) = \frac{K_1 k_3}{k_2 + k_3} \int_0^t C_a(u)du/C_a(t) + \frac{K_1 k_2}{(k_2 + k_3)^2} \tag{10}$$

Eq. 10 has the form of $y = ax + b$ which allows to graphically estimate the slope $\frac{K_1 k_3}{k_2+k_3}$ using the measured data $C_{PET}(t)$ and $C_a(t)$. Usually the last points of the function in Eq. 10 form a straight line which is considered to estimate the slope. By including the values of the glycemia and the lumped constant, MRGlc can be estimated (see Eq. 5).

Because of its simple expression and its rapid calculation, the patlak plot allows to derive parametric images, i.e. the MRGlc can be estimated for each pixel in the image. Some other advantages of this method are the no need to include the blood volume and the time delay to calculate MRGlc. The drawback, however, resides in its supposition that there is no conversion of the FDG-6-PO4 back to ^{18}F-FDG ($k_4 = 0$). This is generally widely accepted for ^{18}F-FDG cardiac studies while k_4 is assumed greater than 0 in ^{18}F-FDG brain studies. Another limitation is the slight instability of the method. The value of the slope depends on the number of data points in Eq. 10 defining the straight line. Since the values of these points are defined experimentally, they always have variation, and then an inclusion or exclusion of a single data point generates an error on MRGlc estimates. It is generally recommended to use the last data points of Eq. 10 forming a straight line, because these last data points correspond to the last frames of the PET data where the tissue has the greater uptake of ^{18}F-FDG, and, at the same time, the values of the input curve present lesser variation (the input curve decreases very slowly). If a rapid bolus injection (less than 30 sec duration) is used, the data corresponding to the interval 25-45 min should be

considered in the graphical method. If the data is collected at times approaching 60 min, the graph is no more linear because of the contribution of the k_4 rate constant.[34]

6.1.3 Autoradiographic method

Also called the single scan method, the autoradiographic method was first introduced by Sokoloff *et al.*[3] Eq. 5 can be approximated by:[4,32]

$$\mathrm{MRGlc} = \frac{\mathrm{gl}}{\mathrm{LC}} \frac{\mathrm{K_1 k_3}}{\mathrm{k_2 + k_3}} \approx \frac{gl}{LC} K_1' \frac{\mathrm{k_3'}}{k_2' + k_3'} \left[\frac{C_{PET}(T) - C_f'(T)}{C_m'(T)} \right] \tag{11}$$

Here the primed rate constants, $K_1' - k_3'$, have values obtained from a sample population of control subjects, and the primed compartment functions C_f' and C_m' are obtained with primed rate constant values applied in Eq. 2 and integrated over the time of measurement T. In these equations k_4 could be counted or could be assumed to be zero. In some cases, Eq. 11 can lead to a negative value of MRGlc. To avoid these negative values, Hutchins *et al.*[40] proposed a modification to Eq. 11:

$$\mathrm{MRGlc} = \frac{\mathrm{gl}}{\mathrm{LC}} \frac{\mathrm{K_1 k_3}}{\mathrm{k_2 + k_3}} \approx \frac{gl}{LC} K_1' \frac{\mathrm{k_3'}}{k_2' + k_3'} \left[\frac{C_{PET}(T)}{C_f'(T) + C_m'(T)} \right] \tag{12}$$

Because K_1 is highly dependent on the initial uptake of the radioactivity in tissue, which makes it sensitive to deviations from the population mean, and because C_f' and C_m' are proportional to K_1 as shown in Eq. 2, estimation of MRGlc using Eq. 12 is therefore independent of K_1 and yields to better estimates of MRGlc.

The input function is still needed in the autoradiographic method while a single PET scan of 30 min collected 30 min after the bolus injection is enough to estimate MRGlc. If the K_4 has to be taken into account, the scan should last more than 60 min after the injection.[41]

The autoradiographic method is also simple and fast in computing which allows calculations of parametric images. It is less sensitive to the variations in the measured data as the graphical method since the experimental data are collected over a longer single scan.

6.1.4 Application in a human brain study

In this section, an analysis of a brain study is presented to emphasize on the methods described above. The ^{18}F-FDG PET scans were acquired in a volunteer with an ECAT EXACT HR scanner (CTI PET Systems, TN, USA) in one hour with the following frames: 10×60 sec and 5×600 sec. The data acquisition protocol was as described in ref.[37] In the example below, only one slice of the brain was considered for simplicity. The input curve was defined with external blood sampling and the samples were cross calibrated with the PET scanner and the units were in counts/pixel/sec. The

input curve and an example of a time-activity curve are displayed in figure 3. The slice considered for this study is shown in figure 4a. Ten regions of interest were drawn on this image to generate the respective TACs to be fitted with the NLLS method as shown in figure 3. These RoIs were chosen in the right (R) and left (L) hemispheres: visual (Vis), parieto-occipital (ParOccip), hemi-cortex (HemiCx), thalamus (Thal), and striatum (Striat). The initial kinetic parameters used were as indicated in figure 3. The same dynamic slice (Figure 4a) was used to generate cerebral metabolic rates of glucose (CMRGlc) parametric images with the graphical and autoradiographic methods.

The image sequences were first thresholded then analyzed using the intervals of data points for the graphical method: 10-14, 10-15 and 11-15, and using the intervals of time for the autoradiographic method: 20–50 min and 20–55 min. Note that these intervals of time refer to the frames included in the calculation, and they are not really integrals of time between 20 min and 50 min, since the radioactivity is already integrated over frame duration.

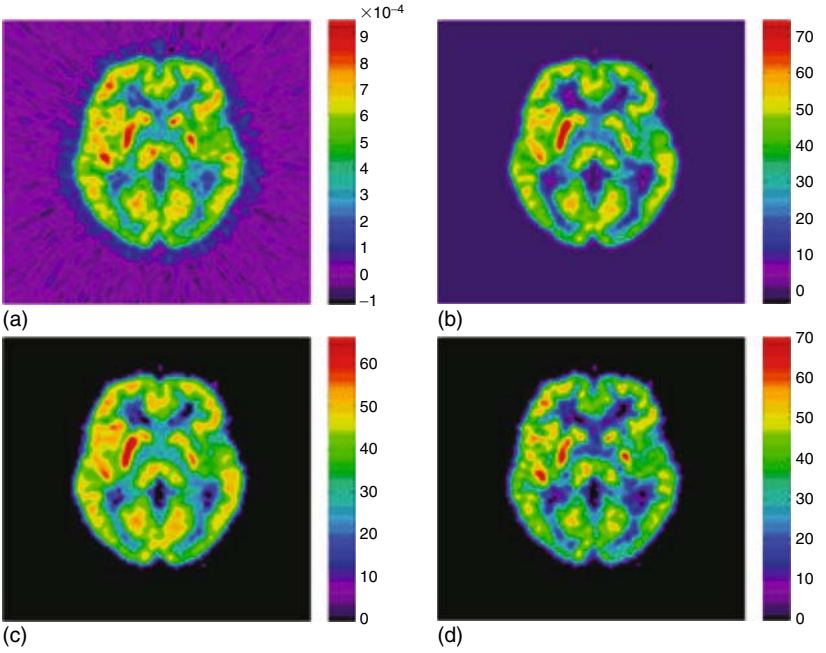


FIGURE 4. Visual comparison of the same plane for frame 15 in reconstructed image (a) reconstructed with filtered backprojection, parametric images obtained with autoradiographic method from 20 min to 55 min using Eq. 11 (b) and using Eq. 12 (c), and with patlak method using the last 5 data points (d). The images were first similarly thresholded to exclude background prior to calculation of parametric images. The color bar for the images in b-d) has units of μmoles/100g/min. Observe the negative intensity in b). See table 1 for comparison of regional CMRGlc values.

Also for the purpose of comparing Eqns. 11 and 12, CMRGlc was calculated for these two methods as can be seen in Figs. 4b-c. Figure 4b shows a wider range of CMRGlc values including negative values for some pixels in the background in comparison to that in Figure 4c. Visually the autoradiographic method (Figure 4d) generated a parametric image showing a distribution over the pixels comparable to the uptake of the tracer (Figure 4a) with a slightly pronounced contrast. Moreover, the RoIs used in the NLLS method were reproduced on the parametric images to extract mean CMRGlc values. These values are reported in Table 1 for the different procedures investigated. If the mean CMRGlc values over the ten RoIs appear nearly constant between the columns, except that generated with NLLS which appears slightly lower, the individual values calculated with these different methods are questionable, such as the case of the right striatum, for example. Recall that for the graphical method the rate constant k_4 was not considered. Finally we mention that the parametric images were relatively fast to compute (1 or 3 minutes per slice depending on computer speed).

The example of this study doesn't show significant differences (see Table 1), because the images have good statistics. The frames duration was long enough to acquire enough counts and the input curve was smooth and finely sampled. This would be different for smaller and/or diseased brain regions. It appears from table 1 that either method can be used to estimate regional CMRGlc values as shown by the low standard deviation in the last column. Also the mean over the 10 regions (row of the mean) appears almost similar for the different methods. As discussed above, the effect of k_4 in the graph-

TABLE 1. Comparison of regional CMRGlc (μmoles/100g/min) calculated with the methods of Patlak (PAT), autoradiographic (ARD) and dynamic fitting (NLLS). The numbers after 'PAT' indicate the number of data points fitted. For ARD method, the numbers indicate the interval of time used in the calculation. Also the methods of Phelps[4] (ph) and Hutchins[40] (hu) were used for comparison.

	PAT 10–14	PAT 10–15	PAT 11–15	ARD ph20–50	ARD hu20–50	ARD ph20–55	ARD hu20–55	NLLS	Mean	SD
VisR	42.16	41.76	41.2	45	44.13	45.31	44.41	37.93	**42.74**	**2.48**
VisL	42.26	41.14	40.69	45.34	44.38	45.57	44.6	38.81	**42.85**	**2.49**
ParOccipR	37.51	36.46	36	37.17	38.43	37.59	38.69	34.98	**37.10**	**1.24**
ParOccipL	33.16	31.88	31.52	32.5	34.76	32.83	34.89	32.54	**33.01**	**1.23**
HemiCxR	43.53	41.79	41.1	44.89	44.06	45.15	44.3	41.42	**43.28**	**1.61**
HemiCxL	36.49	35.1	34.22	37.13	38.08	37.34	38.19	37.76	**36.79**	**1.44**
ThalR	45.56	46.64	48.04	46.13	44.96	47.2	45.81	40.6	**45.62**	**2.25**
ThalL	42.36	41.99	42.16	41.52	41.6	42.25	42.14	37.28	**41.41**	**1.70**
StriatR	58.3	54.76	53.79	62.47	56.86	62.42	57.1	59.71	**58.18**	**3.22**
StriatL	45.03	44.27	44.34	45.73	44.67	46.35	45.18	41.24	**44.60**	**1.53**
Mean	**42.64**	**41.58**	**41.31**	**43.79**	**43.19**	**44.20**	**43.53**	**40.23**		
SD	**6.79**	**6.41**	**6.55**	**8.06**	**5.94**	**7.98**	**5.99**	**7.38**		

ical method doesn't influence the data here since the measurement doesn't exceed 60 min.[41]

6.2 ^{13}N-Ammonia Models

^{15}O-water is the most appropriate radiotracer for blood flow measurements with PET since it is an inert freely diffusible tracer. It is freely permeable across the capillary and cellular membrane, its first pass extraction fraction approaches unity and it is insensitive to flow changes.[42,43] However, this radiotracer has an important limitation, namely the noisy PET images generated due to its short half-life (2.07 min) and its diffusion (no retention by tissue) where no appreciable contrast is present in the images. ^{13}N-ammonia, on the other hand, is avidly extracted by the myocardium and it is physiologically retained for many hours which allows to obtain images with a good contrast and good statistics. Its short half-life (9.96 min) permits repeated PET studies in a single PET session. The good signal in ^{13}N-ammonia PET images also contributes in better defining the input curve from the left ventricle PET images and in correcting for spillover from blood to tissue and from tissue to blood. This tracer moves from the vascular space to tissue by both active transport (sodium-potassium pump[44]) as well as by passive diffusion. Once inside cells, this tracer is primarily metabolized by the glutamic acid-glutamine pathway. At rest, the rate of metabolic conversion of ammonia to glutamine is sufficiently high to convert and trap a major fraction of the delivered ammonia to the myocardium by blood flow. However, for an elevated perfusion, the metabolic conversion rate of ammonia to glutamine is less than that of the delivery resulting in an efflux back out of the tissue to the capillaries which is translated as a non-linearity between myocardial perfusion and metabolism.[45,46] Meanwhile, the kinetic modeling approaches are able to isolate the delivery of ^{13}N-ammonia from its retention. The myocardial blood flow estimated with kinetic modeling approaches has been shown to be well correlated with blood flow measured with microspheres.[47,48] Another drawback of using ^{13}N-ammonia in PET imaging is the metabolites generated by the tracer. During the first 2 min after injection in humans, 90% of the blood activity is made of ^{13}N-ammonia. Between 3 to 5 min, ^{13}N-ammonia accounts for 50% while ^{13}N-glutamine and urea represent about 50% of the blood activity.[49] The effect of the metabolites overestimates the input curve, i.e. underestimate the myocardial blood flow, since the counted metabolites in blood do not penetrate the cells. Care must be taken to either include a correction for metabolites in the kinetic model or to use the first 5 min for the modeling. Finally, ^{13}N-ammonia has been found to be an excellent measure of regional myocardial perfusion in both normal and diseased states, particularly ^{13}N-ammonia studies are often combined with ^{18}F-FDG to compare myocardial blood flow with glucose metabolism in an effort to detect ''mismatch'', an index of viable but compromised tissue.[50-52] The PET-

mismatch is present if the regional ^{18}F-FDG uptake in the myocardium is increased relative to the uptake of ^{13}N-ammonia, while the PET-match is present if the uptake of ^{18}F-FDG and ^{13}N-ammonia are both decreased concordantly.

The compartmental model for ^{13}N-ammonia can be considered similar to that of ^{18}F-FDG diagrammatically represented in figure 2, except that the connection of the k_4 doesn't exist for ^{13}N-ammonia. Its corresponding kinetic model equation is expressed as:[48]

$$C_{PET}(t) = \frac{1}{t_2 - t_1} \int_{t1}^{t2} \left\{ (1 - tbv)\rho \left[\frac{K_1 k_3}{k_2 + k_3} \int_0^T C_a(u)du + \frac{K_1 k_2}{k_2 + k_3} \int_0^T C_a(u) \exp^{-(k2+k3)(T-u)} du \right] + tbv \cdot C_a(T) \right\} dT \tag{13}$$

The terms in this equation represent the tracer concentration in myocardial tissue as measured with PET (C_{PET}), the tracer in arterial blood (C_a), the transport of the tracer from the arterial blood to the tissue (K_1), the washout of the tracer from the extravascular space of the tissue (k_2), the rate of the metabolic transformation of ammonia to glutamine within the myocardial tissue (k_3), and the factor $\rho(= 1.04\text{g/ml})$ is the tissue density. Assuming that the single-pass extraction fraction for the tracer is unity, and by fitting this equation to the measured PET data, K_1 serves as an estimate of myocardial blood flow and is expressed in units of ml of blood/g of tissue/min, while k_2 and k_3 are true rate constants in units of min^{-1}. As stated, Eq. 13 allows to take account of the tissue blood volume (*tbv*). This means that the measured activity in the myocardial tissue always contains a fraction (*tbv*) of the activity as defined in the arterial curve (C_a). This fraction is removed from the tissue activity as indicated by the factor (*1-tbv*).

6.2.1 Non-linear least squares fitting

Similarly to ^{18}F-FDG, NLLS is used to adjust the TACs produced by drawing RoIs on the left ventricle (LV) heart tissue. Usually only the LV is considered in the studies since it is the mostly involved in pumping blood over the whole body, and most of the pathologies affect LV. LV can be divided into longitudinal and transversal segments. The endocardium and epicardium could also be studied separately. These divisions of heart tissue allow to estimate the viability of these regions as a function of the different arteries and veins specifically supplying the heart tissue. The main myocardium longitudinal segments are usually called basal, medium and apical segments. The transaxial segments are called septum, anterior, lateral and inferior, and segments could be further divided to get smaller regions.

Before drawing RoIs on the images of the LV, the images need to be reoriented. This can be done by rotating the 3D images of the heart to obtain

images oriented from base (up) to apex (down) i.e. long axis, and the transaxial slices define sections across the LV (short axis). This operation is necessary because the heart lies in an inclined position relative the PET reconstructed planes. The rotated images should be subsequently resliced to only contain the heart.

6.2.2 Application to a rat cardiac study

Small animals such as rats are widely used in PET research studies in many fields such as biodistribution of pharmacological drugs, oncology, cardiology etc.... Studies involving the brain in small animals are not actually developed mainly because of the relatively small size of the brain compared to the combined low spatial resolution and sensitivity of the scanners to image specific structures of the brain. The sedation of the animals also prevent from some specific brain studies. Another limitation in small animal PET studies is the difficulty to withdraw blood samples from the animals for quantitative analysis. In cardiac studies, the anaesthetic drugs can have an effect on the heart activity thus affecting rest and stress states of the animal.

The application reported here was part of a study described in ref.[53] The reader is referred to this paper for detailed description of the imaging protocol. The input curve was derived from an RoI drawn on the LV blood pool in the images. This input curve was not corrected for spillover from tissue to blood. The tissue RoIs were drawn on the septal, anterior, lateral and inferior regions. Owing to the fact that it is difficult to obtain the input curve from blood sampling, it is safer to estimate the curve from reconstructed images. A simple RoI delimiting the blood pool drawn on the images could be used to generate the input curve (Figure 5). This

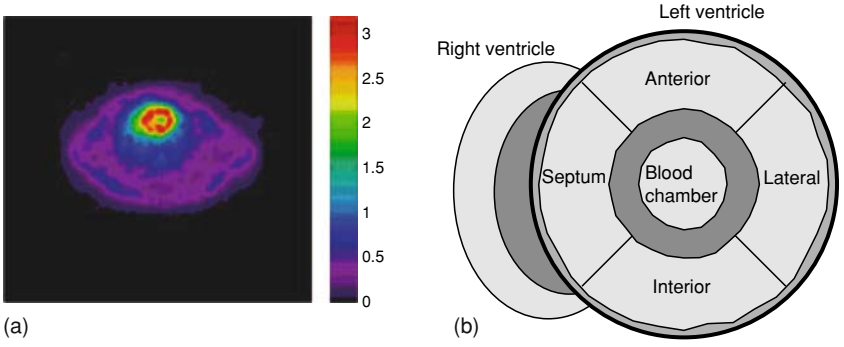


FIGURE 5. a) Example of an image slice obtained with ^{13}N-ammonia PET imaging through the heart of a normal rat with the Sherbrooke small-animal scanner. The image was acquired in 150 sec, 5 min after a bolus injection of nearly 25 sec duration. b) A diagram showing the positioning of the RoIs on the blood pool and on the myocardium in the left ventricle.

approximation might be acceptable in human cardiac studies, while in the rats, the small size of the heart and the movement of both the heart and the lungs introduce contamination from the myocardium tissue to the blood pool. Generally the kinetic modeling doesn't correct for this contamination, however it corrects for the contamination from blood to tissue, since there is always a blood contribution in tissue compartments and this is taken into account by the value of *tbv* in Eq. 13. The error introduced in the input curve is directly (not proportionally) reflected on the values of the perfusion. The input curve could also be obtained by means of factor analysis of dynamic structures (FADS).[54,55] The algorithm of FADS allows to decompose the PET measured images in a predefined number of components. In this work, an RoI was drawn on the dynamic images of the left ventricle, and this RoI was decomposed in blood and tissue components. These two components still form sequences of dynamic images from which blood and tissue TACs can be respectively determined.

Figure 6 depicts an example of image decomposition in rat images (Figure 6a) and in human images (Figure 6b). The data plotted in figure 6 represent TACs obtained from the same RoI drawn on LV blood pool on the measured images and on the decomposed blood and tissue images. The

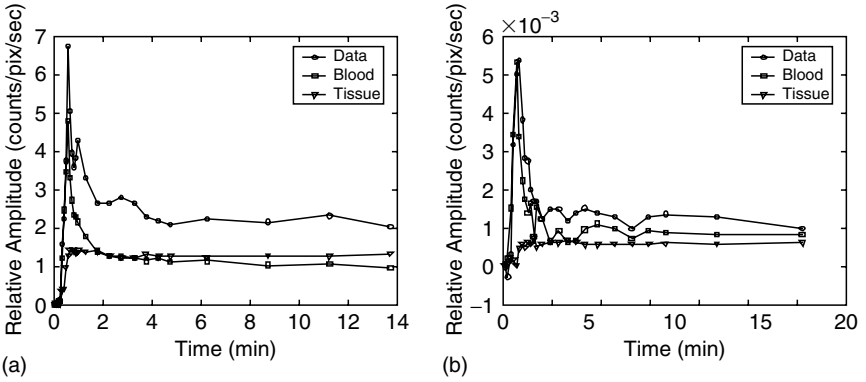


FIGURE 6. Example of input curves defined on LV blood pool images in a rat a) and in a human b). The myocardium image sequences were decomposed by means of FADS, and the same RoI was drawn on the measured image (Data), and on the blood and tissue components as obtained with FADS (Blood and Tissue respectively). In the rat, the TAC obtained from the blood component (Blood) appears smoother and lower than the TAC obtained from the measured data (Data). The difference between the two TACs in terms of amplitude is attributed to the contamination from tissue as represented by the TAC obtained from the same RoI drawn on the blood pool in the tissue component (Tissue). This difference is reduced in human images where the heart is larger and there is less interference of tissue radioactivity concentration in the blood pool.

TABLE 2. K_1 values for the whole heart versus input curves integrals. These values were averaged from 6 normal rats. The fitted TACs were defined on the measured images to show the effect of the input curves only.

	Measured images	FADS blood images	Measured/FADS
Integral of TAC	57.52	41.87	1.37
K_1 (ml/g/min)	2.18	5.03	0.43

blood and tissue components were normalized such that their sum equals that of the measured data. Note that the rat data, in contrary to the human data, were not corrected for scatter and attenuation and the images were not resliced. Also we observed low statistics in the human data. The blood pool RoI on the tissue component still shows counts in the TAC (tissue curves in Figure 6). This could be interpreted as a contamination from tissue regions to blood regions generated by heart and lung movement, by spillover and by contribution from scattered events. The amount of this contamination depends also on the size of the RoI drawn on the LV blood pool. The more the RoI is large the more the contamination is significant. Table 2 presents the values of the perfusion (K_1ml/g/min) as a function of the integrals of the input curves defined on the measured and on the isolated blood images (same RoI applied to measured and blood image sequences). The kinetic modeling was applied to the whole heart TACs obtained from the measured images. The ratio of measured to FADS as depicted in the last column of table 2 doesn't present compensation between the integrals of the input curves and K_1 values, because of the interference from *tbv* factor (see Eq. 13). This table suggests also that missing early blood samples would affect the integral of the input curve which impinges on the K_1 values.

Three other factors that seriously affect K_1 values should be mentioned. These are the size of the RoI that generates the input curve, the partial volume effect and scatter. A small RoI drawn on the *LV* blood pool could contain pixels less contaminated by the tissue radioactivity because of the distance separating the two structures, and also the smaller RoI could contain pixels of higher intensity. For the partial volume, the Sherbrooke PET scanner, a typical small animal scanner, used in this work generates data that are crucially affected by the partial volume effect for the small size of the rat hearts.[56] The RoI drawn on the *LV* blood pool images is about 5 mm which corresponds to a recovery factor of about 60%. The partial volume also introduces errors on the tissue TACs. Finally, scatter contamination can be considered as having a double effect. The first effect concerns the photons originating from the blood radioactivity at early times which is counted as tissue radioactivity, and the second effect is the scatter from tissue to blood at later frames. The first effect could be compensated by an increase in the blood compartment (*tbv*) without affecting K_1 values while the second effect lowers K_1 values since it overestimates the input curve.

Unfortunately FADS could not correct for scatter when decomposing images in blood and tissue components.

7. Summary

This chapter has presented the basic principles and clinical and research applications of the kinetic modeling approach, which can be applied to various dynamic data acquired by planar scintigraphy, SPECT or PET. Two main difficulties can arise when selecting a particular compartmental model: when the number of identifiable components is less than the chosen model (e.g. high noise) or more than the chosen model (e.g. heterogeneity). Different strategies have been suggested to tackle these problems.

The analyses presented above and concerning ^{18}F-FDG and ^{13}N-ammonia are not new *per se* but the aims was to present typical applications and expose some effects which affect PET data quantitation. It should be emphasized that in research settings, there are differences in reported values for the same parameters in the same category of studies, because the protocols of measurements and procedures of image reconstruction and data analysis are different (see chapters 4 and 7).

Acknowledgments. MB holds a scholarship from "Le Fonds de la Recherche en Santé du Québec (FRSQ), Québec, Canada and was supported by the Natural Sciences and Engineering Council of Canada (NSERC/CRSNG) under grant No 238542–01. HZ is supported by the Swiss National Science Foundation under grant SNSF 3152A0–102143.

References

1. Lawson R. S., Application of mathematical methods in dynamic nuclear medicine studies. *Phys Med Biol* **44:** R57-98 (1999).
2. Gunn R. N., Gunn, S. R., Turkheimer, F. E. *et al.*, Positron emission tomography compartmental models: a basis pursuit strategy for kinetic modeling. *J Cereb Blood Flow Metab* **22:** 1425-1439 (2002).
3. Sokoloff L., Reivich, M., Kennedy, C. *et al.*, The [^{14}C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *J Neurochem* **28:** 897-916 (1977).
4. Phelps M. E., Huang, S. C., Hoffman, E. J. *et al.*, Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18)2-fluoro-2-deoxy-D-glucose: validation of method. *Ann Neurol* **6:** 371-388 (1979).
5. Mintun M. A., Raichle, M. E., Kilbourn, M. R. *et al.*, A quantitative model for the in vivo assessment of drug binding sites with positron emission tomography. *Ann Neurol* **15:** 217-227 (1984).

6. Gunn R. N., Gunn, S. R. and Cunningham, V. J., Positron emission tomography compartmental models. *J Cereb Blood Flow Metab* **21:** 635-652 (2001).
7. Gjedde A., Calculation of cerebral glucose phosphorylation from brain uptake of glucose analogs in vivo: A re-examination. *Brain Res Rev* **4:** 237-274 (1982).
8. Patlak C. S., Blasberg, R. G. and Fenstermacher, J. D., Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab* **3:** 1-7 (1983).
9. Cunningham V. and Jones, T., Spectral analysis of dynamic PET studies. *J Cereb Blood Flow Metab* **13:** 15-23 (1993).
10. Logan J., Graphical analysis of PET data applied to reversible and irreversible tracers. *Nucl Med Biol* **27:** 661-670 (2000).
11. Schmidt K. C. and Turkheimer, F. E., Kinetic modeling in positron emission tomography. *Q J Nucl Med* **46:** 70-85 (2002).
12. Carson R., "Tracer kinetic modeling in PET" in: *Positron Emission Tomography: Basic Science and Clinical Practice*., edited by P E Valk, D L Bailey, D W Townsend *et al*. Springer-Verlag, London, (2003), *pp* 147-179.
13. Morris E., Endres, C., Schmidt, K. *et al*., "Kinetic Modeling in PET" in: *Emission Tomography: The Fundamentals of PET and SPECT*., edited by M Wernick and J Aarsvold Academic Press,, San Diego, (2004), *in press*
14. Schwaiger M., *Cardiac Positron Emission Tomography*, Kluwer Academic Publisher, Boston, (1996).
15. Silverman D. and Phelps, M., Application of positron emission tomography for evaluation of metabolism and blood flow in human brain: normal development, aging, dementia, and stroke. *Mol Genet Metab* **74:** 128-138 (2001).
16. Mazière B. and Mazière, M., Where have we got to with neuroreceptor mapping of the human brain? *Eur J Nucl Med* **16:** 817-835 (1990).
17. Delforge J., Syrota, A., Lançon, J. *et al*., Cardiac beta-adrenergic receptor density measured in vivo using PET, CGP 12177, and a new graphical method. *J Nucl Med* **32:** 739-748 Erratum in: J Nucl Med. 1994 May;35(5):921 (1991).
18. Eary J., PET imaging for planning cancer therapy. *J Nucl Med* **42:** 770-771 (2001).
19. Bradley J. D., Perez, C. A., Dehdashti, F. *et al*., Implementing biologic target volumes in radiation treatment planning for non-small cell lung cancer. *J Nucl Med* **45:** 96S-101 (2004).
20. Freedman M. T. N., Sundaram, S. K., Kurdziel, K. *et al*., Comparison of SUV and Patlak slope for monitoring of cancer therapy using serial PET scans. *Eur J Nucl Med Mol Imaging* **30:** 46-53 (2003).
21. Chatziioannou A. F., Molecular imaging of small animals with dedicated PET tomographs. *Eur J Nucl Med Mol Imaging* **29:** 98-114 (2002).
22. Del Guerra A. and Belcari, N., Advances in animal PET scanners. *Q J Nucl Med* **46:** 35-47 (2002).
23. Gambhir S. S., Barrio, J. R., Herschman, H. R. *et al*., Imaging gene expression: principles and assays. *J Nucl Cardiol* **6:** 219-233 (1999).
24. Haberkorn U. and Altmann, A., Functional genomics and radioisotope-based imaging procedures. *Ann Med* **35:** 370-379 (2003).
25. Phelps M. E., PET: the merging of biology and imaging into molecular imaging. *J Nucl Med* **41:** 661-681 (2000).
26. Cherry S., In vivo molecular and genomic imaging: new challenges for imaging physics. *Phys Med Biol* **49:** R13-R48 (2004).

27. Strauss H. W., Grewal, R. K. and Pandit_Taskar, N., Molecular imaging in nuclear cardiology. *Semin Nucl Med* **34:** 47-55 (2004).
28. Zierler K., Whole body glucose metabolism. *Am J Physiol* **276:** E409-426 (1999).
29. Bailey D. and Jones, T., A method for calibrating three-dimensional positron emission tomography without scatter correction. *Eur J Nucl Med* **24:** 660-664 (1997).
30. Bentourkia M., Bol, A., Ivanoiu, A. *et al*., A standardized blood sampling scheme in quantitative FDG-PET studies. *IEEE Trans Med Imaging* **18:** 379-384 (1999).
31. Reivich M., Alavi, A., Wolf, A. *et al*., Glucose metabolic rate kinetic model parameter determination in humans: the lumped constant and rate constants for [^{18}F]fluorodeoxyglucose and [^{11}C]deoxyglucose. *J Cereb Blood Flow Metabol* **5:** 179-192 (1985).
32. Wienhard K., The FDG model and its application in clinical PET studies. *Neural Transm Suppl* **37:** 39-52 (1992).
33. Schmidt K., Lucignani, G. and Sokoloff, L., Fluorine-18-fluorodeoxyglucose PET to determine regional cerebral glucose utilization: a re-examination. *J Nucl Med* **37:** 394-399 (1996).
34. Mori K., Schmidt, K., Jay, T. *et al*., Optimal duration of experimental period in measurement of local cerebral glucose utilization with the deoxyglucose method. *J Neurochem* **54:** 307-319 (1990).
35. Levenberg K., A method for the solution of certain problems in least squares. *Quart Appl Mat* **2:** 164-168 (1944).
36. Marquardt D., An algorithm for least-squares estimation of nonlinear parameters. *SIAM J Appl Math* **11:** 431-441 (1963).
37. Bentourkia M., Bol, A., Ivanoiu, A. *et al*., Comparison of regional cerebral blood flow and glucose metabolism in the normal brain: effect of aging. *J Neurol Sci* **181:** 19-28 (2000).
38. Gjedde A., Wienhard, K., Heiss, W. *et al*., Comparative regional analysis of 2-fluorodeoxyglucose and methylglucose uptake in brain of four stroke patients. With special reference to the regional estimation of the lumped constant. *J Cereb Blood Flow Metab* **5:** 163-178 (1985).
39. Wienhard K., Pawlik, G., Herholz, K. *et al*., Estimation of local cerebral glucose utilization by positron emission tomography of [18F]2-fluoro-2-deoxy-D-glucose: a critical appraisal of optimization procedures. *J Cereb Blood Flow Metab* **5:** 115-125 (1985).
40. Hutchins G., Holden, J., Koeppe, R. *et al*., Alternative approach to single-scan estimation of cerebral glucose metabolic rate using glucose analogs, with particular application to ischemia. *J Cereb Blood Flow Metab* **4:** 35-40 (1984).
41. Lammertsma A. A., Brooks, D. J., Frackowiak, R. S. *et al*., Measurement of glucose utilisation with [18F]2-fluoro-2-deoxy-D-glucose: a comparison of different analytical methods. *J Cereb Blood Flow Metab* **7:** 161-172 (1987).
42. Bergmann S., Fox , K., Rand, A. *et al*., Quantification of regional myocardial blood flow in vivo with H215O. *Circulation* **70:** 724-733 (1984).
43. Schelbert H. and Czernin, J., Noninvasive quantification of regional myocardial blood flow: assessment of myocardial perfusion reserve and collateral circulation. *J Nucl Med* **31:** 271-273 (1990).
44. Therien A. G. and Blostein, R., Mechanisms of sodium pump regulation. *Am J Physiol Cell Physiol* **279:** C541-566 (2000).

45. Krivokapich J., Huang, S. C., Phelps, M. E. *et al.*, Dependence of 13NH3 myocardial extraction and clearance on flow and metabolism. *Am J Physiol* **242:** H536-542 (1982).
46. Hutchins G. D., Schwaiger, M., Rosenspire, K. C. *et al.*, Noninvasive quantification of regional blood flow in the human heart using N-13 ammonia and dynamic positron emission tomographic imaging. *J Am Coll Cardiol* **15:** 1032-1042 (1990).
47. Bergmann S. R., Herrero, P., Markham, J. *et al.*, Noninvasive quantitation of myocardial blood flow in human subjects with oxygen-15-labeled water and positron emission tomography. *J Am Coll Cardiol* **14:** 639-652 (1989).
48. Muzik O., Beanlands, R. S., Hutchins, G. D. *et al.*, Validation of nitrogen-13-ammonia tracer kinetic model for quantification of myocardial blood flow using PET. *J Nucl Med* **34:** 83-91 (1993).
49. Rosenspire K. C., Schwaiger, M., Mangner, T. J. *et al.*, Metabolic fate of [13N]ammonia in human and canine blood. *J Nucl Med* **31:** 163-167 (1990).
50. Czernin J., Porenta, G., Brunken, R. *et al.*, Regional blood flow, oxidative metabolism, and glucose utilization in patients with recent myocardial infarction. *Circulation* **88:** 884-895 (1993).
51. Kofoed K., Hove, J., Freiberg, J. *et al.*, Relationship between regional 18F-fluorodeoxyglucose and 13N ammonia uptake in normal myocardium assessed by positron emission tomography: patterns of mismatch and effects of aging. *Int J Cardiovasc Imaging* **17:** 361-370 (2001).
52. Mesotten L., Maes, A., Herregods, M. C. *et al.*, PET "reversed mismatch pattern" early after acute myocardial infarction: follow-up of flow, metabolism and function. *Eur J Nucl Med* **28:** 466-471 (2001).
53. Bentourkia M., Croteau, E., Langlois, R. *et al.*, Cardiac studies in rats with 11C-acetate and PET: a comparison with 13N-ammonia. *IEEE Trans Nucl Sci* **49:** 2322-2327 (2002).
54. Di Paola R., Bazin, J. P., Aubry, F. *et al.*, Handling of dynamic sequences in nuclear medicine. *IEEE Trans Nucl Sci* **29:** 1310-1321 (1982).
55. Bentourkia M., Lapointe, D., Selivanov, V. *et al.*, "Determination of blood curve and tissue uptake from left ventricle using FADS in rat FDG-PET studies." IEEE Nuclear Science Symposium and Medical Imaging Conference Record, 24-30 October 1999 Seattle, WA USA, Vol. 2; *pp* 1124-1127 (1999).
56. Bentourkia M., Msaki, P., Cadorette, J. *et al.*, Nonstationary scatter subtraction-restoration in high-resolution PET. *J Nucl Med* **37:** 2040-2046 (1996).

13
Methods for Planar Image Quantification

K.F. Koral* and H. Zaidi†

1. Introduction

Planar imaging is being used less and less as its tomographic competitors, SPECT and PET, become more and more established. This is especially true with the advent of dual-modality tomographic scanners. However, a book on quantitative analysis in nuclear medicine imaging would be incomplete without a discussion of methods for planar-image quantification. Moreover, if CT information is employed, the planar method becomes more powerful. Finally, time constraints or economic considerations may argue for planar imaging over SPECT or PET and/or for the need to image the whole body in one acquisition.

We will not attempt to follow the historical development and use of planar-image quantification, but rather we will present methods that, in our opinion, are still of interest. We will also give a few current applications. We will not cover planar quantification in which combined attenuation-scatter correction has been carried out by using an artificially-small linear attenuation coefficient, because that approach has been almost completely superseded by other methods.

Planar-imaging-quantification methods that are currently of interest are strongly influenced by the early use of the geometric mean of opposed (alias "conjugate") views. A brief history of early developments can be found in a book chapter.[1] It is not necessary to take the geometric mean but it does have appeal, is almost always used in newer methods, and will be almost exclusively discussed below except for one study. As explained by van Rensburg *et al.*:[2]

*Dr K.F. Koral, Dept. of Radiology, University of Michigan Medical Center, Ann Arbor, MI 48109-0552, USA
†PD Dr H. Zaidi, Geneva University Hospital, Division of Nuclear Medicine, CH-1211 Geneva, Switzerland

"By acquiring images anteriorly and posteriorly the organ depth need not be determined, while the anterior-posterior attenuation can be determined using a standard source."

The advantage comes, in other words, in not needing to know where the organ (or other target) is located as far as depth when carrying out a correction for the gamma-ray attenuation. Such a correction is needed because 1) the gamma rays emerging from the sources within the patient are attenuated due to surrounding tissue, and 2) the count rate is usually converted to an activity by comparing to a count rate from an unattenuated (measured-in-air) source of known activity.

The authors above continue:

"This relatively simple technique has been widely applied, but has the important drawback that it does not correct for the contribution of scatter to the measurement."

However, in conjugate-view imaging, Van Reenen *et al.*[3] introduced one remedy for this disadvantage. More recently in MIRD Pamphlet Number 16, Siegel *et al.*[4] have discussed various possible corrections, including that of Van Reenen. We discuss conjugate-view scatter correction in a separate subsection on scatter correction technique when it can be covered separately from the conjugate-view method itself. We introduce the correction as we discuss the individual method when it is inherent to the method. For conjugate-view whole-body imaging, we review one recent paper in which two scatter-correction methods were compared.

2. Quantitative Whole-Body Imaging

Quantitative whole-body imaging is employed previous to, or during, some radiopharmaceutical therapies in nuclear medicine.[5,6] Planar whole-body imaging is almost always carried out by translating the patient plus cot in the z direction under and over two heads of a standard scintillation camera. The z axis is defined as the axis that lies along the long dimension of the patient.

Activity quantification during a scan at a given time after injection of a tracer amount of radioactivity can usually be carried out simply by calibrating results by the data at time zero when the whole-body activity equals the injected activity. Whether to carry out attenuation and scatter corrections to improve the quantification appears to be an open question that hasn't been extensively discussed in the literature.

For the case of quantification during radiopharmaceutical therapy, imaging the patient near time zero may not be an option since the patient may be considered to be too radioactive to bring to the imaging suite. Other calibration methods must then be used. We will consider the use of other calibration methods in section 2.2.

2.1 Tracer Amount of Activity

The technique employed by Wahl *et al.*[6] can serve as a prototype of the basic approach. They use a matrix size of at least 128 × 128 and employ a speed for the relative movement of patient to camera of 30 cm/min. Other details:

"Extremities were included in the images, and the arms were not allowed to cross over the body. The camera head(s) were brought as close to the patient as possible; the posterior view was obtained with the camera head directly below the imaging table. The scans were centred on the midline of the patient. A rectangular region of interest was drawn around the entire field of view to obtain separate anterior and posterior counts."

In their case, they obtained a room background count rate before the patient entered the room. From the anterior and posterior images, conjugate-view counts for the whole body were obtained. A separate room background was measured for each camera head and subtracted from the appropriate count rate from the patient to yield C_A and C_P, and then the geometric mean of corrected counts, C, was calculated.

$$C = \sqrt{C_A \times C_P} \tag{1}$$

As mentioned above, the whole body activity at time zero was the injected activity, A(0). It was then assumed that at a later time, τ, when the corrected geometric mean count was $C(\tau)$, the activity, $A(\tau)$, could be obtained from the corrected geometric mean count at time zero, C(0), by the equation:

$$\frac{A(\tau)}{A(0)} = \frac{C(\tau)}{C(0)} \tag{2}$$

Note that throughout this chapter we will suppress the time of acquisition. That is, we will assume that all acquisitions involve the same time. If this is not the case, then count variables like C(t) must be replaced by count/(unit acquisition time) variables.

Advantages: The activity at time zero is guaranteed to be correct.
Disadvantages: Since the activity distribution at times other than time zero is almost surely not identical to that at time zero, the assumption used for calculating $A(\tau)$ has some error. How large the error might be has not been estimated to our knowledge.

Van Reenen *et al.*[3] introduced a method for activity quantification of In-111 in 1982 that is summarized in the next section. They did not attempt to obtain organ uptake in absolute terms, however, but only as a percentage of whole-body uptake, so we have included their results here under quantitative whole-body imaging with administered-activity calibration. They sacrificed baboons to more thoroughly check the accuracy of their quantification methods which are described in the next section. Because of their

choices, their results are not easily compared to those of others. Because of their verification in animals, however, their results should be investigated.

In baboons, for both the 274 keV photopeak window and the 172 keV window, attenuation correction plus geometric mean, scatter correction plus geometric mean, both corrections plus geometric mean, geometric mean, anterior, and posterior estimation all seemed to produce fairly similar values. However, from the measurements in humans, geometric mean did appear to produce consistently different results for relative quantification of both the liver and the spleen than quantification from an anterior view alone, or that from a posterior view alone. It is the prejudice of the authors of this book chapter to assume that this difference implies that in humans conjugate views yielded higher accuracy. Current use of a single view is mostly limited to cases where the target is near the skin surface but not in all cases (see ref.[7]).

2.2 Therapy Amount of Activity

This subsection is concerned with quantitative whole-body imaging that does not require taking an image immediately after a known amount of activity is administered. Delpon *et al.*[8] devised a method which utilizes a reference source that was placed beyond the edge of the patient but within the field of view. Although the authors intended to use both attenuation and scatter correction, they also tested results without one or the other, and without both (labelled 20% in Figure 1). Without any corrections, the calculation is essentially that in equations 1 and 2 except that A(0) is replaced by A_R, the known activity of the reference source at the time of imaging, and C(0) is replaced by the geometric-mean counts within a 5-cm-diameter region of interest (RoI) placed appropriately.

When attenuation correction was employed, a Transmission Attenuation Correction Whole-Body (TAC) prototype developed by Sopha Medical Vision international (SMVi)[8] was called upon:

"A linear source of 370 MBq (10 mCi) of iodine-131 was placed in a lead block on the lower camera detector. The transmission source was collimated using a slit (5mm wide and 470 mm long) in the source block and an electronic window on the upper detector. Acquisitions were performed in step-and-shoot mode, and at each step the source scanned the detector perpendicularly to the direction of gantry movement."

No scatter correction was carried out for the transmission measurement. The transmission was introduced in the usual way for correcting geometric mean values for attenuation (see Eq. 4 below). List-mode acquisition allowed either dual-energy-window[3] or triple-energy window[9] scatter correction as well as no scatter correction.

Delpon *et al.* then tested this method at time zero with tracer imaging for five patients.[5] They compared their calculated whole-body activity based on a reference-source calibration to the administered activity and found that

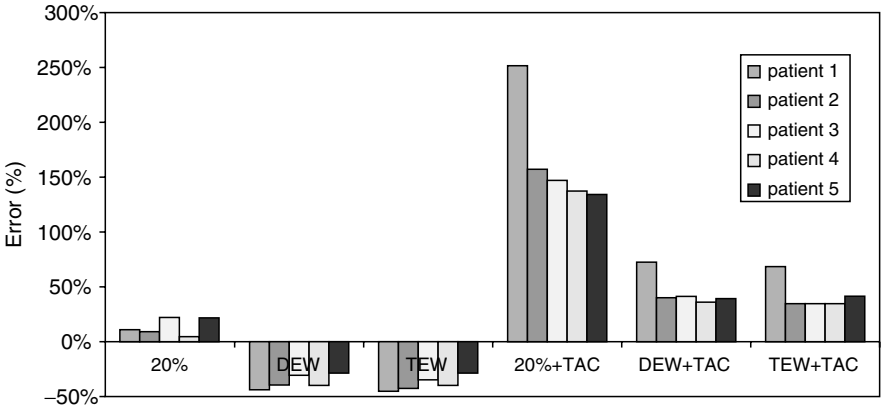


FIGURE 1. Quantification results for five patients. Error in whole-body activity with reference-source calibration is plotted versus method which is with or without transmission-based attenuation correction (TAC) as well as with or without scatter correction. The scatter correction is by either the dual energy window method (DEW) or the triple energy window method (TEW). (Reprinted with permission from ref.[5]).

their approach with only attenuation correction produced values which were considerably too high (see Figure 1). After either scatter correction, values were still overestimates but were improved. With either scatter correction but no attenuation correction, values were underestimates but had about the same absolute value for the error. The values with a simple 20% window and no corrections had the least error (6.9% $\pm$ 10.3%).

Due to the overestimation of activity when attenuation and scatter corrections are both employed, the approach of Delpon *et al.* can arguably be described as a work in progress. Delpon *et al.* are designing a high-energy high-resolution collimator to perhaps remedy the lack of improvement with corrections. A recognized problem with the approach is the lack of correction for source thickness, which is clearly the entire anterior-to-posterior thickness of the body. When an estimate of this correction is employed, the overestimation of activity with both corrections is reduced from 45% to less than 30%.[5] At the moment, using this approach without either attenuation or scatter correction appears the method of choice.

Advantages: The method of choice produced an average error of less than 7%.
Disadvantages: With that method, there was a variation among patients that was large relative to the average error.

Recently, Sjogreen *et al.*[10] devised a method that they applied to not only whole-body activity quantification, but also to quantification of

organs and tumours. The method will be described in section 3.7. We will say here only that in testing the "total image", quantification was in error by 3, 6, −7, and −4% depending on the particular procedure followed. Assuming "total image" can be equated to "whole body," the accuracy is quite good.

3. Activity Quantification with Conjugate Views

Conjugate-view imaging is still the method used by the nuclear medicine community to obtain quantitative activity estimates for organs and/or tumours over a series of times in some instances. One example of the use of these estimates is to calculate radiation absorbed dose to tumours and organs in radioimmunotherapy.[11-14] Besides the use for dosimetric quantification outright, pre-therapy conjugate-view results for a large, composite tumour are also used to estimate the shape of the intra-therapy time-activity curve for the individual tumours that 1) overlap in the anterior-posterior projection, 2) make up the composite tumour, but 3) are resolved by CT/SPECT.[15-17]

Although other methods exist, for example those described in refs.[18-21], we will review seven types of conjugate views for absolute quantification of organ and tumour activity. Note that the conjugate-view method of activity quantification is called the geometric-mean method by some authors.

3.1 Classical Conjugate-View Method

The first reported approach is the classical conjugate-view method. Thomas *et al.* gave one of the older discussions of the use of the geometric mean in their extension of previous discussions of classical conjugate views.[22] Much later, Siegel *et al.* summarized the conjugate-view formalism in a MIRD pamphlet.[4] Although it is not necessary, we will assume the opposed views are an anterior-posterior combination. This is usually the case in practice. Then, define the geometric mean of counts, GM, from those views as:

$$GM = \sqrt{I_A \times I_P} \tag{3}$$

where I_A is the count within the RoI for the target in the anterior view, and I_P is the count within the same RoI appropriately mirrored in the posterior view. Then assuming there is no non-target (alias, background) activity in the region yields the simplest equation for activity of a single target, A:

$$A = \frac{GM}{\sqrt{T}} \times \frac{f}{C} \tag{4}$$

where T is the measured, dimensionless transmission through the body at the object location from anterior to posterior, f is a correction dependent on object thickness that is always less than or equal to 1 and is 1 for small

objects, and C is a conversion factor that converts counts to activity. Usually, 1) the transmission measurement is made using a source of the same radionuclide as contained in the object by simply obtaining the counts within an RoI with, and without, the patient in place; 2) the value for C is derived from a measurement in air with a point source of the same radionuclide.

Note that how to choose the exact extent of the RoI for the target may not be clear in this and many of the other conjugate-view methods. Often, but not always, no comment is made about it in a publication. In a related situation (I-131 quantification from SPECT imaging), the approach used by Gonzales-Trotter *et al*. is to give their results as a function of the size of the RoI.[23]

With the original method of conjugate views, how to choose the extent of the RoI for both the transmission-measurement image and for the calibration-point-source image is again uncertain. The obvious, inexact approach for each RoI is to choose one large enough to enclose most (or "all") of the counts. However, this answer is especially vague in the case of a radionuclide such as I-131 that has a pedestal from septal penetration. In such a case, counts from the source extend, at a low intensity, all the way to the edge of the image. Sometimes, in such a case, the count from the entire image is used, although that usage is somewhat counter-intuitive.

The original derivations by Thomas *et al*. assumed narrow-beam geometry for attenuation while in the usual clinical case scattered photons are corrupting the measurements. However, a correction for the usual clinical case has been proposed and is discussed in the scatter correction subsection below.

The publication of Thomas *et al*. also covered various complicated cases and plots of the dependence of certain parameters on the appropriate variables are presented. Included among the cases was 1) the situation where there were two or more targets located one behind the other from anterior to posterior, 2) the situation where there was a single target plus non-target (alias, background) activity in the region anterior to the source location and/or posterior to it. Originally, these cases and others required a plethora of information about thickness, attenuation coefficient and activity density for multiple regions, and so the method was mostly not implemented for such cases. Since the advent of the availability of CT scans to provide spatial dimensions for the patient's anatomy, and, by energy extrapolation, attenuation coefficients, however, the formalism deserves revisiting. In addition, as pointed out in 1999 in MIRD Pamphlet Number 16,[4] when there is uniform activity throughout the body outside a single target, the activity, A, simply needs to be multiplied by a correction factor, F, dependent on count in a background RoI. In this case, the anterior count rate is the same as the posterior count rate. The number of counts in a background RoI whose size is equal to the target RoI can be called I_B. If the body thickness from anterior to posterior through the target, T, is the same as the body thickness for the background RoI, T_B, the situation is as shown in Figure 2.

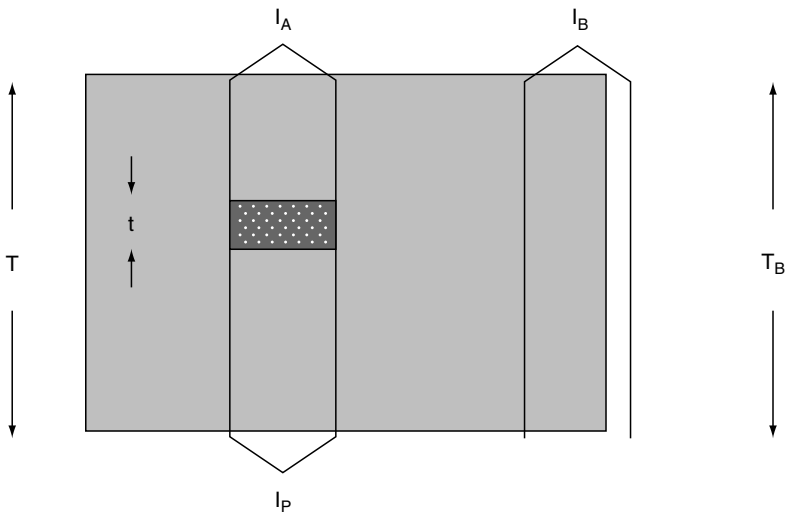


FIGURE 2. Two dimensional cartoon indicating additional measurement of background count in classical conjugate-view imaging. Target, with a uniform activity concentration, is shown in near black and background, with a different activity concentration, is shown in grey. Situation depicted is where body thickness for target is same as body thickness for background measurement.

If, in addition, the target thickness, t, is much smaller than the body thickness, then F is approximately given by

$$F = \sqrt{\left[1 - \frac{I_B}{I_A}\right] \times \left[1 - \frac{I_B}{I_P}\right]} \tag{5}$$

Note that, if the attenuation is uniform, then a measure of the correctness of the assumption that the activity distribution is uniform from anterior to posterior is the degree to which the anterior count rate in the background RoI equals the posterior count rate in that RoI. Note also that, if the background RoI size is different from the target RoI size, or if T_B is not equal to T, one needs simply to compensate for that fact.

Advantage: Extensive phantom testing with specified approaches has been carried out.
Disadvantage: The nature of the interplay between explicit scatter correction (or not), and the size of the required RoIs has not been investigated.

3.2 Conjugate Views with On-Skin Reference Source

The second method is conjugate views with a small reference source placed on the patient's skin during normal acquisition, or possibly during the acquisition of an extra pair of images.[24] To best carry out its function, this

known-activity source would be placed within the body at a position exactly analogous to the target position. In practice, the known-activity source is placed on the patient's anterior or posterior skin in a location likely to have the same transmission as that through the target. Geometric-mean target counts are calculated after an anterior-image-based background count is subtracted from anterior target counts and similarly a posterior-image-based background from posterior counts. The subtraction is primarily to account for activity anterior and posterior to the target. If it is assumed that t is small, then no correction for the difference between T and T − t is needed. It is then shown in the Shulkin *et al.* reference[24] that:

$$\frac{\mathrm{A}}{A_R} = \frac{CGM}{CGM_R} \tag{6}$$

where A is the activity of the target, A_R is the activity of the reference, CGM is the background-corrected geometric-mean count for the target, and CGM_R is the background-corrected geometric mean count for the reference.

Since counts in both the reference-source region of interest as well as in the target region of interest are augmented by patient-scattered gamma rays as well as by higher-energy emissions that deposit only a part of their energy within the crystal, usually no explicit scatter-penetration correction is carried out.

If 1) there is no background, 2) a pair of separate images are acquired with the reference source in place over the target location, and 3) CGM_R is calculated from the difference in count values, anterior and posterior, with and without the reference source in place, then Eq. 4 reduces exactly to Eq. 2.[24] (The extra pair of images acquired with the reference source in place for the reference source method is equivalent to the transmission measurement needed for classical conjugate views.) This fact indicates the close relationship between these two different methods of conjugate views.

There is some testing of different ways to create the background RoI in the original publication[24] and some appear better than others. Unfortunately, a large range of cases is not considered in the investigation.

Advantage: The calibration is carried out in a situation that closely matches the target situation. There is some guidance about choosing the background RoI.
Disadvantage: Location of the position for the reference source on the skin is, at least, somewhat arbitrary.

3.3 Conjugate Views with Two Calibration Curves

The third method was first proposed by Doherty for In-111[25] and requires the determination of two calibration curves. The first curve represents transmission versus thickness and is obtained with a Co-57 flood source and a non-radioactive water bath containing various thicknesses of water.

The second curve represents camera "efficiency" (defined as geometric mean counts/activity) as a function of thickness and is obtained by imaging a known-activity source of the radioisotope of interest centred in various thicknesses of water. No background activity is present, that is, the water is nonradioactive. Because of the lack of background, no correction for counts contributed to the target counts from activity in front of and behind the target is made. Each measured curve is fit with a 2^{nd} order polynomial as shown in (a) and (b) of Figure 3. Then the thickness variable is eliminated between them, yielding a working curve representing counting efficiency versus Co-57 transmission as shown in (c) of Figure 3.

For the radioactive object of interest, a Co-57 transmission scan is taken through it. From the measured value for transmission, the efficiency needed to convert the target geometric-mean counts to an activity estimate is

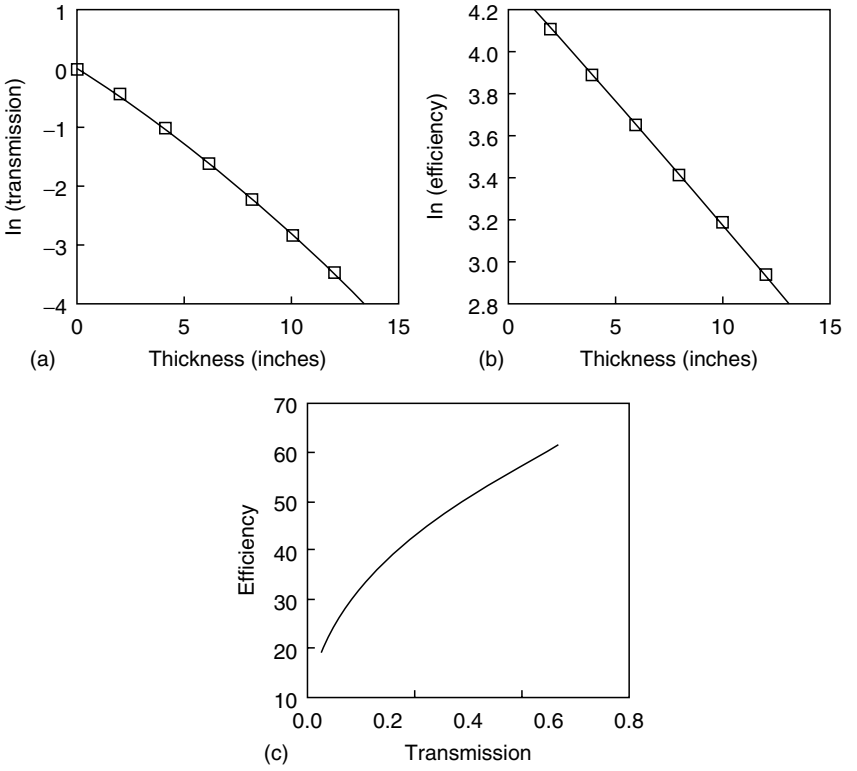


FIGURE 3. Results of phantom measurements for two-calibration-curve conjugate views. In (a), the natural log of Co-57 transmission is plotted versus water thickness. In (b), the natural log of I-131 counting efficiency is plotted versus water thickness. In (c) is shown the working curve of I-131 counting efficiency plotted versus Co-57 transmission. (Unpublished data courtesy of Kenneth R. Zasadny, Denise D. Regan and Richard L. Wahl).

calculated from the working curve (part (c) of Figure 3). Then, for both the anterior and posterior views, an estimate of counts contributed to the target counts from tissue in front of and behind the target is made from a background RoI displaced laterally from the target location. The geometric mean of the counts in this background RoI is called GM_B. We here assume the size of the RoI for the background is the same as that for the target. (If it is not, the GM_B must be corrected by multiplying by the ratio of the sizes.) Next, the thickness fraction, TF, must be estimated from a CT scan. This fraction equals the ratio of the sum of the distance from the posterior edge of the target to the posterior skin surface plus that from the anterior edge of the target to the anterior skin surface, divided by T. After that, the background-corrected geometric mean count, CGM, is given by

$$\mathrm{CGM} = \mathrm{GM} - \mathrm{TF} \times \mathrm{GM_B} \tag{7}$$

This corrected geometric-mean count for the target is then combined with the efficiency from the working curve to yield target activity.

Advantage: Good use is made of CT information. The transmission measurement can be reproduced fairly easily.
Disadvantage: The phantom model used to generate both calibration curves involves a homogeneous medium whereas that is usually not the case for the target.

3.4 Conjugate Views with Dual-Window Scatter Correction

The fourth method is an approach that is very similar to classical conjugate views but is distinguished by its introduction of a dual-window method for scatter correction and a relational constant, k. It was described by Van Reenen *et al*.[3] We have discussed some of the results from its testing in the previous section and defer the description of its method for scatter correction until subsection 3.8 below. Van Reenen *et al*. made a statement about their RoI edge detection, but it seems to indicate both a threshold method and a second-derivative approach.

Disadvantage: The k value is now known to be dependent on the ratio of the background activity density to the target activity density. Therefore, using a single value for a variety of cases must introduce some inaccuracy.[26,27]

3.5 Conjugate Views with Depth-dependent Build-up Factor

The fifth method is called the depth-independent build-up factor method. It was introduced by Siegel *et al*.[28] This method supersedes and somewhat replaces a more complicated method that won't be discussed here but that is

covered in the MIRD pamphlet for which Siegel was the first author.[4] Scatter correction is the driving force for the approach. It uses a phantom measurement to accomplish attenuation and scatter corrections. That phantom measurement produces a curve of transmission, TR, versus thickness, and the method fits the following equation to the curve:

$$TR = 1 - (1 - e^{-\mu t})^{\nu} \tag{8}$$

where t is the thickness of the water attenuator in the phantom measurement, μ is an unknown attenuation coefficient dependent on the source size, and ν is a second unknown. The two unknowns are determined by nonlinear least-squares fitting. Then, for a thin target imaged by conjugate views, one of two appropriate equations is:

$$I_A = C_{air} \times \left\{1 - (1 - e^{-\mu d})^{\nu}\right\} \tag{9}$$

where C_{air} is a target count to be determined, d is the target depth from the anterior skin surface that is to be determined, ν is the value found from the phantom measurement using equation 8, and μ is the linear attenuation coefficient for water, in lieu of tissue. A similar equation exists for I_P after d is replaced by $T - d$. The ratio of I_A over I_P is then independent of C_{air} and is solved numerically for d under the constraint that d must be greater than zero but less than T. Once d is known, C_{air} is obtained from either equation; then the activity is obtained by rationing C_{air} to the counts for a known activity source imaged in air. The rationing is exactly analogous to Eq. 6.

Although source thickness and background were not considered in the original publications, a work in 1993 addressed both.[29] Details can be found in that reference as well as a certain amount of testing.

Advantage: For a thin source, a single ν value is sufficient "... independent of source size and depth for a given window setting..." for Tc-99m.[4,30]
Disadvantage: For a thin source, the μ value is dependent on the cross sectional area of the source.[30] Also, the ν value may vary with the distribution of the background activity and with non-homogeneities in the attenuation-coefficient distribution.

3.6 *Conjugate Views with a Registered CT Scan Yielding Spatially-Variable Linear-Activity Density*

The sixth method is the CT-Assisted Matrix Inversion (CAMI) method. It is well described in a publication that tested its use for In-111.[31] This method relies on the availability of a CT scan as do several of the methods reviewed here. It is different than the others in two important respects: 1) a pair of conjugate views are not fundamentally required, and 2) the activity density per unit distance in the anterior-posterior direction, a, can have different values across the target. The later fact means the method can go a step

toward tomographic imaging, compared to classical conjugate views. On the other hand, one can assume there is no variation and then simply solve for an average a. With the method, there is, at the minimum, a VoI and an a for the target (perhaps a tumour), for each organ with a significant uptake, and for the remaining background tissue. Note that given objects, for example the liver and the right kidney can partially overlap and the method still function. This capability is a strong feature of the CAMI method.

The approach requires the registration of the CT scan with the projection image or image pair. After the registration, that CT scan then determines what length of each unknown a is associated with a given count in a projection pixel (or anterior-posterior pixel pair). The magnitude of each a is obtained from the projection data using a least-squares approach. The solution by least squares is the matrix inversion that leads to the method's name. (The system of equations is overdetermined when there are more projection pixels than the number of unknown a values.) A calibration curve from a disk source placed at various depths of uniform attenuator is used to correct for attenuation and scatter. This latter feature is similar to the depth-independent build-up factor method. As there and in some of the other methods, the necessity to let a uniform attenuator model the true system is a possible drawback.

Advantage: Good use is made of CT information. Additional information beyond total organ and/or total tumour activity can be obtained.
Disadvantage: An accurate superimposition of the three-dimensional space of the CT scan and the two-dimensional space of the projection image is required. (In the case of conjugate views, matching the CT to both projection images is required but, since the two opposed views have a known alignment, this requirement is no more problematic than the previous one.)

3.7 Conjugate Views with a Registered CT Scan and a Transmission Scan

The seventh method is an ambitious approach that uses a geometric-mean count image, a registered CT image set, as well as a transmission measurement with a Co-57 flood source.[10] The three-dimensional CT image set is used to define maximum extents in the anterior-posterior projection for organs and tumours. Both the Co-57 transmission measurement and the CT image set integrated to two-dimensions are energy extrapolated to yield patient attenuation maps appropriate for I-131 (or for the particular radionuclide of interest). In tests, attenuation correction of the emission data was carried out with both attenuation maps with about equal accuracy for organ activities. The method also employs scatter-penetration correction by deconvolution using a point-spread function which describes the response to a source in tissue.[32] Point-spread functions from different depths are deter-

mined and the most appropriate for a given patient is employed. The activity analysis is performed using regions of interest in an image derived from the coronally-sliced CT scan. Image registration is required in several places in the processing. A background subtraction is also performed. The corrected counts within a RoI are converted to activity using a camera sensitivity measured in air. This sensitivity is derived from a series of thin disks with known diameters for which the count density at the centre is determined and then multiplied by the known area. The resulting sensitivities are extrapolated to one for zero diameter which is then employed for quantification.

The components of the input data interact in a rather complicated way which Sjögreen *et al.* describe using a flow chart. As an example of the interaction, the scatter-penetration correction requires a water-equivalent patient thickness. The appropriate value was chosen as that which yielded the same attenuation as the maximum value in the Co-57 based attenuation map.

In the implementation, organ outlines "... were manually segmented in slices where outer boundaries of organs were best visualized." This single-plane procedure would preclude accuracy for an object of interest that had greatest extent in one plane for one direction and in a different plane for another direction. This might be especially important if the same approach were to be applied to tumours with convoluted shapes. However, one can envision a more robust approach designed to get the extreme extent if that were most appropriate.

In testing, the activity estimates for various organs were usually accurate within 10% and were never off by more than 21%. For four tumours placed in various locations, the activity was always underestimated, by as little as 6% and as much as 47%. The authors believe this underestimation to be due to the partial volume effect but offer no proof. Since the tumour diameters were from 2 to 2.5 times the FWHM of a point source, that cause can reasonably be questioned.

Advantage: A wealth of information, both from calibrations and also of the patient-specific type is brought to bear on the quantification.
Disadvantages: As in the CAMI method, an accurate superimposition is required. Also, with the method's many facets, it is difficult to know where to look to improve the accuracy of the tumour activity estimates. Finally, the method faces some uncertainty about the appropriate projected region of interest for a tumour with a convoluted shape. This disadvantage is probably inherent to the two-dimensionality of conjugate-view quantification, however, rather than being particular to this method.

3.8 Scatter Correction

In methods where an activity calibration source is imaged under nearly the same conditions as the target (for example the method of Shulkin *et al.*[24]), no explicit scatter correction is carried out. The assumption is that whatever

counts are added to the count total for the target are added proportional to the count total for the calibration source. When the ratio of count totals is taken (as in equation 4), the proportionality fraction cancels out independent of its value, leaving one with the correct answer.

When a point source in air is employed for activity calibration, such as in the method of Thomas *et al*., an explicit correction for scatter of gamma rays within the patient is needed to avoid a positive bias in the activity estimate for the target. One method to compensate for scatter was introduced and championed by Thomas *et al*.[1,4,33] and uses a ''pseudoextrapolation number,'' n. This number is determined from a curve of transmission versus water-phantom thickness. In the curve, transmission initially doesn't decrease linearly with thickness because scattered photons are being counted in the region of interest. The value for n, the intercept of the extrapolated straight-line portion of the curve, is 2.48 for Ga-67 with summed counts from the 93 and 184 keV photopeaks using 20% windows and a medium-energy collimator. It multiplicatively increases the T value in equation 4. Since n has a value greater than 1, the activity estimate is reduced. Thus, the correction goes in the correct direction since counting scattered gamma rays would otherwise increase the activity. The modified equation is to be applied only to targets that lie at a depth below the skin surface greater than the extent of the curve's ''shoulder.'' To our knowledge, this correction has not yet been widely employed nor investigated for other radionuclides.

In 1982, Van Reenen *et al*.[3] introduced the idea of employing two energy windows—photopeak and scatter—and a constant k to relate the scatter corruption count of the photopeak window to the count in the scatter window. They did this for planar In-111 imaging for both energy windows (172 keV and 247 keV). The anterior count as well as the posterior count is corrected before the geometric mean count is calculated. Van Reenen *et al*. chose their k value from optimizing results in a phantom experiment with the source at different depths, but they didn't reveal the k value obtained.

3.9 Independent Testing

Recently, Tang[20] carried out testing of the on-skin-reference-source method using Tc-99m. With tumours in the liver of an anthropomorphic phantom, using four different observers to independently choose the regions of interest, he found that the error increased as the background-over-tumour activity concentration increased. This dependence occurred only in some cases.

Longer ago, I-131 phantom measurements validated classical conjugate-views in certain cases, but indicated less accuracy in other specific situations.[34,35] The use of Monte Carlo simulation, where the true activity is known, would be ideal for the comparison of the bias from different methods of measurement, but so far hasn't been extensively used in such comparisons. A comparative noise analysis for different methods, though relevant, has also not as yet been carried out, to our knowledge. Recently,

head to head comparisons between tumour-activity estimates from conjugate-views and those from SPECT have been reported.[36]

4. Clinical Applications

Emerging clinical applications of quantitative planar imaging encompassed different areas including quantification of renal activity[37] and many other organs such as lungs, liver, spleen and thyroid,[7,38,39] myocardial perfusion[40] and tumour size and/or uptake in oncology studies.[6] The latter will be used as an example in this chapter to illustrate clinical implementation aspects and usefulness of the approach in a clinical and research setting.

4.1 Quantitative Analysis of Oncology Studies

Wahl *et al.*[6] have employed quantitative tracer whole-body imaging to compute whole-body radiation dose in order to ascertain the administered activity needed for the I-131 Bexxar™ therapy of an individual patient with non-Hodgkin's lymphoma. They choose the administered activity with the goal of delivering 75 cGy of whole-body radiation dose. They have found that the clearance of activity varies enough from patient to patient that this is the optimal way to choose the administered activity. The method requires carrying out tracer whole-body imaging at three time points distributed over seven days. Figure 4 shows the number of patients that were administered a

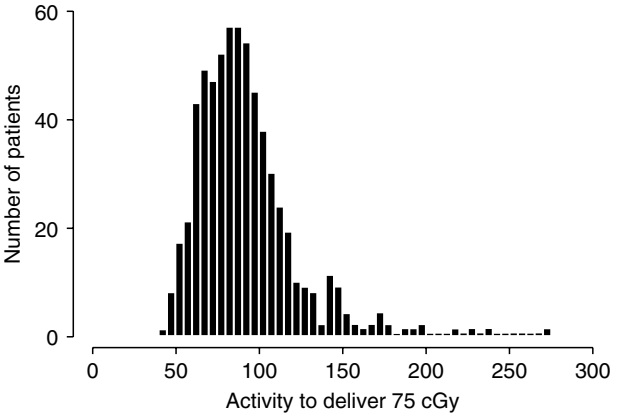


FIGURE 4. Histogram of the administered dose in millicuries estimated to be necessary to deliver a total-body dose of 75 cGy to each patient. Results from 634 patients with non-Hodgkin's lymphoma treated with tositumomab and I-131 tositumomab (I-131 Bexxar™.). Each total-body-dose estimate was based on tracer whole-body imaging and quantified as described in section 2.1. (Reprinted from ref.[41] with permission from Elsevier).

TABLE 1. Tumour self-absorbed radiation dose (TSARD) and overall disease response.

	Overall Disease Response				
TSARD (cGy)	N	PR	MR	SD	PD
> 7,000	9	7	2	0	0
1700–7000	9	3	1	5	0
< 1700	9	3	0	3	3

PR, partial response; MR, mixed response; SD, stable disease; PD, progressive disease.

particular activity in millicuries to receive 75 cGy of whole body dose.[41] It is seen that the administered activity had to vary by 50 millicuries to cover the majority of patients, and the total range of administered activity was more than 200 millicuries.

Matthay *et al.* have applied the on-skin reference source conjugate-view method (section 3.2) to accomplish tumour dosimetry of neuroblastoma patients employing I-131.[14] Tumour radiation dose was shown to have a statistically significantly correlation with both tumour volume reduction and with patient response in the results from the phase I and phase II neuroblastoma study using a single dose of I-131 meta-iodobenzylguanidine (MIBG). The median 131I-MIBG dose administered was 310 mCi (range 90 – 837 mCi). The median whole body irradiation dose achieved was 228 cGy (range 37 – 650 cGy). The median tumour self absorbed radiation dose (TSARD) was 3300 cGy (range 312 – 30,000 cGy). The TSARD correlated significantly with tumour volume decrease ($p = 0.02$). In addition, the TSARD correlated with overall disease response as shown in Table 1.

Koral *et al.*[42] have applied the two-calibration curve conjugate-view method (section 3.3) to assist in accomplishing tumour dosimetry of lymphoma patients who had not received any previous therapy but who underwent I-131 Bexxar™ therapy. They found that the patient's average radiation dose didn't correlate with the degree of response, that is, partial remission (PR) versus complete remission (CR), when the dose was low. However, a high radiation dose value was correctly associated with a CR in 9 of 12 patients (75%) when a conjugate view alone was employed for the tumour-dose estimation. Adding a single intra-therapy SPECT evaluation to improve the dose estimation increased the percentage to 94% (15 of 16 patients).

Also, a statistically significantly correlation between radiation absorbed dose and tumour volume reduction was observed for individual tumours at 12 weeks post therapy with the SPECT-supplemented dose estimation. For the study, the data set was restricted to the patients who went on to a PR. Volume reduction was calculated from repeat CT scans evaluated by the same radiologist. A 100% reduction corresponded to "disappearance" of the tumour by a criterion that it be less than 1 cm in diameter. The

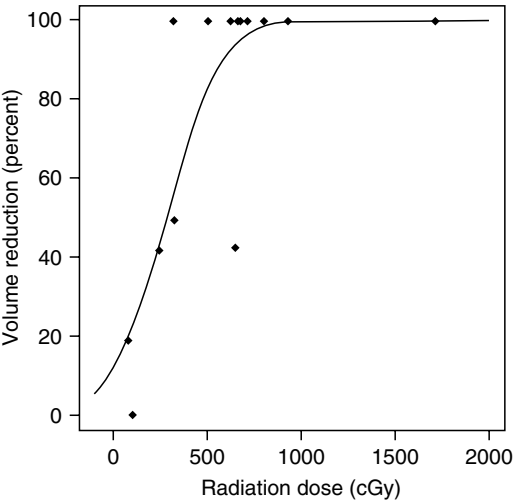


FIGURE 5. Volume reduction at 12 weeks post therapy versus radiation dose. Values are shown for 15 individual tumours after therapy with I-131 Bexxar™. Dose calculation was based on the two-calibration curve conjugate-view method supplemented by intra-therapy SPECT imaging. COPYRIGHT © 2002 American Cancer Society. (Reprinted from ref.[43] by permission of Wiley-Liss, Inc., a subsidiary of John Wiley & Sons, Inc).

significance occurred for the best-fit sigmoidally-shaped relationship between tumour volume reduction and radiation dose compared to no dose-response relationship; that is, compared to the volume reduction being constant. For the data subset consisting of tumours with an initial mass less than or equal to 10 g, a statistically significant P value of 0.029 was determined for 15 tumours from 6 patients (Figure 5).[43]

For the entire data set, irrespective of initial mass, the correlation was also statistically significant in the same sense as above with $P = 0.0496$. This data set contained 41 tumours from 9 patients.[15]

5. Summary

It is gratifying to see in overview the progress that quantitative planar imaging has made, from simple conjugate-view imaging, through conjugate-views combined with different correction techniques, and most recently towards more sophisticated approaches based on patient-specific transmission and CT images. Challenges remain, particularly in deriving the optimal and less constraining imaging protocol for applications in clinical routine and defining an objective approach for determining the calibration curves and correction factors needed for accurate quantification.

Acknowledgments. One of the co-authors (KFK) acknowledges the support of Grant R01 CA87955 awarded by the National Cancer Institute, United States Department of Health and Human Services and expresses his appreciation for an insightful critique of an early draft of the chapter by Larry Williams, City of Hope Medical Center. HZ was supported by the Swiss National Science Foundation under grant SNSF 3152A0-102143. The contents are solely the responsibility of the authors and do not necessarily represent the official view of the sponsoring agencies.

References

1. Thomas S. R., Maxon, H. R., III and Kereiakes, J. G., "Techniques for quantitation of in vivo radioactivity" in: *Effective use of computers in nuclear medicine*, edited by Michael J. Gelfand and Stephen R. Thomas McGraw-Hill, New York, (1988), *pp* 468-484.
2. van Rensburg A. J., Lotter, M. G., Heyns, A. D. *et al.*, An evaluation of four methods of ^{111}In planar image quantification. *Med Phys* **15:** 853-861 (1988).
3. van Reenen P. C., Lotter, M. G., Heyns, A. D. *et al.*, Quantification of the distribution of ^{111}In-labelled platelets in organs. *Eur J Nucl Med* **7:** 80-84 (1982).
4. Siegel J. A., Thomas, S. R., Stubbs, J. B. *et al.*, MIRD pamphlet no. 16: Techniques for quantitative radiopharmaceutical biodistribution data acquisition and analysis for use in human radiation dose estimates. *J Nucl Med* **40:** 37S-61S (1999).
5. Delpon G., Ferrer, L., Lisbona, A. *et al.*, Impact of scatter and attenuation corrections for iodine-131 two-dimensional quantitative imaging in patients *Cancer Biother Radiopharm* **18:** 191-199 (2003).
6. Wahl R. L., Kroll, S. and Zasadny, K. R., Patient-specific whole-body dosimetry: principles and a simplified method for clinical implementation. *J Nucl Med* **39:** 14S-20S (1998).
7. Macey D. J., DeNardo, G. L. and DeNardo, S. J., Planar gamma camera quantitation of ^{123}I, ^{99m}Tc or ^{111}In in the liver and spleen of an abdominal phantom. *Cancer Biother Radiopharm* **14:** 299-306 (1999).
8. Delpon G., Ferrer, L., Lenta, C. *et al.*, Comparison of four scatter correction methods for patient whole-body imaging during therapeutic trials with iodine-131. *Cancer* **94:** 1224-1230 (2002).
9. Ogawa K., Harata, Y., Ichihara, T. *et al.*, A practical method for position-dependent Compton-scatter correction in single-photon emission CT. *IEEE Trans Med Imaging* **10:** 408-412 (1991).
10. Sjogreen K., Ljungberg, M. and Strand, S. E., An activity quantification method based on registration of CT and whole-body scintillation camera images, with application to 131I. *J Nucl Med* **43:** 972-982 (2002).
11. O'Donnell R. T., DeNardo, G. L., Kukis, D. L. *et al.*, A clinical trial of radioimmunotherapy with 67Cu-2IT-BAT-Lym-1 for non-Hodgkin's lymphoma. *J Nucl Med* **40:** 2014-2020 (1999).
12. DeNardo G. L., DeNardo, S. J., Shen, S. *et al.*, Factors affecting 131I-Lym-1 pharmacokinetics and radiation dosimetry in patients with non-Hodgkin's lymphoma and chronic lymphocytic leukemia. *J Nucl Med* **40:** 1317-1326 (1999).

13. DeNardo S. J., DeNardo, G. L., Kukis, D. L. *et al.*, 67Cu-2IT-BAT-Lym-1 pharmacokinetics, radiation dosimetry, toxicity and tumor regression in patients with lymphoma. *J Nucl Med* **40:** 302-310 (1999).
14. Matthay K. K., Panina, C., Huberty, J. *et al.*, Correlation of tumor and whole-body dosimetry with tumor response and toxicity in refractory neuroblastoma treated with (131)I-MIBG. *J Nucl Med* **42:** 1713-1721 (2001).
15. Koral K. F., Kaminski, M. S. and Wahl, R. L., Correlation of tumor radiation-absorbed dose with response is easier to find in previously untreated patients. *J Nucl Med* **44:** 1541-1543 (2003).
16. Koral K. F., Dewaraja, Y., Li, J. *et al.*, Initial results for Hybrid SPECT–conjugate-view tumor dosimetry in 131I-anti-B1 antibody therapy of previously untreated patients with lymphoma. *J Nucl Med* **41:** 1579-1586 (2000).
17. Koral K. F., Dewaraja, Y., Clarke, L. A. *et al.*, Tumor-absorbed-dose estimates versus response in tositumomab therapy of previously untreated patients with follicular non-Hodgkin's lymphoma: preliminary report. *Cancer Biother Radiopharm* **15:** 347-355 (2000).
18. Macey D. J., Grant, E. J., Bayouth, J. E. *et al.*, Improved conjugate view quantitation of I-131 by subtraction of scatter and septal penetration events with a triple energy window method. *Med Phys* **22:** 1637-1643 (1995).
19. Miller C., Filipow, L., Jackson, S. *et al.*, Planar imaging quantification using 3D attenuation correction data and Monte Carlo simulated buildup factors. *Phys Med Biol* **41:** 1401-1423 (1996).
20. Tang H. R., "A combined x-ray-CT scintillation camera system for monitoring radionuclide uptake in tumors" Ph.D Dissertation, University of California, 1998.
21. Nicolini G., Butti, I., Perna, L. *et al.*, A practical approach for accurate quantitative measurements involving planar imaging by gamma-camera. *Physica Medica* **18:** 115-120 (2002).
22. Thomas S. R., Maxon, H. R. and Kereiakes, J. G., In vivo quantitation of lesion radioactivity using external counting methods. *Med Phys* **3:** 253-255 (1976).
23. Trotter D. E. G., Bowsher, J. E. and Jaszczak, R. J., Absolute quantitation of a spherical I-131 activity distribution using a high-resolution rotating collimator: A phantom study. *IEEE Trans Nucl Sci* **48:** 65-73 (2001).
24. Shulkin B. L., Sisson, J. C., Koral, K. F. *et al.*, Conjugate view gamma camera method for estimating tumor uptake of iodine-131 metaiodobenzylguanidine. *J Nucl Med* **29:** 542-548 (1988).
25. Doherty P., Schwinger, R., King, M. *et al.*, "Distribution and dosimetry of Indium-111 labeled F(ab′)2 fragments in humans." Fourth International Radiopharmaceutical Dosimetry Symposium, Oak Ridge, TN, *pp* 464-476 (1985).
26. Koral K. F., "Monte Carlo in SPECT Scatter Correction" in: *Monte Carlo Calculations in Nuclear Medicine*, edited by M. Ljungberg, S.-E. Strand, and M.A. King IOP Publishing Ltd., Bristol and Philadelphia, (1998), *pp* 165-181.
27. Zaidi H. and Koral, K. F., Scatter modelling and compensation in emission tomography. *Eur J Nucl Med Mol Imaging* **31:** 761-782 (2004).
28. Siegel J. A., Wu, R. K. and Maurer, A. H., The buildup factor: effect of scatter on absolute volume determination. *J Nucl Med* **26:** 390-394 (1985).
29. Kojima A., Takaki, Y., Matsumoto, M. *et al.*, A preliminary phantom study on a proposed model for quantification of renal planar scintigraphy. *Med Phys* **20:** 33-37 (1993).

30. Siegel J. A., The effect of source size on the buildup factor calculation of absolute volume. *J Nucl Med* **26:** 1319-1322 (1985).
31. Liu A., Williams, L. E. and Raubitschek, A. A., A CT assisted method for absolute quantitation of internal radioactivity. *Med Phys* **23:** 1919-1928 (1996).
32. Yanch J. C., Flower, M. A. and Webb, S., A comparison of deconvolution and windowed subtraction techniques for scatter compensation in SPECT. *IEEE Trans Med Imaging* **7:** 13-20 (1988).
33. Thomas S. R., Gelfand, M. J., Burns, G. S. *et al*., Radiation absorbed-dose estimates for the liver, spleen, and metaphyseal growth complexes in children undergoing gallium-67 citrate scanning. *Radiology* **146:** 817-820 (1983).
34. Hammond N. D., Moldofsky, P. J., Beardsley, M. R. *et al*., External imaging techniques for quantitation of distribution of I-131 F(Ab)2 fragments of monoclonal-antibody in humans. *Med Phys* **11:** 778-783 (1984).
35. Eary J. F., Appelbaum, F. L., Durack, L. *et al*., Preliminary validation of the opposing view method for quantitative gamma camera imaging. *Med Phys* **16:** 382-387 (1989).
36. Tang H. R., Da Silva, A. J., Matthay, K. K. *et al*., Neuroblastoma imaging using a combined CT scanner-scintillation camera and 131I-MIBG. *J Nucl Med* **42:** 237-247 (2001).
37. Kojima A., Takaki, Y., Tsuji, A. *et al*., Quantitative renography with the organ volume method and interporative background subtraction technique. *Ann Nucl Med* **10:** 401-407 (1996).
38. Forge N. I., Mountford, P. J. and O'Doherty, M. J., Quantification of technetium-99m lung radioactivity from planar images. *Eur J Nucl Med* **20:** 10-15 (1993).
39. Zaidi H., Comparative methods for quantifying thyroid volume using planar imaging and SPECT. *J Nucl Med* **37:** 1421-1426 (1996).
40. Prigent F. M., Maddahi, J., Van Train, K. F. *et al*., Comparison of thallium-201 SPECT and planar imaging methods for quantification of experimental myocardial infarct size. *Am Heart J* **122:** 972-979 (1991).
41. Wahl R. L., The clinical importance of dosimetry in radioimmunotherapy with tositumomab and iodine I 131 tositumomab. *Semin Oncol* **30:** 31-38 (2003).
42. Koral K. F., Dewaraja, Y., Li, J. *et al*., Update on hybrid conjugate-view SPECT tumor dosimetry and response in 131I-tositumomab therapy of previously untreated lymphoma patients. *J Nucl Med* **44:** 457-464 (2003).
43. Koral K. F., Francis, I. R., Kroll, S. *et al*., Volume reduction versus radiation dose for tumors in previously untreated lymphoma patients who received iodine-131 tositumomab therapy. Conjugate views compared with a hybrid method. *Cancer* **94:** 1258-1263 (2002).

14
Quantitative Analysis in Functional Brain Imaging

K. Van Laere* and H. Zaidi†

1. Introduction

Molecular genetics and brain mapping have created two totally different revolutionary approaches in medicine and neuroscience. Microscopic and macroscopic approaches to understand brain function both *in vitro* and *in vivo* have provided an overwhelming amount of data concerning its operation in health and disease. Modern functional brain mapping techniques such as PET (positron emission tomography), SPE(C)T (single-photon emission computed tomography), fMRI (functional magnetic resonance imaging), EEG (electro-encephalography), MEG (magneto-encephalography), optical imaging and neuroanatomical tools have been used for assessing the functional organization of the human brain.[1] Through these techniques, neuroscience has progressed greatly in the study of physiology, sensory and motor systems, vision, attention, memory, language and emotion in normal and pathological neurological or psychiatric conditions.[2]

The specific role of nuclear medical imaging techniques (PET and SPECT) in the expansion of our understanding of the pathophysiological mechanisms of neurological and psychiatric diseases and in the clinical management of patients is steadily progressing. PET and SPECT allow *in vivo*, non-invasive 3D imaging of regional cerebral blood flow, metabolism and neuroreceptor binding in the brain, amino acid synthesis or amyloid plaques. Since functional disturbances occur often earlier than structural once, a faster and more sensitive detection is possible.[3]

During recent years, non-invasive tomographic 'molecular' mapping of brain function with PET and SPECT has improved markedly through the development of dedicated instrumentation and the synthesis of new radiopharmaceuticals. Triple-headed gamma cameras with fanbeam collimation

*Prof. K. Van Laere, Leuven University Hospital, Division of Nuclear Medicine, 3000 Leuven, Belgium
†PD Dr H. Zaidi, Geneva University Hospital, Division of Nuclear Medicine, CH-1211 Geneva, Switzerland

have been in use since the early nineties,[4] allowing acquisition of 7-8 mm resolution images of the brain with improved detection efficiency. The development of high resolution PET devices dedicated for brain imaging in clinical use[5] or under development in research laboratories[6] is already on a high level. Through advanced techniques such as coregistration of functional and/or anatomical images, as well statistical parametric analysis techniques, synergistic information is obtained about cerebral functional or biochemical processes that contribute importantly in the refining of the diagnostic process.[7] Many new radioligands have been developed and some have been installed in daily clinical practice such as dopamine transporter and receptor ligands.[8]

It should be clarified that the term *quantification* has been used in several meanings in the functional brain imaging literature:

1. as (inappropriate) synonym for semi-quantification where relative measurements are compared to a disease-free or receptor-void region;
2. as 'absolute' physical quantification, where 'true' *activity* quantification is indicated, mostly incorporating corrections for physical degradation factors inherent to emission tomography, such as photon scatter, attenuation and resolution loss (partial volume effects); and
3. true, *physiological* 'absolute' quantification where corrected activity measurements (as in 2.) are transformed into physiological meaningful parameters, for example perfusion (e.g. in ml/min/100 g tissue), glucose metabolic rate (rGMCglc), receptor occupancy (B_{max}) and avidity (the inverse of the dissociation constant: $1/K_D$).

It should be kept in mind that only the third is of true physiological interest and therefore kinetic modelling of the activity measurements is mandatory. Nevertheless, approaches to derive approximations of physiological parameters from semi-quantitative measures have enabled simplified but clinically achievable and highly relevant measurements. Measures of group discrimination and disease progression will not necessarily benefit from the acquisition of parameters associated with improved absolute quantitation. For example, large intersubject variability of rCMRglc measurements in the normal population have always been a major problem for the detection of pathology[9,10] governed not only by physiological factors, but also large number of error sources that can affect the reproducibility of quantitative rCMRglc.[11]

Nevertheless, many other authors have agreed that absolute physiological quantification can be of clinical importance for evaluation of global cerebral involvement as in cerebrovascular disease or traumatic brain injury, in activation studies and pharmacological testing, as well as in long-term follow-up.[12-17] Moreover, since arterial sampling has been a major restricting factor in the clinical application of absolute quantification, non-invasive techniques have been investigated to avoid traumatic invasive sampling both for PET[18-24] and SPECT.[25-29]

The possibility of (non-invasive) physiologically quantitative monitoring of brain function is still one of the strengths of emission tomography. Newer techniques such as functional magnetic resonance imaging using arterial spin labelling (ASL) currently also offer prospects of quantitative determination of cerebral blood flow that are in accordance to PET measurements.[30] However, the technique has not found widespread clinical applications because of methodological difficulties e.g. in tagging multislice volumes.

Physiological quantitation will not be the main aim of this chapter, for this the reader is referred to the principles of kinetic modelling in chapter 12. The content of this current chapter has been divided into three main parts. In the first section, the general approach of functional human brain mapping is outlined as based upon anatomical standardization. It can be assumed that the brain surrounded by skull possesses sufficient anatomical rigidity so that linear or moderate warping transformations can be devised to map each individual's skull into a common stereotactic framework. From this approach, the generation of functional imaging databases can be devised. The transfer of normal databases between different cameras and the (automated) quantitative analysis techniques such as volume-of-interest (VOI)- and voxel-based approaches (statistical parametric mapping and stereotactic surface projections) are described and compared. In the next subchapters, the current status of available normal databases which can be used for various pathophysiological and clinical applications is assessed both for PET and SPECT. As a conclusion, some new prospects in functional brain imaging and associated quantitative data analysis are discussed.

2. Functional Human Brain Mapping Using PET and SPECT

Variable quality and operator dependence, lack of suitable (universally) normal data, limited quantitative possibilities and spatial/temporal resolution characteristics are methodological factors that have hampered the clinical applications of brain SPECT and PET. Moreover, few large, well controlled clinical studies have been described, with relatively small numbers of patients, with selection problems, possible medication interference etc. Also, competing modalities have emerged such as perfusion and diffusion MR, which especially for cerebrovascular applications have gained widespread use.[31] Although a general problem for neurological and psychiatric patients has been a relative lack of therapeutic options, during the last years new medication (e.g. acetylcholinesterase blockers for cognitive impairment) and neurosurgical techniques (e.g. neurostimulation, neuroablation) have become available and are being investigated (e.g. gene transfer, implantation of foetal cells). The lack of a golden standard (clinical diagnosis with long-term follow-up) versus pathological confirmation is a common problem to most studies regarding brain pathology.

Whereas current analysis methods in many nuclear medicine departments are based on mere visual assessment or on a semi-quantitative approach based on operator-dependent and time-consuming manual VOI techniques, thereby increasing observer variance, anatomical standardization allows automated and operator-independent VOI- or voxel-based quantitative analysis of whole-brain data. The technique dates back at the late 1980's[32] but has come into use mainly since the second half of the 1990's. Such automatic quantitative assessment of brain functional data (perfusion, metabolism, receptor binding, ...) is attractive since it can also lower variability across institutions and may enhance the consistency of image interpretation independent of reader experience.

2.1 Anatomical Standardization of Functional Brain Data

Image registration algorithms (see chapter 9) are an essential tool to standardize 3D functional or structural measurements in a common orientation, e.g. stereotactic Talairach space. Such anatomical or stereotactic standardization, is an attractive methodology to compare intra- as well as intermodality *in vivo* brain information and allows the (automated) analysis of information covering the whole brain at a resolution even better than the intrinsic camera capabilities or generation of specific functional or structural brain maps. This brain mapping approach or anatomic standardization was originally used to localize activation foci in PET activation studies and afterwards to find hypometabolic areas in FDG PET images. However, in the last few years, it has become the rule rather than an exception in the scientific literature, as mapping techniques using PET and SPECT have increased almost exponentially, especially after 1997, as can be seen from a PUBMED literature search shown in Figure 1.

There is a difference between anatomic standardization (also called spatial normalization) and coregistration in the narrow sense of the word. As anatomic standardization is a technique to transform brain images of individual subjects into a standard brain, it is different from coregistration which matches one image to another, usually of a different modality or tracer, but of the same subject through a rigid-body transformation (see chapter 9).

Mathematically, both processes transform an image to match another so that a cost function is minimized, but the algorithmic difference is whether a nonlinear or rigid-body transformation is used (although a non–rigid-body transformation may be used for PET–MRI coregistration to account for distortion in MRI). Anatomic standardization analyzes subject groups, for which individual variations are treated statistically. The examination of a patient image in the standardized coordinate system to detect pixels above or below the normal range is a special case of a group analysis, because the patient is compared with a group of healthy volunteers and the result depends on the selection of the volunteers.

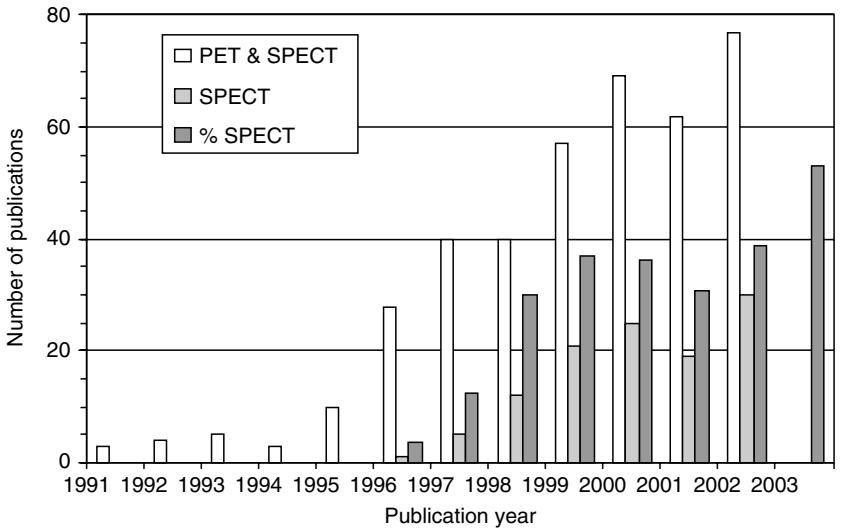


FIGURE 1. Number of peer reviewed publications of parametric mapping using PET and SPECT in a PUBMED search between 1991 and 06/2003. There is a steep increase after 1997. SPECT seems to show a lag-time of about two to three years compared to PET, but the relative proportion of SPECT (X pattern) is increasing. (Search terms: (PET OR SPECT) AND (brain OR cerebr*) AND (SPM OR parametr* map* OR SSP OR Neurostat OR voxel-based OR anatomical standard* OR spatial normali* OR brain mapping)).

Another important difference is that the true solution exists for registration but not for anatomic standardization. Every brain is different from another both morphologically and functionally. The gyral pattern is known to be topologically different between subjects and functional differentiation is also different between subjects, even when it is tracked down to the neural network at the microscopic level. Using caution when standardizing patients is therefore reasonable.[33]

Coregistration methods are classified into 2 major approaches. In the landmark approach, where the brain of each subject is transformed so that the landmarks match those in the standard brain. This includes the human brain atlas (HBA), which uses morphologic information provided by PET-registered MRI.[34] In this method, operator dependency must be considered. In the second (automated approach), the radioactivity distribution (or CBF or CMRglc) that is to be matched, uses a voxel-similarity measure. Count difference and mutual information cost schemes are among those most widely used (see also chapter 9). This latter cost function is used in the software package of statistical parametric mapping (SPM), which is an academic freeware developed by the Functional Imaging Laboratory, London (see *http://www.fil.ion.ucl.ac.uk/spm/* for download and latest information) and the Michigan 3D-SSP method developed by Minoshima *et al.*[35]

2.1.1 Perfusion and Metabolism

In view of the close relation between brain function and morphology, it is very intriguing how precisely normalized brains coincide with each other. The anatomic precision of spatial normalization was validated in several studies. Both for PET perfusion and metabolism studies and high-resolution anatomical images, spatial normalization is most effective when non-linear methods are used.[36] Template registration accuracy for high-resolution perfusion SPECT or metabolic PET is of the order 2 mm.[37] The anatomic precision of spatial normalization has been validated with several methods. Large variance in structures other than the anterior commissure (AC) and size of the brain in the linear method suggests the necessity of nonlinear transformations for effective spatial normalization.[36]

The algorithms for automated voxel-similarity based anatomical standardization are well characterized for SPECT and PET perfusion or metabolism[35,38,39] and newer software is mostly based upon mutual information.[40]

It is important to consider some clinical aspects with respect to anatomical standardization. In the presence of a focal perfusion or metabolism lesions, automated algorithms attempt to reduce image mismatch between template and image at the site of the lesion. This can lead to significant and inappropriate image distortion, especially when nonlinear transforms are used. One solution is to use cost-function masking, hereby excluding the lesion areas used in the calculation of image difference, so that this lesion does not bias the transformations.[41] First studies with this technique have shown excellent results,[42] when presuming a priori knowledge of the lesion location. Automated determination of the masked-out regions, when no *a priori* lesion location information is present, remains to be investigated further.

2.1.2 Receptor Studies

Whereas many of the aspects regarding spatial normalization have been optimised for perfusion and metabolism imaging, the issue of brain mapping of receptor ligands is more difficult. For highly specific ligands with a poorly defined anatomical distribution, the techniques by which spatial normalization could be performed automatically on the functional images need to be utilized with care. Very specific radioligands (e.g. dopamine transporters with high signal-to-noise ratio) or very abnormal receptor binding images (e.g. medication blockade, severe endogenous pathology, ...) can contain very poor anatomical landmark information which is inherently needed for the automated algorithms to function properly (Figure 2). Especially nonlinear registration, which is not uniquely defined, is tricky the least, if not impossible to perform reliably.

For PET, where dynamic imaging is usually performed, the spatial normalization problem can be reduced to the one for perfusion as early images, which to a fair degree represent influx/blood flow, can be used. All other dynamic frames can then undergo the same transformation subsequently.

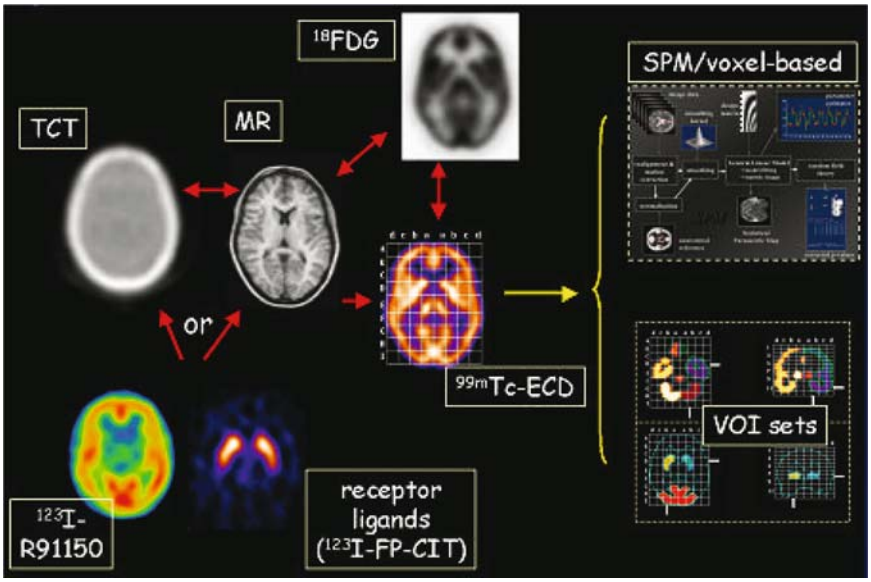


FIGURE 2. The concept of anatomical standardization for brain PET and SPECT. Receptor ligands can be spatially normalized either by intermediary transmission scan or (preferably) rigid MR coregistration. This can also be done for metabolism or perfusion, although direct automated spatial normalisation is also highly accurate. Upon spatial normalisation (e.g. in stereotaxic space), statistical parametric mapping or other voxel-based approaches can be conducted, or predefined volume-of-interest (VOI) analysis may be performed.

Another approach is the intermediary (affine) spatial normalisation of a simultaneously (e.g. SPECT with ^{99mTc} emission and ^{153}Gd transmission) or sequentially acquired transmission CT (TCT) scan to a stereotactic transmission template, and again subsequent application of the transformation parameters to the emission image. Intermediate TCT coregistration is clinically feasible and more accurate than direct manual or automated spatial normalization based upon the emission information for several receptor studies.[37] Limitations of this approach exist in situations where incomplete transmission data is acquired as the discriminative regions such as the sinuses, oronasal cavity and skull base should be completely imaged to avoid elongation along the long body axis in the automated coregistration procedure. Such effects have also been described in brain PET/MR registrations.[43] Furthermore, non-linear warping is difficult since the information included in the low-resolution transmission images is small.

The third and most generally used approach is the dual-phase approach of a rigid registration to the individual subject's high-resolution MR scan, and a non-linear MR transformation which can be performed accurately to anatomically standardize the data subsequently using the same

transformation. Many recent studies have applied the latter approach for a variety of (mainly) PET radioligands to create a variety of ligand-specific templates such as for [^{11}C] WAY-100635 (serotonin (5HT)-1-A receptors),[44] [^{11}C] raclopride (striatal dopamine D2 receptors),[45,46], [^{11}C]FLB 457 (extrastriatal distribution of D2 receptors),[47], [^{11}C] flumazenil (central benzodiazepine receptors)[48] (see also further).

2.2 *Quantitative Data Analysis Techniques*

Once anatomically standardized data have been obtained, a number of options exist to extract relevant patient information from the comparison to control (or normal) data. This comparison can be conducted by VOIs, which can be predefined in the standardized space to sample activity or parametric distributions automatically.[37,49-51]

Also voxel-based approaches are implemented within several commercial and non-commercial software packages, and are currently the most widely used method for the analysis of functional (activation) images. Statistical Parametric Mapping (SPM) is well-documented, freely available and strongly supported by many brain imaging researchers. It offers quantitative voxel-by-voxel analysis both in functional[7,52] and structural brain studies (voxel-based morphometry).[53] Neurostat (3D-SSP) is based upon the projection of the data along radial rays by which atrophy effects are minimized.[35] Among the commercial packages, BRASS (Brain Registration and Automated SPECT Semi-quantification, Nuclear Diagnostics, Hägersted, Sweden) is specifically oriented towards routine clinical brain SPECT and PET applications, and allows voxel-wise comparison of individual studies by means of statistical intensity differences, compared to a mean and standard deviation image from a control group, based upon region-growing of deviating voxels.

2.2.1 Voxel-based Analysis

2.2.1.a Statistical Parametric Mapping (SPM)

SPM is the most popular method in voxel-based data analysis of functional brain data, for PET, SPECT as well as fMRI. SPM has been used to reveal which parts of the brain are significantly activated by a task for a group of subjects or to find out which parts of the brain have a significant difference in cerebral blood flow or metabolism between healthy volunteers and patients.[7] SPM offers a collection of non-interactive techniques that deal with (a) the spatial normalization of brain images into a standard stereotactic space to facilitate the anatomical accuracy of subsequent voxel-based analyses; (b) removal of the confounding effect of intersubject global metabolic differences or other confounding variables; (c) protection from false positives in statistical inference and (d) data-led assessment at a spatial resolution which is higher than that allowed by ROI analyses.[10]

Spatial normalization uses a two-step approach. Firstly the affine transformation that best maps the image to a template image in a standard space is calculated. In SPM, a rapid and automatic method for performing this coregistration is utilized which uses a Bayesian scheme to incorporate prior knowledge of the variability in the shape and size of heads. The affine transformations derived using the Bayesian scheme were shown to be more robust and more rapidly converging when compared to affine registrations without incorporating the prior knowledge.[54] Secondly, the nonlinear warps are described by a linear combination of low spatial frequency basis functions in order to reduce the number of parameters to be fitted. Whereas most approaches assume that the variance associated with each voxel is already known and that there is no covariance between neighbouring voxels, the SPM normalization approach attempts to estimate this variance from the data and corrects for the correlations between neighbouring voxels. This makes the approach suitable for the spatial normalization of both magnetic resonance images and low-resolution or noisy PET or SPECT images.[55,56]

After spatial normalization, the studies are convolved with a Gaussian smoothing kernel to correct for inter-individual differences in underlying brain structure and to be able to apply the general linear model approach for consecutive statistical analysis.[7] Newer versions of SPM use a modified statistical analysis approach, but the theoretical details were still in press at the time of writing this chapter.

Since spatial normalization in SPM is done onto the MNI (Montreal Neurological Institute) template, which is similar but not identical to the Talairach system,[57] a correction algorithm for the conversion of these data needs to be included, to enable defining activation locations in terms of Brodmann cytoarchitectonic areas (BAs) as defined on the Talairach atlas. Differences are largest in the inferior temporal cortex and a transformation between both systems has been published.[58] Web-based software is available for this conversion (the so-called Talairach Daemon (*http://ric.uthscsa.edu/projects/talairachdaemon.html*).

2.2.1.b Validation

Although increasingly used, validation and sensitivity determinations for PET and SPECT have been performed by a limited number of studies, in simulation and experimental designs. In simulation studies, activation spheres were created on a digital 3D Hoffman brain phantom,[59] which represents the perfusion pattern of the human brain. For selected combinations of study sizes and activation focus diameters, minimal activation levels required for detection were determined, by which theoretical guidelines for SPECT group activation studies could be put forward. It was demonstrated that the minimal activity difference for focus detection shows a marked decrease with both increasing focus diameter (also due to partial volume effects) and study size (due to loss of degrees of freedom

hence lower statistical power).[60] A similar approach has been described recently for PET studies using realistic Monte Carlo simulations to investigate design parameters.[61] In this study, fifteen data sets were generated for different foci activation levels, and localizations using five spheres located within the digital Hoffman 3D brain phantom with diameters corresponding to 6 and 8 mm, close to the resolution limit of the ECAT ART PET scanner (CTI/Siemens). The foci were defined in structures of particular interest in cognitive neuroscience. In addition, a coordinate search of the resulting maximum cluster was carried out, which was found to be close to 10 mm around the true value on the Talairach atlas.[57]

Because of the limitations of simulation studies, including absence of scatter effects, built-in perfect study alignment, and no optimization of the pre-processing steps (spatial smoothing) prior to statistical analysis, an experimental set-up which allows to incorporate realistic acquisition, physical corrections and reconstruction schemes was also investigated. A sub-resolution 'sandwich' or stacked phantom was constructed using the same high-resolution digital Hoffman phantom and a $^{99m}TcO_4^-$ ink mixture. The impact of activation characteristics (size, intensity and location), study size, physical degradation factors and reconstruction technique were studied. As for minimal activation detection, for a 15×15 design all 24 mm activations of 5 % intensity were detected for the fan-beam data using filtered back projection, with attenuation and scatter correction. Decreased focus or study size, lower SPECT resolution, absence of scatter and attenuation correction resulted in an increase in minimally detectable activation. For the clinical application of a single study referred to 15 control studies, only 24 mm - 10 % activation foci were detected at levels corrected for multiple comparisons.[62] As for experimental design predictions, from these studies it seems that for cognitive activation studies, where activations of only a few per cent may be expected, not only optimal experimental setup is required, but minimal sample sizes of at least 10 subjects are necessary.

Improvements in activation analysis procedures optimising the SPM approach are currently being investigated for fMRI and PET data. Anatomically Informed Basis Functions (AIBF) offer the feature of removing anatomical variability by isotropic stationary smoothing followed by anatomically informed spatial deconvolution using a canonical cortical surface.[63,64] The subsequent activation analyses are expected to be more sensitive. Improvement in detection sensitivity can be also obtained by applying an atlas-based adaptive smoothing method instead of the usual Gaussian filtering method.[65] Requiring only minimal assumptions for validity, nonparametric permutation testing provides a flexible and intuitive methodology for the statistical analysis of data from functional neuroimaging experiments, at some computational expense. The permutation approach readily accounts for the multiple comparison problem implicitly in the standard voxel-by-voxel framework. When the appropriate assumptions hold, the statistical non-parametric mapping gives results similar to those

obtained from a comparable SPM approach using a general linear model with multiple comparisons correction derived from random field theory. For analyses with low degrees of freedom, such as single subject PET/SPECT experiments or multi-subject PET/SPECT or fMRI designs assessed for population effects, the nonparametric approach employing a locally pooled (smoothed) variance estimate might outperform SPM approach, but this needs to be studied further. Although the theory and relative advantages of permutation approaches have been discussed by various authors, there has been no accessible explanation of the method, and only since recently there has been freely distributed software accessible.[66]

2.2.1.c Stereotactic Surface Projections

The University of Michigan with Minoshima and coworkers developed a method of projecting the cortical activity visualized in a 3-dimensional volume image onto the brain surface to create a surface representation of the cortical activity distribution.[35] This method has been combined with the previously developed standardization method, and the entire process is named 3-dimensional Stereotactic-Surface Projections (SSP-Neurostat). Anatomic standardization by 3-dimensional SSP is a combination of volume image standardization and surface projection. Many important contemporary PET and SPECT studies have been published using this method.[38,48,67-75]

Practically, surface projection has the advantage of erasing the radial mismatch between subjects in radioactivity distribution, which remains after standardization on volume images. Because of the partial-volume effect, the activity distribution within the cortical rim of healthy volunteers reflects the distribution of grey matter more than that of radioactivity per millilitre of grey matter. Therefore, surface projection erases the mismatch in gray matter distribution in the radial direction. Surface projection is also beneficial in reducing data size and forming an explanatory display for clinicians or nonprofessionals.[33]

The surface projection loses information about radial profiles within the cortical rim and deep structures. Surface projection is based on the notion that the cerebral cortex is essentially a 2-dimensional sheet of laminar structure bent and folded in 3-dimensional space and that, therefore, the radial profile information is trivial. This notion may be true for gyral areas but not for sulcal areas. On an area covering a sulcus, the algorithm of 3-dimensional SSP projects the maximum activity inside the sulcus. To strictly realize the notion of a 2-dimensional sheet, sulci must be unfolded or flattened, requiring far higher resolution than PET images provide, thus applicable only to fMRI.

2.2.2 ROI and VOI Approaches

At this time, manual ROI analyses are no longer of much clinical interest since a long time is required to draw accurately ROIs needed to cover an entire brain and the variability associated with the operator-dependent is

large compared to inter-individual physiological variability. The artificial boundaries from the ROI or VOI region map and the relatively large search volumes constitute a known inherent disadvantage of VOI techniques in the sense that they imply a preconception about the topography of the functional deficits and that the size of the VOI imposes a spatially smoothing effect.[76] Smaller focal defects can be observed by the voxel-based techniques and a brain region can be reported as abnormal even when only part of the underlying VOI was hypoperfused (dilution effect). Additional tools for subject-to-group statistical comparisons should be available with automatic VOI software and these should take into account the special problems of correction for multiple comparisons.[10,77].

However, if sufficient variance from the VOI measurements is included, the overall performance of automated stereotactic VOI-based analysis can be similar to that of the voxel-based analysis for the same discrimination task. In some studies, VOI-based analysis performed poorly at low false-positive fraction and was less tolerant to noise than the voxel-based analysis.[78] Another study found that under clinical conditions (in traumatic brain injury and cognitive impairment), classification of brain SPECT studies can greatly be aided by anatomic standardization techniques and that under the investigated circumstances, SPM was found to have a lower sensitivity than VOI or voxelwise region-growing techniques, at low false-positive fractions, in contrast to the former study.[79] It is expected that SPM will replace conventional VOI-based analysis in comparison studies of image correction and reconstruction techniques.[80]

2.2.3 Other Techniques

Several other quantification techniques exist to investigate the functional specialisation as well as the functional segregation and networking of the human brain. Because of their availability only in specific centres and relatively low general use, they are not further discussed here and the reader is referred to the relevant literature as indicated. Automated procedures which may improve diagnostic accuracy aside from atlas-based methods or template-based approaches as described above include the domain of artificial intelligence-based methods.[81] Also the implementation of neural networks for analysis of perfusion and receptor data may enhance clinical analysis of datasets as has been shown by several authors.[82-85]

Spatial activation patterns can also be investigated by ANCOVA-adjusted voxel-based principal component analysis (PCA) and PCA with a Scaled Subprofile Model (SSM)[86] where all obtainable information about regional ratios can, in theory, be derived from the SSM regional covariance patterns,[87] allowing the generation of disease-specific metabolic or perfusion signatures or covariance patterns.[88-91]

2.3 The Impact of Physical Correction Factors

Voxel-based or VOI analysis of emission tomography designs is intrinsically based on a large number of variables which all affect the final result to a greater or lesser extent. Acquisition and reconstruction options rely on isotope, acquisition time/injected activity, camera distance, collimator type for SPECT and randoms for PET, scatter and attenuation correction, reconstruction algorithm, pre- and/or post-filtering. Subsequent analysis is done after (intra-individual) coregistration, anatomical standardization (spatial normalization) and smoothing. Therefore, it is clear that a multiparameter optimization procedure is necessary to determine the optimal conditions for measuring defects or activations.[62,92]

As it has been shown extensively in some of the previous chapters in this book, physical factors degrade the actual image and thus the true activity obtained by SPECT or PET measurements.[93] Therefore, it is not surprising that to improve brain mapping, it is imperative to investigate the impact of correction techniques for physical degradation factors such as attenuation (see chapter 6), scatter (see chapter 7) and resolution loss or partial volume effect (see chapter 8). The reader is referred to these chapters for detailed description of methods and techniques and their impact on quantitative analysis of functional brain images.

3. Brain Imaging Using Normal Databases

3.1 PET and SPECT Databases of the Normal Brain

3.1.1 Concept

The combination of normal functional studies that are spatially normalized using the techniques outlined above can result in the creation of a probabilistic atlas, whereby to every voxel a mean and standard deviation of the functional ligand activity is assigned. Probabilistic atlases are more representative of population anatomy than single brain atlases and allow anatomical labelling of the results of group studies in stereotaxic space and automated anatomical labelling of individual brain imaging datasets.[50] Normal ranges may be determined for each voxel (or other appropriate region) from a representative normal sample using univariate analysis, obtaining mean and standard deviation images or multivariate analysis, which accounts also for normal patterns of variation (represented as principal components). For a new image, the variation from normality (in *SDs*) for each region may be determined.[94]

3.1.2. Normal Databases for PET and SPECT

^{99m}Tc-hexamethylene-propylene amine oxime (HMPAO) and ^{99m}Tc-ethyl cysteine dimmer (ECD) have been used for regional cerebral blood flow (rCBF) studies using SPECT. HMPAO-SPECT and ECD-SPECT images of normal individuals have slightly different perfusion patterns from rCBF-PET images, reflecting differences in the mechanism of accumulation of each agent in the brain. For clinical diagnoses, these patterns must be taken into consideration.[95]

^{18}F-fluorodeoxyglucose (FDG) is a specific glucose analogue that is trapped intracellularly since further metabolisation through hexokinase cannot occur. Physiological glucose consumption as in the brain and the heart, but also pathophysiological increase such as in tumours (through increased cell division and glycolisis) and infection/inflammation, forms the basis for different functional clinical imaging applications. The energy metabolism of the adult human brain is practically completely dependent on oxidative glucose burning. Glucose metabolism reflects neuronal activity and more specifically (pre)synaptic activity, where the energy is mostly needed for membrane potentials and the renewal of ionic gradients.[96] The rest of the energy is consumed by glial processes.

The functional coupling between regional cerebral blood flow and glucose metabolism has been demonstrated in many animal experimental and human studies in basal circumstances,[97] but is not always valid in pathophysiological conditions (e.g. acute stroke). Perfusion changes therefore also indicate changes in synaptic activity and cells, which can be caused by both excitation and inhibition. Since inhibition also requires higher energy consumption, a decreased activity in perfusion of metabolism thus relates to neuronal dysfunction or neuronal loss.[98]

By means of PET, perfusion can be determined with ^{15}O labelled substances (water and carbon dioxide), but this radionuclide has a very short half-life (2 minutes). Therefore, in the clinical situation predominantly SPECT with the ^{99m}Tc-labeled radioligands HMPAO or ECD is used to determine the regional blood flow (physical half-life ^{99m}Tc = 6 hours).[99]

An important *conditio sine qua non* for the application of brain mapping techniques in group and especially clinical studies is the availability of a well-characterized reference group.[100,101] In most studies, there is consensus that at least a minimal of variables needs to be controlled to define the state of normality. This includes history (e.g. exclusion of disorders of the central nervous system such as head trauma or epilepsy; of psychiatric disease in the first- or second-degree relatives, of substance abuse or use of psychoactive medication), in a fewer number also (para)clinical factors such as neurological, neuropsychological or psychiatric testing, internal heart, lung, liver or kidney disease, diabetes mellitus, the presence of abnormalities on CT or MR scans. Apart from these, numerous other covariates or confounders exist which influence the result of these analyses: age, gender, education,

race, smoking, caffeine use, (un)controlled hypertension, oestrogen status, etc. It is virtually impossible to construct a sufficiently general dataset which accounts for these variables with enough degrees of freedom left (and this for all radioligands). Therefore, it is of utmost importance for the progression of the field that sharing of well-defined normal data is advocated (see section 3.1.3).

Table 1 gives an overview of the larger normal databases (more than 20 subjects) and their characteristics as has been described in the literature the last 15 years. It can be seen that during the past years, resolution has improved, and more studies have included structural imaging.

Table 1. Normal perfusion SPECT and PET perfusion and metabolism databases in resting condition (n > 20).

Ligand	Resolution FWHM (mm)	# subjects (M/F)	Age range	Structural imaging	Reference
SPECT					
HMPAO	9–10	50 (25,25)	31 – 78	no	[106]
^{133}Xe	15	187 (136,51)	18 – 71	no	[107]
ECD	8	89 (43/46)	20 – 81	MRI	[49]
ECD	8	48 (22/26)	22 – 95	no	[108]
ECD	8–15	28 (14/14)	57 – 81	MRI	[109]
HMPAO	7–8	152 (67/ 85)	50 – 92	CT or MRI	[110]
HMPAO	11	27 (17/10)	26 – 71	no	[111]
HMPAO	(8–10)	44 (26/18)	20 – 73	CT or MRI (partial)	[112]
ECD	7	33 (15/18)	4–15 and 27–56	CT	[113]
HMPAO	16	68 (34/34)	19–45 and 65–79	CT	[114]
HMPAO	9.7	20 (11/9)	21–34 and 70–76	no	[115]
HMPAO	7.5	33 (20/13)	20 –76	no	[116]
HMPAO	10	53 (27/26)	21–83	CT	[117]
^{133}Xe	> 15	76 (38/38)	18–72	no	[118]
PET					
FDG	6	110	57 ± 13	CT or MRI	[119]
FDG	5	66 (38/28)	20–69	no	[120]
FDG	5	36 (21/ 15)	19–42 and 55–77	no	[121]
FDG + raclopride	6.5	37 (32/5)	24–86	no	[122]
$H_2^{15}O$	(8)	37 (21/16)	19–50	MRI	[123]
FDG	5.5	24 (15/9)	20–67	MRI	[124]
FDG	8	150 (72 / 78)	21–90	no	[89]
FDG	6	120 (55/65)	21–91	MRI (partial)	[125]
FDG	5.5	61 (37/24)	27.5 ± 7	MRI	[126]
FDG	5.5	120 (64/56)	19–79	no	[127]
$H_2^{15}O$	6	25 (14/11)	20–68	CT	[128]
$H_2^{15}O$	8.5	30 (15/15)	30–85	no	[129]
$C^{15}O_2$	16	34	22–82	no	[130]
FDG	15	76 (39/37)	21–84	MRI (partial)	[131]
FDG	17	49 (49/0)	21–83	CT	[132]

An important aspect of functional data analysis compared to normal data and in the construction of the normal database itself is the choice of the reference region. This problem has been dealt with by various authors in specific studies, indicating the need to use a standard reference region depending on the pathology.[102,103] Whereas normalization of physiological measures can significantly reduce intersubject variation, interpretation of normalized results can be more complex.[104] In SPM, usually normalization is carried out on the total brain activity implicitly by using a proportional scaling approach. Whereas this may not be the optimal approach for specific disease entities and large defects or activations, it is generally conceived as the working solution for the wide range of clinical applications.

Apart from indirect cerebral parameters such as metabolism and perfusion, nuclear techniques can also offer the unique possibility to image a large set of receptors. Nerve signals pass through synapse between different neurons. This process is a chemical event in which neurotransmitters from the presynaptic nerve terminals are released to the synaptic cleft, and act on the postsynaptic neurotransmitters to induce excitation or inhibition on the target neuron. Through reuptake sites (transporter) on the presynaptic membrane, the surplus in neurotransmitters can be recycled. Changes of important neurotransmitter systems such as the dopamine, GABA (gamma aminobutyric acid), benzodiazepine, choline and serotonin neurotransmitter systems have been demonstrated in many neurological and psychiatric disease stages. For a detailed overview, the reader is referred to Tikofsky *et al.*[105]

Although currently the clinical indications for receptor imaging are limited to predominantly movement disorders (see section 3.2.4), they do show potential for clinical dose titration of medication, early diagnosis of neuropsychiatry disorders, as well as more specific localization of cerebral functional disturbances. Specific normal databases where the brain mapping approach has been described have been given in section 2.1.2.

3.1.3 Interdepartmental Transfer and Sharing of Databases

In view of the high number of potential covariates, widespread use of normal, anatomically standardized databases for comparison of pathological states, would largely benefit from transferability of these data between cameras hence multiple centres using different apparatus. A concept can be described to perform such transfer, based on intrinsic known SPECT or PET camera properties. There are a few studies that have validated this for human data. One study aimed at investigating transfer of three normal perfusion databases obtained in the same large population of healthy volunteers that underwent sequential scanning in 3 multi-head gamma cameras with different resolution. The highest reduction in between-camera variability was achieved by resolution adjustment in combination with a 3D Hoffman brain phantom-based correction.[133] The difficulties in the inverse

transfer of absolute metabolic PET data from lower to higher resolution was reported by Grady et al. as early as in 1989, with data in 9 subjects of relatively low resolution, different attenuation and operator dependent ROIs.[134]

Intrinsically, the same inference could be made for receptor studies, although the actual outcome measures (e.g. regional versus aspecific uptake) may have to be verified both by phantom data and on the specific equipment used. Longitudinal studies were evaluated on two PET tomographs in a FDOPA study. Uniform resolution correction resulted in comparable data.[135]

In other nuclear imaging domains, such as in nuclear cardiology where database-driven algorithms are being used in clinical practice, it was only until recently that a full methodological validation study was published on comparison of normal PET data concluding that after resolution adjustment cardiac studies from multiple PET systems can be pooled for statistical analysis.[136]

Moreover, studies that have been performed over several years with a change from old tomographs to new machines with better sensitivity and resolution can be made more comparable by such an approach. Matching semi-quantitative results obtained with different tomographs through an automated analysis procedure, would also be an effective way to improve interdepartmental collaboration and enables a better appreciation of results published by different groups.

Together with the issue of database transfer comes database sharing. The issues concerning data sharing for neuroimaging databases have been addressed by the Organization for Human Brain Mapping.[137] Whereas potential benefits have been realized in other fields such as genomics, there are important factors that differentiate brain imaging from other fields, including the rapid pace of change in brain imaging technologies, the complexity of variables that must be specified (method of acquisition, behavioral design, subject characteristics) and concerns about participant confidentiality for structural data. At the rate of the number of brain studies conducted each year and the data size, especially in fMRI and MRI, data sharing is an enormous project. Data contents, metadata descriptors, data import and export procedures, data quality issues, access, ownership, credit, confidentiality are all but easily solved parameters. Specifically for SPECT and PET data, the construction and quality control of the database through peer-reviewed literature or independent data analysis is necessary and maintenance is of crucial importance for which industrial support might be mandatory.

3.2 Pathophysiological and Clinical Applications

It is impossible to shortly summarize the vast literature of PET and SPECT of the brain with respect to the brain mapping approach, but a number of relevant studies will be discussed here with a specific aim at characterizing the potential clinical applications where one subject is compared to a suitable reference/normal group.

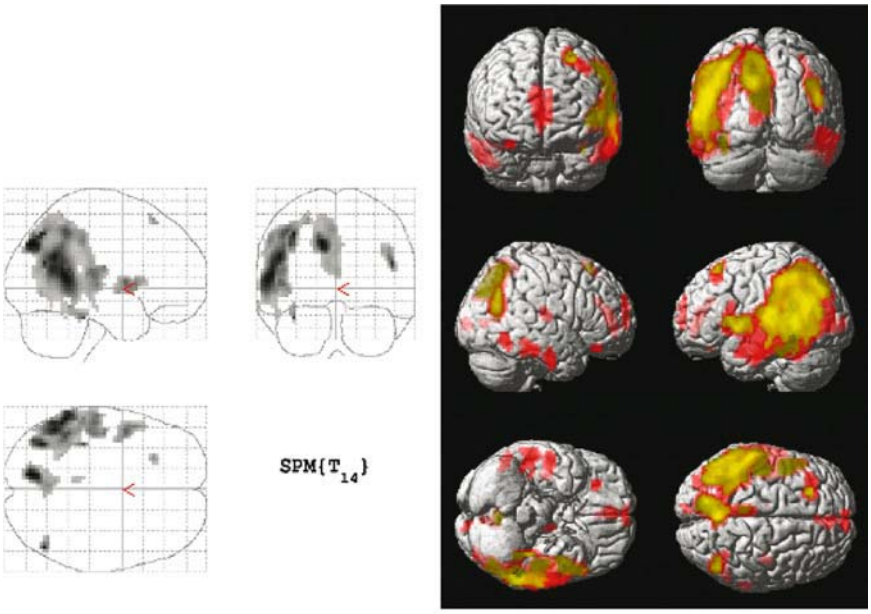
3.2.1 Dementia

Dementia is an increasing problem for which diagnosis still is based on clinical and exclusion criteria. Where CT and MRI only show atrophy and ventricular enlargement, functional imaging reflects more specific changes. The different entities such as dementia of Alzheimer's type (DAT) (40 to 60% of all senile dementias), Lewy Body dementia, multi-infarct dementia and frontotemporal dementia, can show the same initial symptoms, nevertheless therapy and prognosis will vary considerably according to the exact diagnosis and clearly different metabolic or perfusion patterns are observed.[138,139]

In very early DAT, it has been shown by automated analysis techniques that flow or metabolism reduces first in the posterior cingulate gyrus and precuneus.[71] This reduction may arise from functional deafferentation caused by primary neural degeneration in the remote area of the entorhinal cortex that is the first to be pathologically affected in DAT. Then medial temporal structures and parietotemporal association cortex show flow or metabolic reduction as disease progresses (Figure 3). The reason why flow or metabolism in medial temporal structures shows delay in starting to reduce in spite of the earliest pathological affection remains to be elucidated. Reduction of flow or metabolism in the posterior cingulate gyrus and precuneus has been reported even in presymptomatic nondemented subjects who were cognitively normal and had at least a single epsilon4 allele. PET and SPECT imaging has become to be useful for assessing therapeutic effects of newly introduced treatments for DAT where investigations observed significant regional flow increase after donepezil treatment.[140] Most of these observations have been made by applying computer assisted analysis of three-dimensional stereotactic surface projection or statistical parametric mapping instead of conventional regions of interest technique.

Histopathological data in DAT correlate to these patterns as well as *in vivo* diagnostics.[141] Currently, ^{18}FDG-PET has the highest diagnostic accuracy (up to 96%) for DAT[142,143] and is cost-effective.[144] This is the main reason why reimbursement for this indication is likely in the near future. The largest voxel-based study so far in DAT comprising 110 normal controls and 395 patients with probable DAT is a multicenter (eight centres) trial. The indicator of metabolic abnormality in DAT-related regions provided 93% sensitivity and specificity for distinction of mild to moderate probable DAT from normals, and 84% sensitivity at 93% specificity for detection of very mild probable DAT. All regions related to DAT severity were already affected in very mild DAT, suggesting that all vulnerable areas are already affected to a similar degree at disease onset.[69,119]

Somewhat lower although comparable values are found for high resolution SPECT.[145] Correspondence between ^{18}F-FDG PET and HMPAO SPECT is limited to the main finding of temporoparietal and posterior cingulate functional impairment in mild to moderate DAT. The distinction between healthy volunteers and patients is less sensitive to threshold selection with PET than



Statistics: *volume summary (p-values corrected for entire volume)*

set-level		cluster-level			voxel-level				x, y, z [mm]
p	c	$p_{corrected}$	k_E	$p_{uncorrected}$	$p_{corrected}$	T	(Z_δ)	$p_{uncorrected}$	
0.003	7	0.000	316	0.000	0.016	8.97	(5.09)	0.000	-12 -78 36
					0.759	5.19	(3.82)	0.000	-3 -43 48
					0.856	4.91	(3.69)	0.000	0 -43 15
		0.000	1159	0.000	0.026	8.50	(4.97)	0.000	-57 -48 9
					0.037	8.15	(4.87)	0.000	-48 -57 27
					0.050	7.88	(4.79)	0.000	-36 -69 45
		0.090	50	0.011	0.098	7.26	(4.60)	0.000	48 -66 21
		0.522	20	0.087	0.324	6.16	(4.22)	0.000	-36 -33 -24
					0.791	5.08	(3.76)	0.000	-36 -45 -15
		0.007	100	0.001	0.339	6.11	(4.20)	0.000	-48 3 6
					0.570	5.55	(3.97)	0.000	-51 12 3
					0.709	5.26	(3.84)	0.000	-42 0 0

FIGURE 3. SPM analysis of moderate dementia of Alzheimer's type with ^{99mTc}-ECD SPECT. Comparison of a subject with 15 age-matched healthy volunteers. Upper left: glass brain representation of significant decreases in perfusion. Upper right: 3D rendering on a template MR image, the decreases in perfusion are shown for two thresholds $p_{height} = 0.01$ (red) and $p_{height} = 0.001$ (yellow). Bottom: statistics table output.

with SPECT, and findings in the frontal, temporobasal, and temporomesial cortices and in the cerebellum may differ between the 2 techniques.[146] The frontal regions of the brain, in addition to parietal and temporal lobes, may show reduced CBF in patients with DAT even at an early stage of dementia. The reduced rCBF in the cortical regions cannot be explained entirely by advanced atrophy and fast aging process.[147]

In studies of DAT, the cerebellum or pons is often used as a reference region, assuming that it is spared any major pathological involvement. SPM may enhance the evaluation of SPECT scans in DAT patients. An

easily applied, objective, and reproducible method for determining average cerebellar tracer uptake has been made that can be scaled specifically to cerebellar activity prior to the performance of SPM analysis. Cerebellar normalization produced more extensive abnormalities in SPM analyses of DAT patients than global normalization. Furthermore, cerebellar or pons normalization produced marginally more accurate diagnostic results in single-scan SPM analysis of DAT patients than did global normalization.[68,148]

3.2.2 Epilepsy

About 0,5 to 1% of the population suffers from epilepsy. Imaging is important for the determination of epileptogenic foci in a surgical setting for patients that have become refractory to drug treatment. Non-invasive imaging allows limitation of the proportion of patients that need invasive EEG-monitoring for seizure focus localisation.

Focal epilepsy is accompanied with moderate to severe metabolic and perfusion related disturbances. Both SPECT and PET can demonstrate these abnormalities, even in the absence of anatomical lesions such as hippocampal gliosis. ^{18}FDG-PET shows an interictal hypometabolism which is more extensive than the epileptic focus, probably due to distance effects by diaschisis.[142] It has a sensitivity of 70 to 85% and a specificity of 85% for lateralisation of the epileptic focus in temporal lobe epilepsy. In extratemporal lobe epilepsy, these figures are much lower (< 50 %). However, using voxel-based techniques and 3D-SSP increases up to 70 % sensitivity were demonstrated.[149,150]

Due to the long physical half life of ^{99m}Tc and the "frozen image" properties of the ^{99m}Tc-labeled SPECT perfusion ligands, ictal or early postictal injection can provide a diagnostic perfusion image. Brain perfusion shows a sharp (up to 300%) rise during an epileptic seizure at the position of the epileptogenic focus, while interictally most epileptic foci show a diminished perfusion. By means of ictal/interictal subtraction studies, with subsequent coregistration onto MRI (SISCOM: Subtraction Ictal SPECT Coregistered to MRI), a predictive value up to 97% for the correct localization of an epileptic focus can be obtained,[151] this is higher than any other modality. SPM has been validated for localizing seizure onset. There is evidence supporting SPM difference analysis in assessing regions of significant CBF change from baseline in concordance with the clinically used technique of SPECT ictal-interictal difference imaging in epilepsy patients. Difference analysis using SPM could serve as a useful diagnostic tool in the evaluation of seizure focus in temporal lobe epilepsy.[152] With these techniques, it has been stated that interictal SPET could be omitted if a normal brain SPET database were to be established which would reduce the medical cost of seizure localization.[153]

3.2.3 Cerebrovascular Disease

In more than 80% of the patients, the diagnosis of acute stroke can be made clinically without paraclinical investigations. In some of the remaining patients, CT or MRI can define the etiology of the symptoms such as in brain tumours, whereby stroke is put as exclusion diagnosis. However, during the first days of the clinical event, it is frequently hard to estimate and investigate the size of the stroke by means of only structural imaging. Functional techniques correlate more accurately to the size of the lesions and provide important information for planning and rehabilitation, as well as prognosis.[154]

Recent studies have demonstrated that in acute ischemic stroke within 6 hours after the symptomatic start, perfusion SPECT is predictive to outcome (positive predictive value of 92%) and provides potential to risk stratification on treatment complications in the decision to thrombolysis.[155] Cerebral perfusion imaging can be used to diagnose stroke immediately and define other areas in the brain at risk due to decreased perfusion. Absolute physiological quantification is important as regions with a residual regional perfusion above 35 ml/100g/min show a low probability for infarction without need for recanalisation (local intra-arterial thrombolysis). Regions where this perfusion has been maintained to 25 ml/100g/min could be repaired by early recanalisation. However for regions where the perfusion has been lowered below 20 ml/100g/min, a severe risk of intracerebral haemorrhage and oedema is present.[156]

Apart from perfusion SPECT, the combination diffusion/perfusion MRI or perfusion CT are of increasing importance in the diagnosis of hyperacute ischemic stroke and identification of tissue viability.[157] By means of currently progressing quantitative studies, the role and prognostic value of these different modalities is being determined further.

Of the hemorrhagic strokes (20 % of all strokes), subarachnoid haemorrhage and rupture of aneurysms can be treated surgically by clipping or obliteration to reduce the risk of rebleeding, which is the most lethal event in the patient follow-up. The perfusion effect of cerebral vasospasm after subarachnoidal haemorrhage can be investigated with SPECT by repeated investigations. When SPECT shows a new perfusion decrease, more aggressive therapies such as end vascular or hemodynamic therapy can be necessary.[158]

In patients with transient ischemic attacks (TIAs), a pharmacological vasodilatory intervention with acetazolamide, can show a relative hypoperfusion of brain areas distal to a stenotic lesion with a high sensitivity. Acetazolamide results in arterial or arteriolar vasodilatation of the cerebrovascular structures. Since a compensatory maximal vasodilatation is already present in these cases, the perfusion distally of the stenoses will remain relatively normal only where sufficient collateral circulation exists. In this way, the patient selection for revascularisation can be made objectively since

these patients with diagnostic relative hypoperfusion after acetazolamide administration are potentially the best candidates.[159] SPECT after the administration of acetazolamide can also image the cerebrovascular reserve carotid stenosis. An abnormal perfusion SPECT is predictive for a poor collateral circulation and can aid in the decision for endarterectomy and to identify risk patients, as well as to detect bilateral carotid stenoses and play a role in the follow-up.[160] Objective and quantitative evaluation of abnormalities of CBF and CVR with 3D-SSP identifies patients at risk for postoperative hyperperfusion with high predictive value.[75]

3.2.4 Movement Disorders

Idiopathic Parkinson's disease (IPD) is the most prevalent neurodegenerative disease after dementia, with an estimated prevalence of 150 per 100000 and incidence of 20 per 100000. The typical clinical symptoms such as rigidity, hypokinesia, tremor and postural instability, are associated to the degeneration of dopaminergic neurons in the substantia nigra of the brainstem. Although IPD is the most prevalent cause of parkinsonism, there are other diseases that present with one or more similar symptoms. From recent studies it has been shown that of the patients that deceased under the diagnosis IPD, up to 25% actually did not have the disease.[161] Various radioligands are currently commercially available to evaluate the pre- and postsynaptic status of the dopaminergic system.[162] Both ^{123}I labelled iodobenzamide (IBZM) and epidepride, as well as ^{11}C-raclopride can be used as a sensitive differentiator between IPD and other parkinsonian syndromes, and are predictive for the response on dopaminergic therapy.[163] In Parkinson's disease, the postsynaptic system is normal to upregulated, while in the most Parkinsonian syndromes, of which multi system atrophy (MSA) and progressive supranuclear palsy (PSP) are the most common, there is a decreased postsynaptic binding. The cocaine-analogue ^{123}I-FP-CIT (ioflupane) has a clinical role in the early diagnosis of Parkinson's disease (since abnormalities are present even 5 years before clinical symptomatology), differentiation between IPD and essential tremor, as well as in the differentiation between Alzheimer and Lewy body dementia.[164] This tracer is very sensitive for presynaptic damage to the striatum and the uptake in the contralateral striatum is correlated to the severity of clinical disease.[165]

Despite controversial clinicopathological distinctions between Parkinson's disease with dementia (PDD) and DAT, similar patterns of metabolic reduction in the posterior brain were reported previously using PET with ^{18}F-FDG. More specific regional differences were measured between PDD and DAT using accurate and objective brain mapping techniques. PDD showed greater metabolic reduction in the visual cortex and relatively preserved metabolism in the medial temporal cortex.[72]

Metabolic or perfusion brain imaging is useful in the differential diagnosis of extrapyramidal disorders due to the application of anatomical standard-

ization.[166,167] Progressive supranuclear palsy (PSP) and corticobasal degeneration (CBD) are neurodegenerative disorders that may be accompanied by dementia and Parkinsonism as clinical symptoms. Measurement of glucose metabolism by PET and a voxel-based analysis is useful to understand the pathophysiology and differentiate IPD and corticobasal degeneration with cognitive impairments.[168]

3.2.5 Brain Tumours

Structural imaging modalities are a crucial part of the diagnostic work up in the rather heterogenous group of brain tumours. A major question in the evaluation of patients with brain tumours is the grading of the tumour (related to prognosis), extension in surrounding tissue and tumour heterogeneity. SPECT radiopharmaceuticals, such as ^{201}TI or ^{99mTc}-MIBI, but in particular ^{123}I-IMT, are slightly correlated to grading of the tumour and allow some differentiation.[169] Also the PET ligands ^{18}F-FDG, ^{11}C-methionine, ^{18}F-FET, allow similar evaluation and allow a sensitive differentiation between necrosis or fibrosis and tumour recurrence in postoperative or post-radiotherapeutical setting. These properties contrast to the lack of specificity with gadolinium-enhanced MRI which necessitates blood-brain barrier breakage and overestimates viability.[170] Image fusion, as described above, is highly recommended in tumours with heterogeneous uptake and has proven synergistic value. It is worth emphasizing that Statistical parametric mapping has not yet been fully evaluated in the study of brain tumours for individual cases.

3.2.6 Traumatic brain injury

Traumatic brain injury has its highest incidence in the younger population and is the most prevalent cause of cognitive disturbances in this age group. The sequels of a closed cerebral trauma frequently occur as inability to concentrate, loss of attention or changes in activity on the long term. With functional imaging, earlier and more lesions are visualized than is with structural imaging. Also in the absence of structural abnormalities, metabolism and perfusion can be decreased, and many of the functional deficits show a good correlation with neuropsychiatric scores.[171] The literature findings are not only indicative for rehabilitation potential, but also show a high negative predictive value.[172] Therefore, as rule of thumb it can be stated that a negative perfusion SPECT investigation virtually excludes functional sequels. SPM has a role in SPECT image interpretation because it allows better visualization than other methods of quantitative analysis of the spatial distribution of abnormalities in focal and diffuse head injury. Frontal lobe blood flow abnormality (particularly anterofrontal regions and mesiofrontal areas) is common after head injury.[173]

3.2.7 Psychiatry

Current clinical indications for nuclear imaging within the domain of psychiatry are less clearly defined and as yet not clearly evidence based. In contrast to the recommendations by neurological societies, neither the European nor the American psychiatric societies have yet defined guidelines for the role of brain PET and SPECT in psychiatry clinics. This is mostly due to inconsistencies in current literature, caused by small populations, variations in age and sex distributions, diagnostic heterogeneity, different severity of disease, study circumstances and medication status. Moreover, many major mental diseases are complex entities, and probably due to multiple etiological factors, which have influence on various and distinct regions in the human brain.

The diagnosis of depression can be severe, especially in the elderly populations and in cases of possible associated dementia. Apart from a consistent global perfusion reduction, the most classical pattern observed in bipolar phasic disturbances (major depression) is a hyperperfusion of the frontal lobes during the manic phase and a hypoperfusion during the depressive phase.[174] Functional perfusion abnormalities mostly have a symmetrical character and improve with treatment.[175] In schizophrenia, a disease with substantial morbidity and mortality, already in the seventies the negative symptoms have been connected to frontal hypoperfusion. Current research in this area suggests that this is a disease of spread brain circuits that are involved in screening, integration and co-ordination of a large number of mental tasks.[176] In obsessive-compulsive syndrome, the most consistently reported deviations are an increased activity orbitofrontal and in the anterior cingulus gyrus, during the symptomatic condition.[177]

It can be expected that by the improved, automated and objective techniques and analysis methods, as well as by better judgement of the higher mentioned limitations in study design, functional image research within psychiatry will increase, as can be observed from the strongly growing psychiatric and nuclear medicine professional literature in the domain of biological psychiatry.[178-180]

The results of PET and SPECT brain mapping studies on personality dimensions indicate the existence of a relationship between behavioral and neurobiological factors and support the concept that the variability of PET data may be explained by neurochemical differences related to the prevalence of specific personality traits.[181,182]

4. Concluding Remarks and Future Prospects

It is highly likely that automated operator independent whole-brain analysis of high-resolution functional emission images will lead to largely reduced analysis variability, improved availability as well as increased diagnostic

sensitivity in pathophysiological cases, independent of rater performance. It should also be stressed, however, that in the first place 'clinical diagnosis' cannot be easily automated. Whereas automated image analysis methods are important, they should be regarded as complementary tools to diagnosis. Final interpretation of changes seen in a SPECT or PET scan, as in any other instrumental examination, needs a vast repertoire of knowledge, involving physiology, pathophysiology, and clinical science, which is hard to incorporate in a 'simple' decision making algorithm or general computer expert system.

Quantitative nuclear imaging of the brain will continue to play an important role in the development of a common basis and language for psychiatry, neurology and psychology.[183,184] Apart from imaging of indirect markers (perfusion and metabolism), and molecular markers such as membrane receptors (nicotine, glutamate, adenosine . . .), new radiopharmaceutical developments are also focused to suitable tracers for postreceptor signal transduction second messenger systems and gene expression, in parallel with current molecular biological progression. The triad molecular neuroscience, genetics and imaging will probably expand the understanding of the pathophysiology within the psychiatric and neurological clinical diseases in the coming decennia, hopefully leading to improved diagnosis and treatment, and eventually to the implementation of preventive techniques.

Acknowledgments. KVL is supported by a Clinical Research Grant of the Leuven University Hospital. HZ is supported by the Swiss National Science Foundation under grant SNSF 3152A0-102143. The authors would like to thank Dr Marie-Louise Montandon, Geneva University Hospital, for proofreading of this chapter and adding comments and suggestions.

References

1. Gilman S., Imaging the brain. *New Engl J Med* **338:** 812-820 (1998).
2. Mazziota J. C., Toga, A. W. and Frackowiak, R. S., *Brain Mapping: The Disorders*, Academic Press, San Diego, (2000).
3. Costa D. C., Pilowsky, L. S. and Ell, P. J., Nuclear medicine in neurology and psychiatry. *Lancet* **354:** 1107-1111 (1999).
4. Kouris K., Clarke, G. A., Jarritt, P. H. *et al*., Physical performance evaluation of the Toshiba GCA-9300A triple-headed system. *J.Nucl Med.* **34:** 1778-1789 (1993).
5. Wienhard K., Schmand, M., Casey, M. E. *et al*., The ECAT HRRT: performance and first clinical application of the new high resolution research tomograph. *IEEE Trans Nucl Sci* **49:** 104-110 (2002).
6. Braem A., Chesi, E., Joram, C. *et al*., "Novel design of high-resolution parallax-free Compton enhanced PET scanner dedicated to brain research" *Conf. Proc*. of the First International meeting on Applied Physics, 15-18th October 2003, Badajoz, Spain, Vol. *pp* 86-87 (2003).

7. Friston K., Holmes, A., Worsley, K. *et al*., Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping* **2:** 189-210 (1995).
8. Kung H. F., Kung, M. P. and Choi, S. R., Radiopharmaceuticals for single-photon emission computed tomography brain imaging. *Semin Nucl Med* **33:** 2-13 (2003).
9. Herzog H., Basic ideas and principles for quantifying regional blood flow with nuclear medical techniques. *Nuklearmedizin* **35:** 181-185 (1996).
10. Signorini M., Paulesu, E., Friston, K. *et al*., Rapid assessment of regional cerebral metabolic abnormalities in single subjects with quantitative and non-quantitative [^{18}F]FDG PET: A clinical validation of statistical parametric mapping. *Neuroimage*. **9:** 63-80 (1999).
11. Alavi A., Smith, R. and Duncan, D., What are the sources of error in measuring and calculating cerebral metabolic rates with fluorine-18-fluorodeoxyglucose and PET? [editorial; comment]. *J Nucl Med* **35:** 1466-1470 (1994).
12. Lassen N. A., Cerebral blood flow quantitation in clinical routine studies: how far have we now come? *J Nucl Med* **36:** 2343-2344 (1995).
13. Rosenthal M. S., Cullom, J., Hawkins, W. *et al*., Quantitative SPECT imaging: a review and recommendations by the Focus Committee of the Society of Nuclear Medicine Computer and Instrumentation Council. *J Nucl Med* **36:** 1489-1513 (1995).
14. Zaidi H., Quantitative SPECT: Recent developments in detector response, attenuation and scatter correction techniques. *Phys Med* **12:** 101-117 (1996).
15. Heiss W. D., Graf, R., Grond, M. *et al*., Quantitative neuroimaging for the evaluation of the effect of stroke treatment. *Cerebrovasc Dis* **8 Suppl 2:** 23-29 (1998).
16. Meyer M. A., Quantitative brain SPECT and the NIH Stroke Scale. *Stroke* **29:** 1480 (1998).
17. Venneri A., Shanks, M. F., Staff, R. T. *et al*., Evaluating improvements in rCBF in patients with Alzheimer's disease treated with rivastigmine using SPM. *Neuroimage* **13:** S1112 (2001).
18. Nelson A. D., Miraldi, F., Muzic, R. F., Jr. *et al*., Noninvasive arterial monitor for quantitative oxygen-15-water blood flow studies. *J Nucl Med* **34:** 1000-1006 (1993).
19. Takikawa S., Dhawan, V., Spetsieris, P. *et al*., Noninvasive quantitative fluorodeoxyglucose PET studies with an estimated input function derived from a population-based arterial blood curve. *Radiology* **188:** 131-136 (1993).
20. Watabe H., Itoh, M., Cunningham, V. *et al*., Noninvasive quantification of rCBF using positron emission tomography. *J Cereb Blood Flow Metab* **16:** 311-319 (1996).
21. Eberl S., Anayat, A. R., Fulton, R. R. *et al*., Evaluation of two population-based input functions for quantitative neurological FDG PET studies. *Eur J Nucl Med* **24:** 299-304 (1997).
22. Shiozaki T., Sadato, N., Senda, M. *et al*., Noninvasive estimation of FDG input function for quantification of cerebral metabolic rate of glucose: optimization and multicenter evaluation. *J Nucl Med* **41:** 1612-1618 (2000).
23. Chen K., Bandy, D., Reiman, E. *et al*., Noninvasive quantification of the cerebral metabolic rate for glucose using positron emission tomography, 18F-fluoro-2-

deoxyglucose, the Patlak method, and an image-derived input function. *J Cereb Blood Flow Metab* **18:** 716-723 (1998).
24. Iida H., Miura, S., Shoji, Y. *et al*., Noninvasive quantitation of cerebral blood flow using oxygen-15-water and a dual-PET system. *J Nucl Med* **39:** 1789-1798 (1998).
25. Matsuda H., Higashi, S., Tsuji, S. *et al*., A new noninvasive quantitative assessment of cerebral blood flow using N-isopropyl-(iodine 123)p-iodoamphetamine. *Am J Physiol Imaging* **2:** 49-55 (1987).
26. Matsuda H., Tsuiji, S., Shuke, N. *et al*., A quantitative approach to technetium-99m hexamethylpropylene amine oxime. *Eur J Nucl Med* **19:** 195-200 (1992).
27. Iida H., Itoh, H., Nakazawa, M. *et al*., Quantitative mapping of regional cerebral blood flow using iodine-123-IMP and SPECT. *J Nucl Med* **35:** 2019-2030 (1994).
28. Odano I., Ohkubo, M. and Yokoi, T., Noninvasive quantification of cerebral blood flow using 99mTc-ECD and SPECT. *J Nucl Med* **40:** 1737-1744 (1999).
29. Van Laere K., Dumont, F., Koole, M. *et al*., Non-invasive methods for absolute cerebral blood flow measurement using 99mTc-ECD: a study in healthy volunteers. *Eur J Nucl Med* **28:** 862-872 (2001).
30. Yang Y., Frank, J. A., Hou, L. *et al*., Multislice imaging of quantitative cerebral perfusion with pulsed arterial spin labeling. *Magn Reson Med* **39:** 825-832 (1998).
31. Hennig J., Speck, O., Koch, M. A. *et al*., Functional magnetic resonance imaging: a review of methodological aspects and clinical applications. *J Magn Reson Imaging* **18:** 1-15 (2003).
32. Bohm C., Greitz, T. and Eriksson, L., A computerized adjustable brain atlas. *Eur J Nucl Med* **15:** 687-689 (1989).
33. Senda M., Anatomic standardization, although controversial, finds yet another application. *J Nucl Med* **41:** 1888-1891 (2000).
34. Roland P. E. and Zilles, K., Brain atlases–a new research tool. *Trends Neurosci* **17:** 458-467 (1994).
35. Minoshima S., Koeppe, R. A., Frey, K. A. *et al*., Anatomic standardization: linear scaling and nonlinear warping of functional brain images. *J Nucl Med* **35:** 1528-1537 (1994).
36. Sugiura M., Kawashima, R., Sadato, N. *et al*., Anatomic validation of spatial normalization methods for PET. *J Nucl Med* **40:** 317-322 (1999).
37. Van Laere K., Koole, M., D'Asseler, Y. *et al*., Automated stereotactic standardization of brain SPECT receptor data using single-photon transmission images. *J Nucl Med* **42:** 361-375 (2001).
38. Minoshima S., Koeppe, R. A., Frey, K. A. *et al*., Stereotactic PET atlas of the human brain: aid for visual interpretation of functional brain images. *J Nucl Med* **35:** 949-954 (1994).
39. Friston K. J., Ashburner, J., Frith, C. D. *et al*., Spatial registration and normalisation of images. *Human Brain Map* **2:** 165-189 (1995).
40. Maes F., Collignon, A., Vandermeulen, D. *et al*., Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* **16:** 187-198 (1997).
41. Brett M., Leff, A. P., Rorden, C. *et al*., Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* **14:** 486-500 (2001).
42. Stamatakis E. A., Wilson, J. T. and Wyper, D. J., Using lesion masking to facilitate nonlinear alignment of SPECT images. *NeuroImage* **13:** S42 (2001).

43. Studholme C., Hill, D. L. and Hawkes, D. J., Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Med Phys* **24:** 25-35 (1997).
44. Rabiner E. A., Messa, C., Sargent, P. A. *et al.*, A database of [(11)C]WAY-100635 binding to 5-HT(1A) receptors in normal male volunteers: normative data and relationship to methodological, demographic, physiological, and behavioral variables. *Neuroimage* **15:** 620-632 (2002).
45. Ito H., Okubo, Y., Halldin, C. *et al.*, Mapping of central D2 dopamine receptors in man using [11C] raclopride: PET with anatomic standardization technique. *NeuroImage* **9:** 235-242 (1999).
46. Meyer J. H., Gunn, R. N., Myers, R. *et al.*, Assessment of spatial normalization of PET ligand images using ligand- specific templates. *Neuroimage* **9:** 545-553 (1999).
47. Okubo Y., Olsson, H., Ito, H. *et al.*, PET mapping of extrastriatal D2-like dopamine receptors in the human brain using an anatomic standardization technique and [11C]FLB 457. *Neuroimage* **10:** 666-674 (1999).
48. Frey K. A., Minoshima, S., Koeppe, R. A. *et al.*, Stereotaxic summation analysis of human cerebral benzodiazepine binding maps. *J Cereb.Blood Flow Metab* **16:** 409-417 (1996).
49. Van Laere K., Versijpt, J., Audenaert, K. *et al.*, 99mTc-ECD brain perfusion SPET: variability, asymmetry and effects of age and gender in healthy adults. *Eur J Nucl Med* **28:** 873-887 (2001).
50. Hammers A., Allom, R., Koepp, M. J. *et al.*, Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Hum Brain Mapp* **19:** 224-247 (2003).
51. Zaidi H., M-L, M. and Slosman, D. O., Magnetic resonance imaging-guided attenuation correction in 3D brain positron emission tomography. *Med Phys* **30:** 937-948 (2003).
52. Friston K. J., Holmes, A., Poline, J. B. *et al.*, Detecting activations in PET and fMRI: levels of inference and power. *Neuroimage* **4:** 223-235 (1996).
53. Ashburner J. and Friston, K., Voxel-based morphometry: the methods. *Neuroimage* **11:** 805-821 (2000).
54. Ashburner J., Neelin, P., Collins, D. L. *et al.*, Incorporating prior knowledge into image registration. *Neuroimage* **6:** 344-352 (1997).
55. Ashburner J. and Friston, K. J., Nonlinear spatial normalization using basis functions. *Hum Brain Mapp* **7:** 254-266 (1999).
56. Salmond C. H., Ashburner, J., Vargha-Khadem, F. *et al.*, The precision of anatomical normalization in the medial temporal lobe using spatial basis functions. *Neuroimage* **17:** 507-512 (2002).
57. Talairach J. and Tournoux, P., *Co-planar stereotactic atlas of the human brain*, Thieme Medical Publishers, Stuttgart, (1988).
58. Brett M., Christoff, K., Cusack, R. *et al.*, Using the Talairach atlas with the MNI template. *Neuroimage* **13:** S85 (2001).
59. Hoffman E. J., Cutler, P. D., Guerrero, T. M. *et al.*, Assessment of accuracy of PET utilizing a 3-D phantom to simulate the activity distribution of [18F]fluorodeoxyglucose uptake in the human brain. *J Cereb Blood Flow Metab* **11:** A17-25 (1991).

60. Lahorte P., Vandenberghe, S., Van Laere, K. *et al*., Assessing the performance of SPM analyses of SPECT neuroactivation studies. *Neuroimage* **12:** 757-764 (2000).
61. Schoenahl F., Montandon, M.-L., Slosman, D. *et al*., ''Assessment of the performance of SPM analysis in PET neuroactivation studies: A Monte Carlo investigation.'' *Conf. Proc*. of the VIIth International Meeting on Fully Three-dimensional Image Reconstruction in Radiology and Nuclear Medicine, 29 June-4 July 2003, Saint-Malo, France. Available on CDROM.
62. Van Laere K. J., Versijpt, J., Koole, M. *et al*., Experimental performance assessment of SPM for SPECT neuroactivation studies using a subresolution sandwich phantom design. *Neuroimage* **16:** 200-216 (2002).
63. Kiebel S. J., Goebel, R. and Friston, K. J., Anatomically informed basis functions. *Neuroimage* **11:** 656-667 (2000).
64. Kiebel S. J. and Friston, K. J., Analysis of multisubject neuroimaging data using anatomically informed basis functions. *Neuroimage* **13:** S172 (2001).
65. Davatzikos C., Li, H. H., Herskovits, E. *et al*., Accuracy and sensitivity of detection of activation foci in the brain via statistical parametric mapping: a study using a PET simulator. *Neuroimage*. **13:** 176-184 (2001).
66. Nichols T. E. and Holmes, A. P., Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Map* **15:** 1-25 (2002).
67. Kuhl D. E., Koeppe, R. A., Fessler, J. A. *et al*., In vivo mapping of cholinergic neurons in the human brain using SPECT and IBVM. *J Nucl Med* **35:** 405-410 (1994).
68. Minoshima S., Frey, K. A., Foster, N. L. *et al*., Preserved pontine glucose metabolism in Alzheimer disease: a reference region for functional brain image (PET) analysis. *J Comput Assist Tomogr* **19:** 541-547 (1995).
69. Burdette J. H., Minoshima, S., Vander Borght, T. *et al*., Alzheimer disease: improved visual interpretation of PET images by using three-dimensional stereotaxic surface projections. *Radiology* **198:** 837-843 (1996).
70. Kuhl D. E., Minoshima, S., Fessler, J. A. *et al*., In vivo mapping of cholinergic terminals in normal aging, Alzheimer's disease, and Parkinson's disease. *Ann Neurol* **40:** 399-410 (1996).
71. Minoshima S., Giordani, B., Berent, S. *et al*., Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann Neurol* **42:** 85-94 (1997).
72. Vander B. T., Minoshima, S., Giordani, B. *et al*., Cerebral metabolic differences in Parkinson's and Alzheimer's diseases matched for dementia severity. *J Nucl Med* **38:** 797-802 (1997).
73. Bartenstein P., Minoshima, S., Hirsch, C. *et al*., Quantitative assessment of cerebral blood flow in patients with Alzheimer's disease by SPECT. *J Nucl Med* **38:** 1095-1101 (1997).
74. Cross D. J., Minoshima, S., Nishimura, S. *et al*., Three-dimensional stereotactic surface projection analysis of macaque brain PET: development and initial applications. *J Nucl Med* **41:** 1879-1887 (2000).
75. Hosoda K., Kawaguchi, T., Ishii, K. *et al*., Prediction of hyperperfusion after carotid endarterectomy by brain SPECT analysis with semiquantitative statistical mapping method. *Stroke* **34:** 1187-1193 (2003).
76. Acton P. D. and Friston, K. J., Statistical parametric mapping in functional neuroimaging: beyond PET and fMRI activation studies [editorial]. *Eur J Nucl Med* **25:** 663-667 (1998).

77. Weeks R. A., Cunningham, V. J., Piccini, P. *et al*., 11C-diprenorphine binding in Huntington's disease: a comparison of region of interest analysis with statistical parametric mapping. *J Cereb Blood Flow Metab* **17:** 943-949 (1997).
78. Liow J. S., Rehm, K., Strother, S. C. *et al*., Comparison of voxel- and volume-of-interest-based analyses in FDG PET scans of HIV positive and healthy individuals. *J Nucl Med* **41:** 612-621 (2000).
79. Van Laere K. J., Warwick, J., Versijpt, J. *et al*., Analysis of clinical brain SPECT data based on anatomic standardization and reference to normal data: an ROC-based comparison of visual, semiquantitative, and voxel-based methods. *J Nucl Med* **43:** 458-469 (2002).
80. Montandon M.-L., Slosman, D. O. and Zaidi, H., Assessment of the impact of model-based scatter correction on ^{18}F-[FDG] 3D brain PET in healthy subjects using statistical parametric mapping. *Neuroimage* **20:** 1848-1856 (2003).
81. Magnotta V., Andreasen, N., O'Leary, D. *et al*., Brain structures defined by an artificial neural network. *Neuroimage* **13:** S29 (2001).
82. Page M. P., Howard, R. J., O'Brien, J. T. *et al*., Use of neural networks in brain SPECT to diagnose Alzheimer's disease. *J Nucl Med* **37:** 195-200 (1996).
83. Kippenhan J. S., Barker, W. W., Nagel, J. *et al*., Neural-network classification of normal and Alzheimer's disease subjects using high-resolution and low-resolution PET cameras. *J Nucl Med* **35:** 7-15 (1994).
84. Munley M. T., Floyd, C. E., Jr., Bowsher, J. E. *et al*., An artificial neural network approach to quantitative single photon emission computed tomographic reconstruction with collimator, attenuation, and scatter compensation. *Med Phys* **21:** 1889-1899 (1994).
85. Hamilton D., O'Mahony, D., Coffey, J. *et al*., Classification of mild Alzheimer's disease by artificial neural network analysis of SPET data. *Nucl Med Commun* **18:** 805-810 (1997).
86. Strother S. C., Anderson, J. R., Schaper, K. A. *et al*., Principal component analysis and the scaled subprofile model compared to intersubject averaging and statistical parametric mapping: I. ''Functional connectivity'' of the human motor system studied with [15O]water PET. *J Cereb Blood Flow Metab* **15:** 738-753 (1995).
87. Strother S. C., Liow, J. S., Moeller, J. R. *et al*., Absolute quantitation in neurological PET: do we need it? *J Cereb Blood Flow Metab* **11:** A3-16 (1991).
88. Eidelberg D., Moeller, J. R., Dhawan, V. *et al*., The metabolic topography of parkinsonism. *J Cereb Blood Flow Metab* **14:** 783-801 (1994).
89. Moeller J. R., Ishikawa, T., Dhawan, V. *et al*., The metabolic topography of normal aging. *J Cereb.Blood Flow Metab* **16:** 385-398 (1996).
90. Antonini A., Moeller, J. R., Nakamura, T. *et al*., The metabolic anatomy of tremor in Parkinson's disease. *Neurology* **51:** 803-810 (1998).
91. Moeller J. R., Nakamura, T., Mentis, M. J. *et al*., Reproducibility of regional metabolic covariance patterns: comparison of four populations [In Process Citation]. *J Nucl Med* **40:** 1264-1269 (1999).
92. Links J. M., Special issues in quantitation of brain receptors and related markers by emission computed tomography. *Q J Nucl Med* **42:** 158-165 (1998).
93. Zaidi H. and Sossi, V., Correction for image degrading factors is essential for accurate quantification of brain function using PET. *Med Phys* **31:** 423-426 (2004).

94. Houston A. S., Kemp, P. M. and Macleod, M. A., Optimization of factors affecting the state of normality of a medical image. *Phys Med Biol* **41:** 755-765 (1996).
95. Koyama M., Kawashima, R., Ito, H. *et al.*, SPECT imaging of normal subjects with Technetium-99m-HMPAO and Technetium-99m-ECD. *J Nucl Med* **38:** 587-592 (1997).
96. Jueptner M. and Weiller, C., Review: does measurement of regional cerebral blood flow reflect synaptic activity? Implications for PET and fMRI. *Neuroimage* **2:** 148-156 (1995).
97. Lou H., Edvinsson, L. and MacKenzie, E. T., The concept of coupling blood flow to brain function: revision required? *Ann Neurol* **22:** 289-297 (1987).
98. Waldvogel D., van Gelderen, P., Muellbacher, W. *et al.*, The relative metabolic demand of inhibition and excitation. *Nature* **406:** 995-998 (2000).
99. Asenbaum S., Brücke, T., Pirker, W. *et al.*, Imaging of cerebral blood flow with technetium-99m-HMPAO and technetium-99m-ECD: a comparison. *J Nucl Med* **39:** 613-618 (1998).
100. Houston A. S., Kemp, P. M. and Macleod, M. A., A method for assessing the significance of abnormalities in HMPAO brain SPECT images. *J Nucl Med* **35:** 239-244 (1994).
101. Newberg A. B. and Alavi, A., Normal patterns and variants in single-photon emission computed tomography and positron emission tomography brain imaging. *Semin Nucl Med* **33:** 42-55 (2003).
102. Syed G. M., Eagger, S., Toone, B. K. *et al.*, Quantification of regional cerebral blood flow (rCBF) using 99Tcm-HMPAO and SPECT: choice of the reference region. *Nucl Med Commun* **13:** 811-816 (1992).
103. Talbot P. R., Lloyd, J. J., Snowden, J. S. *et al.*, Choice of reference region in the quantification of single-photon emission tomography in primary degenerative dementia. *Eur J Nucl Med* **21:** 503-508 (1994).
104. Carson R. E., Precision and accuracy considerations of physiological quantitation in PET. *J Cereb Blood Flow Metab* **11:** A45-A50 (1991).
105. Tikosfky R. S., Ichise, M., Seibyl, J. P. *et al.*, Functional brain SPECT imaging: 1999 and beyond. *Semin Nucl Med* **29:** 193-238 (1999).
106. Pagani M., Salmaso, D., Jonsson, C. *et al.*, Regional cerebral blood flow as assessed by principal component analysis and (99m)Tc-HMPAO SPET in healthy subjects at rest: normal distribution and effect of age and gender. *Eur J Nucl Med Mol Imaging* **29:** 67-75 (2002).
107. Slosman D. O., Chicherio, C., Ludwig, C. *et al.*, (133)Xe SPECT cerebral blood flow study in a healthy population: determination of T-scores. *J Nucl Med* **42:** 864-870 (2001).
108. Tanaka F., Vines, D., Tsuchida, T. *et al.*, Normal patterns on Tc-99m-ECD brain SPECT scans in adults. *J Nucl Med* **41:** 1456-1464 (2000).
109. Lobaugh N. J., Caldwell, C. B., Black, S. E. *et al.*, Three brain SPECT region-of-interest templates in elderly people: normative values, hemispheric asymmetries, and a comparison of single- and multihead cameras. *J Nucl Med* **41:** 45-56 (2000).
110. Jones K., Johnson, K. A., Becker, J. A. *et al.*, Use of singular value decomposition to characterize age and gender differences in SPECT cerebral perfusion. *J Nucl Med* **39:** 965-973 (1998).

111. Krausz Y., Bonne, O., Gorfine, M. *et al*., Age-related changes in brain perfusion of normal subjects detected by 99mTc-HMPAO SPECT. *Neuroradiology* **40:** 428-434 (1998).
112. Mozley P. D., Sadek, A. M., Alavi, A. *et al*., Effects of aging on the cerebral distribution of technetium-99m hexamethylpropylene amine oxime in healthy humans. *Eur J Nucl Med* **24:** 754-761 (1997).
113. Barthel H., Wiener, M., Dannenberg, C. *et al*., Age-specific cerebral perfusion in 4- to 15-year-old children: a high-resolution SPET study using ^{99m}Tc-ECD. *Eur J Nucl Med* **24:** 1245-1252 (1997).
114. Catafau A. M., Lomena, F. J., Pavia, J. *et al*., Regional cerebral blood flow pattern in normal young and aged volunteers: a ^{99m}Tc-HMPAO study. *Eur J Nucl Med* **23:** 1329-1337 (1996).
115. Markus H. S., Ring, H., Kouris, K. *et al*., Alterations in regional cerebral blood flow, with increased temporal interhemispheric asymmetries, in the normal elderly: an HMPAO SPECT study. *Nucl Med Commun* **14:** 628-633 (1993).
116. Matsuda H., Tsuji, S., Shuke, N. *et al*., Noninvasive measurement of regional cerebral blood flow using technetium-99m hexamethylpropylene amine oxime. *Eur J Nucl Med* **20:** 391-401 (1993).
117. Waldemar G., Hasselbalch, S. G., Andersen, A. R. *et al*., 99mTc-d,l-HMPAO and SPECT of the brain in normal aging. *J Cereb.Blood Flow Metab* **11:** 508-521 (1991).
118. Rodriguez G., Warkentin, S., Risberg, J. *et al*., Sex differences in regional cerebral blood flow. *J Cereb Blood Flow Metab* **8:** 783-789 (1988).
119. Herholz K., Salmon, E., Perani, D. *et al*., Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *Neuroimage* **17:** 302-316 (2002).
120. Willis M. W., Ketter, T. A., Kimbrell, T. A. *et al*., Age, sex and laterality effects on cerebral glucose metabolism in healthy adults. *Psychiatry Res* **114:** 23-37 (2002).
121. Lee J. S., Lee, D. S., Kwon, J. S. *et al*., Changes in regional cerebral glucose metabolism with healthy aging. *Neuroimage* **11:** S805 (2000).
122. Volkow N. D., Logan, J., Fowler, J. S. *et al*., Association between age-related decline in brain dopamine activity and impairment in frontal and cingulate metabolism. *Am J Psychiatry* **157:** 75-80 (2000).
123. Schultz S. K., O'Leary, D. S., Ponto, L. L. B. *et al*., Age-related changes in regional cerebral blood flow among young to midlife adults. *Neuroreport* **10:** 2493-2496 (1999).
124. Petit-Taboue M. C., Landeau, B., Desson, J. F. *et al*., Effects of healthy aging on the regional cerebral metabolic rate of glucose assessed with statistical parametric mapping. *Neuroimage* **7:** 176-184 (1998).
125. Murphy D. G., DeCarli, C., McIntosh, A. R. *et al*., Sex differences in human brain morphometry and metabolism: an in vivo quantitative magnetic resonance imaging and positron emission tomography study on the effect of aging. *Arch Gen Psychiatry* **53:** 585-594 (1996).
126. Gur R. C., Mozley, L. H., Mozley, P. D. *et al*., Sex differences in regional glucose metabolism during the resting state. *Science* **267:** 528-531 (1995).

127. Loessner A., Alavi, A., Lewandrowski, K. U. *et al*., Regional cerebral function determined by FDG-PET in healthy volunteers: normal patterns and changes with age. *J Nucl Med* **36:** 1141-1149 (1995).
128. Marchal G., Rioux, P., Petit-Taboue, M. C. *et al*., Regional cerebral oxygen consumption, blood flow, and blood volume in healthy human aging. *Arch Neurol* **49:** 1013-1020 (1992).
129. Martin A. J., Friston, K. J., Colebatch, J. G. *et al*., Decreases in regional cerebral blood flow with normal aging. *J.Cereb.Blood Flow Metab* **11:** 684-689 (1991).
130. Leenders K. L., Perani, D., Lammertsma, A. A. *et al*., Cerebral blood flow, blood volume and oxygen utilization. Normal values and effect of age. *Brain* **113:** 27-47 (1990).
131. Yoshii F., Barker, W. W., Chang, J. Y. *et al*., Sensitivity of cerebral glucose metabolism to age, gender, brain volume, brain atrophy and cerebrovascular risk factors. *J Cereb Blood Flow Metab* **8:** 654-661 (1988).
132. Slageter N. L., Horwitz, B., Creasey, H. *et al*., Relation of cerebral glucose utilization and cerebral atrophy in man. *J Neurol Neurosurg Psy* **50:** 779-785 (1987).
133. Van Laere K., Koole, M., Versijpt, J. *et al*., Transfer of normal 99mTc-ECD brain SPET databases between different gamma cameras. *Eur J Nucl Med* **28:** 435-449 (2001).
134. Grady C. L., Berg, G., Carson, R. E. *et al*., Quantitative comparison of cerebral glucose metabolic rates from two positron emission tomographs. *J Nucl Med* **30:** 1386-1392 (1989).
135. Ribeiro M. J., Remy, P., Bendriem, B. *et al*., Comparison of clinical data sets acquired on different tomographs using 6-^{18}F-L-dopa. *Eur J Nucl Med* **27:** 707-712 (2000).
136. Nuyts J., Mortelmans, L., Van de, W. F. *et al*., Cardiac phantom measurement validating the methodology for a cardiac multi-centre trial with positron emission tomography. *Eur J Nucl Med Mol Imaging* **29:** 1588-1593 (2002).
137. OHBM T. G. C. o. t. O. f. H. B. M. O., Neuroimaging databases. *Science* **292:** 1673-1676 (2001).
138. Okamura N., Arai, H., Higuchi, M. *et al*., [18F]FDG-PET study in dementia with Lewy bodies and Alzheimer's disease. *Prog Neuropsychopharmacol Biol Psy* **25:** 447-456 (2001).
139. Varrone A., Pappata, S., Caraco, C. *et al*., Voxel-based comparison of rCBF SPET images in frontotemporal dementia and Alzheimer's disease highlights the involvement of different cortical networks. *Eur J Nucl Med Mol.Imaging* **29:** 1447-1454 (2002).
140. Staff R. T., Gemmell, H. G., Shanks, M. F. *et al*., Changes in the rCBF images of patients with Alzheimer's disease receiving Donepezil therapy. *Nucl Med Comm* **21:** 37-41 (2000).
141. Raichle M., Imaging the mind. *Sem Nucl Med* **28:** 279-289 (1998).
142. Barrington S. F., Clinical use of PET in neurology. *Nucl Med Comm* **21:** 237-240 (2000).
143. Minoshima S., Frey, K. A., Koeppe, R. A. *et al*., A diagnostic approach in Alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *J Nucl Med* **36:** 1238-1248 (1995).

144. Silverman D. H., Gambhir, S. S., Huang, H. W. *et al*., Evaluating early dementia with and without assessment of regional cerebral metabolism by PET: a comparison of predicted costs and benefits. *J Nucl Med* **43:** 253-266 (2002).
145. Jobst K. A., Barnetson, L. P. and Shepstone, B. J., Accurate prediction of histologically confirmed Alzheimer's disease and the differential diagnosis of dementia: the use of NINCDS-ADRDA and DSM-III-R criteria, SPECT, X-ray CT, and APO E4 medial temporal lobe dementias. The Oxford Project to Investigate Memory and Aging. *Int Psychogeriatr* **9 Suppl 1:** 191-222 (1997).
146. Herholz K., Schopphoff, H., Schmidt, M. *et al*., Direct comparison of spatially normalized PET and SPECT scans in Alzheimer's disease. *J Nucl Med* **43:** 21-26 (2002).
147. Imran M. B., Kawashima, R., Awata, S. *et al*., Parametric mapping of cerebral blood flow deficits in Alzheimer's disease: a SPECT study using HMPAO and image standardization technique. *J Nucl Med* **40:** 244-249 (1999).
148. Soonawala D., Amin, T., Ebmeier, K. P. *et al*., Statistical parametric mapping of (99m)Tc-HMPAO-SPECT images for the diagnosis of Alzheimer's disease: normalizing to cerebellar tracer uptake. *Neuroimage* **17:** 1193-1202 (2002).
149. Drzezga A., Arnold, S., Minoshima, S. *et al*., 18F-FDG PET studies in patients with extratemporal and temporal epilepsy: evaluation of an observer-independent analysis. *J Nucl Med* **40:** 737-746 (1999).
150. Noachtar S., Arnold, S., Yousry, T. A. *et al*., Ictal technetium-99m ethyl cysteinate dimer single-photon emission tomographic findings and propagation of epileptic seizure activity in patients with extratemporal epilepsies. *Eur J Nucl Med* **25:** 166-172 (1998).
151. O'Brien T. J., So, E. L., Mullan, B. P. *et al*., Subtraction ictal SPECT co-registered to MRI improves clinical usefulness of SPECT in localizing the surgical seizure focus. *Neurology* **50:** 445-454 (1998).
152. Chang D. J., Zubal, I. G., Gottschalk, C. *et al*., Comparison of statistical parametric mapping and SPECT difference imaging in patients with temporal lobe epilepsy. *Epilepsia* **43:** 68-74 (2002).
153. Lee J. D., Kim, H. J., Lee, B. I. *et al*., Evaluation of ictal brain SPET using statistical parametric mapping in temporal lobe epilepsy. *Eur J Nucl Med* **27:** 1658-1665 (2000).
154. AAN A. A. o. N. A. A. N., Assessment of brain SPECT. Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. *Neurology* **46:** 278-285 (1996).
155. Ueda T., Sakaki, S., Yuh, W. T. *et al*., Outcome in acute stroke with successful intra-arterial thrombolysis and predictive value of initial single-photon emission-computed tomography. *J Cereb Blood Flow Metab* **19:** 99-108 (1999).
156. Umemura A., Suzuka, T. and Yamada, K., Quantitative measurement of cerebral blood flow by 99mTc-HMPAO in acute ischaemic stroke: usefulness in determining therapeutic options. *J Neurol Neurosury Psychiatry* **69:** 472-478 (2000).
157. Ostergaard L., Sorensen, A. G., Chesler, D. A. *et al*., Combined diffusion-weighted and perfusion-weighted flow heterogeneity magnetic resonance imaging in acute stroke. *Stroke* **31:** 1097-1103 (2000).

158. Elliott J. P., Newell, D. W., Lam, D. J. *et al*., Comparison of balloon angioplasty and papaverine infusion for the treatment of vasospasm following aneurysmal subarachnoid hemorrhage. *J Neurosurgery* **88:** 277-284 (1998).
159. Ramsay S. C., Yeates, M. G., Lord, R. S. *et al*., Use of technetium-99m-HMPAO to demonstrate changes in flow reserve following carotid endartrectomy. *J Nucl Med* **32:** 1382-1386 (1991).
160. Cikrit D. F., Dalsing, M. C., Harting, P. S. *et al*., Cerebral vascular reactivity assessed with acetazolamide single photon emission computer tomography scans before and after carotid endarterectomy. *Am J Surg* **174:** 193-197 (1997).
161. Lang A. E. and Lozano, A. M., Parkinson's disease. First of two parts. *N Engl J Med* **339:** 1044-1053 (1998).
162. Booij J., Tissingh, G., Winogrodzka, A. *et al*., Imaging of the dopaminergic neurotransmission system using single- photon emission tomography and positron emission tomography in patients with parkinsonism. *Eur.J.Nucl.Med.* **26:** 171-182 (1999).
163. Schwarz J., Tatsch, K., Gasser, T. *et al*., [123]IBZM binding predicts dopaminergic responsiveness in patients with parkinsonism and previous dopaminomimetic therapy. *Mov Disord* **12:** 898-902 (1997).
164. Costa D., Walker, Z., Walker, R. *et al*., Dementia with Lewy bodies: preliminary data on clinical, pathological and FP-CIT SPECT correlations. *Nucl Med Comm* **20:** 467-468 (1999).
165. Seibyl J. P., Marek, K., Quinlan, D. *et al*., Decreased 123Ibeta-CIT striatal uptake correlates with symptom severity in idiopathic Parkinson's disease. *Ann Neurol* **38:** 589-598 (1995).
166. Spetsieris P. G., Moeller, J. R., Dhawan, V. *et al*., Visualizing the evolution of abnormal metabolic networks in the brain using PET. *Comput.Med Imaging Graph* **19:** 295-306 (1995).
167. Bosman T., Van Laere, K. and Santens, P., Anatomically standardised (99m)Tc-ECD brain perfusion SPET allows accurate differentiation between healthy volunteers, multiple system atrophy and idiopathic Parkinson's disease. *Eur J Nucl Med Mol Imaging* **30:** 16-24 (2003).
168. Hosaka K., Ishii, K., Sakamoto, S. *et al*., Voxel-based comparison of regional cerebral glucose metabolism between PSP and corticobasal degeneration. *J Neurol Sci* **199:** 67-71 (2002).
169. Weber W., Bartenstein, P., Gross, M. W. *et al*., Fluorine-18-FDG PET and iodine-123-IMT SPECT in the evaluation of brain tumors. *J Nucl Med* **38:** 802-808 (1997).
170. Bartenstein P., Grünwald, F., Kuwert, T. *et al*., Clinical applications of single photon emission tomography in neuromedicine. Part 1: Neuro-oncology, epilepsy, movement disorders, cerebrovascular disease. *Nuklearmedizin* **39:** 180-195 (2000).
171. Jacobs A., Put, E., Ingels, M. *et al*., One-year follow-up of technetium-99m-HMPAO SPECT in mild head injury. *J Nucl Med* **37:** 1609 (1996).
172. Ichise M., Chung, D. G., Wang, P. *et al*., Technetium-99m-HMPAO SPECT, CT and MRI in the evaluation of patients with chronic traumatic brain injury: a correlation with neuropsychological performance. *J Nucl Med* **35:** 217-226 (1994).

173. Stamatakis E. A., Wilson, J. T., Hadley, D. M. *et al*., SPECT imaging in head injury interpreted with statistical parametric mapping. *J Nucl Med* **43:** 476-483 (2002).
174. Holman L. B. and Devous, M. D., Functional brain SPECT: the emergence of a powerful clinical method. *J Nucl Med* **33:** 1888-1904 (1995).
175. Goodwin M. A., Functional imaging, affective disorder and dementia. *Br Med Bull* **1996:** 495-512 (1996).
176. Laruelle M., Abi-Dargham, A., Gil, R. *et al*., Increased dopamine transmission in schizophrenia: relationship to illness phases. *Biol Psychiatr* **46:** 56-72 (1999).
177. Busatto G. F., Zamignani, D. R., Buchpiguel, C. A. *et al*., A voxel-based investigation of regional cerebral blood flow abnormalities in obsessive-compulsive disorder using single photon emission computed tomography (SPECT). *Psychiatry Res* **99:** 15-27 (2000).
178. Bonne O., Louzoun, Y., Aharon, I. *et al*., Cerebral blood flow in depressed patients: a methodological comparison of statistical parametric mapping and region of interest analyses. *Psychiatry Res* **122:** 49-57 (2003).
179. Siessmeier T., Nix, W. A., Hardt, J. *et al*., Observer independent analysis of cerebral glucose metabolism in patients with chronic fatigue syndrome. *J Neurol Neurosurg Psy* **74:** 922-928 (2003).
180. Imran M. B., Kawashima, R., Sato, K. *et al*., Detection of CBF deficits in neuropsychiatric disorders by an expert system: a 99Tcm-HMPAO brain SPET study using automated image registration. *Nucl Med Commun* **20:** 25-32 (1999).
181. Moresco F. M., Dieci, M., Vita, A. *et al*., In vivo serotonin 5HT(2A) receptor binding and personality traits in healthy subjects: a positron emission tomography study. *Neuroimage* **17:** 1470-1478 (2002).
182. Turner R. S., Grafton, S. T., McIntosh, A. R. *et al*., The functional anatomy of parkinsonian bradykinesia. *Neuroimage* **19:** 163-179 (2003).
183. Price B. H., Adams, R. D. and Coyle, J. T., Neurology and psychiatry. Closing the great divide. *Neurology* **54:** 8-14 (2000).
184. Fujita M. and Innis, R. B., Nuclear psychiatry: specialty for the twenty-first century? *Eur J Nucl Med* **27 (Suppl):** S39-S42 (2000).

15
Quantitative Analysis in Myocardial SPECT Imaging

G. Germano*

1. Introduction

Nuclear cardiology is a mature, yet still expanding non-invasive imaging modality that has maintained an annual growth rate of approximately 20% over the past 20 years. Its continuing success parallels clinical cardiology's evolution from simple detection and diagnosis of coronary artery disease (CAD) to cost-effective risk stratification of patients with respect to specific adverse outcomes (cardiac death, myocardial infarction, etc.), at the same time fulfilling a need for effective monitoring of the effects of medical or surgical therapies. Single-photon emission computed tomography (SPECT) of the myocardium accounts for the great majority of nuclear cardiology procedures performed, thanks to its wide availability, conclusively established and clinically validated diagnostic and prognostic accuracy, and ability to perform reliable and reproducible assessment of myocardial perfusion and function through the use of automated software algorithms. This chapter will list the various cardiac parameters amenable to quantitative analysis, describe the methods through which their quantitation is performed, and investigate the validation and limitation of such quantitative measurements as they have been reported in the published literature.

2. Quantitation of Myocardial Perfusion

The typical processing steps associated with myocardial perfusion quantitation are illustrated in Figure 1. In essence, the myocardium is isolated ("segmented") from other organs or areas of extracardiac uptake, with a myocardial mask being automatically or manually defined for the purpose of guiding the quantitative algorithm. Myocardial uptake is sampled at regular intervals and parametrically displayed in a polar map, whose pixels

*Prof. G. Germano, Cedars-Sinai Medical Center, Artificial Intelligence Program, 90048 Los Angeles, California, USA

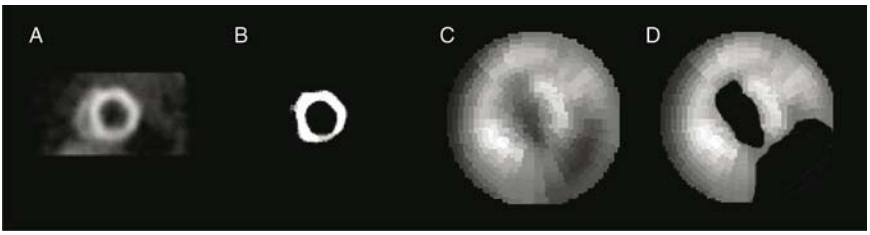


FIGURE 1. Perfusion quantification. A. Original short axis image; B. Segmented myocardial mask; C. Parametric myocardial perfusion polar map; D. "Blacked out" polar map highlighting areas of abnormal perfusion.

are marked (usually blacked out) whenever they contain fewer counts than what is considered normal for the particular patient/protocol/radioisotope combination studied.

While most quantitative techniques follow the above processing sequence, they differ as to the manner in which the sequence's steps are implemented. For example, traditional perfusion quantitation approaches are based on the extraction of a maximal count "circumferential profile" from each short axis image in the tomographic data set, according to a hybrid sampling scheme that models the left ventricular (LV) myocardium as having cylindrical shape in its most basal two thirds, and spherical shape at the apex.[1] The circumferential profiles (each comprising 36-60 equally spaced maximal count samples) constitute a condensed or parametric representation of myocardial perfusion, and can be combined in 2-dimensional "polar maps" or bullseye maps.[2] A polar map consists of a variable number of circumferential profiles (proportional to the number of short axis slices, i.e., to the size of the myocardium), which are displayed as concentric annuli or rings having either the same thickness (distance-weighted polar maps) or thicknesses representative of the volume of the myocardium in the individual slices (volume-weighted polar maps).[3] Polar maps from normal patients (or patients with a low likelihood of coronary artery disease)[4] are pooled together, and a fixed number of "normal circumferential profiles" (12 in the hybrid cylindrical/spherical approach)[5] are taken to represent the average radioactivity uptake (radioisotope- as well as gender-specific) in a normal patient's myocardium. Criteria for abnormality of a specific region of the myocardium are established, and limits of normality derived based on the mean and standard deviation of the normal uptake in that region. As a result, the circumferential profiles for every patient can be matched to these normal limits, and the extent by which uptake falls below them can be measured on a regional and global basis.

More recent perfusion quantitation approaches have been developed that are not based on circumferential profiles, in the sense that they do not generate a number of profiles equal to the number of myocardial short axis slices for a particular patient. Rather, every LV is treated as a

3-dimensional structure, and a standard number of longitudinally and latitudinally equi-spaced myocardial perfusion data samples is extracted, regardless of LV size.[6,7] With this approach, homologous sample points in different patients' myocardia are intrinsically registered, and can be pooled for the purpose of normal limits generation. The sampling model determines the directions along which sampling rays are drawn, and can be optimized to ensure consistency with the expected ellipsoidal shape of the LV; also, the average of all perfusion data points comprised between the endocardial and the epicardial surface along each sampling ray can be extracted, rather than just the maximal count pixel.[7] Finally, each sample point can be weighted by the area of the myocardial surface patch to which it corresponds, in order to reflect the curvature of the specific myocardium analyzed at different sampling locations. Generally speaking, there is a good number of commercially available algorithms for the quantitative assessment of myocardial perfusion, each with its own characteristics as to mathematical techniques used, degree of automation and clinical validation.[6-10]

There are several ways in which myocardial perfusion defects can be measured and expressed. Perhaps the most intuitive approach involves counting the number of pixels with abnormal perfusion in a polar map (as determined by comparison to homologous pixels in a "normal" map), and ratioing that number to the total number of pixels in the polar map. The resulting percent value is referred to as "global myocardial perfusion defect extent", or "defect extent" for short. Similarly, it is possible to determine the number of standard deviations by which each abnormal pixel is below the normal threshold for the myocardial location it represents, and add those numbers to obtain a "global myocardial perfusion severity" ("defect severity") quantity.

While perfusion defect extent and severity are straightforward parameters, their quantitative estimates can only be generated by software, are dependent on the particular algorithm used, and are not easily modifiable to reflect, for example, attenuation artefacts not accounted for in the normal databases but detected by clinical visual assessment. These limitations are compounded by the following factors: 1) essentially no published guidelines exist defining diagnostic and prognostic thresholds for quantitatively determined defect extent and severity, and 2) the two measurements are not usually combined into a global quantity assessing overall myocardial perfusion.

An alternative to pixel-based quantitation makes use of "categorical" perfusion scores assigned to a discrete number of myocardial regions or "segments". In one popular model, the LV myocardium is divided into 20 segments, each of which roughly corresponds to 5% of the LV and is assigned a number from 0 to 4 [0 =normal; 1 =slight reduction of uptake (equivocal); 2 = moderate reduction of uptake (usually implies a significant abnormality); 3 =severe reduction of uptake; 4 =absence of radioactive uptake][11] (Figure 2). Perfusion defects with scores of three or four can be

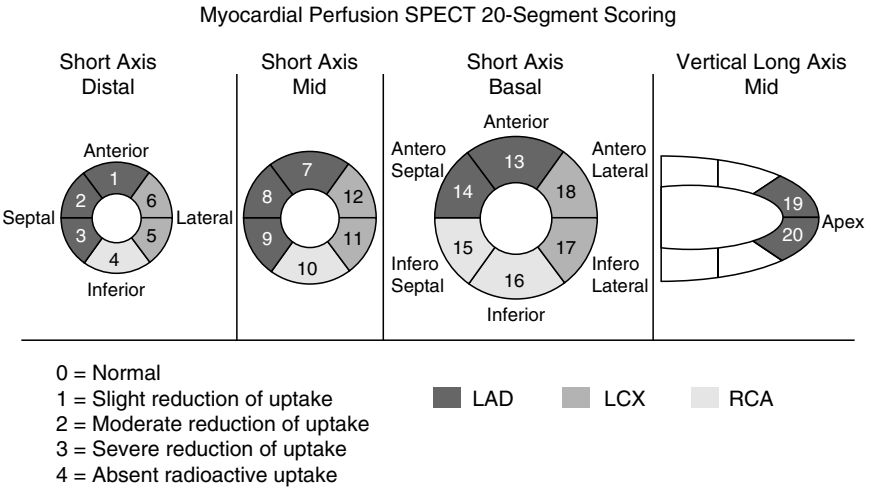


FIGURE 2. Diagrammatic representation of the segmental division of the SPECT slices and assignment of individual segments to individual coronary arteries using the 20-segment, 5-scores model. LAD= left anterior descending coronary artery; LCX= left circumflex coronary artery; RCA= right coronary artery. (reproduced with permission from ref.[17]).

reported as consistent with a critical (> 90%) coronary stenosis.[12-13] Segmental scores can be automatically determined by software calibrated on expert readers' visual assessment,[7,14] and can be easily reviewed and adjusted to compensate for various artefacts. A remarkable feature of categorical perfusion scores is that they can be used to generate global indices of myocardial perfusion (''summed scores'') combining perfusion defect extent and severity. In particular, the summed stress score (SSS) is defined as the sum of the stress scores for the 20 segments. The summed rest score (SRS) is defined as the sum of the rest scores or redistribution scores, and the summed differences score (SDS), measuring the degree of reversibility, is defined as the difference between the summed stress score and the summed rest score. Risk groups may be defined using the SSS,[15,16] where a score < 4 is considered normal or nearly normal, scores of 4-8 are mildly abnormal, scores of 9-13 moderately abnormal, and summed stress scores > 13 severely abnormal.

A 17-segment model has been recently proposed and endorsed by a number of professional organizations, based on its more balanced weighting of the myocardium as well as its potential consistency with non-nuclear (echocardiography, magnetic resonance imaging, etc) segmental models.[18] In this approach, the smaller size of the apical short axis slice is accounted for by dividing it into 4, rather than 6, segments, while the apex is considered to comprise a single segment. Given that models other than the 17- and 20-segment ones do exist, and in order to make quantitation model-

independent, summed scores can also be expressed as a percentage of the worst possible score associated with the particular categorical model chosen.[19] Thus, the concept that SSS = 40 in the 20-segment model (max SSS = 80) has the same diagnostic and prognostic meaning as SSS = 34 in the 17-segment model (max SSS = 68) is straightforwardly conveyed by a ''percentage summed stress score'' SS% = 40/80 = 34/68 = 50%.

Figure 3 gives an example of a perfusion quantitation display based on the QPS algorithm, with pixel-based extent, stress/rest/reversibility scores and summed scores, as well as ''normalized'' or percentage summed scores.

While several commercially available algorithms exist that quantitatively measure myocardial perfusion,[6-10] their outputs have been reported to be non-interchangeable, whether for determination of simple infarct extent,[20] more complex perfusion defect size,[21] or perfusion defect reversibility - the latter measurement is particularly vulnerable to error propagation, because it is determined as the difference of the rest and stress perfusion abnormalities. As was previously discussed, most perfusion quantitation algorithms are based on the development and validation of normal databases and regional criteria (thresholds) for abnormality. This process can be cumbersome if performed as originally described by Van Train *et al.*[5,22], in that it requires 1) two gender-specific populations of 20-30 normal or low-likelihood patients each, from which regional means and standard deviations of

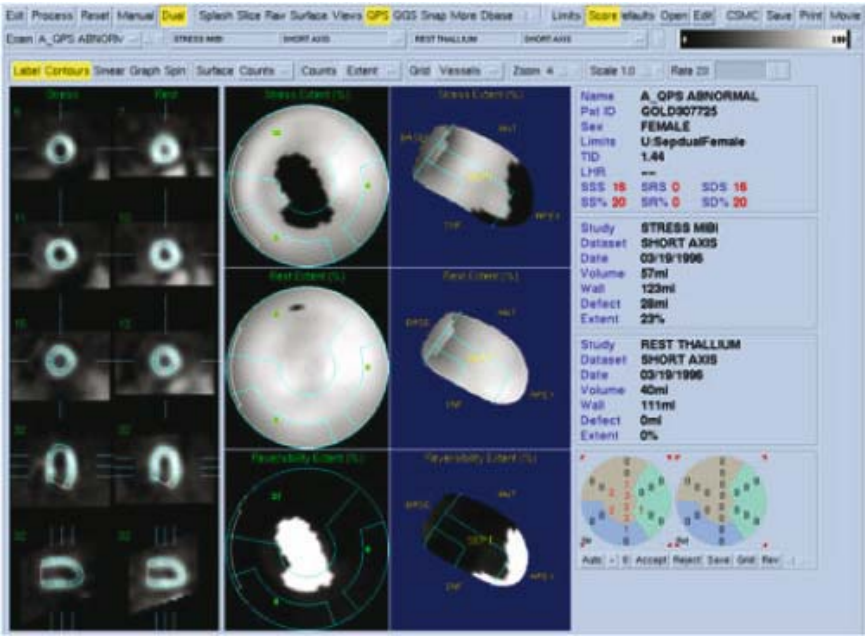


FIGURE 3. Quantitative perfusion SPECT display for a patient with ischemia in the mid-left anterior descending coronary artery territory.

radionuclide uptake are calculated; 2) two gender-specific "pilot" populations of at least 60 patients each, usually gathered from a much larger patient pool so as to comprise a statistically adequate range of perfusion defect location and severity, and 3) two gender-specific "validation" populations similar to, but distinct from, the pilot populations, and whose patients also underwent coronary angiography. If one considers that, ideally, different sets of normal limits would be required based on patient's gender, body habitus and stance (supine, prone, upright), type of radioisotope (^{201}Tl, ^{99m}Tc-sestamibi, ^{99m}Tc-tetrofosmin, etc.), protocol (same-day, 2-day, dual isotope, etc.), type of stress (exercise, pharmacologic), collimator type, acquisition orbit and zoom, processing filter (active/passive, cutoff frequency), reconstruction (iterative, filtered backprojection), and possibly a host of other variables, it is understandable that end-users might hesitate to fully embrace perfusion quantitation until such time when either a small number of "basic" normal limits is agreed upon to well address most clinical situations, and/or a simpler way to develop and validate normal limits is implemented.

3. Quantitation of Myocardial Function

Myocardial function is generally assessed with respect to the global LV (LV cavity volumes, stroke volume and ejection fraction) and/or one of its regions (regional myocardial wall motion and wall thickening).

3.1 End-Diastolic and End-Systolic Volumes

Determination of LV cavity volumes from cardiac SPECT images is generally performed by identifying the endocardial surface and the valve plane in the 3-dimensional space, counting the voxels bound by those two surfaces and multiplying the result by the individual voxel's volume. If this process is repeated for all the intervals (frames) of a gated SPECT acquisition, a "time-volume curve" is produced from which the end-diastolic volume (EDV) and the end-systolic volume (ESV) can be isolated as the largest and smallest value (Figure 4).

Of note, the actual volumetric values measured will depend on the temporal sampling of the cardiac cycle, with a larger number of frames acquired generally corresponding to greater dynamic range, i.e., slightly larger EDVs and smaller ESVs. Even if fine sampling is used, the relatively low resolution of nuclear cardiology images can lead to the apparent shrinkage or even obliteration of the LV cavity in patients with small ventricles, with consequent underestimation of EDV, and, particularly, ESV.[23,24] Finally, volume measurements can be compromised by incorrect listing of the pixel size in the image header. Pixel size is usually automatically calculated by modern cameras, based on knowledge of field of view and zoom information.

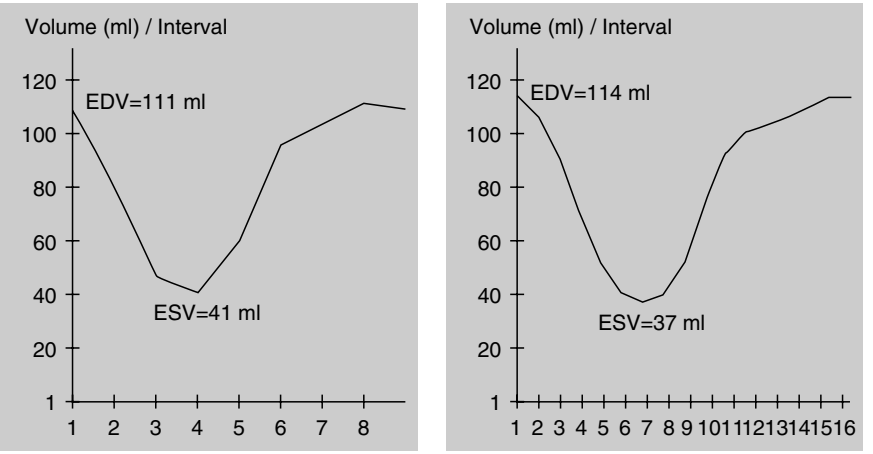


FIGURE 4. Quantitative time-volume curves derived from an 8-frame (left – EDV is at frame 1, ESV at frame 4) and a 16-frame (right – EDV is at frame 1, ESV at frame 7) gated SPECT acquisition. Note the larger dynamic range of the 16-frame measurements.

However, older cameras or ''hybrid'' systems (where one manufacturer's camera is interfaced to another manufacturer's computer) may not be set up to transfer pixel size information from the gantry, or may take a ''standard'' size (i.e., 1 cm) as default. In these cases, a correction factor should be manually calculated by imaging a known pattern (for example, two line sources separated by an exact distance), and counting the number of pixels between the lines' centroids in the reconstructed transaxial image.

As with perfusion quantitation, there is a good number of commercially available algorithms for quantitative assessment of end-diastolic and end-systolic volumes, each with its own characteristics as to mathematical techniques used, degree of automation and clinical validation.[9,25-27] Table 1 presents an aggregate summary of published validations, demonstrating good to excellent agreement of gated SPECT EDV and ESV measurements with results obtained from a variety of reference imaging techniques.

3.2 Ejection Fraction

Unlike planar blood pool approaches (which are, for the most part, count-based), quantitative measurements of left ventricular ejection fraction (LVEF) from gated SPECT images are usually derived from the EDV and ESV values, as follows:

$$\%LVEF = (EDV\text{-}ESV)/EDV^{*}100 \quad (1)$$

Of note, errors in the determination of end-diastolic and end-systolic volumes would be expected to occur in the same general direction, and

Table 1. Validation of quantitative measurements of EDV and ESV from gated myocardial perfusion SPECT (reproduced with permission from ref.[28]).

"Gold standard"	# reports	# patients	Spearman's r (EDV)	Spearman's r (ESV)	Radiopharmaceutical	Reference
2-D echo	11	564	0.70–0.92	0.71–0.94	^{99m}Tc-sestamibi, ^{201}Tl	29–35
MRI	9	185	0.81–0.97	0.90–0.99	^{99m}Tc-sestamibi, ^{99m}Tc-tetrofosmin, ^{201}Tl	9, 36–41
MUGA	4	123	0.7–0.88	0.7–0.95	^{99m}Tc-sestamibi, ^{99m}Tc-tetrofosmin, ^{201}Tl	42–44
contrast ventriculography	3	117	0.86–0.89	0.90–0.95	^{99m}Tc-sestamibi, ^{201}Tl	45–47
3-D echo	2	26	0.94–0.99	0.97–0.99	^{99m}Tc-sestamibi, ^{201}Tl	48, 49
first pass	2	99	0.85–0.93	0.91–0.92	^{99m}Tc-sestamibi	9, 50
thermodilution	2	45	0.86–0.89	0.94	^{99m}Tc-sestamibi	51, 52
TOTAL	33	1,159	0.87	0.90		

therefore would at least partially cancel out when the volumes are ratioed, making LVEF measurements comparatively more reliable than absolute volumes.[25] Nevertheless, similar considerations apply to volumes and ejection fractions with respect to temporal sampling and small LVs. Specifically, 8-frame gating leads to mild underestimation of the LVEF (3-4 LVEF percentage points, uniform over a wide range of ejection fractions) compared to 16-frame gating,[25,53-56] and LVEF will likely be overestimated in patients with small hearts.[57]

The thickness of the myocardium is also an issue, in the sense that quantitative gated SPECT algorithms are either calibrated for the range of thicknesses most typically encountered in clinical practice,[25] or assume a fixed myocardial thickness in the normal range;[9] consequently, gated SPECT LVEFs measured in patients with left ventricular hypertrophy are likely to be underestimated.[58] On the other hand, several published reports found accurate gated SPECT LVEFs measured in patients with large perfusion defects and/or low LVEF.[36,45,55,59-64] While most quantitative algorithms tend to "cut off" aneurysms by assuming a regular or "smooth" LV shape in areas with severely reduced or absent perfusion, the consequent error in measured LVEF is generally mild.

As with perfusion quantitation, there is a good number of commercially available algorithms for quantitative assessment of LVEF, each with its own characteristics as to mathematical techniques used, degree of automation and clinical validation.[9,25-27] Table 2 presents an aggregate summary of all published validations, demonstrating good to excellent agreement of gated SPECT LVEF measurements with results obtained from a variety of reference imaging techniques.

TABLE 2. Validation of quantitative measurements of LVEF from gated myocardial perfusion SPECT (reproduced with permission from ref.[28]).

"Gold standard"	# reports	# patients	Spearman's r (LVEF)	Radiopharmaceutical	Reference(s)
MUGA	15	585	0.7–0.94	^{99m}Tc-sestamibi, ^{99m}Tc-tetrofosmin, ^{201}Tl	42–44, 46, 56, 60, 65–70
first pass	12	533	0.82–0.92	^{99m}Tc-sestamibi, ^{201}Tl, ^{123}I BMIPP	9, 25, 50, 68, 71–76
2-D echo	11	643	0.72–0.90	^{99m}Tc-sestamibi, ^{99m}Tc-tetrofosmin, ^{201}Tl	29–33, 35, 59, 77, 78
MRI	11	238	0.71–0.94	^{99m}Tc-sestamibi, ^{99m}Tc-tetrofosmin, ^{201}Tl	9, 36–41, 61, 79
contrast ventriculography	8	357	0.26–0.97	^{99m}Tc-sestamibi, ^{99m}Tc-tetrofosmin, ^{201}Tl	45–47, 76, 78, 80–82
3-D echo	1	18	0.80	^{201}Tl	48
3-D MUGA	1	10	0.97	^{99m}Tc-tetrofosmin	81
thermodilution	1	21	0.84	^{99m}Tc-sestamibi	51
EBCT	1	10	0.94	^{99m}Tc-sestamibi	83
TOTAL	61	2,415	0.86		

As with absolute LV cavity volumes, LVEF measurements derived by different quantitative algorithms applied to the same patient images are not necessarily interchangeable. Indeed, although cross-algorithm reproducibility studies have shown strong linear correlations,[26,60,84-87] it has been suggested that systematic differences between the measurements may exist, preventing the direct merging of differently analyzed data in the context of multi-center trials.[35,84,87-89]

Normal limits for 8-frame LVEF and LV volumes have been reported,[90-92] and appear to be gender-specific[91-94] but independent of radioisotope or type of camera used.[92] As would be expected, normal limits have been reported to be dependent on the specific quantification algorithm used to determine them.[91,93,95] With respect to the QGS algorithm,[25] normal thresholds of 45% (LVEF), 120 ml (EDV) and 70 ml (ESV) had originally been determined in a mixed patient population,[90] and more recently have been refined to 45% and 50% (LVEF), 142 ml and 100 ml (EDV), 70 ml/m^2 and 56 ml/m^2 (EDV, indexed to body surface), 65 ml and 42 ml (ESV), and 32 ml/m^2 and 25 ml/m^2 (ESV, indexed to body surface) for males and females, respectively.

3.3 Myocardial Wall Motion and Thickening

The same 3-dimensional endocardial surface of the LV used for volumes and LVEF calculation can be followed throughout the cardiac cycle to derive quantitative measurements of regional myocardial wall motion (WM),

generally with respect to a coordinate-less or ''centerline'' reference system.[96-98] Measurements of regional myocardial wall thickening (WT) would theoretically require tracking both the endocardial and the epicardial surfaces throughout the cycle - however, an alternative approach is to assume a linear relationship between the increase in maximal counts in a myocardial segment from diastole to systole and the physical thickening of that segment in the same time interval, consistent with the phenomenon termed ''partial volume effect''.[99]

While WM and WT measurements are essentially equivalent, because in order to thicken the myocardium must necessarily move (and vice versa), there are exceptions. For example, small infarcted portions of the myocardium can be ''tethered'' by surrounding healthy cardiac muscle, thus appearing as having normal motion but absent thickening. On the other hand, it is well known that typically post-cardiac surgery patients have septal dyskinesis in the presence of normal septal thickening.[100,101] Therefore, it is important to evaluate *both* regional motion and thickening.

Of note, Figure 5 demonstrates that neither myocardial wall motion nor thickening are uniform across the myocardium of a normal patient;[102] therefore, unlike with myocardial perfusion, dark areas in the motion or thickening polar maps cannot be considered a visual clue that myocardial function is abnormal in those areas. These physiologic findings have already been described for echocardiography: ''*normal segmental cavity shrinkage has been reported to vary from 0 percent (i.e., akynesis) to 100 percent and segmental wall thickening to vary from 0 to 150 percent*'',[103] and for magnetic resonance imaging (MRI), where the ''normal'' thickening range of 18%-100% depends of the specific myocardial region considered.[104]

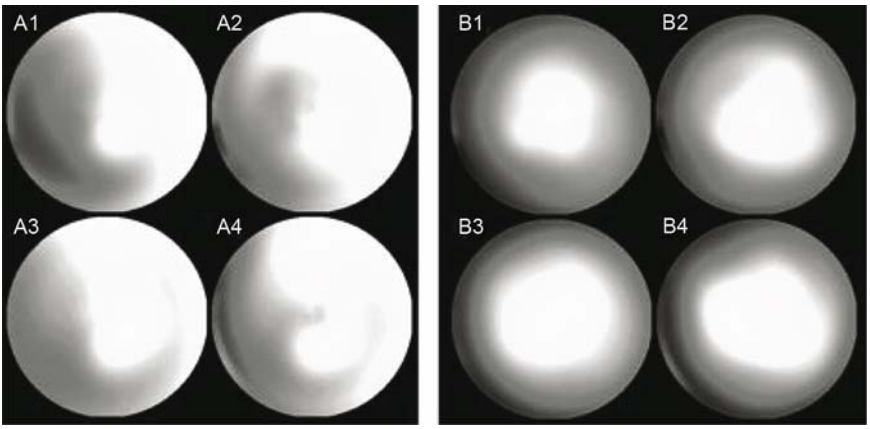


FIGURE 5. (left) Normal myocardial wall motion patterns for (A1) males, ^{99m}Tc-sestamibi, (A2) males, ^{201}Tl, (A3) females, ^{99m}Tc-sestamibi, and (A4) females, ^{201}Tl. (right) Normal myocardial wall thickening patterns for (B1) males, ^{99m}Tc-sestamibi, (B2) males, ^{201}Tl, (B3) females, ^{99m}Tc-sestamibi, and (B4) females, ^{201}Tl.

Absolute measurements of WM and WT have greater clinical value if translated into an assessment of function abnormality, through the development of region-specific normal limits[105-109] and criteria for abnormality.[110] Normal databases are usually quite independent of radioisotope type and patient gender (Figure 5), and are developed based on an expert observer's classification of segmental motion and thickening using a categorical scale and a multi-segment model of the myocardium. Typical models are the 20- or 17-segment one already described for perfusion assessment, where each segment is scored on a 6-point (WM) or 4-point scale (WT). Semi-quantitative scoring of myocardial wall motion and thickening can now be generated automatically by computer algorithms based on normal limits, and has been validated against reference semi-quantitative visual assessment.[110-112]

3.4 Diastolic Function

While it is generally agreed that 8-frame gated SPECT imaging does not allow for meaningful measurement of LV diastolic function,[113] it has been reported that 12-frame imaging may be somewhat adequate,[114] 16-frame imaging quite effective,[115] and 32-frame imaging feasible and resulting in excellent agreement with a MUGA standard.[116-119] Parameters of LV diastolic function that can be quantitatively measured from gated myocardial perfusion SPECT images include the peak filling rate (PFR), the time at which the PFR occurs (TPFR), the peak ejection rate (PER), the time at which the PER occurs (TPER), and the mean filling fraction (1/3 FF), each of these parameters corresponding to a specific feature of the time-volume curve and its derivative curve.

4. Quantitation of Other LV Parameters

4.1 Transient Ischemic Dilation of the LV

Transient ischemic dilation (TID) of the LV is considered present when the left ventricular cavity appears to be significantly larger in the post-stress images than in the resting images, and a TID index can be easily calculated as the ratio of the post-stress to the rest LV cavity volume, whether gated or ungated. It should be noted that what is referred to as transient ischemic dilation may actually be an apparent cavity dilation secondary to diffuse subendocardial ischemia (obscuring the endocardial border). This phenomenon is likely to explain why "transient ischemic dilation" may be seen for several hours following stress, when true cavity dilation is probably no longer present. The pattern of transient ischemic dilation of the left ventricle has similar clinical implications whether it is observed on exercise or pharmacologic stress studies,[120] in that patients who have TID index above a

certain threshold are likely to have severe and extensive coronary artery disease (> 90% stenosis of the proximal left anterior descending coronary artery, or of multiple vessels).[121]

The actual threshold value that identifies an abnormal TID index depends on the imaging protocol used. For example, with dual-isotope protocols, the greater Compton scatter associated with ^{201}Tl causes the myocardial walls to appear intrinsically thicker (and the cavity smaller) in ^{201}Tl rest images compared to post-stress Tc-99m sestamibi images.[121] Therefore, a greater degree of transient enlargement must be evident for a dual-isotope study to be considered to demonstrate transient ischemic dilation. Published thresholds include TID index = 1.22-123 for a rest ^{201}Tl/ post-exercise stress ^{99m}Tc-sestamibi protocol,[121,122] TID index = 1.14 for a same-day post-exercise stress/ rest ^{99m}Tc-sestamibi protocol,[122] and TID index = 1.27-1.40 for a rest ^{201}Tl/ post-pharmacologic stress ^{99m}Tc-sestamibi or ^{99m}Tc-tetrofosmin protocol using common pharmacologic stressors.[123] It has recently been suggested that TID values may be gender-dependent and higher in males,[124] while the use of end-systolic volumes[125] or end-diastolic volumes in the calculation of the TID index is likely to be further explored in the future.

4.2 Lung/Heart Ratio

The degree of lung uptake of myocardial perfusion tracers has traditionally been assessed from visual inspection of the raw projection images, based on the fact that there is a strong linear relationship between the degree of lung uptake and the pulmonary capillary wedge pressure at the time of injection.[126,127] Additionally, a lung/heart ratio (LHR) can be quantified by automatically[128] or manually placing two regions of interest over the most normal cardiac zone and a representative pulmonary region of an anterior or LAO 45° projection image, and ratioing the average as maximum pixel counts in the lung and heart regions (Figure 6).

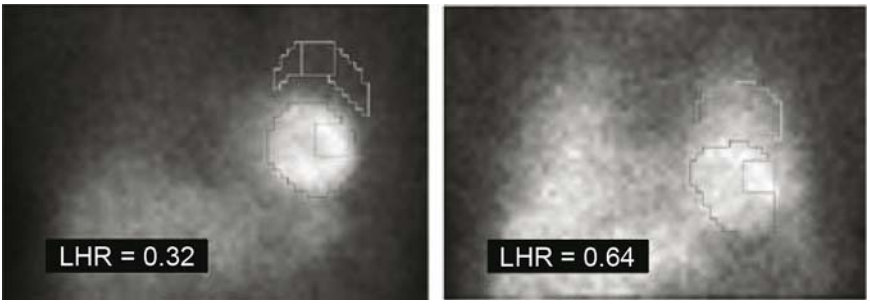


FIGURE 6. Circular regions of interest (ROIs) outlining the LV myocardium and crescent-shaped ROIs sampling the pulmonary area for (left) a normal volunteer and (right) a patient with severe and extensive coronary artery disease. Square ROIs denote maximal count pixels within the larger ROIs.

In general, quantitative lung/heart ratios for ^{201}Tl have been shown to have an upper limit of normal of 0.54,[129-131] while preliminary data for Tc-99m sestamibi suggests an upper limit of 0.44.[132] As with the TID index, abnormal lung/heart ratios are associated with severe and extensive coronary artery disease - however, a study involving 4,618 consecutive patients who underwent exercise/redistribution ^{201}Tl SPECT found that, while 351 patients (7.8%) had an abnormal lung/heart ratio (defined as LHR$>$ 0.5) and 322 patients had an abnormal TID index (defined as $>$ 1.1), only 40 patients (0.9%) had both measurements abnormal.[133] These findings suggest that TID and LHR are not correlated, and that quantitation of both parameters offers incremental knowledge over quantitation of either one alone.

4.3 Myocardial Mass

Quantitative measurement of LV myocardial mass is conceptually possible by counting the voxels bound by the 3-dimensional endocardial surface, epicardial surface and valve plane, then multiplying that value by an individual voxel's volume as well as by the myocardial density. As seen with respect to the measurement of myocardial wall thickening, the relatively low resolution of nuclear cardiology images generally prevents us from measuring "small" structures such as myocardial thickness with a high degree of accuracy,[98] and LV myocardial mass determinations will likely be best accomplished using high resolution modalities like MRI.[134] Nevertheless, SPECT algorithms measuring LV mass based on the assumption of variable[135] or constant[9] myocardial thickness have been reported to correlate reasonably well with MRI standards, and can represent a useful adjunctive tool in our quantitative armamentarium.

4.4 Eccentricity

A parameter that is not currently of wide use in nuclear cardiology, but that can be straightforwardly and automatically quantified from gated and ungated SPECT images, is the sphericity or eccentricity of the LV. Since pathologic ventricular remodelling produces alterations of the LV geometry and can be caused not only by ischemic heart disease and cardiomyopathy, but also by valvular heart disease and hypertension,[136] being able to quantitatively evaluate LV shape from SPECT images would theoretically expand the diagnostic and prognostic range of nuclear cardiology techniques into what traditionally has been the province of echocardiography.

As Figure 7 shows, the same model previously described and developed for the optimal sampling of myocardial perfusion can be used to derive the ellipsoid that best fits the maximal count (or mid-myocardial) surface of the LV, and the minor and major axes of that ellipsoid ratioed to assess LV sphericity. Given the high reproducibility and automation of nuclear cardiology algorithms, serial SPECT studies might represent a practical and

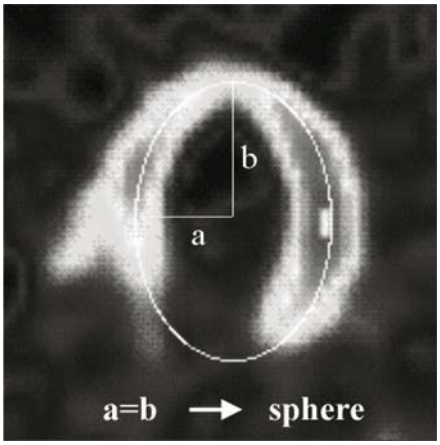


FIGURE 7. Sphericity of the LV is directly related to the relative size of the minor and major axes of its "best fit" ellipsoid.

efficient way to detect even small evolutions in LV remodelling during the course of medical therapy.

5. Summary

A wide range of cardiac parameters can be quantitatively measured using automated and commercially available algorithms applied to myocardial SPECT images. Numerous published validation studies have shown that those quantitative measurements are accurate and reproducible, and provide incremental diagnostic and prognostic value compared to visual assessment alone. Quantitative analysis currently represents one of the key advantages of nuclear cardiology vis-à-vis competing imaging modalities, and will likely further contribute to its growth in the near future.

References

1. E. V. Garcia, C. D. Cooke, K. F. Van Train, et al., Technical aspects of myocardial SPECT imaging with technetium-99m sestamibi. *Am J Cardiol* **66:** 23E-31E (1990).
2. E. E. DePasquale, A. C. Nody, E. G. DePuey, et al., Quantitative rotational thallium-201 tomography for identifying and localizing coronary artery disease. *Circulation* **77:** 316-327 (1988).
3. G. Germano, K. Van Train, H. Kiat, and D. Berman, in: *Diagnostic nuclear medicine,* edited by M. P. Sandler (Williams & Wilkins, Baltimore, 1995), pp. 347-386.
4. A. Rozanski, G. A. Diamond, J. S. Forrester, et al., Alternative referent standards for cardiac normality. Implications for diagnostic testing. *Ann Int Med* **101:** 164-171 (1984).

5. K. F. Van Train, J. Areeda, E. V. Garcia, et al., Quantitative same-day rest-stress technetium-99m-sestamibi SPECT: definition and validation of stress normal limits and criteria for abnormality. *J Nucl Med* **34:** 1494-1502 (1993).
6. E. Ficaro, J. Kritzman, and J. Corbett, Development and clinical validation of normal Tc-99m sestamibi database: comparison of 3D-MSPECT to CEqual. (abstract) **40:** 125P (1999).
7. G. Germano, P. B. Kavanagh, P. Waechter, et al., A new algorithm for the quantitation of myocardial perfusion SPECT. I: technical principles and reproducibility. *J Nucl Med* **41:** 712-719 (2000).
8. P. J. Slomka, G. A. Hurwitz, J. Stephenson, and T. Cradduck, Automated alignment and sizing of myocardial stress and rest scans to three-dimensional normal templates using an image registration algorithm. *J Nucl Med* **36:** 1115-1122 (1995).
9. T. L. Faber, C. D. Cooke, R. D. Folks, et al., Left ventricular function and perfusion from gated SPECT perfusion images: an integrated method. *J Nucl Med* **40:** 650-659 (1999).
10. S. Kirac, F. J. Wackers, and Y. H. Liu, Validation of the Yale circumferential quantification method using ^{201}Tl and ^{99m}Tc: a phantom study. *J Nucl Med* **41:** 1436-1441 (2000).
11. D. S. Berman, H. Kiat, K. Van Train, E. Garcia, et al., Technetium 99m sestamibi in the assessment of chronic coronary artery disease [see comments]. *Sem Nucl Med* **21:** 190-212 (1991).
12. S. Reisman, D. Berman, J. Maddahi, and H. J. Swan, The severe stress thallium defect: an indicator of critical coronary stenosis. *Am Heart J* **110:** 128-134 (1985).
13. L. Matzer, H. Kiat, K. Van Train, et al., Quantitative severity of stress thallium-201 myocardial perfusion single-photon emission computed tomography defects in one-vessel coronary artery disease. *Am J Cardiol* **72:** 273-279 (1993).
14. T. Sharir, G. Germano, P. B. Waechter, et al. A new algorithm for the quantitation of myocardial perfusion SPECT. II: validation and diagnostic yield. *J Nucl Med* **41:** 720-727 (2000).
15. R. Hachamovitch, D. S. Berman, H. Kiat, et al., Exercise myocardial perfusion SPECT in patients without known coronary artery disease: incremental prognostic value and use in risk stratification. *Circulation* **93:** 905-914 (1996).
16. R. Hachamovitch, D. S. Berman, L. J. Shaw, et al., Incremental prognostic value of myocardial perfusion single photon emission computed tomography for the prediction of cardiac death: differential stratification for risk of cardiac death and myocardial infarction. *Circulation* **97:** 535-543 (1998).
17. D. Berman, L. Shaw, and G. Germano, Nuclear cardiology, In *Hurst's the heart,* edited by V. Fuster (McGraw-Hill Medical Pub. Division, New York, 2001), pp. 525-565.
18. M. D. Cerqueira, N. J. Weissman, V. Dilsizian, et al., Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: A statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. *J Nucl Cardiol* **9:** 240-245 (2002).
19. D. S. Berman, X. P. Kang, A. Abidov, et al., Prognostic value of myocardial perfusion SPECT comparing 17-segment and 20-segment scoring systems. (abstract) *J Am Coll Cardiol* **41:** 445A (2003).

20. J. De Sutter, C. Van de Wiele, Y. D'Asseler, et al., Automatic quantification of defect size using normal templates: a comparative clinical study of three commercially available algorithms. *Eur J Nucl Med* **27:** 1827-1834 (2000).
21. J. Krasnow, I. Trask, S. Dahlberg, et al., Automatic determination of SPECT perfusion defect size and reversibility; comparison of four quantitative software programs. *J Nucl Cardiol* **8:** S129 (abstract) (2001).
22. K. F. Van Train, E. V. Garcia, J. Maddahi, et al., Multicenter trial validation for quantitative analysis of same-day rest-stress technetium-99m-sestamibi myocardial tomograms. *J Nucl Med* **35:** 609-618 (1994).
23. K. Nakajima, J. Taki, T. Higuchi, et al., Gated SPET quantification of small hearts: mathematical simulation and clinical application. *Eur J Nucl Med* **27:** 1372-1379 (2000).
24. P. V. Ford, S. N. Chatziioannou, W. H. Moore, and R. D. Dhekne, Overestimation of the LVEF by quantitative gated SPECT in simulated left ventricles. *J Nucl Med* **42:** 454-459 (2001).
25. G. Germano, H. Kiat, P. B. Kavanagh, et al., Automatic quantification of ejection fraction from gated myocardial perfusion SPECT. *J Nucl Med* **36:** 2138-2147 (1995).
26. K. Nakajima, T. Higuchi, J. Taki, et al., Accuracy of ventricular volume and ejection fraction measured by gated myocardial SPECT: Comparison of 4 software programs. *J Nucl Med* **42:** 1571-1578 (2001).
27. P. T. Lam, F. J. T. Wackers, and Y. H. Liu, Validation of a new method for quantification of left ventricular function from ECG-gated SPECT. *J Nucl Med* **42:** 93P-94P (abstract) (2001).
28. G. Germano, and D. Berman, Gated single-photon emission computed tomography. in: *Nuclear cardiac imaging: principles and applications. Ed 3,* edited by A. E. Iskandrian, and M. S. Verani (Oxford University Press, New York, 2003), pp. 121-136.
29. D. Zanger, A. Bhatnagar, E. Hausner, et al., Automated calculation of ejection fraction from gated Tc-99m sestamibi images - comparison to quantitative echocardiography. (abstract) *J Nucl Cardiol* **4:** S78 (1997).
30. T. Bateman, A. Magalski, C. Barnhart, et al., Global left ventricular function assessment using gated SPECT: comparison with echocardiography. (abstract) **31:** 441A (1998).
31. E. Cwajg, J. Cwajg, Z. He, et al., Comparison between gated-SPECT and echocardiography for the analysis of global and regional left ventricular function and volumes. (abstract) **31:** 440A-441A (1998).
32. E. Cwajg, J. Cwajg, Z. X. He, et al., Gated myocardial perfusion tomography for the assessment of left ventricular function and volumes: comparison with echocardiography. *J Nucl Med* **40:** 1857-1865 (1999).
33. D. Mathew, Y. Zabrodina, and F. Mannting, Volumetric and functional analysis of left ventricle by gated SPECT: a comparison with echocardiographic measurements. (abstract) *J Am Coll Cardiol* **31:** 44A (1998).
34. K. Nichols, D. Lefkovitz, T. Faber, et al., Ventricular volumes compared among three gated SPECT methods and echocardiography. (abstract) **33:** 409A (1999).
35. K. Nichols, D. Lefkowitz, T. Faber, et al., Echocardiographic validation of gated SPECT ventricular function measurements. *J Nucl Med* **41:** 308-1314 (2000).

36. Z. He, G. Vick, P. Vaduganathan, and M. Verani, Comparison of left ventricular volumes and ejection fraction measured by gated SPECT and by cine magnetic resonance imaging. (abstract) J Am Coll Cardiol **31:** 44A (1998).
37. D. Atsma, H. Kayser, C. Croon, P. et al., Good correlation between left ventricular ejection fraction, endsystolic and enddiastolic volume measured by gated SPECT as compared to magnetic resonance imaging. (abstract) **33:** 436A (1999).
38. P. Vaduganathan, Z. X. He, G. W. Vick, et al., Evaluation of left ventricular wall motion, volumes, and ejection fraction by gated myocardial tomography with technetium 99m-labeled tetrofosmin: a comparison with cine magnetic resonance imaging. *J Nucl Cardiol* **6:** 3-10 (1999).
39. E. Tadamura, T. Kudoh, M. Motooka, et al., Assessment of regional and global left ventricular function by reinjection T1-201 and rest Tc-99m sestamibi ECG-gated SPECT: comparison with three-dimensional magnetic resonance imaging. *J Am Coll Cardiol* **33:** 991-997 (1999).
40. J. Vansant, R. Pettigrew, T. Faber, J. Galt, U. Bilkay, M. Blais, G. Chatzimavroudis, L. Kimble, N. Chronos, and E. Garcia, Comparison and accuracy of two gated-SPECT techniques for assessing left ventricular function defined by cardiac MRI. (abstract) *J Nucl Med* **40:** 166P (1999).
41. E. Tadamura, T. Kudoh, M. Motooka, et al., Use of technetium-99m sestamibi ECG-gated single-photon emission tomography for the evaluation of left ventricular function following coronary artery bypass graft: comparison with three-dimensional magnetic resonance imaging. *Eur J Nucl Med* **26:** 705-712 (1999).
42. D. Daou, B. Helal, P. Colin, et al., Are LV ejection fraction (EF), end diastolic (EDV) and end systolic volumes (ESV) measured with rest Tl-201 gated SPECT accurate? (abstract) *J Nucl Cardiol* **6:** S31 (1999).
43. J. Yoshioka, S. Hasegawa, H. Yamaguchi, et al., Left ventricular volumes and ejection fraction calculated from quantitative electrocardiographic-gated ^{99m}Tc-tetrofosmin myocardial SPECT. *J Nucl Med* **40:** 1693-1698 (1999).
44. T. Chua, L. C. Yin, T. H. Thiang, et al., Accuracy of the automated assessment of left ventricular function with gated perfusion SPECT in the presence of perfusion defects and left ventricular dysfunction: correlation with equilibrium radionuclide ventriculography and echocardiography. *J Nucl Cardiol* **7:** 301-311 (2000).
45. K. Nichols, J. Tamis, E. G. DePuey, et al., Relationship of gated SPECT ventricular function parameters to angiographic measurements. *J Nucl Cardiol* **5:** 295-303 (1998).
46. P. Vera, A. Manrique, V. Pontvianne, et al., Thallium-gated SPECT in patients with major myocardial infarction: effect of filtering and zooming in comparison with equilibrium radionuclide imaging and left ventriculography. *J Nucl Med* **40:** 513-521 (1999).
47. P. Adiseshan, and J. Corbett, Quantification of left ventricular function from gated tomographic perfusion imaging: development and testing of a new algorithm. (abstract) *Circulation* **90:** I-365 (1994).
48. O. Akinboboye, L. El-Khoury Coffin, R. Sciacca, et al., Accuracy of gated SPECT thallium left ventricular volumes and ejection fractions: comparison with three-dimensional echocardiography. *J Am Coll Cardiol* **31:** 85A (abstract) (1998).
49. C. Cittanti, D. Mele, P. Colamussi, et al., Determination of left ventricular volume and ejection fraction by g-SPECT myocardial perfusion scintigraphy. A

comparison with quantitative 3-D echocardiography. (abstract) *J Nucl Cardiol* **6:** S34 (1999).
50. R. Schwartz, C. Thompson, L. Mixon, et al., Gated SPECT analysis with 3-D wall parametrization method: accurate and reproducible evaluation of left ventricular volumes and ejection fraction. (abstract) *Circulation* **92:** I-449 (1995).
51. G. Germano, W. VanDecker, R. Mintz, et al., Validation of left ventricular volumes automatically measured with gated myocardial perfusion SPECT. (abstract) *J Am Coll Cardiol* **31:** 43A (1998).
52. A. E. Iskandrian, G. Germano, W. VanDecker, et al., Validation of left ventricular volume measurements by gated SPECT ^{99m}Tc-labeled sestamibi imaging. *J Nucl Cardiol* **5:** 574-578 (1998).
53. K. Imai, Y. Azuma, S. Nakajima, et al., Frames a cardiac cycle in quantitative gated SPECT (QGS) for clinical use: 8 versus 16. (abstract) **6:** S17 (1999).
54. C. Cohade, R. Taillefer, A. Gagnon, et al., Effect of the number of frames per cardiac cycle and the amount of injected dose of radionuclide on the determination of left ventricular ejection fraction (LVEF) with gated SPECT myocardial perfusion imaging (GS). (abstract) *J Nucl Med* **41:** 154P (2000).
55. A. Manrique, P. Vera, A. Hitzel, et al., 16-interval gating improves thallium-201 gated SPECT LVEF measurement in patients with large myocardial infarction. (abstract) *J Am Coll Cardiol* **33:** 436A-437A (1999).
56. A. Manrique, R. Koning, A. Cribier, and P. Véra, Effect of temporal sampling on evaluation of left ventricular ejection fraction by means of thallium-201 gated SPET: comparison of 16- and 8-interval gating, with reference to equilibrium radionuclide angiography. *Eur J Nucl Med* **27:** 694-699 (2000).
57. J. Case, S. Cullom, T. Bateman, et al., Overestimation of LVEF by gated MIBI myocardial perfusion SPECT in patients with small hearts. *J Am Coll Cardiol* **31:** 43A (abstract) (1998).
58. M. Santos, H. Lewin, S. Hayes, et al., A potential cause for underestimation of LVEF by QGS. (abstract) *J Nucl Cardiol* **8:** S130 (2001).
59. R. Schwartz, J. Eckdahl, and C. Thompson, 3-D wall parametrization method for quantitative LVEF of gated SPECT sestamibi with LV dysfunction and severe perfusion defects. (abstract) *J Nucl Cardiol* **2:** S114 (1995).
60. H. Everaert, A. Bossuyt, and P. R. Franken, Left ventricular ejection fraction and volumes from gated single photon emission tomographic myocardial perfusion images: comparison between two algorithms working in three-dimensional space. *J Nucl Cardiol* **4:** 472-476 (1997).
61. J. C. Stollfuss, F. Haas, I. Matsunari, et al., Regional myocardial wall thickening and global ejection fraction in patients with low angiographic left ventricular ejection fraction assessed by visual and quantitative resting ECG-gated ^{99m}Tc-tetrofosmin single-photon emission tomography and magnetic resonance imaging. *Eur J Nucl Med* **25:** 522-530 (1998).
62. F. Al-Khori, P. McNelis, and W. Van Decker, Reliability of gated SPECT in assessing left ventricular ejection fraction in ventricles with scarred myocardium. (abstract) *J Nucl Cardiol* **6:** S26 (1999).
63. R. Giubbini, A. Terzi, P. Rossini, and E. Milan, Gated myocardial perfusion single photon emission tomography (GSPECT) in the evaluation of left ventricular ejection fraction in CAD patients with previous myocardial infarction. (abstract) *J Nucl Cardiol* **6:** S58 (1999).

64. J. Bax, H. Lamb, P. Dibbets, et al., Comparison between LV volumes and LVEF assessed by MRI and gated SPECT in patients with severe ischemic LV dysfunction. (abstract) *J Nucl Med* **40:** 45P (1999).
65. M. Moriel, G. Germano, H. Kiat, et al., Automatic measurement of left ventricular ejection fraction by gated SPECT Tc-99m sestamibi: a comparison with radionuclide ventriculography. (abstract) *Circulation* **88:** I-486 (1993).
66. T. Bateman, J. Case, M. Saunders, et al., Gated SPECT LVEF measurements using a dual-detector camera and a weight-adjusted dosage of thallium-201. (abstract) *J Am Coll Cardiol* **29:** 263A (1997).
67. P. Carpentier, H. Benticha, P. Gautier, and C. Sulman, Thallium 201 gated SPECT for simultaneous assessment of myocardial perfusion, left ventricular ejection fraction and qualitative regional function. (abstract) *J Nucl Cardiol* **6:** S39 (1999).
68. K. Nichols, E. G. DePuey, and A. Rozanski, Automation of gated tomographic left ventricular ejection fraction. *J Nucl Cardiol* **3:** 475-482 (1996).
69. H. Everaert, P. R. Franken, P. Flamen, et al., Left ventricular ejection fraction from gated SPET myocardial perfusion studies: a method based on the radial distribution of count rate density across the myocardial wall. *Eur J Nucl Med* **23:** 1628-1633 (1996).
70. D. A. Calnon, R. J. Kastner, W. H. Smith, et al., Validation of a new counts-based gated single photon emission computed tomography method for quantifying left ventricular systolic function: comparison with equilibrium radionuclide angiography. *J Nucl Cardiol* **4:** 464-471 (1997).
71. Z. He, J. Mahmarian, J. Preslar, and M. Verani, Correlations of left ventricular ejection fractions determined by gated SPECT with thallium and sestamibi and by first-pass radionuclide angiography. (abstract) *J Nucl Med* **38:** 27P (1997).
72. M. Inubushi, E. Tadamura, T. Kudoh, et al., Simultaneous assessment of myocardial fatty acid utilization and LV function using I-123 BMIPP gated SPECT (GSPECT). (abstract) *J Nucl Cardiol* **6:** S66 (1999).
73. M. Inubushi, E. Tadamura, T. Kudoh, et al., Simultaneous assessment of myocardial free fatty acid utilization and left ventricular function using 123I-BMIPP-gated SPECT. *J Nucl Med* **40:** 1840-1847 (1999).
74. Z. X. He, E. Cwajg, J. S. Preslar, et al., Accuracy of left ventricular ejection fraction determined by gated myocardial perfusion SPECT with Tl-201 and Tc-99m sestamibi: comparison with first-pass radionuclide angiography. *J Nucl Cardiol* **6:** 412-417 (1999).
75. K. Nichols, E. G. DePuey, A. Rozanski, et al., Image enhancement of severely hypoperfused myocardia for computation of tomographic ejection fraction. *J Nucl Med* **38:** 1411-1417 (1997).
76. K. A. Williams, and L. A. Taillon, Left ventricular function in patients with coronary artery disease assessed by gated tomographic myocardial perfusion images. Comparison with assessment by contrast ventriculography and first-pass radionuclide angiography. *J Am Coll Cardiol* **27:** 173-181 (1996).
77. C. Bacher-Stier, S. Müller, O. Pachinger, et al., Thallium-201 gated single-photon emission tomography for the assessment of left ventricular ejection fraction and regional wall motion abnormalities in comparison with two-dimensional echocardiography. *Eur J Nucl Med* **26:** 1533-1540 (1999).

78. C. Di Leo, A. Bestetti, L. Tagliabue, et al., ^{99m}Tc-tetrofosmin gated-SPECT LVEF: correlation with echocardiography and contrastographic ventriculography. (abstract) **4:** S56 (1997).
79. J. C. Stollfuss, F. Haas, I. Matsunari, et al., ^{99m}Tc-tetrofosmin SPECT for prediction of functional recovery defined by MRI in patients with severe left ventricular dysfunction: additional value of gated SPECT. *J Nucl Med* **40:** 1824-1831 (1999).
80. D. Atsma, C. Croon, P. Dibbets-Schneider, et al., Good correlation between left ventricular ejection fraction and wall motion score assessed by gated SPECT as compared to left ventricular angiography. (abstract) J Am Coll Cardiol **33:** 409A (1999).
81. A. Paul, S. Hasegawa, J. Yoshioka, et al., Left ventricular volume and ejection fraction from quantitative gated SPECT: comparison with gated pool SPECT and contrast ventriculography. *J Nucl Med* **40:** 178P (abstract) (1999).
82. E. Ficaro, R. Quaife, J. Kritzman, and J. Corbett, Accuracy and reproducibility of 3D-MSPECT for estimating left ventricular ejection fraction in patients with severe perfusion abnormalities. (abstract) *Circulation* **100:** I-26 (1999).
83. M. Toba, Y. Ishida, K. Fukuchi, et al., Application of ECG-gated Tc-99m sestamibi cardiac imaging to patients with arrhythmogenic right ventricular dysplasia (ARVD). (abstract) *J Nucl Cardiol* **6:** S41 (1999).
84. K. Nichols, R. Folks, D. Cooke, et al., Comparisons between "ECTb" and "QGS" software to compute left ventricular function from myocardial perfusion gated SPECT data. (abstract) *J Nucl Cardiol* **7:** S20 (2000).
85. P. Véra, R. Koning, A. Cribier, and A. Manrique, Comparison of two three-dimensional gated SPECT methods with thallium in patients with large myocardial infarction. *J Nucl Cardiol* **7:** 312-319 (2000).
86. J. Krasnow, I. Trask, S. Dahlberg, G. Margulis, and J. Leppo, Automatic determination of left ventricular function (LVEF) by gated SPECT; comparison of four quantitative software programs. (abstract) *J Nucl Cardiol* **8:** S138 (2001).
87. T. Lewis, K. Grewal, and D. Calnon, Discrepancies in estimating left-ventricular volumes and ejection fraction by two commercially available gated SPECT algorithms: comparison to echocardiography. (abstract) *J Nucl Cardiol* **8:** S18 (2001).
88. P. R. Franken, H. Everaert, A. Momen, and C. Vanhove, Comparison of three automatic software to measure left ventricular cavity volume and ejection fraction from perfusion gated tomograms. (abstract) *Eur J Nucl Med* **26:** 1076 (1999).
89. M. R. Boussaha, G. Storto, C. Antonescu, and A. B. Delaloye, Ejection fraction evaluation by gated myocardial perfusion SPECT: Comparison between gated spect quantification (GSQ) and emory cardiac tool box (ECTB). (abstract) *Eur J Nucl Med* **28:** OS240 (2001).
90. T. Sharir, G. Germano, P. B. Kavanagh, et al., Incremental prognostic value of post-stress left ventricular ejection fraction and volume by gated myocardial perfusion single photon emission computed tomography. *Circulation* **100:** 1035-1042 (1999).
91. A. Rozanski, K. Nichols, S. S. Yao, et al., Development and application of normal limits for left ventricular ejection fraction and volume measurements from ^{99m}Tc-sestamibi myocardial perfusion gates SPECT. *J Nucl Med* **41:** 1445-1450 (2000).

92. A. A. Ababneh, R. R. Sciacca, B. Kim, and S. R. Bergmann, Normal limits for left ventricular ejection fraction and volumes estimated with gated myocardial perfusion imaging in patients with normal exercise test results: influence of tracer, gender, and acquisition camera. *J Nucl Cardiol* **7:** 661-668 (2000).
93. T. Sharir, G. Germano, J. Friedman, et al., Prognostic value of gated myocardial perfusion single photon emission computed tomography in women versus men. (abstract) *Circulation* **102:** II-544 (2000).
94. P. De Bondt, C. Van de Wiele, J. De Sutter, et al., Age- and gender-specific differences in left ventricular cardiac function and volumes determined by gated SPET. *Eur J Nucl Med* **28:** 620-624 (2001).
95. C. A. Santana, E. V. Garcia, R. Folks, et al., Comparison of normal values of left ventricular function between two programs: QGS and emory cardiac toolbox (ECTB). (abstract) *J Nucl Med* **42:** 166P (2001).
96. F. H. Sheehan, H. T. Dodge, D. Mathey, et al., *Application of the centerline method: analysis of change in regional left ventricular wall motion in serial studies*. In *Computers in Cardiology. Ninth Meeting of Computers in Cardiology*. Seattle, WA, USA: IEEE Computer Society Press, *pp* 97-100 (1983).
97. T. L. Faber, M. S. Akers, R. M. Peshock, and J. R. Corbett, Three-dimensional motion and perfusion quantification in gated single-photon emission computed tomograms. *J Nucl Med* **32:** 2311-2317 (1991).
98. G. Germano, J. Erel, H. Lewin, et al., Automatic quantitation of regional myocardial wall motion and thickening from gated technetium-99m sestamibi myocardial perfusion single-photon emission computed tomography. *J Am Coll Cardiol* **30:** 1360-1367 (1997).
99. W. H. Smith, R. J. Kastner, D. A. Calnon, et al., Quantitative gated single photon emission computed tomography imaging: a counts-based method for display and measurement of regional and global ventricular systolic function. *J Nucl Cardiol* **4:** 451-463 (1997).
100. D. De Nardo, Q. Caretta, C. Mercanti,et al., Effects of uncomplicated coronary artery bypass graft surgery on global and regional left ventricular function at rest. Study by equilibrium radionuclide angiocardiography. *Cardiology* **76:** 285-292 (1989).
101. S. Canclini, P. Rossini, A. Terzi, et al., Gated SPECT (GSPECT) evaluation of septal wall motion after cardiac surgery. (abstract) *J Nucl Med* **41:** 125P (2000).
102. I. Adachi, K. Morita, M. B. Imran, et al., Heterogeneity of myocardial wall motion and thickening in the left ventricle evaluated with quantitative gated SPECT. *J Nucl Cardiol* **7:** 296-300 (2000).
103. A. Katz, T. Force, E. Folland, et al., in: *Marcus cardiac imaging: a companion to Braunwald's Heart disease,* edited by M. L. Marcus, and E. Braunwald (Saunders, Philadelphia), *pp* 297-324 (1996).
104. U. Sechtem, B. A. Sommerhoff, W. Markiewicz, et al., Regional left ventricular wall thickening by magnetic resonance imaging: evaluation in normal persons and patients with global and regional dysfunction. *Am J Cardiol* **59:** 145-151 (1987).
105. C. Cooke, E. Garcia, R. Folks, and J. Ziffer, Myocardial thickening and phase analysis from Tc-99m sestamibi multiple gated SPECT: development of normal limits. (abstract) *J Nucl Med* **33:** 926-927 (1992).

106. S. Shirakawa, N. Hattori, N. Tamaki, et al., [Assessment of left ventricular wall thickening with gated ^{99m}Tc-MIBI SPECT-value of normal file]. *Kaku Igaku* **32:** 643-650 (1995).
107. Y. Itoh, I. Adachi, T. Kohya, et al., Heterogeneity in myocardial perfusion, wall motion and wall thickening with Tc-99m-sestamibi quantitative gated SPECT in normal subjects. **40:** 165P (1999).
108. S. Fujino, K. Masuyama, S. Kanayama, et al., Early and delayed technetium-99m labeled sestamibi myocardial ECG-gated SPECT by QGS program in normal volunteers. *J Nucl Med* **40:** 180P (1999).
109. H. Everaert, C. Vanhove, and P. R. Franken, Effects of low-dose dobutamine on left ventricular function in normal subjects as assessed by gated single-photon emission tomography myocardial perfusion studies. *Eur J Nucl Med* **26:** 1298-1303 (1999).
110. T. Sharir, D. S. Berman, P. B. Waechter, et al., Quantitative analysis of regional motion and thickening by gated myocardial perfusion SPECT: Normal heterogeneity and criteria for abnormality. *J Nucl Med* **42:** 1630-1638 (2001).
111. E. P. Ficaro, J. N. Kritzman, and J. R. Corbett, Automatic segmental scoring of myocardial wall thickening and motion: validation of a new semi-quantitative algorithm. *J Nucl Med* **42:** 171P (2001).
112. D. Berman, and G. Germano, An approach to the interpretation and reporting of gated myocardial perfusion SPECT. In: *Clinical gated cardiac SPECT,* edited by G. Germano, and D. Berman (Futura Publishing Company, Armonk, NY), pp. 147-182 (1999).
113. Y. Damrongpipatkij, F. Mohammed, E. Brown, et al., Quantitative cardiac SPECT: measuring diastolic function. (abstract) *J Nucl Med* **41:** 154P (2000).
114. T. Higuchi, J. Taki, T. Yoneyama, et al., Diastolic and systolic parameters obtained by myocardial ECG-gated perfusion study. (abstract) *J Nucl Med* **41:** 160P (2000).
115. K. Nakajima, J. Taki, M. Kawano, et al., Diastolic dysfunction in patients with systemic sclerosis detected by gated myocardial perfusion SPECT: an early sign of cardiac involvement. *J Nucl Med* **42:** 183-188 (2001).
116. M. Kikkawa, T. Nakamura, K. Sakamoto, et al., Assessment of left ventricular diastolic function from quantitative electrocardiographic-gated (99)mTc-tetrofosmin myocardial SPET (ERRATA **28:** 1579; 2001). *Eur J Nucl Med* **28:** 593-601 (2001).
117. S. Kumita, K. Cho, H. Nakajo, et al., Assessment of left ventricular diastolic function with electrocardiography-gated myocardial perfusion SPECT: Comparison with multigated equilibrium radionuclide angiography. *J Nucl Cardiol* **8:** 568-574 (2001).
118. T. Higuchi, K. Nakajima, J. Taki, et al., The accuracy of left-ventricular time volume curve derived from ECG-gated myocardial perfusion SPECT. (abstract) **8:** S18 (2001).
119. T. Higuchi, K. Nakajima, J. Taki, et al., Assessment of left ventricular systolic and diastolic function based on the edge detection method with myocardial ECG-gated SPET. *Eur J Nucl Med* **28:** 1512-1516 (2001).
120. P. Chouraqui, E. A. Rodrigues, D. S. Berman, and J. Maddahi, Significance of dipyridamole-induced transient dilation of the left ventricle during thallium-201 scintigraphy in suspected coronary artery disease. *Am J Cardiol* **66:** 689-694 (1990).

121. M. Mazzanti, G. Germano, H. Kiat, et al., Identification of severe and extensive coronary artery disease by automatic measurement of transient ischemic dilation of the left ventricle in dual-isotope myocardial perfusion SPECT. *J Am Coll Cardiol* **27:** 1612-1620 (1996).
122. J. N. Kritzman, E. P. Ficaro, and J. R. Corbett, Post-stress LV dilation: The effect of imaging protocol, gender and attenuation correction. (abstract) *J Nucl Med* **42:** 50P (2001).
123. K. Williams, C. Schnieder, and A. Jain, Transient ischemic dilatation (TID) with pharmacological stress dual isotope SPECT. (abstract) *Circulation* **102:** II-546 (2000).
124. S. Madison, M. Dalipaj, and T. Ruddy, Effects of gender and stress on transient ischemic dilation ratios in normals. (abstract) *J Nucl Cardiol* **10:** S85 (2003).
125. A. Bestetti, C. Di Leo, A. Alessi, et al., Transient left ventricular dilation during myocardial perfusion gated-SPECT in hypertensive patients. (abstract) *Eur J Nucl Med* **28:** OS239 (2001).
126. P. Liu, M. Kiess, R. D. Okada, et al., Increased thallium lung uptake after exercise in isolated left anterior descending coronary artery disease. *Am J Cardiol* **55:** 1469-1473 (1985).
127. E. E. Martinez, S. F. Horowitz, H. J. Castello, et al., Lung and myocardial thallium-201 kinetics in resting patients with congestive heart failure: correlation with pulmonary capillary wedge pressure. *Am Heart J* **123:** 427-432 (1992).
128. G. Germano, P. B. Kavanagh, and D. S. Berman, An automatic approach to the analysis, quantitation and review of perfusion and function from myocardial perfusion SPECT images. *Int J Card Imaging* **13:** 337-346 (1997).
129. S. V. Aksut, C. Mallavarapu, J. Russell, et al., Implications of increased lung thallium uptake during exercise single photon emission computed tomography imaging. *Am Heart J* **130:** 367-373 (1995).
130. D. Jain, B. Thompson, F. J. Wackers, and B. L. Zaret, Relevance of increased lung thallium uptake on stress imaging in patients with unstable angina and non-Q wave myocardial infarction: results of the Thrombolysis in Myocardial Infarction (TIMI)-IIIB Study. *J Am Coll Cardiol* **30:** 421-429 (1997).
131. R. A. Vaccarino, L. L. Johnson, M. L. Antunes, et al., Thallium-201 lung uptake and peak treadmill exercise first-pass ejection fraction. *Am Heart J* **129:** 320-329 (1995).
132. C. Bacher-Stier, T. Sharir, P. B. Kavanagh, et al., Postexercise lung uptake of ^{99m}Tc-sestamibi determined by a new automatic technique: validation and application in detection of severe and extensive coronary artery disease and reduced left ventricular function. *J Nucl Med* **41:** 1190-1197 (2000).
133. C. L. Hansen, R. Sangrigoli, E. Nkadi, and M. Kramer, Comparison of pulmonary uptake with transient cavity dilation after exercise thallium-201 perfusion imaging. *J Am Coll Cardiol* **33:** 1323-1327 (1999).
134. C. B. Higgins, Which standard has the gold? [editorial; comment]. *J Am Coll Cardiol* **19:** 1608-1609 (1992).
135. T. Mochizuki, K. Murase, H. Tanaka, et al., Assessment of left ventricular volume using ECG-gated SPECT with technetium-99m-MIBI and technetium-99m-tetrofosmin. *J Nucl Med* **38:** 53-57 (1997).
136. M. Maytin, and W. Colucci, Molecular and cellular mechanisms of myocardial remodeling. *J Nucl Cardiol* **9:** 319-327 (2003).

16
Quantitative Analysis in Nuclear Oncologic Imaging

D.A. Mankoff*, M. Muzi* and H. Zaidi†

1. Introduction

Diagnosis, staging, treatment, prognosis and follow-up are the principal elements in the management of cancer, and nuclear medicine plays an important role in all these elements. Among all diagnostic and therapeutic procedures, nuclear medicine is unique in that it is based on molecular and pathophysiological mechanisms, and employs radioactively labelled biological molecules as tracers to study the pathophysiology of the tumour *in vivo* to direct treatment and assess response to therapy.

Historically, nuclear medicine played an important role in the diagnosis and therapy of cancer starting from the pioneering applications of ^{131}I in the treatment of thyroid cancer. The use of planar (2D) and tomographic (3D) single-photon imaging for lesion detection, dosimetry and planning prior to therapy was revitalized with the renewed interest in radioimmunodiagnosis and radioimmunotherapy.

While PET was originally used as a research tool, in recent years, it has gained an increasingly important clinical role. The largest current area of clinical use of PET is in oncology, where the most widely used tracer is ^{18}F-fluorodeoxyglucose (FDG). FDG-PET has already had a large beneficial effect on cancer treatment and its use in clinical oncology practice continues to evolve. The development of radiopharmaceuticals targeted to other aspects of tumour biology, including cell growth, cell death, oncogene expression, drug delivery, and tumour hypoxia will greatly expand the ability of clinical investigators to characterize tumours and are likely to be used to guide treatment decisions for patients in coming years.[1] The contribution of PET to understanding the clinical biology of cancer and to guiding targeted, individualized therapy will continue to grow with these new developments. Central to this expanding role in oncology will be the ability to make

*Dr D. A. Mankoff and Dr M. Muzi, Division of Nuclear Medicine, Box 346113, University of Washington Medical Center, Seattle, WA 98195, USA
†PD Dr H. Zaidi, Division of Nuclear Medicine, Geneva University Hospital, CH-1211 Geneva, Switzerland

quantitative interpretations of the PET imaging data. Taking into account the latest developments in the field, PET will be given greater weight in this chapter. Emphasis is given to rigorous quantitative analysis of nuclear oncologic images. The last section discusses practical approaches to image quantification that may facilitate routine clinical PET studies.

2. Quantitative Planar Imaging and SPECT

A detailed overview of progress made in the design of gamma cameras operating in planar imaging and SPECT modes is provided in chapter 1 of this book. Likewise, the quantitative procedures used in planar conjugate-view imaging are described in chapter 13, while chapters 5-8 address image correction techniques required to compensate physical degrading factors hampering the quantitative accuracy in both SPECT and PET.

Quantification of tumour size or volume through segmentation of patient data into volumes of interest (VoIs) is ofen desired in diagnostic and oncologic nuclear imaging. This parameter is of paramount importance for accurate radiation dose estimates when using, for example, the Medical Internal Radiation Dose (MIRD) committee formalism. Chapter 10 presents the methodological basis of image segmentation and summarizes computational algorithmic developments to solve this complex task. If a registration between the radiopharmaceutical distribution image set (SPECT or PET) and structural (CT or MRI) images is available, the VoI outlines defined on the latter can be easily transferred to the 3D space of the reconstructed activity distribution (see Chapter 9). The advent of dual-modality imaging systems offering the possibility of correlated anatomical and functional imaging has revolutionized the practice of nuclear medicine and simplified to a great extent the coregistration procedure (see chapter 2). Conceptually, absorbed dose assessment requires the determination of spatial distribution of mass and radioactivity distribution (spatial analysis) in addition to cumulated activity assessment from the necessarily limited number of imaging and sampling times (time analysis). Chapter 17 summarizes the computational models used for absorbed dose estimates using simple tools and refined models involving the use of patient-specific anatomy and sophisticated Monte Carlo modelling.

Different single-photon emitting radiotracers have been used for diagnosis and therapy in clinical oncology. However, the spatial resolution of planar scintigraphy and SPECT prohibits imaging of small tumours. Moreover, the quantitative accuracy of both methods is limited. This has stimulated the development of positron-emitting analogs to improve detection of small lesions and achieve more accurate quantification of tumour size and uptake. Typical problems related to the use of high energy collimators in connexion with ^{131}I SPECT imaging including object scatter, partial volume errors, scatter and septal penetration in the collimator have been addressed in a

number of publications and proved to be difficult to solve.[2] Quantitative imaging of beta emitters is important for the *in vivo* management of antibody therapy using either ^{32}P or ^{90}Y. Quantitative bremsstrahlung SPECT imaging has been developed and used with limited success for imaging of pure beta emitters by bremsstrahlung detection.[3]

3. Quantitative Imaging for PET

One of the important advantages of PET over other forms of radiotracer imaging is its ability to measure absolute regional radionuclide concentration.[4] As a result, PET can generate quantitative dynamic images of regional radiopharmaceutical uptake, resulting in regional measurements of tracer kinetics. Quantitative PET imaging requires a variety of skills, resources and personnel beyond what is needed for routine clinical imaging. Imaging centres specializing in oncologic PET research are usually populated by multi-disciplinary groups of investigators with backgrounds in tumour biology, statistics, bioengineering, physics, and computer science. The common goal of these investigators is to develop, test, and validate tools for producing quantitative regional estimates of physiologic parameters from dynamic radiotracer studies.

By its very nature, PET is a quantitative imaging process. The stoichiometry of biochemical reactions of well-designed imaging agents permits quantitative estimates of the rate of metabolism from specific interactions of the imaging substrates. The amount of tracer in a tissue at any time depends on its rate of delivery, local biochemical reactions, and both its biological and physical clearance. Quantitative PET accounts for each of these factors to give a numerical assessment of discrete physiological characteristics of tissues without invasive procedures, subjective interpretation, and operator bias. These measures are used to provide quantitative estimates of many dynamic physiologic variables, for example estimating the rate at which a tumour is metabolizing glucose (MRglc). These data are then correlated to clinical variables such as response to therapy or tumour progression. In this way, quantitative disease physiology can be related to clinical disease behaviour. These quantitative measurements can also serve as early surrogate endpoints in clinical therapy trials. Quantitative estimates of tissue uptake of the PET radiotracer has the decided advantages of 1) comparing the current analysis of a patient to a study population with similar disease progression, and 2) assessing a series of analyses of an individual patient through the course of therapy. Assessing the tissue uptake in a quantitative way allows to better manage the therapy for an individual patient and eventually assess overall response of a therapy in a population of patients.

In order to gain as much information from the PET imaging signal, a variety of data must be collected from each study, in addition to the imaging data. This is illustrated in Figure 1. The following section describes these

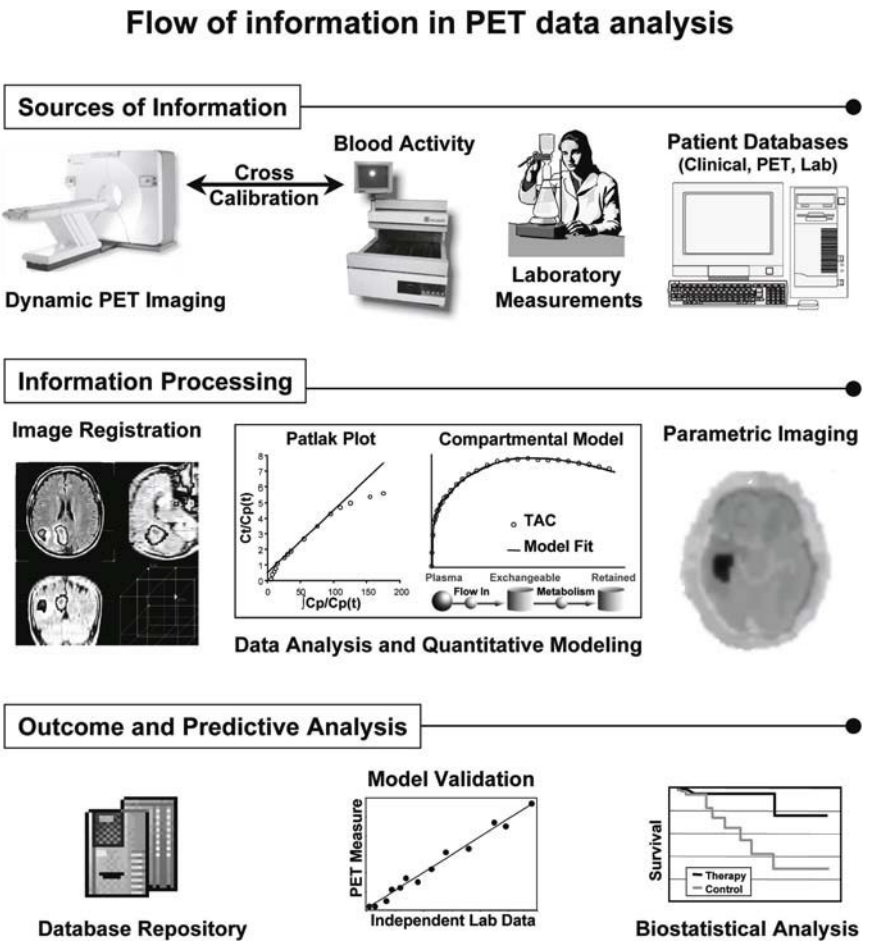


FIGURE 1. The flow of information from acquisition to interpretation of dynamic PET imaging involves an array of procedures to recover the maximum amount of quantitative information from each imaging session. Many decisions in quantitative PET involve tomograph acquisition and reconstruction, cross-calibration between counting devices, blood processing and data management that all depend on the behaviour of the imaging agent and desired study endpoints. Source information is processed through techniques such as multi-modality registration and implementation of physiological models of tracer metabolism required to determine significant metabolic parameters. Parametric imaging is just an extension of modelling, in which a regional quantitative map of a particular parameter or combination of parameters is generated as an image. The database of processed information can then provide the basis of biostatistical analysis on groups of patients using procedures such as predictive outcome analysis.

types of information, their associated errors, and the methods for which they are obtained.

3.1 PET Scanner and Calibration Issues

Maintaining the primary source of tracer uptake information, the PET scanner, is essential for quantitating PET information. The stability and quality of the acquired images relies on support personnel performing routine testing and tuning of the PET scanner. Daily quality assurance tests of the tomograph allows for adjusting the detector gain to maintain calibration. In addition, careful attention to image acquisition protocols, data correction procedures including transmission scanning for attenuation correction, and image reconstruction parameters is important in order to obtain consistent and quantitatively accurate imaging data.

Many quantitative oncologic PET studies require blood samples, both to measure radiotracer blood clearance and also for specific biological assays, such as measuring glucose concentration for FDG studies. To assess blood radioactivity, a sample aliquot is counted in an external counting system, such as a gamma counter, or possibly using an on-line counting device in conjunction with automated blood sampling. Any external counting system must be calibrated with the tomograph in order to relate the resulting activity information to the PET image data. Calibration between the gamma counter and the PET scanner can be performed, typically by imaging a phantom cylinder of known activity. The activity injected into the phantom is measured using a dose calibrator. After PET imaging of the phantom, several aliquots of the solution from the phantom are counted in the gamma counter. The values obtained from PET images and gamma counter are used to provide a calibration factor between the scanner, the well counter, and the dose calibrator. With these calibration factors, all measurements can be converted to standard units of activity concentration such as MBq/cc or μCi/cc.

3.2 Patient Preparation

Prior to imaging, a variety of information should be obtained from the patient. This includes simple data such as height and weight, in addition to factors that may affect the study, for example, dietary conditions prior to the study and any potential interfering medications. Clinical data such as medical condition, history of prior illnesses, and possibly blood analysis should be reviewed to make sure the study can be done safely. Patient comfort is also a consideration for long studies or studies requiring extensive or specialized blood sampling with indwelling catheters.

The patient is positioned in the scanner, immobilized to restrict movement, after which catheter lines are placed in the patient. These lines provide not only access for delivery of the tracer, but also provide for periodic blood

sampling during the scan to assess radioactive delivery to the tissues. Several centres have automated blood sampling devices coordinated with the PET scanner. The coordinated timing of the blood samples with the PET scanner is essential for quantitative assessment. These devices can be programmed to acquire blood samples usually on a time resolution similar to or greater than the PET image acquisition protocol.

The radiotracer is then injected IV, ideally using a syringe pump or similar device to provide a known and repeatable infusion schedule. Similarity in radiotracer delivery minimizes the errors of the blood time activity curve acquired from the blood samples. A common infusion schedule for patients on a protocol also allows for a historical assessment of a population of patients receiving the same radiotracer, as will be discussed below.

During the emission scans, some additional patient requirements may apply. For example, when imaging brain with ^{18}FDG or H^{15}O, any external stimulation, such as noise, activity or lights, should be limited. These factors stimulate the brain causing variation in tracer uptake. Talking to the patient during these brain scans should be done only when necessary. For many quantitative studies, dynamic PET imaging data from a series of scans over time provide an immense amount of information. The timing and duration of these scans depends on the half-life of the nuclide, the kinetics of the tracer, and the analysis model for quantitation.

Dynamic imaging provides some advantages over static imaging, such as 1) tracer uptake and retention is a dynamic process better assessed through a time series of images, and 2) dynamic imaging reduces the bias that may be encountered in choosing a single static time frame to represent tissue uptake. Simple static images possess only a fraction of the information in a dynamic study. However for some situations, a single static image may suffice. Temporal sampling may be reduced later in the course of an investigation given that a more limited set of uptake measures provide the desired information within acceptable error limits (see section 5). The choice of a trade-off between an image protocol complexity and the quantity of data collected will depend on the disease, metabolism of the radiotracer, the level of significance in a population of patients with similar disease, and the desired endpoints of patient assessment. If a static imaging protocol has been validated to assess tracer metabolism, the time after injection and the duration of the imaging frame should be fixed for subsequent patient studies. Variation in the scan start and duration can significantly alter uptake patterns and lead to erroneous conclusions about the imaging study. A good general principle is to begin a new investigation by collecting and processing as much data as possible, and then design simpler, more clinically feasible protocols based upon a thorough understanding of tracer kinetics and how they relate to the disease process under study.

3.3 Blood Sampling – Arterial, Venous, or Image-based

3.3.1 Blood Sampling Approaches

Sampling the patient blood during the imaging study to measure the blood clearance curve gives a direct measure of the circulating levels of radioactivity, and therefore an indication of how much tracer is available for tissue uptake. Capillary concentrations available for tissue uptake are assumed to be identical to peripheral blood measurements. Arterial blood provides the best measure of radiotracer availability to the tissue and is essential in the initial evaluation of radiopharmaceuticals with high first pass extraction. Arterial blood is usually sampled by an automated blood sampler at a time resolution similar to the dynamic series of PET scans. Depending on the tracer, either whole blood (for $H^{15}0$, ^{15}O, ^{11}CO, ^{18}FMISO) or plasma (for ^{18}FDG, ^{18}FLT, ^{11}C-glucose, ^{11}C-thymidine) are collected and counted. The samples are decay corrected to the injection time and converted to conventional units of radioactive concentration (MBq/cc).

Because arterial blood sampling is not always feasible in cancer patients, other blood sampling schemes may be employed to assess the blood clearance curve. Some methods use image-based sampling from an region of interest (RoI) placed over a cardiac chamber or the aorta to acquire a time course of tracer blood concentration.[5] This has the advantage of providing the arterial curve without blood sampling. To account for some of the effects of partial volume and spillover, the image-derived blood time-activity curve can be scaled using late venous samples. The late samples should be acquired at a time after equilibration between arterial and venous blood concentrations, when arterial and venous curves have identical clearance. Rescaling the image-derived blood activity by this method will aid in offsetting potential influence of reconstruction and partial volume that may alter the image-derived blood curve.

For some tracers, venous blood sampling may suffice, especially if first-pass extraction is low. Immersion of the hand in warm water increases shunting through the extremity and may provide an 'arterialized' venous sample, closer to an arterial blood sample.[6] In other applications, a hybrid imaging sequence or 'ping-pong' approach may be employed to assess blood activity. In this approach, the early scans are centred over the heart to measure the early arterial curve from the cardiac blood pool and matched to late venous sampling; the imaging field-of-view is then moved to view tumour uptake.[7] In general, data from a few early studies comparing the arterial and venous blood clearance curves will provide guidance on how to most practically obtain a blood clearance curve while still obtaining the data necessary for kinetic analysis.

3.3.2 Population-based Blood Clearance Curves

Typically, about 25 arterial blood samples are acquired for routine modelling of a dynamic FDG scan following a 1 minute infusion of FDG. However, the shape of the blood clearance curve is quite similar for most patients. An alternate approach to detailed blood sampling is to use a standard, population-based blood clearance curve together with more limited blood sampling. To generate a standardized blood time-activity curve (TAC), the infusion rate, sampling schedule, and blood clearance data from a population of patients are accumulated and combined to form a population curve. Olshen and O'Sullivan[8] proposed a statistical method of combining the blood data from these patients into a population curve, which is then scaled by a few late venous samples to produce an individual patient blood time activity curve. The population based input function can reduce blood sampling to as few as 2 venous samples and incur approximately a 5% error in estimating metabolic rate.[8]

3.3.3 Labelled Metabolites

The radioactivity measured in blood samples is usually assumed to be in a form chemically identical to the injected agent. This is true for radiotracers such as FDG that have minimal metabolism prior to clearance from the blood, but is not for many others such as ^{11}C-thymidine, ^{1}8F-FLT, and ^{11}C-glucose that have notable levels of circulating metabolites.[9] Labelled metabolites result from the biochemical breakdown of the imaging agent and return to the circulation. Labelled metabolites may or may not be taken up in tissues and may therefore confound quantitative image analysis. For tracers with labelled metabolites, independent assays of the precursor and products in the blood may need to be performed to determine the nature of the labelled blood components. Using the total blood activity curve from any of the methods above and the component profile of the imaging agent from an independent assay, an accurate model input function of the radiotracer can be generated for further quantitative analysis. Approaches to modelling tracers with circulating metabolites are given in section 4.4.

3.4 Image Coregistration

Another aspect of quantitative imaging is the registration of PET images to a conventional imaging modality such as CT or MRI. This is often necessary for quantitative image analysis, for example to know where to draw RoIs for tumour and normal tissue. Image coregistration is covered in detail in chapter 9. Many institutions have installed image archive and retrieval systems in their Radiology departments. These services make it quite easy

to obtain MRI or CT images from a digital patient folder. Many software packages exist that perform basic image registration for rigid structures, such as the head. Non-rigid registration algorithms for areas of the body, such as the thorax, are currently under development. With the advent of dual-modality PET/CT scanners, anatomical placement of RoIs has been facilitated. However registration with MRI scans or mis-aligned CT scans will still require software registration. After alignment, regions from the conventional image can be applied directly to the registered PET scan. Aligning PET images to conventional scans allows for accurate assessment of the quantitative PET information in terms of conventional tissue structures.

3.5 Precision of the Data (test/re-test)

Reproducibility of quantitative PET measures is essential for the interpretation of patient images over the course of therapy or even to assess a patient value relative to a population of patients with similar disease. The precision of the PET measure must also be known for activation studies that measure resting and stimulated states, since the error in the estimate may obscure the stimulated effect. The effect of a therapy on normal and pathologic tissue may also be mis-represented due to measurement error. For quantitative analysis, it is essential to estimate the error due to biologic variability through repeated measures of PET radiotracer uptake. The general method is to scan patients or volunteers on several occasions and statistically analyze the change in the quantitative estimate. The rates of development of different pathologies suggest that the error associated with a quantitative PET measure also varies. So for any disease, such as brain cancer, breast cancer, and sarcoma, the variation in the PET measurement would have to be independently assessed by test-retest methods to determine the reproducibility for each cancer.

Some limitations to the test-retest procedure are worth commenting on. Since the radiotracer emits radiation, the injected activity may limit the number of times a patient can be scanned within a period of time. Patients with pathologic states of disease receiving therapy may have drastic metabolic changes in a short period of time. They may also be unable or unwilling to return for a second imaging procedure that is not clinically indicated even if the cost is covered by the investigator. To reduce the physiological effect of therapy and disease, repeat scans should be performed as close to each other as possible and separated in time from chemotherapy radiation therapy or surgical resection. Mechanistically, the scans should have a similar injected activity, infusion schedule, blood sampling, image reconstruction and quantitative analysis methods.

Reports of reproducibility, as expressed by the percent coefficient of variation ($COV = 100x(SD/mean)$, in the literature vary from 5.5%[6] and 10%[10] to 20%[11] in thoracic cancers and 13%[12] in normal patients. The

coefficient of variation allows for the development of criteria of assessment for pathologic states and possible directed therapy based on physiological PET parameters.

4. Kinetic Analysis in Oncologic Imaging

The key to interpreting dynamic PET data is a model of the kinetic behaviour of the tracer, which can be used to estimate biologically relevant parameters (see chapter 12). In considering the approach to kinetic analysis, it is important to remember the goal – to use imaging to measure cancer biology. Kinetic models should not be an endpoint in and of themselves, but rather a tool for measuring how tumours behave. Oncologic PET provides a unique opportunity to apply the power of applied mathematics to clinical tumour biology. It is important to keep sight of the endpoint and not get lost in an internally elegant model with no relevance to the biology of the tumour or the mechanism by which the tracer appears in the tissues. Model formulation and kinetic analysis therefore involve a compromise between a complete mathematical description of tracer biochemistry and kinetics and the practical limitations associated with patient imaging. The process of kinetic analysis is therefore an iterative one, refining (and simplifying, if necessary) the model as more experience is gained for a particular tracer applied to a specific disease process.

4.1 Constructing the Model Based Upon Tracer Biochemistry

4.1.1 Model Considerations

The first step in developing an approach to kinetic analysis for radiopharmaceutical imaging in cancer is formulating the model. Although the biochemistry under study is assumed to be at equilibrium during the PET scan, the time course of uptake and clearance of the radiopharmaceutical is dynamic and complex. Compartmental modelling offers a method to condense vast amounts of data into biologically meaningful parameters. The compartmental model should be based upon a consideration of the key steps and kinetic barriers in biochemical pathway under study. Important questions in constructing a model are as follows:

(1) Initial transport: how does the radiopharmaceutical get to the tumour? Are there barriers between the capillaries and the tumour site? This will depend upon the nature of the radiopharmaceutical and the site of the tumour. For somatic tumours and small-molecule radiopharmaceuticals, transit across capillaries is rarely a rate-limiting step. In this case, initial transport is primarily limited by blood flow. For larger molecules, for

example labelled antibodies, transit across capillaries and through the interstitial space may limit initial transport and may need to be taken into account. Another special case is brain tumour imaging. In the normal brain, the blood-brain barrier (BBB) limits the trans-capillary transport of many substances and may provide a significant barrier to initial transport. This effect needs to be accounted for and can be particularly tricky in brain tumour imaging, where both tumour and treatment can damage the blood-brain barrier and increase apparent capillary permeability. The approach to image analysis needs to separate the effects of enhanced tracer delivery due to BBB damage from increased retention of tracer due to flux through the biochemical pathway under study.

(2) Cellular transport: how does the radiopharmaceutical get from the interstitial space to the inside of the cell? Is transport purely passive diffusion, as is the case for many lipophilic compounds? Is there a specialized transport system, either passive or active? For cell-surface receptors, intracellular transport may not be important. In most cases, the time resolution and data quality from PET imaging will not be sufficient to separate initial capillary transport effects from transport across tumour cell membranes. It is therefore important to understand the cellular transport of the compound under consideration to know how to interpret the estimates of initial blood-to-tissue transport that are provided by compartmental analysis.

(3) Key steps in the biochemical pathway: most cancer radiopharmaceuticals are targeted to particular biochemical processes, and therefore understanding the steps in these biochemical pathways is key to interpreting the data that arise from PET imaging. Typically, when the target pathway is active in the tumour, the retention of the tracer will be enhanced. For example, high glycolysis in a tumour will lead to high rates of phosphorylation of FDG, in parallel to the first step in glycolyisis, and therefore increased tracer trapping in the tumour. In formulating the model, it is most important to know the rate limiting step(s) in the targeted pathway. The rate limiting reaction will most strongly influence retention of the tracer, and estimates of tracer flux through the pathway will, in essence, be estimates of the flux through the rate-limiting step. For example, in imaging ^{11}C-thymidine, the rate limiting step is incorporation into DNA; therefore ^{11}C-thymidine provides an indication of DNA synthesis by measuring the flux of thymidine along the salvage pathway into newly formed DNA.[13] For receptor agents, the constants for binding to and release from the receptor are key, and the PET data will provide an indication of receptor binding that reflects both binding constants and the quantity of receptors expressed. In most cases, *in vitro* studies will provide much of the information necessary to formulate the aspects of the model related to the biochemical pathway. The rest must

be gathered from pre-clinical (*in vitro* and animal) studies and preliminary patient studies.

4.1.2 Model Example

As an example of how to construct and interpret parameters in the model, consider the biochemical pathway illustrated in Figure 2. A hypothetical compound (X) which is incorporated into a biochemical pathway is considered, where the rate limiting step is phosphorylation, resulting in a compound, X-P that is retained in the cell either by direct trapping or entry into other pathways (this could easily represent the glucose biochemistry, for example). There is slow de-phosphorylation of X-P back to X. A potential compartmental model using a radioactive form, X is also shown in the Figure. The term X_n is used to indicate the native (unlabeled compound) and X_r to indicate the radiolabelled tracer. In this model, the blood concentration of tracer, C_b, is typically given in units μmole/cc for the native compound and MBq/g for the radiolabelled tracer. The concentration in tissue compartments for X_r (exchangeable compartment, E and trapped X-P compartment, T) are considered as tissue contents, Q_E and Q_T, in units μmole/g for the native compound and MBq/g for the radiolabelled compound.

Let us assume that X is a small molecule that readily transports across capillaries. The first rate-limiting step is delivery of X to the tissue. The parameter, K_1 (cc/min/g) describes this step. K_1 includes the effect of both tracer bulk delivery, via tissue blood flow, trans-capillary transport, movement through the interstitial fluid, and transport across the cell membrane. K_1 can be approximated by FxE, where F is blood flow (cc/min/g) and E is the first-pass extraction fraction (unitless), which accounts for all barriers between the capillary and the intra-cellular space.[14] While the use of a single

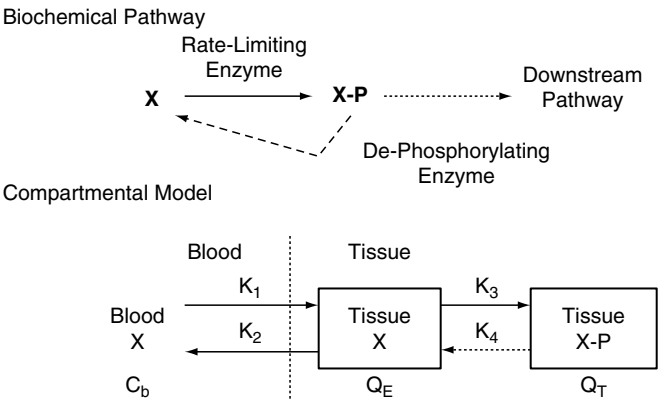


FIGURE 2. Hypothetical biochemical pathway and model for compound X for both the native and labelled compound.

parameter K_1 is a significant over-simplification of the transport process, it is often the most that can be gleaned from typical dynamic PET data. The parameter, k_2 (1/min) describes transport from the exchangeable compartment back to the blood. It is important to note that k_2 is dependent not only upon the absolute rate of back-transport, but also the size of the exchangeable pool, Q_E. The parameter k_3 (1/min) describes the rate-limiting biochemical reaction and therefore movement from the exchangeable compartment to the trapped compartment. De-phosphorylation, which implies passage back to the exchangeable compartment is described by k_4 (1/min). Note that in this formulation, K_1 has different units than the other rate constants. Blood concentration of X is measured in concentration units (for example, MBq/cc), while the compartments (Q_E and Q_T) are tissue compartments, described as tissue contents (MBq/g).

The equations describing the behaviour of the model are given below (units for the variables for the radiolabelled tracer are given in Figure 3:

$$\frac{dQ_E}{dt} = K_1 C_b - k_2 Q_E - k_3 Q_E + k_4 Q_T \tag{1}$$

$$\frac{dQ_T}{dt} = k_3 Q_E - k_4 Q_T \tag{2}$$

$$C_t = (V_b C_b + (Q_E + Q_T))\rho \tag{3}$$

The first two equations describe the compartmental model kinetics, driven by the blood activity clearance curve (C_b). Where V_b is the fractional volume of blood per gram of tissue and ρ is the tissue density. The third equation relates model compartments to the total tissue time-activity curve (C_t), which is equivalent to what is measured by dynamic PET imaging. In this equation, V_b is the fractional volume of blood per gram of tissue and ρ is the tissue density. V_b accounts for the contribution of tracer in the blood, both due to blood physically present in the imaged tissue and also due to "spill-over" – counts from adjacent blood pool structures contributing to the tissue time-activity curve owing to the finite spatial resolution of the tomograph. An example is spillover of counts from the abdominal aorta to the adjacent spine.

Formulation of the model implies a number of assumptions. Perhaps the most important is that the native (i.e. unlabeled) substance X_n is in steady state. This implies that the blood concentration of the native substance is

K_1	cc/min/g	Q_E, Q_T	MBq/g
K_2, k_3, k_4	1/min	V_b	cc/g
C_t, C_b	MBq/cc	ρ	g/cc

FIGURE 3. Typical units for compartmental models.

static and the flux through the biochemical pathway is constant. This is obviously not true for the injected radioactive moiety, X_r, but must be true for the substance under study. For example, studying glucose metabolism under conditions where the blood glucose concentration is rapidly changing, for example immediately following a meal, would not be appropriate. A second important assumption is that the tracer does not perturb the biochemical and physiological system under study. This means that the peak chemical blood concentration of the tracer injected should be at least 1-2 orders of magnitude less than blood levels of the native compound.

4.1.3 Estimating Parameters (optimizing the model)

Compartmental analysis entails the use of dynamic imaging data to estimate the compartmental model's parameters. The inputs to the process are the measured blood activity clearance curve (C_b), the measured tissue time-activity curve (Y), and starting guesses for the model parameters, including V_b. Since closed-form solutions to most sets of compartmental model differential equations do not exist, estimating parameters involves adjusting parameter guesses until the model output, C_t, best matches the measured tissue curve, Y (Figure 4). Since the model output does not depend upon the parameters in a linear fashion, this requires a non-linear optimization approach, for example the Marquardt-Levenberg or Powell algorithm.[15] Non-linear optimization searches for the best match between the model estimate of the tissue uptake curve and the measured tissue curve by minimizing the chi-square function:

$$\chi^2 = \sum_{i=0}^{N} \left(\frac{(C_{t_i} - Y_i)}{\omega_i} \right)^2 \tag{4}$$

where N is the total number of tissue uptake curve measurements over the time course of imaging, C_{ti} and Y_i are the model estimate and measured tissue uptake value at the ith time point, and ω_i is a weighting factor for the ith time point. Typically the weight for each tissue uptake time point is based

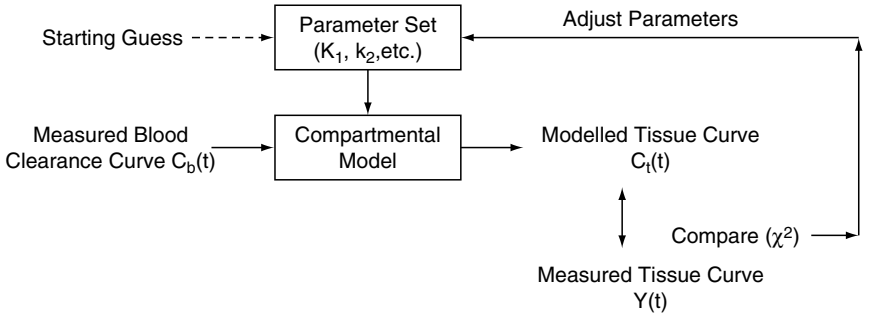


FIGURE 4. Estimating parameters in compartmental analysis.

on the variance of the measurement. As described below, the variance of the measurement is usually approximated by:

$$\sigma_i^2 = \frac{Y_i}{\Delta T_i} \tag{5}$$

where ΔT_i is the length of the time bin for the ith tissue uptake measurement. Section 4.2 describes how model behaviour can be tested to characterize the accuracy with which non-linear optimization can estimate model parameters.

4.1.4 Interpretation of Model Parameters

Some discussion is in order on how to interpret model parameters. We return to the hypothetical radiopharmaceutical, labelled X_r, which traces the native compound X_n down a biochemical pathway where phosphorylation is the rate-limiting step. The goal of compartmental analysis is to delineate the kinetics through the pathway and specifically to estimate flux through the pathway. Flux describes the rate at which X_n leaves the blood and moves through the biochemical pathway in tissue, typically given in units μmole/min/g. The first temptation upon viewing the model is to equate flux to k_3; however k_3 is a rate constant that is dependent upon the pool size in the exchangeable compartment, Q_E. The flux of substance X through the pathway is given by

$$Flux = Q_E k_3 \tag{6}$$

It can be seen from this equation that k_3 is dependent upon both the flux and exchangeable compartment content and not simply equal to flux. To relate flux to measurable quantities, we rely upon the methods described by Sokoloff for deoxyglucose[16] and recall that the native substance, X_n, is in steady state. For the moment, the small contribution of de-phoshorylation (k_4) is ignored. Steady state implies that for the native substance, the exchangeable compartment content, Q_E, is constant, and thus

$$\frac{dQ_E}{dt} = K_1 C_b - k_2 Q_E - k_3 Q_E = 0 \tag{7}$$

Re-arranging terms yields

$$Q_E = \frac{C_b K_1}{k_2 + k_3} \tag{8}$$

For the native substance, X_n, the blood concentration is constant. Referring to the constant native compound concentration as $[X_n]$ and substituting for C_b, the flux equation now becomes

$$Flux = [X_n]\frac{K_1 k_3}{k_2 + k_3} \tag{9}$$

The fraction term describes what is often termed the flux constant, K_i, which appears in the flux term as a rate constant describing the flux from the blood through the pathway:

$$Flux = [X_n]K_i \tag{10}$$

where

$$K_i = \frac{K_1 k_3}{k_2 + k_3} \tag{11}$$

For FDG, this leads to the familiar form describing tumour metabolic rate:

$$MRFDG = [Glu\cos e]\frac{K_1 k_3}{k_2 + k_3} = [Glu\cos e]K_i \tag{12}$$

where [*Glucose*] is the plasma glucose concentration (μmole/cc). The combination of rate constants in the fraction term, which describes a rate constant for flux from the blood to the phosphorylated state, is often termed the flux constant, K_i (cc/min/g). Equation 12 differs from the classic work of Reivich and Phelps[6,17] for brain imaging in that there is no attempt to relate FDG kinetics directly to glucose kinetics. Instead, the term metabolic rate of FDG (MRFDG) is used to indicate glucose flux estimated by FDG-PET and to reflect the fact that there are differences between FDG and glucose metabolism, especially in tumours.[18] The use of a "lumped constant" to convert MRFDG to the glucose metabolic rate is not appropriate for tumour imaging where the relationship between glucose and FDG can vary considerably.

4.2 *Testing the Model – Parameter Sensitivity, Covariance, and Precision*

Once the model is formulated and the approximate parameter range estimated, the next step is to determine how accurately the model parameters can be estimated. Biologic and kinetic considerations may dictate a particular format for a compartmental model; however, biologic validity does not guarantee that model parameters can be accurately estimated. For example, a complex model may have parameters that affect model behaviour but cannot be reasonably estimated from the type and quality of data obtained in a PET imaging protocol. The final output of kinetic analysis in this case is therefore a compromise between modelling all the biologic effects for the radiopharmaceutical and making assumptions that allow the investigator to obtain robust estimates of the important kinetic parameters. Methods for understanding how a particular model behaves mathematically and assessing how accurately the parameters can be estimated are the topic of this section.

It is important to note that a compartmental model's mathematical behaviour is dependent not only upon the format of the model, but also the

range of expected parameters, or in other words where the particular application falls in the multi-dimensional parameter space of the model. Model behaviour is not necessarily linear and models may behave quite differently in different parts of the parameter space. An example is ^{11}C-thymidine imaging,[13,19] where the same compartmental model behaved differently for somatic tumours, where there are no significant transport barriers, versus brain tumours, where the BBB can form a significant barrier to thymidine transport for normal brain and tumours with an intact BBB. The transport parameter (K_1) is therefore much lower in brain than for somatic tumours.

Equally important in ascertaining model behaviour is the input function, in this case the blood clearance curve, which will strongly affect the shape of tissue time-activity curves and have significant implications for parameter estimation. Therefore, prior to testing model behaviour, it is important to:

(1) Estimate the expected range and typical values for representative tissue types for model parameters. This can be done using information from a combination of pre-clinical studies of tracer biodistribution and kinetics, from review of studies of radiopharmaceuticals with similar properties (for example, 2-deoxyglucose and 2-fluorodeoxyglucose), and from initial experience with studies in patients. Estimating the range of parameter values may be an iterative process; initial exploration of the model may suggest alternate parameter formulations and/or different expected ranges.
(2) Identify up to a few sets of representative blood clearance curves that can be used as inputs for testing the model. Ideally this should come from preliminary patient studies; however, preliminary animal model studies or blood clearance of a similar tracer may suffice.

Once the model is formulated, the range of parameters determined, and sample blood clearance curves obtained, a series of calculations and simulations will quantify model behaviour under conditions typically encountered in quantitative PET imaging. Many of these calculations can be performed on standard spreadsheet software such as Excel (Microsoft, Redmond, WA, USA); however, a number of software packages are available and helpful for this type of analysis. The specialized packages typically have graphical interfaces specifically designed for modelling analyses and often have modules for analysis and simulations specific to model testing. All approaches share the key common element of a module to calculate model output for a given input function (blood clearance curve) and parameter set. Since the differential equations for most physiologically relevant models do not have a closed from solution, model output is typically generated using approximate methods to integrate the differential equations.

The steps involved in model analysis are fundamental to the process of non-linear optimization of the model.[15] These have been reviewed for radiotracer applications by several authors.[14,20] A similar approach is described here. These steps are as follows (summarized also in Table 1):

Table 1. Methods for testing compartmental model behaviour.

Analysis	Model Feature Tested
Sensitivity analysis	Effect of change in a parameter on model output
Identifiability analysis	Ability to estimate parameters independently
Simulations, including noise	Expected accuracy of parameter estimates (bias and precision)

(1) calculate the sensitivity of the model output (i.e. the tissue time-activity curve) to changes in the individual parameters;
(2) estimate parameter covariance and identifiability;
(3) estimate the expected accuracy of parameter estimates – bias and precision.

These are discussed individually below.

4.2.1 Sensitivity

Sensitivity analysis describes how a small change in an individual parameter changes the output of the model. Parameters that strongly affect the model are more likely to be estimated accurately in the process of model optimization. Parameters that have a relatively small impact on the model may be difficult to estimate. The sensitivity function that quantifies this behaviour is calculated as the fractional change in the model output per fractional change in the parameter, given by ref.[15]

$$Sens_{k_\alpha}(t) = \frac{\delta C_t(t)/C_t(t)}{\delta k_\alpha / k_\alpha} \tag{13}$$

where $Sens_{k\alpha}(t)$ describes the sensitivity as a function of time after injection for the model parameter, k_α;$C_t(t)$ is the tissue time-activity curve that is the output of the model; and δ describes a small incremental change. In practice, sensitivity functions are calculated as the difference in model output for a small change of the parameter, typically +/− 1% to 5%. Note that the sensitivity function is time-varying. The ability to estimate a parameter depends upon the magnitude of the change it induces in the model output and also the extent to which the effect on the model output has a different time-dependence than other parameters. Sensitivity analysis can identify those parameters where accurate parameter estimation is unlikely. The choice then becomes whether to fix the parameters to reasonable physiologic values or re-formulate the model parameters to generate new parameters with greater sensitivity.

4.2.2 Identifiability

Parameter identifiability analysis is a method to quantify the extent to which each model parameter can be estimated independent of the other parameters. It is based upon the parameter covariance matrix that arises in many of the non-linear optimization methods.[15] A practical approach is to use the

parameter sensitivity functions to generate a correlation matrix which indicates the extent to which parameter estimates covary. Pairs of parameters with high correlation coefficients in this analysis are unlikely to be able to be estimated independently. Identifiability analysis provides only an estimate of parameter covariance; more precise information on how parameters covary and how accurate parameter estimates requires more detailed model simulations. Nevertheless, identifiability analysis provides a useful tool for devising parameter formulations and for determining the number of model parameters that can be independently estimated. When identifiability analysis suggests that not all parameters can be estimated independently, it may be necessary to make simplifying assumptions, for example fixing a particular parameter to a physiologically reasonable value, to assure robust estimation of the parameters of greatest interest.

To generate the parameter correlation matrix, the first step is to calculate the matrix of second partial derivatives of the model error function (chi-square function), known as the Hessian, from the integration of sensitivity functions.[15]

$$D_{\alpha\beta} = \int_0^{t_{end}} Sens_{k_\alpha}(\tau) Sens_{k_\beta}(\tau) d\tau \tag{14}$$

Where $D_{\alpha\beta}$ is the Hessian matrix element for parameters, k_α and k_β, and t_{end} is the time for which dynamic imaging data are obtained after injection. The next step is to invert the Hessian matrix to generate the covariance matrix. Finally, normalization of the covariance matrix to yield diagonal elements equal to *1* yields the correlation matrix. These last two steps are described by the following:

$$R_{\alpha\beta} = \left| \left[D_{\alpha\beta} \right]^{-1} \right| \tag{15}$$

where $R_{\alpha\beta}$ is the correlation matrix element for parameters, k_α and k_β. Values of $R_{\alpha\beta}$ close to zero indicate parameters with little covariance, while parameters close to 1 or -1 indicate closely covarying parameters that are unlikely to be independently estimated.

4.2.3 Accuracy of Parameter Estimation

To estimate the bias and precision of the model parameter estimates, it is necessary to perform simulations. This generally involves the generation of simulated tissue time-activity curves, estimation of model parameters, and comparison of the estimated parameters to the ''true'' parameters used to generate the simulated curve. The addition of statistical noise to the simulated curves is required to match the conditions of PET imaging, and without the generation of ''noisy'' curves, this would be a trivial exercise. The simulation process is illustrated in Figure 5.

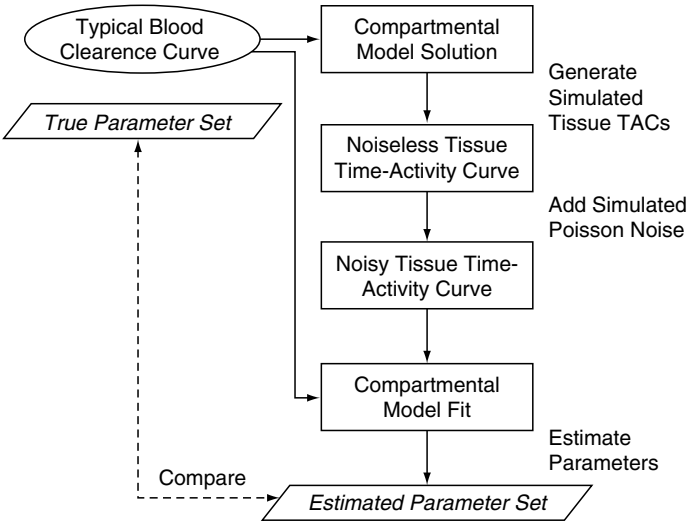


FIGURE 5. Simulation to determine likely parameter estimation accuracy.

The simulated statistical noise should lead to variability in the time activity curve that matches the variability encountered in the PET imaging protocol. This is a key element of the simulation; the more realistic the noise in the simulation, the closer the estimates of parameter accuracy will be to the real world. The variability in the PET image data depends upon a number of factors, including those listed below:

(1) the biological chemistry of the imaging agent;
(2) the amount and type of radioactivity injected;
(3) the size of the patient;
(4) the characteristics of the PET tomograph;
(5) scattered and random coincidence rates and the type of correction algorithms used;
(6) the length of the time bins used in the data acquisition protocol;
(7) the size of the region-of-interest (RoI) used in the image analysis.

In practice it is very difficult to make an *a priori* estimate of noise; therefore most investigators make a few simplifying assumptions and then match the variability of the simulated tissue time-activity curves to the variability of representative patient data. It remains a reasonable assumption that the data are approximately Poisson in nature; therefore the standard deviation of the average count rate in any time bin will be proportional to the image counts in that bin. For any given time bin, the image counts per unit time will be proportional to the tissue tracer concentration for that time bin:

$$\frac{Cnts_i}{T_i} \propto C_t(t_i) \tag{16}$$

where $Cnts_i$ is the total counts in the ith time bin, t_i is the mid-time of the ith time bin, and T_i is the length of the ith time bin. For Poisson noise:

$$\sigma_{C_t(t_i)} \propto \frac{\sqrt{Cnts_i}}{T_i} \propto \sqrt{\frac{C_t(t_i)}{T_i}} \tag{17}$$

This formula can be used along with a Poisson random number generator to add simulated noise to each time bin as follows:

$$C_t'(t_i) = C_t(t_i) + fR\sqrt{\frac{C_t(t_i)}{T_i}} \tag{18}$$

where the prime indicates the simulated noisy curve, R is a unit-normal random number generator (varying between -1 and $+1$) and f is a proportionality factor. To adjust the simulated noise, f is adjusted empirically until the error function (chi-square) for the noisy curve versus the noiseless curve matches the chi-square obtained when comparing clinical data to a model fit. This approach makes the assumption that the difference between the measured curve and the fit curve is simply due to statistical noise and not a systematic mismatch of the model and the data. This is unlikely to be true, but is a reasonable approach for estimating noise levels for simulations.

Once the simulation has been set up and the noise in the simulation has been adjusted, the simulation module can then be used to characterize the accuracy of the parameters. Two approaches are helpful:

(1) Repeated iterations generate a series of parameter estimates for a single representative set of parameters used to generate the noisy simulated tissue curves. The distribution of parameter estimates for the noisy curves will give an excellent estimate of the precision of parameter estimates (standard deviation of the parameter estimates from repeated iterations) and any bias that the parameter estimation introduces (difference between ''true'' parameter used to generate the curves and the mean of the estimated parameters).
(2) Iterations using different sets of parameters spanning the expected range of parameters will characterize the behaviour of the model and parameter estimation over the range of values likely to be encountered in oncologic imaging. To vary the 'true' parameters used to generate curves, one can either use a grid of regularly spaced parameter values through parameter space or perhaps more efficiently using a random number generator to generate random combinations of parameters spanning the parameter space. This will generate a series of tissue uptake curves, one for each parameter set, to which simulated noise can be added. The result of repeated iterations, optimizing the model to estimate parameters for each noisy curve, is a plot for each parameter of the estimated versus true value. From these plots, the fidelity of the parameter estimates (for example, using the correlation coefficient), bias (difference in mean true

versus mean estimated parameter) and precision (standard error of the estimate of the line fit for estimated versus true) can be estimated for conditions matching the range of values encountered in patient studies. If uncorrelated random numbers are used to generate the parameters, the correlation between parameter estimates provides another estimate of parameter covariance for comparison with the calculated correlation matrix from identifiability analysis.

By the end of sensitivity and identifiability calculations and simulations, the investigator will have a good idea of which parameters can be reasonably estimated and the expected accuracy of the estimates. Initial analysis may suggest fixing difficult-to-estimate parameters to reasonable values or tie them to some function of other estimated parameters. Simulations can then be done to test the effect of fixing these parameters on the bias and precision of the estimates of the parameters of interest. A good example of this kind of analysis is shown in the testing of the model for ^{11}C-thymidine imaging of somatic and brain tumours.[13,19]

The mathematical characterization of the model provides estimate of the bias and precision of parameters. This is important in designing animal or clinical research protocols; in particular for performing sample-size determinations from study power calculations. Furthermore, it provides a check on parameter error estimates for individual data sets and indicates how the model optimization process may bias parameter estimates. All these pieces of information are key in being able to draw biologically valid conclusions from quantitative PET imaging.

4.3 Implementing the Model

Once the model has been formulated and its mathematical behaviour tested, the model is now ready for implementation and validation. This typically involves a combination of pre-clinical studies using animal models and then preliminary testing in patients. Typically, animal studies precede human studies; however, preliminary testing humans may reveal unanticipated issues in the approach to image analysis. This may, in turn, spur further animal studies to test specific aspects of the model and image analysis.

4.3.1 Pre-clinical (animal studies)

Some pre-clinical data from animal models nearly always precede model formulation. These data typically include biodistribution data and some information on blood and tissue clearance. Additional data may be necessary to infer the range of certain model parameters. For example, the classic studies of Oldendorf and colleagues[21,22] measuring first-pass transport of a variety of substances into the brain are often used to estimate the range of K_1 for compounds used for brain or brain tumour imaging. Pre-clinical data

may be necessary to estimate the expected rate of loss from the trapped compartment, for example the de-phosphorylation of FDG-6P.[6]

Animal models typically provide the most rigorous test of the compartmental model, owing to the ability to sample tissue at various time points over the course of imaging thus allowing the user to directly test model predictions. For example, suppose a model describes the rate at which a substance is trapped in tissue in a particular form, for example FDG-6P for FDG or ^{11}C-labeled DNA for ^{11}C-thymidine. One can then measure the time course of uptake after tracer injection, fit the model to the total uptake and test whether the model correctly predicts how much of the label is in the trapped species. This can be done by sacrificing the animal at the end of the study, and analyzing the labelled species in the tissue of interest. An even more rigorous test of the model is to test model predictions of the time course of label trapping after injection; this requires tissue samples taken at multiple times after injection. The validation of a model of ^{11}C-thymidine incorporation into DNA provides a good example of this approach.[23] If initial studies show that the model predictions are incorrect, the model assumptions should be re-examined, including the expected range of the parameters, and the model re-adjusted to provide a better prediction of *in vivo* kinetics.

4.3.2 Early Patient Studies

The ultimate test of a compartmental model is its application to patient studies. Direct validation in patients avoids some potential pitfalls in extrapolating from pre-clinical models:

1. There may be significant species differences in biochemical pathways and substance transport. An example is PET oestrogen receptor (ER) imaging. Pre-clinical studies in rodents showed high uptake in ER-rich tissues but relatively little uptake in ER-expressing tumours in patients, presumably due to significant differences in sex steroid transport proteins for rodents versus humans.[24]
2. Even if biochemical pathways are similar, there may be quantitative differences in enzyme kinetics for animals versus humans. For example, the normal brain lumped constant for FDG versus glucose is different for rodents versus humans.[25]

The first step in testing a model in patients is simply to show that it "fits" the data. More precisely, can the model fit tissue-time activity curves taken from a range of target tissue types (for example, tumours with a range of growth rates) and yield biologically plausible parameter estimates within the expected range. This is a necessary first test, but not sufficient proof of model accuracy. The next step is to relate the kinetic measurements to an external "gold standard". Examples include comparison of breast tumour uptake of ^{18}F-fluoroestradiol (FES) to *in vitro* radioligand binding assays

performed on tumour biopsy material[26] and comparison of ^{18}F-FLT kinetics versus *in vitro* assay of cellular proliferation using the Ki-67 immunohistochemistry.[27] In some cases, this type of comparison may not be possible. For example, there is no way to stain biopsy material for glucose consumption. Instead Bos and colleagues[28] compared estimates of FDG kinetics to a variety of potentially related biologic processes to help guide the interpretation of the results of the analysis of FDG-PET tumour studies.

Perhaps the ultimate test of a compartmental model is its ability to predict patient outcome. This returns to the theme set forth at the beginning of the section, namely, that the ultimate value of a compartmental model lies in its ability to provide new insights into tumour behaviour in patients. In this light, once initial patient studies and model validation are complete, well-designed clinical trials should compare the model's parameter estimates to outcome measures such as response to treatment and survival. A general approach to such studies is represented by the test of the following statistical model (O'Sullivan, F, personal communication).

$$Outcome = function\ of(clinical\ data,\ PET\ parameters) \tag{19}$$

where clinical parameters is the standard clinical information available for a patient, and the PET parameters are kinetic parameters from PET imaging and compartmental analysis. If the PET parameters are useful in describing clinical tumour behaviour, then they will statistically significantly improve the ability to account for variability in the outcome variable. This is the most rigorous test of the application of quantitative imaging to clinical oncology and requires good clinical study design, in addition to rigorous quantitative analysis. Such studies are, by nature, complex and remain one of the great challenges in applying quantitative radiotracer imaging to the care of cancer patients.

4.4 *Special Cases – Paired or Sequential Tracer Studies, Labelled Metabolites*

Two situations frequently encountered in oncologic PET imaging deserve special comment: (1) tracers with labelled metabolites and (2) sequential paired studies. The former arises anytime the injected tracer is metabolized and labelled metabolites circulate in the blood. The second scenario arises when two aspects of tumour biology are measured at the same time with two different radiopharmaceuticals. This is possible when the first tracer uses a short-lived radionuclide (for example ^{15}O or ^{11}C) followed by an ^{18}F tracer, but it is impractical for two longer-lived tracers (for example two ^{18}F radiopharmaceuticals). Sequential studies may arise from the need to characterize the behaviour of both a tracer and its labelled metabolite, which is often the case for tracers resulting in $^{11}CO_2$.[9,19]

4.4.1 Labelled Metabolites

Circulating labelled metabolites are often ignored in quantitative analysis and represent a potentially significant source of error if ignored. Metabolites most frequently have kinetic properties quite different than the parent compound. While the intact radiopharmaceutical targets a particular biologic process, the labelled metabolites may be non-specifically taken up in tissue or even targeted to a different process. In this case, kinetic models that assume uniform behaviour of all labelled species will be incorrect and subject to unpredictable errors in parameter estimation.

In some applications, for example neuroreceptor imaging in the brain, metabolite subtraction can be done using a reference region that is similar to the tissue of interest but contains only non-specific uptake and not the specific target uptake. This is generally not appropriate for tumour imaging, where the properties of the tumour are unpredictable and most frequently not like any normal reference tissue. Oncologic applications may therefore requires analysis of blood samples to determine the fraction of radioactivity in the blood (i.e. – the input function) present as intact tracer versus labelled metabolites and then using this data as an additional input to the model to account for the contribution of labelled metabolites to the total tissue tumour time-activity curve measured by imaging. If the time course of the concentration of labelled metabolites in the blood is fairly predictable, curve-fitting approaches can be used to limit the number of samples needed.[29]

The profile of labelled metabolites in the blood clearance curve can then be used to account for metabolites in the tissue. This requires some knowledge about the kinetics of metabolite uptake and retention in tissue. Some metabolites may be conjugates which have little access to tissues outside the vascular space. In this case, the metabolite activity can simply be subtracted from the blood clearance curve to provide a better input function for the compartmental model (the total blood activity used in the term representing partial blood volume, multiplied by V_b, must include labelled metabolites, however). Metabolites that have access to tissue but are not trapped can be subtracted using a compartmental model with separate inputs from intact tracer and metabolites as shown in Figure 6. If the intact tracer is trapped and the metabolite is not, the parameters for the intact tracer can be estimated without undo bias from the metabolites,[13,30] although this approach may require some metabolite parameters to be fixed or highly constrained during optimization.[13] When both the intact tracer and metabolites are retained in tissue, sequential injections of the labelled metabolite and intact tracer may be necessary.

4.4.2 Sequential Paired Studies

Sequential studies can provide a very powerful tool for characterizing cancer. Comparing two aspects of tumour biology simultaneously can provide more information than the sum of the individual measures. Examples

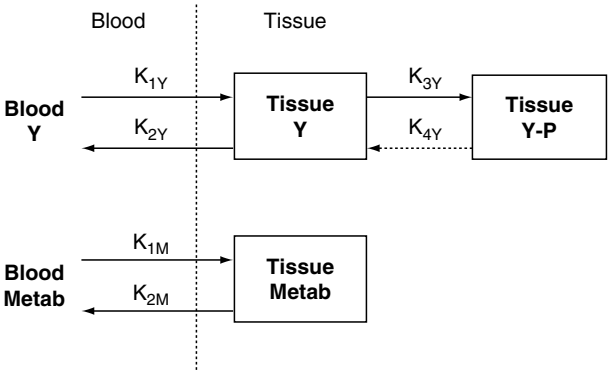


FIGURE 6. Potential model for radiotracer with circulating labelled metabolites.

include comparison of glucose and FDG metabolism to yield the relative hexose utilization or lumped constant[18] and simultaneous measurement of glucose metabolism and either blood flow or oxygen consumption to infer patterns of tumour metabolism.[31,32] If the second imaging study follows closely on the heals of the first study, then it may be necessary to account for the leftover activity from the first injection in analyzing the results from the second injection. For the blood clearance curve, if the two tracers have different half lives, it may be possible to determine the blood concentration of the leftover first tracer simply by counting twice. An alternate approach[25] is to fit the tail of the first blood clearance curve to an empiric function to estimate the contribution of remaining tracer from the first injection to the early portion of the second blood clearance curve.

When analyzing sequential injection studies, the temptation is to break the parameter estimation task into smaller pieces by dealing with each injection separately. However, there may be significant advantages to using a model that incorporates both injections and estimates parameters using the combined model.[33] The combined model can explicitly account for the presence of two tracers in the tissue during the overlap period. In addition, some parameters may be biologically coupled, for example blood flow and tracer delivery (K_1), and the sequential model can incorporate these factors. Simulations can be used to determine the best optimization strategy – simultaneous estimation of all parameters, a two-pass approach with some parameters fixed on the first pass, etc. A good example of the approach to sequential studies is seen in the studies of ^{11}C-glucose and ^{18}F-FDG by Spence and colleagues.[18,25]

4.5 Parametric Imaging

In quantitative radiotracer imaging, the typical analysis paradigm is as follows:

1. perform dynamic imaging with blood sampling;
2. examine raw images summed over time to identify structures of interest; perhaps with the aid of registered anatomic imaging such as CT or MRI;
3. draw regions-of-interest (RoIs) around the desired structures;
4. generate tissue time-activity curves from application of the RoIs to the dynamic image data;
5. process blood and tissue activity curves converting them to similar units of activity;
6. perform compartmental model analysis using the RoI-based tissue curves and measured blood input function.

While this approach is simple and robust, it has two related disadvantages:

1. Tumours are often heterogeneous. An RoI may include sections of a tumour with a variety of biologic properties. RoI analysis assesses an average effect over a range of tissue types in a region and may fail to represent some important properties of the tumour.
2. Boundaries defined by summed images or anatomic boundaries from CT or MRI may not be appropriate for quantitative biochemical imaging and may bias the result. For example, RoIs based upon the "hot spot" in the summed raw image may select the most active portion of the tumour and not reflect the full extent of variability in the tumour.

An alternate approach to regional estimates of tracer kinetics is termed parametric imaging (Figure 7).[34] The approach consists in estimating model parameters on a pixel-by-pixel basis to generate an image of kinetic parameters. This approach depicts the heterogeneity of tumour down to the image sampling distance (typically a few mm) and avoids the bias inherent in defining tissue RoIs from summed or anatomic images. The chief limitations are computational expense and image noise. This is especially true in the approach to parametric imaging which treats each image pixel as a separate data point, generating an entire volume of pixel-wise tissue time-activity curves. The number of image counts contributing to the individual pixel's curve is small, generating noisy curves and often creating considerable inaccuracy and unpredictability in parameter estimation. A more statistically robust approach is to consider the image over both time and space (4-D) and decompose the image into a limited number of components which can be used to describe the 4-dimenional image. Examples include spectral analysis, factor analysis and mixture analysis.[35,36] For mixture analysis,[33] the image is segmented to group together pixels with similar time courses of biological uptake in the initial step. The segmentation process is then iteratively refined to define a new set of time-activity curves and a set of constants for each pixel describing the pixel as a linear combination of the time-activity curves. This yields a modest number of time-activity curves (typically 50 or

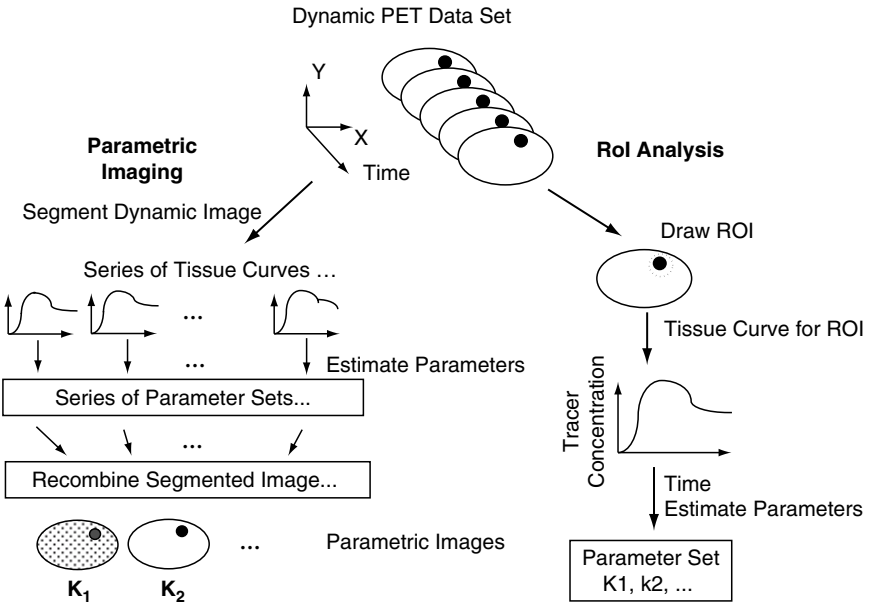


FIGURE 7. Diagram of parametric imaging.

less) with much better counting statistics than an individual pixel time-activity curve. Compartmental analysis is applied to each individual time-activity curve, and the final parametric image is formed as a linear combination of the parameter estimates for each curve, using the linear combination constants calculated for each pixel in the segmentation process. Because it combines all image data into a relatively small number of time-activity curves, mixture analysis yields low-noise tissue time-activity curves, with resulting reliable parameter estimates. This in essence provides a kinetically correct approach to image smoothing. The result on image quality can be dramatic, as seen in the ^{11}C-thymdine brain tumour images in shown Figure 8.

An alternate approach to parametric imaging is spectral analysis.[37] In this method, the image is decomposed into a linear combination of time-activity curves arising from single-compartment models with a range of rate constants, yielding series of single-exponential clearance curves. This leads directly to pixel-wise images of the linear combination of the constants for each clearance curve in the basis set, or in other words, the average clearance constant for each pixel. This provides a robust, model-independent approach to pixel-wise image analysis; however, the features of individual biochemical pathways cannot be incorporated into spectral analysis, and it may therefore be difficult to relate spectral analysis results directly to biologic parameters of interest.

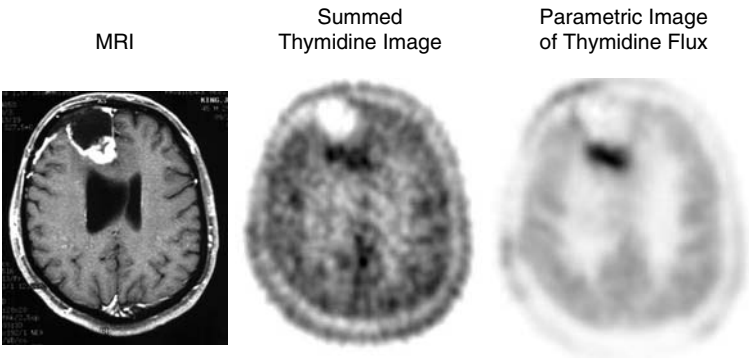


FIGURE 8. Example of parametric imaging for a patient with a recurrent high-grade brain tumour imaged using ^{11}C-thymidine PET. Summed thymidine images indicate increased tracer accumulation at the tumour site; however, contrast is low owing to the presence of unincorporated thymidine and thymidine metabolites in addition to thymidine retained in newly synthesized DNA. The parametric image of the flux constant for thymidine incorporation into DNA, obtained using mixture analysis and a compartmental model for thymidine and metabolites, clearly delineates actively growing tumour.

5. Practical Considerations and Approaches for Clinical Research and Clinical Imaging

5.1 Constraints of Clinical Research and Clinical Imaging

While pre-clinical studies and early pilot human studies can be performed in a sufficiently detailed fashion to allow rigorous kinetic analysis, such detail is not always possible in larger scale clinical research studies and in routine clinical practice. The reasons for this include the following: (1) many patients, especially those with more advanced cancers, may not tolerate long procedures and may be sufficiently anemic to limit blood sampling that can be performed; (2) many patients involved in serial studies will be unwilling to return for follow -up studies if the first study was too long and/or uncomfortable; (3) long procedure times may not be practical for large-volume studies with a large number of subjects; and (4) centres primarily focused on clinical imaging may not have the expertise to perform more complex quantitative analyses. These considerations have prompted the development of simplified quantitative approaches for more routine use in clinical research or especially in clinical practice. These approaches have best been tested for FDG-PET, which has seen by far the most widespread clinical use of the PET radiopharmaceuticals. Because the simplified approaches invariably make approximations regarding the kinetic behaviour of the radiopharmaceutical, they are inherently more prone to bias than full kinetic analysis. Therefore, the decision to use a simplified quantitative analysis

approach always calls for some trade-off between clinical convenience versus quantitative rigor. Where to aim in the balance between these competing considerations depends upon the clinical and biologic question addressed by the PET imaging and the clinical research study or clinical diagnostic application for which PET is being used.

An important general principle of research involving new radiopharmaceuticals is that the approach should start with a detailed rigorous analysis and use the results of that analysis to suggest alternate, simpler approaches. In other words, it is inappropriate for new imaging approaches to start with overly simplified quantitative approaches only to find later that sources of variability and bias that would have been apparent in more detailed studies were missed.

From the standpoint of clinical convenience, the following properties are desirable for a practical quantitative measure:

1. obtained using short imaging times, preferably one hour or less;
2. no blood sampling, or minimal blood sampling, including avoiding the need for multiple venous access sites;
3. easy to implement, with few details outside those required for standard PET image acquisition;
4. applicable to a wide variety of PET tomographs with the requirement only to be able to calibrate images in units of MBq/cc or MBq/g.

At the same time, a good approach should retain the following quantitative features:

1. good precision – i.e., minimal variance on test/re-test studies;
2. robust to variations in imaging parameters – injected dose, imaging times, etc;
3. robust to variations in patient preparation;
4. robust to slight variations in scanner performance.

No ideal method exists and typically, since many of these desirable features are not mutually compatible, trade-offs must be made. In this section, alternate, simplified approaches to quantitative analysis are reviewed with focus on graphical approaches and static uptake measures. Most of the work in this area has arisen from the FDG literature, and we focus on approaches that have been developed for FDG-PET. Broad classes of approaches are summarized in Table 2.

5.2 Graphical Analysis

One of the milestones in the quantitative analysis of radiotracer imaging studies, seen in early papers by Patlak[38] and Gjedde,[39] was the recognition that certain classes of radiopharmaceuticals reach near-equilibrium and lend themselves to a graphical approach for estimating the flux of the test substance from the blood to the biochemical target. Graphical analysis

Table 2. Approximate methods of image quantification for FDG-PET in oncology.

Method	Required Inputs	Uptake Measure	Metabolic Rate Estimate*	Estimation Method
Compartmental Analysis	Tissue Uptake Curve and Blood Clearance Curve (injection to ~ 60 minutes)	Flux constant, $K_i = K_1 k_3/(k_2 + k_3)$ (cc/min/g)	$MRFDG = [Gluc]K_i$ (μmole/min/g)	Non-linear Regression
Graphical Analysis	Blood Clearance Curve (injection to ~60 minutes) and Tissue Uptake Curve (~20 to ~60 minutes)	Flux constant, K_i, estimated from graphical approach (cc/min/g)	$MRFDG = [Gluc]K_i$ (μmole/min/g)	Linear Regression
Static Measures	Average Tissue Uptake over a short time period (5–10 minutes)	*SUV* (unitless or cc/g)	–	Simple Algebra

applies to classes of compounds which can be modelled as having a compartment of irreversible or nearly irreversible binding, as seen in Figure 9. In this case, a graphical approach can be used to estimate the flux constant, K_i, describing movement from the blood into the trapped compartment. The constant describes the flux through the biochemical pathway as previously described for kinetic analysis:

$$Flux_x = [X_n]K_i \tag{20}$$

where $Flux_x$ is the flux of substance x through the pathway to the trapped compartment (μmole/min/g), K_i is the flux constant (cc/min/g), and $[X_n]$ is the blood or plasma concentration of the native (unlabeled) substance X (μmole/cc). The flux constant is in the same units as blood flow and can be thought of as a constant describing the "flow" of a substance through the pathway. As discussed in earlier sections of this chapter, the flux can be directly related to kinetic parameters in the compartmental model:

$$K_i = \frac{K_1 k_3}{k_2 + k_3} \tag{21}$$

For graphical analysis to apply, the compartmental model should be able to be described by one or more reversible compartments in series with the trapped compartment[38] as shown in Figure 9. In addition, at some finite time after injection, the reversible compartment must be in pseudo-equilibrium with the blood. In other words, the time constant describing the rate of change of the blood input function, must be much less than the time constant describing the reversible compartment. For the standard 2-compartment 3- or 4-parameter model, this is described by the following:

$$\tau_b << (k_2 + k_3) \tag{22}$$

*MRFDG = metabolic rate of FDG

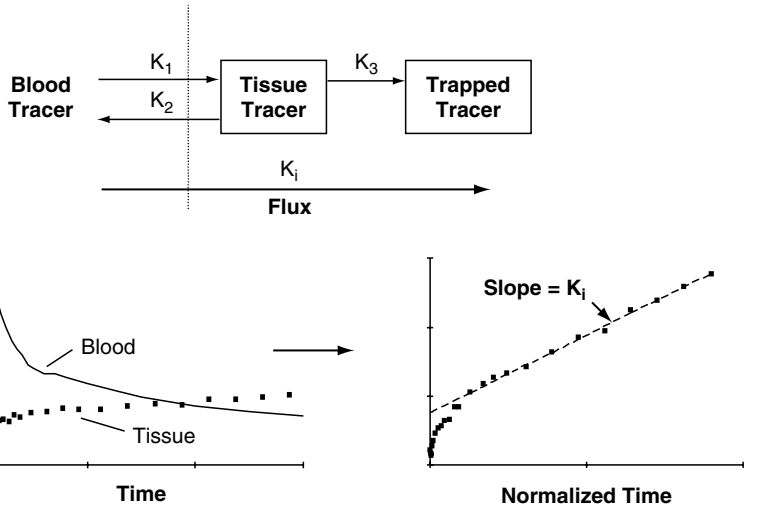


FIGURE 9. Compartmental model amenable to graphical analysis and example of transformation of blood and tissue time-activity curves to graphical variables for estimation of the flux constant.

where τ_b is a time constant describing the mono-exponential clearance of the tracer late after injection, when the transients of initial clearance have passed. When such conditions are met, the following linear equation holds:

$$\frac{C_t(t)}{C_b(t)} = K_i \frac{\int_0^t C_b(\tau)d\tau}{C_b(t)} + V_d \tag{23}$$

where $C_t(t)$ and $C_b(t)$ are the time-varying tissue and blood time-activity curves, respectively (units of MBq/g and MBq/cc, respectively), K_i is the flux constant, and V_d is a constant (cc/g) describing the tracer distribution volume. The flux constant can be estimated using linear regression as the slope of the line of the tissue and blood time-activity curves transformed to "Patlak space":

$$\frac{C_t(t)}{C_b(t)} \text{ Vs } \frac{\int_0^t C_b(\tau)d\tau}{C_b(t)} \tag{24}$$

The simplicity of the graphical approach is that it translates a problem requiring the solution of a non-solvable differential equation by non-linear optimization to a linear equation amenable to linear regression. This avoids many of the problems inherent in non-linear optimization, namely sensitivity to noise in the tissue time-activity curve, parameter covariance, local minima in the approximate solution to the differential equations, and dependence of parameter estimates on starting guesses. The trade-off is that graphical analysis estimates only the flux constant, but not the individual kinetic

parameters. This may be a limitation in some applications, for example, brain tumour imaging, where it may be important to ascertain the contribution of transport versus retention in tracer uptake studies.[19]

One other advantage of graphical analysis is that it affords some flexibility in the imaging protocol. Since the early portion of the tissue time-activity curve is not necessary, one can take a ''ping-pong'' approach to limit the need for blood sampling. For example, in imaging sarcomas located in the extremities, Eary and Mankoff[7] imaged the cardiac blood pool to measure the early blood input function and then drew a limited number of venous samples while viewing the tumour site from 21 to 60 minutes after injection. This reduced the number of blood samples needed, which may be important for patients who have undergone chemotherapy and have reduced blood counts, and avoided the need for arterial sampling without compromising the quality of the graphical estimate of FDG flux.

Modifications of the graphical approach have been made for different radiopharmaceuticals and different quantitative imaging applications. For tracers with some washout from the trapped compartment (k_4), the release from the trapped compartment will lead to a concave curve in Patlak space and an underestimate of K_i if the curvature is not explicitly accounted for. Patlak[40] suggested a modification of his original graphical analysis as an approximate method for compensating for a finite k_4. For tracers with highly reversible binding, say receptor agents, it may be more appropriate to focus on the tracer distribution volume or binding potential than the flux constant. For such applications, an alternate graphical formulation known as Logan graphical analysis can be used.[41] This approach is used frequently for neuroreceptor imaging, but less frequently in nuclear oncologic applications. Labelled metabolites may confound the standard graphical approaches. When the metabolites have been measured through analysis of blood samples, an alternate formulation of the graphical analysis which explicitly accounts for labelled metabolites can be used, provided that none of the labelled metabolites are trapped.[42,43] These alternate approaches may apply to radiopharmaceuticals whose kinetics and/or metabolism are more complex than FDG and may offer the simplifying advantages of linear estimation inherent in graphical analysis.

5.3 Uptake Ratios (static measures)

While graphical analysis simplifies the estimation of flux constants compared to compartmental analysis, it does not necessarily simplify the imaging procedure. Like compartmental analysis, dynamic curves, including the blood clearance function are necessary, and blood sampling may be required if there is no large blood pool structure in the imaging field. To avoid these requirements, some simpler measures of tracer uptake have been devised. One approach is to compare tumour uptake to a reference normal tissue uptake, yielding a tumor:normal or tumor:background uptake ratio. This

approach has two principal disadvantages for oncology: (1) In tumour images, a suitable "normal" tissue for reference may not exist. This is different from other applications such as brain imaging, where a tissue such as the cerebellum may provide a useful reference for the rest of the brain.[44] (2) The tumour typically has more tracer uptake than the normal or background tissue. Therefore the denominator in the uptake ratio (normal or background uptake) can have high statistical noise, leading to a noisy uptake ratio. Therefore while tumour-to-normal ratios can be helpful, their use in oncologic applications is prone to the pitfalls of variable normal tissue uptake. These ratios are widely used in SPECT, where it is difficult to measure absolute uptake; however, PET affords alternate measures that take advantage of its ability to measure absolute regional tracer concentration.

An alternate approach is to compare uptake in the tissue to the injected dose of radiotracer. This has been the traditional approach in pre-clinical animal studies, using the measure, % injected dose per gram (%ID/g), given by the following expression:[45]

$$\%ID/g = \frac{C_t}{ID} \times 100 \tag{25}$$

where C_t is the tissue tracer content (μCi/g) and ID is the injected dose (μCi). In patient studies, $\%ID$/g measure can be quite variable due to variability in the patient size. A large patient has a larger initial distribution volume for any given tracer and therefore, on the average, lower tissue uptake of tracer per unit injected dose. To compensate for differences in patient size, a variant of $\%ID$/g, typically known as the Standardized Uptake Value (SUV), is used:

$$SUV = \frac{C_t}{ID/wt} \tag{26}$$

where wt refers to the patient weight. With the typical units for tissue concentration (C_t) obtained from the PET image (decay corrected to scan start time) of MBq/cc, MBq for the ID and kg for the patient weight, the SUV has units of g/cc. Preferably, C_t can be divided by tissue density (1 g/cc) to yield tissue content (MBq/kg), in which case the SUV is unitless.

Figure 10 depicts a thought experiment that illustrates the principle of SUV. Consider a W=80 kg patient that is a water bucket, injected with ID=296 MBq of radiotracer, yielding an activity concentration of 3.7 MBq/cc or a content of 3.7 MBq/g. Using the formula given above:

$$SUV = \frac{3.7}{296/80} = 1 \tag{27}$$

In this case, completely uniform distribution of tracer in a subject results in an SUV equal to 1. $SUVs$ less than one imply low uptake or clearance by an

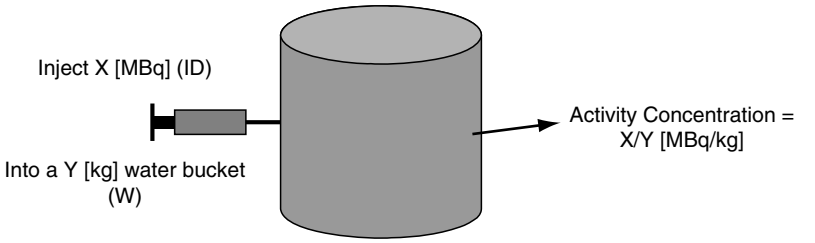


FIGURE 10. Example illustrating the principles of the SUV calculation. Evenly distributed tracer, as in the example of a water bucket, results in an SUV equal to 1.

excretory organ such as the kidneys. An $SUV > 1$ implies tracer concentration or retention by the tissue or tumour site. Zasadny[46] and others have reported typical SUV values for FDG-PET. Note that the value of SUV is strongly dependent upon tracer clearance from the blood and tracer biodistribution, so the range of expected $SUVs$ will vary for different radiopharmaceuticals.

A number of variants on the SUV have been proposed. Zasadny and Wahl and others[46,47] have emphasized that the weight is not always a good measure of initial tracer distribution volume, especially for obese patients, and have suggested SUV using body surface area (BSA) or lean -body mass (lean) in place of patient weight in Eq. 26, yielding SUV_{BSA} and SUV_{lean}, respectively. Most investigators have preferred SUV_{lean} since it retains the same units and approximately the same range as SUV. To account for variability due to difference in blood clearance, Hunter *et al.*[48] proposed the simplified kinetic method, using a single late blood sample to refine the estimate of tracer availability, the denominator in the SUV. While this approach reduces bias and variability in the SUV, it requires blood sampling after injection and may require a second intravenous line, since tracer is often retained at the injection site and may affect blood samples drawn through the same site after injection.

5.4 *Approximations and Sources of Error Compared to Compartmental Analysis*

While uptake ratios are simpler and more clinically feasible than more rigorous kinetic analysis, there are a number of approximations implicit in the use of uptake ratios that may lead to variability and bias. This led one investigator to call the SUV a "silly useless value".[49] However, for any given tracer, the sources of bias and variability can be investigated, and armed with this knowledge; the clinician can make appropriate use of these simplified quantitative measures. This has been best studied for FDG[50] and some aspects of this analysis are reviewed here. While the SUV can be an approximate and accurate substitute for a more rigorous assessment of FDG flux[7]

without attention to detail in patient preparation and image acquisition, *SUV* may be misleading.

The most fundamental approximations for the *SUV* are in the numerator and denominator terms of the expression:

(1) in the numerator, the average tissue uptake over a finite time interval serves as an estimate of tracer uptake and washout from the tissue site;
(2) in the denominator, the injected dose per unit patient weight is an approximation for the blood clearance curve and serves as an indicator of tracer availability to the tumour.

For FDG-PET, this results in the following ensuing sources of variability:[50]

1. alterations in tissue uptake due to competition from blood glucose;
2. changes in tissue uptake as a function of time after injection;
3. alterations in FDG blood clearance;
4. apparent alterations in uptake due to partial volume effects as a result of finite tomograph resolution.

These are reviewed separately below.

5.4.1 Competition Between FDG and Glucose

Most tracers, including FDG directly trace a particular biochemical normally found in the body, which is glucose in the case of FDG. As blood (and therefore plasma) levels of glucose rise, tissues utilizing glucose can use a lower fraction of available blood glucose and still maintain the same flux of glucose through metabolic pathways. This is apparent viewing the equation defining flux constant above (Eq. 12). For a given flux, each unit rise in glucose will result in a proportional fall in the flux constant. This will directly affect the uptake of the glucose analog, FDG, in an approximately inverse relationship. This is described by the following:[50]

$$SUV = SUV_0 \frac{\overline{C_{glu}}}{C_{glu}} \tag{28}$$

where *SUV* refers to the measured FDG *SUV* at a glucose concentration (C_{glu}), SUV_0 indicates the *SUV* that would have been measured for a typical or mean glucose, indicated by $\overline{C_{glu}}$. An approximate correction for changes in glucose is given by the following:[7]

$$SUV' = SUV_m \frac{C_{glu}}{\overline{C_{glu}}} \tag{29}$$

where SUV' refers to the glucose-corrected *SUV*, and SUV_m is the measured *SUV* for blood glucose concentration, C_{glu}. The mean glucose, in typical units used in most laboratories, is 100 mg/dL. This correction yields values

which correlate well with measures of FDG flux over a modest range of plasma glucose levels, up to approximately 150-200 mg/dL.[7]

5.4.2 Changes in Tissue Uptake with Time after Injection

For FDG, the blood clearance time is modest, with finite levels of FDG in the blood for several hours after injection. For tissues that are not metabolically active, where most of the label in tissue is in the form of FDG as opposed to FDG-6P, FDG *SUV* will fall in parallel to the blood curve. Metabolically active tissues will continue to trap and accumulate FDG-6P as long as FDG is still present in the blood and will therefore have rising tissue time-activity curves. Tissues with high levels of G6P phosphorylase, for example the liver and brain, may have modest k_4, leading to a late fall in the tissue time-activity curve.[18,25] This time-variability in *SUV* was recognized by Hamberg[51] who noted significant changes in *SUV* over a period of up to 4 hours after FDG injection. More recently, Thie[52] and Beaulieu[53] have shown significant changes in highly metabolically active tissues with time differences as small as 15-20 minutes. This is not a problem for kinetic analysis or graphical analysis, which explicitly use the shape of the tissue time-activity curve to estimate FDG flux. However, for *SUV* the time-dependence of the FDG uptake can be a significant source of variability when comparing FDG uptake for different patients imaged at different times after injections, or perhaps more importantly, for the same patient imaged at serial time points over the course of treatment. The best approach to eliminating this source of variability in clinical practice is to consistently image at the same time after injection. In a busy clinical practice, this may not always be possible. The correction scheme described by Beaulieu[53] offers an approximate method for compensating for time effects when imaging times vary.

5.4.3 Alterations in Blood Clearance

The variability of the *SUV* depends in large part on the consistency of FDG clearance from the blood. Since FDG availability to tissues is simply the injected dose divided by patient weight, variability in the blood clearance or biodistribution will affect the accuracy of *SUV* as an approximation of glucose flux to the tissues. Since FDG is largely excreted through the kidneys, factors affecting renal clearance will affect FDG clearance. This includes the significant renal dysfunction, altered volume status resulting in a pre-renal state, and hyperglycemia. Significant hyperglycemia is also a problem due to direct competition with FDG, as noted above, and is one of the reasons that *SUV* measures are unreliable in severe or brittle diabetics, even with an empiric correction for glucose. Some approaches such as the simplified kinetic method[48] are more robust to altered FDG blood clearance, but not to the extent of the direct measurement and calculation of flux with compartmental or graphical analysis.

5.4.4 Partial Volume Effects

The effect of object size on the apparent tracer concentration measured by PET has been studied ever since PET instrumentation was first developed,[54] and has been reviewed in chapter 8 of this book. Due to finite resolution, the tomograph may not recover the true concentration of objects that are smaller than 2-3 times the spatial resolution full-width at half-maximum (FWHM). This has been described using the recovery coefficient as follows:[54]

$$RC(d_o,d_{RoI},r) = \frac{C_{meas}}{C_{true}} \tag{30}$$

where RC is the recovery coefficient, C_{meas} is the tissue tracer concentration estimated from the PET scan and C_{true} is the true tissue concentration. The equation explicitly describes recovery coefficient as a function of the object diameter (d_o), the size of the region-of-interest over which uptake is measured (d_{RoI}), and the system spatial resolution (r), all of which affect the RC. The "r" term refers to the spatial resolution of the reconstructed image, which includes the effect of tomograph spatial resolution and also any blurring resulting from the reconstruction algorithm. Changes in reconstruction, for example the reconstruction filter, will affect the RC. Estimates of RC for given d_o, d_{RoI}, and r can be made using "hot sphere" phantom studies with phantoms containing fillable spherical inserts of varying size.[55] Such studies should be done using tracer concentration levels, RoI sizes, and reconstruction algorithms similar to those used in patient image acquisition and analysis.

The effect of partial volume correction can be approximately corrected for if the lesion size can be approximated and the recovery coefficients for the system have been measured. This correction is given by the following:[54]

$$SUV' = \frac{SUV_{lesion} - SUV_{bck}}{RC} + SUV_{bck} \tag{31}$$

where SUV' is the partial volume corrected SUV, RC is the recovery coefficient for the given lesion size, SUV_{lesion} is the SUV measured for the lesion and the SUV_{bck} is the SUV for background tissue adjacent to the lesion. This formula approximately accounts for partial volume count loss for a "hot" lesion adjacent to a cold or "warm" background. Note that when there is no partial volume loss (RC=1), SUV_{bck} terms cancel and SUV' equals SUV_{lesion}.

The correct size RoI to estimate tracer uptake has been the subject of some controversy. Many investigators use an RoI formed by the isocontour of pixels with counts that are a fixed percentage (typically 50%) of the maximum lesion counts. To minimize partial volume loss, some investigators use small regions at the centre of lesions. The smallest region is the lesion pixel with maximum counts ($SUV_{\max}$), and this measure is frequently used in clinical FDG-PET. Although $SUV_{\max}$ is more robust to partial volume

effects than average *SUVs*, partial volume effects still occur for lesions less than 2 times the spatial resolution FWHM and can be severe ($RC = 0.5$ or less) for lesions smaller than 1 FHWM. The best choice depends upon the clinical or biological application and the noise tolerance of the analysis method to be used.

To overcome the limitations of the simple *RC* approach, Chen *et al.*[56] developed a model-based optimization method to simultaneously recover the size and the activity concentration of small spheroid tumours. The results obtained provided evidence that their method improved estimation of radioactivity concentration over that corrected by the *RC* method and that made without any correction. The method still requires further development and validation in clinical setting especially for small lesions, which is the aim of some ongoing studies.[57]

6. Summary

The role of PET during the past decade has evolved rapidly from that of a pure research tool to a methodology of enormous clinical potential. FDG-PET is widely used in the diagnosis, staging, and assessment of tumor response to therapy, since metabolic changes generally precede the more conventionally measured parameter of change in tumor size. Data are accumulating rapidly to validate the efficacy of FDG imaging in a wide variety of malignant tumors with sensitivities and specificities often in the high 90-percentile range. Although metabolic imaging is an obvious choice, the design of specific clinical protocols is still under development. The tracers or combinations of tracers to be used, when the imaging should be done after therapy, the selection of optimal acquisition and processing protocols, the method of accurately performing quantitative or semi-quantitative analysis of data are still undetermined. Moreover, each tumor–therapy combination may need to be independently optimized and validated.

We expect that whole-body FDG-PET-based techniques may be accurate and cost-effective for staging or restaging of different cancer types and can contribute to the improvement of cancer patients' management and monitoring with respect to medical cost. In addition to improving overall scanner performance, further work will focus on implementing practical corrections for patient-related perturbations such as non-homogeneous scatter and photon attenuation. During the next few years, it is believed that such sophisticated techniques will become more widely available in clinical settings and not only limited to research studies in nuclear medicine departments with advanced scientific and technical support. Therefore, it is expected that commercial hardware and software for accurate quantitative analysis will undertake major revisions in the future to be capable of facing emerging clinical applications and challenging research perspectives

Acknowledgments. DAM and MM are supported in part by US NIH/NCI grants P01CA42045, R01CA72064, and R01CA90771. HZ acknowledges the support of Grant SNSF 3152A0-102143 awarded by the Swiss National Science Foundation. The authors wish to acknowledge the helpful discussions and mentorship of Drs. Kenneth Krohn, Michael Graham, and Finbarr O'Sullivan which form the basis for much of the material presented in this chapter.

References

1. Sullivan D. C., Challenges and opportunities for in vivo imaging in oncology. *Technol Cancer Res Treat* **1:** 419-422 (2002).
2. Dewaraja Y. K., Ljungberg, M. and Koral, K. F., Characterization of scatter and penetration using Monte Carlo simulation in ^{131}I imaging. *J Nucl Med* **41:** 123-130 (2000).
3. Siegel J. A., Zeiger, L. S., Order, S. E. *et al.*, Quantitative bremsstrahlung single photon emission computed tomographic imaging: use for volume, activity, and absorbed dose calculations. *Int J Radiat Oncol Biol Phys* **31:** 953-958 (1995).
4. Hoffman E., van-der-Stee, M., Ricci, A. *et al.*, Prospects for both precision and accuracy in positron emission tomography. *Ann Neurol* **15:** S25-34 (1984).
5. Weinberg I. N., Huang, S. C., Hoffman, E. J. *et al.*, Validation of PET-acquired input functions for cardiac studies. *J Nucl Med* **29:** 241-247 (1988).
6. Phelps M. E., Huang, S. C., Hoffman, E. J. *et al.*, Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18)2-fluoro-2-deoxy-D-glucose: validation of method. *Ann Neurol* **6:** 371-388 (1979).
7. Eary J. F. and Mankoff, D. A., Determination of tumor metabolic rates in sarcoma using FDG PET: practical approaches. *J Nucl Med* **39:** 250-254 (1998).
8. Olshen A. B. and O'Sullivan, F., Camouflaged deconvolution with application to blood curve modeling in FDG PET studies. *J Am Stat Assoc* **92:** 1293-1303 (1997).
9. Shields A. F., Graham, M. M., Kozawa, S. M. *et al.*, Contribution of labeled carbon dioxide to PET imaging of carbon-11-labeled compounds. *J Nucl Med* **33:** 581-584 (1992).
10. Minn H., Zasadny, K. R., Quint, L. E. *et al.*, Lung cancer: reproducibility of quantitative measurements for evaluating 2-[F-18]-fluoro-2-deoxy-D-glucose uptake at PET. *Radiology* **196:** 167-173 (1995).
11. Weber W. A., Ziegler, S. I., Thodtmann, R. *et al.*, Reproducibility of metabolic measurements in malignant tumors using FDG PET. *J Nucl Med* **40:** 1771-7 (1999).
12. Maquet P., Dive, D., Salmon, E. *et al.*, Cerebral glucose utilization during sleep-wake cycle in man determined by positron emission tomography and [18F]2-fluoro-2-deoxy-D-glucose method. *Brain Res* **513:** 136-143 (1990).
13. Mankoff D. A., Shields, A. F., Graham, M. M. *et al.*, Kinetic analysis of 2-[C-11]-thymidine PET imaging studies: compartmental model and mathematical analysis. *J Nucl Med* **39:** 1043-1055 (1998).
14. Bassingthwaighte J. B. and Chaloupa, J. M., Sensitivity functions in the estimation of parameters of cellular exchange. *Fed Proc* **43:** 181-184 (1984).

15. Press W., Flannery, B., Teukolsky, S. *et al*., *Numerical Recipes in C*., Cambridge University Press, New York, (1988).
16. Sokoloff L., Reivich, M., Kennedy, C. *et al*., The [^{14}C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *J Neurochem* **28:** 897-916 (1977).
17. Reivich M., Alavi, A., Wolf, A. *et al*., Glucose metabolic rate kinetic model parameter determination in humans: the lumped constant and rate constants for [^{18}F]fluorodeoxyglucose and [^{11}C]deoxyglucose. *J Cereb Blood Flow Metabol* **5:** 179-192 (1985).
18. Spence A., Muzi, M., Graham, M. *et al*., Glucose metabolism in human malignant gliomas measured quantitatively with PET, 1-[C-11]glucose and FDG: analysis of the FDG lumped constant. *J Nucl Med* **39:** 440-448 (1998).
19. Wells J. M., Mankoff, D. A., Muzi, M. *et al*., Kinetic analysis of 2-[C-11]-thymidine PET imaging studies of malignant brain tumors: Compartmental model investigation and mathematical analysis. *Mol Imaging* **3:** 145-150 (2002).
20. Graham M. M., Kinetic evaluation using sensitivity functions and correlation matrices. *J Nucl Med* **36:** 268P (1995).
21. Oldendorf W. H. and Szabo, J., Amino acid assignment to one of three blood-brain barrier amino acid carriers. *Am J Physiol* **230:** 94-98 (1976).
22. Cornford E. M. and Oldendorf, W. H., Independent blood-brain barrier transport systems for nucleic acid precursors. *Biochim Biophys Acta* **394:** 211-219 (1975).
23. Mankoff D. A., Shields, A. F., Link, J. M. *et al*., Kinetic analysis of 2-[C-11] thymidine PET imaging studies: validation studies. *J Nucl Med* **40:** 614-624 (1999).
24. Jonson S. D., Bonasera, T. A., Dehdashti, F. *et al*., Comparative breast tumor imaging and comparative in vitro metabolism of 16α-[^{18}F]fluoroestradiol-17β and 16β-[^{18}F]fluoromoxestrol in isolated hepatocytes. *Nucl Med Biol* **26:** 123-130 (1999).
25. Graham M. M., Muzi, M., Spence, A. M. *et al*., The FDG lumped constant in normal human brain. *J Nucl Med* **43:** 1157-1166 (2002).
26. Mintun M. A., Welch, M. J., Siegel, B. A. *et al*., Breast cancer: PET imaging of estrogen receptors. *Radiology* **169:** 45-48 (1988).
27. Vesselle H., Grierson, J., Muz, i. M. *et al*., In vivo validation of 3′deoxy-3′-[(18)F]fluorothymidine ([(18)F]FLT) as a proliferation imaging tracer in humans: correlation of [(18)F]FLT uptake by positron emission tomography with Ki-67 immunohistochemistry and flow cytometry in human lung tumors. *Clin Cancer Res* **8:** 3315-3323 (2002).
28. Bos R., van Der Hoeven, J. J., van Der Wall, E. *et al*., Biologic correlates of (18)fluorodeoxyglucose uptake in human breast cancer measured by positron emission tomography. *J Clin Oncol* **20:** 379-387 (2002).
29. Shields A. F., Mankoff, D. A., Graham, M. M. *et al*., Analysis of 2-[C-11]-thymidine blood metabolites in PET imaging. *J Nucl Med* **37:** 290-296 (1996).
30. Huang S.-C., Yu, D.-C., Barrio, J. R. *et al*., Kinetics and modeling of L-6-[F-18]Fluoro-DOPA in human Positron Emission Tomographic studies. *J Cereb Blood Flow Metab* **11:** 898-913 (1991).

31. Mankoff D. A., Dunnwald, L. K., Gralow, J. R. *et al*., Blood flow and metabolism in locally advanced breast cancer: relationship to response to therapy. *J Nucl Med* **43:** 500-509 (2002).
32. Zasadny K. R., Tatsumi, M. and Wahl, R. L., FDG metabolism and uptake versus blood flow in women with untreated primary breast cancers. *Eur J Nucl Med Mol Imaging* **30:** 274-280 (2003).
33. O'Sullivan F., Metabolic images from dynamic positron emission tomography studies. *Stat Methods Med Res* **3:** 87-101 (1994).
34. Lassen N. A., Iida, H. and Kanno, I., Parametric imaging in nuclear medicine. *Ann Nucl Med* **9:** 167-170 (1995).
35. Cunningham V. and Jones, T., Spectral analysis of dynamic PET studies. *Cereb Blood Flow Metab* **13:** 15-23 (1993).
36. O'Sullivan F., Imaging radiotracer model parameters in PET: a mixture analysis approach. *IEEE Trans Med Imag* **12:** 399-412 (1993).
37. Cunningham V., Ashburner, J., Byrne, H. *et al*., "Use of spectral analysis to obtain images from dynamic PET studies." in: *Quantification of Brain Function: Tracer Kinetics and Image Analysis in PET*, edited by Uemura K, NA Lasson, T Jones *et al*. Elsevier, Amsterdam, (1993), *pp* 101-108.
38. Patlak C. S., Blasberg, R. G. and Fenstermacher, J. D., Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab* **3:** 1-7 (1983).
39. Gjedde A., Calculation of cerebral glucose phosphorylation from brain uptake of glucose analogs in vivo: A re-examination. *Brain Res Rev* **4:** 237-274 (1982).
40. Patlak C. S. and Blasberg, R. G., Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Generalizations *J Cereb Blood Flow Metab* **5:** 584-590 (1985).
41. Logan J., Graphical analysis of PET data applied to reversible and irreversible tracers. *Nucl Med Biol* **27:** 661-670 (2000).
42. Patlak C. S., Dhawan, V., Takikawa, S. *et al*., "Estimation of striatal uptake rate constant of FDOPA using PET: methodologic issues." in: *Quantification of Brain Function. Tracer Kinetics and Image Analysis in Brain PET*, edited by K. Uemera et al Elsevier Science Publishers, (1993).
43. Mankoff D. A., Shields, A. F., Graham, M. M. *et al*., A graphical analysis method to estimate blood-to-tissue transfer constants for tracers with labeled metabolites. *J Nucl Med* **37:** 2049-2057 (1996).
44. Zaidi H., Montandon, M.-L. and Slosman, D. O., Attenuation compensation in cerebral 3D PET: effect of the attenuation map on absolute and relative quantitation. *Eur J Nucl Med Mol Imaging* **31:** 52-63 (2004).
45. Woodward H. Q., Bigler, R. E., Freed, B. *et al*., Expression of tissue isotope distribution. *J Nucl Med* **16:** 958-959 (1975).
46. Zasadny K. R. and Wahl, R. L., Standardized uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-D-glucose: variations with body weight and a method for correction. *Radiology* **189:** 847-850 (1993).
47. Kim C. K., Gupta, N., Chandramouli, B. *et al*., Stardized uptake values of FDG: body surface area correction is preferable to body weight correction. *J Nucl Med* **35:** 164-167 (1994).
48. Hunter G., Hamberg, L., Alpert, N. *et al*., Simplified measurement of deoxyglucose utilization rate. *J Nucl Med* **37:** 950-955 (1996).

49. Keyes J. W. J., SUV: standard uptake or silly useless value? *J Nucl Med* **36:** 1836-1839 (1995).
50. Huang S.-C., Anatomy of SUV. *Nucl Med Biol* **27:** 643-646 (2000).
51. Hamberg L. M., Hunter, G. J., Alpert, N. M. *et al.*, The dose uptake ratio as an index of glucose metabolism: useful parameter or oversimplification? *J Nucl Med* **35:** 1308-1312 (1994).
52. Thie J. A., Hubner, K. F. and Smith, G. T., Optimizing imaging time for improved performance in oncology PET studies. *Mol Imaging Biol* **4:** 238-244 (2002).
53. Beaulieu S., Kinahan, P., Tseng, J. *et al.*, SUV varies with time after injection in FDG PET of breast cancer: characterization and method to adjust for time differences. *J Nucl Med* **44:** 1044-1050 (2003).
54. Hoffman E. J., Huang, S. C. and Phelps, M. E., Quantitation in positron emission computed tomography: 1. Effect of object size. *J Comput Assist Tomogr* **3:** 299-308 (1979).
55. Geworski L., Knoop, B. O., de_Cabrejas, M. L. *et al.*, Recovery correction for quantitation in emission tomography: a feasibility study. *Eur J Nucl Med* **27:** 161-169 (2000).
56. Chen C. H., Muzic, R. F., Nelson, A. D. *et al.*, Simultaneous recovery of size and radioactivity concentration of small spheroids with PET data. *J Nucl Med* **40:** 118-130 (1999).
57. Schoenahl F. and Zaidi, H., "Towards optimal model-based partial volume effect correction in oncological PET." *Proc.* IEEE Nuclear Science Symposium and Medical Imaging Conference, Oct. 19-22, Rome, Italy (2004) *in press*

17
Quantitative Imaging-Based Dosimetry and Treatment Planning in Radionuclide Therapy

MG Stabin and AB Brill[†]

1. Introduction

One of the important applications for quantitative analyses in nuclear medicine is the measurement of activity in human and animal organs and tissues, for the purposes of calculating radiation dose to these tissues. Knowledge of radiation dose is needed for making appropriate prescriptions of activity to patients for optimal therapeutic benefit, while assuring patient safety. Once reasonably accurate estimates of activity have been obtained, as has been outlined in the other chapters of this book, such data are applied to mathematical models that describe the kinetics of the radionuclide distribution *in vivo* and that estimate the amount of energy deposited in different regions of the body (via simulations of radiation transport and interaction). In this chapter, we outline the basic methods and models currently in use for making these calculations, and for relating calculated dose to biological effects, and discuss some of the uncertainties and limitations of these approaches.

2. Overview of Current Calculational Methods

To define the task of calculating internal doses, we must define the quantities we wish to estimate. The principal quantity of interest in internal dosimetry is the *absorbed dose*, or the *equivalent dose*. Absorbed dose (D) is defined as:[1]

$$D = \frac{d\varepsilon}{dm} \tag{1}$$

where $d\varepsilon$ is the mean energy imparted by ionizing radiation to matter of mass dm. The units of absorbed dose are typically erg/g or J/kg. The special units are

[†]Dr MG Stabin and Dr AB Brill, Department of Radiology and Radiological Sciences, Vanderbilt University, Nashville, TN, USA

Table 1. Radiation Weighting Factors Recommended in ICRP 60.

Alpha particles	20
Beta particles (+/−)	1
Gamma rays	1
X-rays	1

the rad (100 erg/g) or the gray (Gy) (1 J/kg = 100 rad = 10^4 erg/g). Equivalent dose (H) is the absorbed dose multiplied by a 'radiation weighting factor' (w_R, formerly called a 'quality factor, Q'), the latter accounting for the effectiveness of different types of radiation in causing biological effects:

$$H = Dw_R \tag{2}$$

Because the radiation weighting factor is in principle dimensionless, the pure units of this quantity are the same as absorbed dose (i.e. erg/g or J/kg). However, the special units have unique names, the rem and sievert (Sv). Values for the radiation weighting factor have changed as new information about radiation effects has become available. Current values, recommended by the International Commission on Radiological Protection,[2] are given in Table 1.

The quantity equivalent dose was originally derived for use in radiation protection programs. The development of the effective dose equivalent (*EDE*) (to be defined later) by the ICRP in 1979, and the effective dose (*ED*), in 1991,[2] however, allowed nonuniform internal doses to be expressed as a single value, representing an equivalent whole body dose.

2.1 MIRD and RADAR Calculation Systems

2.1.1 MIRD System

The equation for absorbed dose in the MIRD system[3] is given as:

$$D_{r_k} = \sum_h \tilde{A}_h S(r_k \leftarrow r_h) \tag{3}$$

In this system, r_h is a source region and r_k is a target region. The cumulated activity is as defined above; all other terms needed for calculating internal dose are combined in the factor S:

$$S(r_k \leftarrow r_h) = \frac{k \sum_i n_i \; E_i \phi_i(r_k \leftarrow r_h)}{m_k} \tag{4}$$

where n_i is the number of radiations with energy E_i emitted per nuclear transition, E_i is the energy per radiation (MeV), $\phi(r_k \leftarrow r_h)$ is the fraction of energy emitted in source region r_h that is absorbed in target region r_k, m_k is the mass of target region r_k (g or kg) and k is a proportionality constant (rad-g/μCi-hr-MeV or Gy-kg/MBq-sec-MeV).

In the MIRD equation, the factor k is 2.13, which gives doses in rad, from activity in microcuries, mass in grams, and energy in MeV. The MIRD

Committee has published an extensive amount of literature in the area of internal dose assessment in medicine, including 'pamphlets', 'dose estimate reports' and books.

2.1.2 MIRD Literature - The MIRD Pamphlets

A MIRD Pamphlet is a document which generally contains material needed to implement the MIRD schema for internal dose calculations, including equations, data, methods, etc. Several of the MIRD Pamphlets were issued and then revised and reissued; therefore, some of the MIRD Pamphlet titles contain the word 'revised'. Most of the important MIRD Pamphlets are shown below (Table 2). MIRD 1, revised has been superseded by the MIRD Primer (see books, below). MIRD 5 and 5, revised are not much in use, as

Table 2. Selected MIRD pamphlets.

Pamphlet	Publication Date	Main Information	Comments
1, 1 revised	1968, 1976	Discussion of MIRD internal dose technique	Superseded by the MIRD Primer[3]
3	1968	Photon absorbed fractions for small objects	Superseded by Stabin and Konijnenberg[5]
5, 5 revised	1969, 1978	Description of anthropomorphic phantom representing Reference Man, photon absorbed fractions for many organs	Superseded by availability of Cristy/Eckerman phantom series[4]
7	1971	Dose distribution around point sources, electron, beta emitters	Good data, difficult to use; use of Monte Carlo codes like MCNP, EGS is generally preferred
8	1971	Photon absorbed fractions for small objects	Same as Pamphlet 3, smaller objects, also superseded by Stabin and Konijnenberg[5]
11	1975	*S*-values for many nuclides	Newer *S* values available, see RADAR dose factor page
12	1977	Discussion of kinetic models for internal dosimetry	
13	1981	Description of model of the heart, photon absorbed fractions	
14, 14 revised	1992, 1999	Dynamic urinary bladder for absorbed dose calculations	Software to be made available, see RADAR software page
15	1996	Description of model for the brain, photon absorbed fractions	
16	1999	Outline of best practices and methods for collecting and analyzing kinetic data	Widely cited, useful document
17	1999	*S*-values for voxel sources	
18	2001	Administered activity for xenon studies	
19	2003	Multipart kidney model with absorbed fractions	

the Cristy/Eckerman phantom series is widely used.[4] MIRD Pamphlets 3 and 8 are also not much in use, as new absorbed fractions for spheres were calculated,[5] and problems with the MIRD values were pointed out.

2.1.3 MIRD Literature - The MIRD Dose Estimate Reports

The MIRD Dose Estimate Reports contain radiation dosimetry for particular radiopharmaceuticals, along with the kinetic model employed. They were published as separate articles in the Journal of Nuclear Medicine (as are the reports after number 12) and are mostly quite similar. The first table in the article is usually a summary of the dose estimates for all significant organs for unit administrations of the pharmaceutical. Later tables and figures show some of the developmental material used to calculate the dose estimates. All of these estimates contain some useful information, except perhaps in cases in which the pharmaceutical is no longer in use (Table 3). The kinetic models are presented in various ways which must be studied individually.

2.1.4 The MIRD Literature - MIRD Books

The MIRD Committee has published a number of books as well:

1. The MIRD Primer,[3] described above.
2. A set of decay data was published in 1988,[24] which replaced old MIRD compendia of decay data (MIRD pamphlets 4, 6, and 10). These data are, however, also now somewhat dated.

Table 3. Selected MIRD dose estimate reports.

Dose Estimate Report Number	Publication Reference	Compound or Pharmaceutical Studied
1	6	Se-75-L-Selenomethionine
2	7	Ga-66-, Ga-67-, Ga-68-, and Ga-72-Citrate
3	8	Tc-99m-Sulfur Colloid in Various Liver Conditions
4	9	Au-198-Colloidal Gold in Various Liver Conditions
5	10	I-123, I-124, I-125, I-126,I-130,I-131, and I-132 as Sodium Iodide
6	11	Hg-197- and Hg-203-Labeled Chlormerodrin
7	12	I-123, I-124, I-126, I-130, and I-131 as Sodium Rose Bengal
8	13	Tc-99m as Sodium Pertechnetate
9	14	Radioxenons in Lung Imaging
10	15	Albumin Microspheres labelled with Tc-99m
11	16	Fe-52, Fe-55, and Fe-59 Used to Study Ferrokinetics
12	17	Tc-99m Diethylenetriaminepentaacetic Acid
13	18	Tc-99m Labeled Bone Imaging Agents
14	19	Tc-99m Labeled Red Blood Cells
15	20	Radioindium-Lableled Autologous Platelets
16	21	Tc-99m Diethylenetriaminepentaacetic Acid Aerosol
17	22	Inhaled Kr-81m Gas in Lung Imaging
18	23	Indium 111 B72.3 (IgG Antibody to Ovarian and Colorectal Cancer)

3. In 1998, a tabulation of *S* values was published for cellular sources and targets, with activity being on the surface of a cell (of various dimensions), in the cytoplasm, or uniformly throughout the entire cell, with the target being the whole cell or the nucleus.

2.1.5 The RADAR System

In an effort to provide data needed for dose calculations to the user community rapidly and in electronic form, the RAdiation Dose Assessment Resource (RADAR) group was formed.[25] The group maintains an internet web site (current address www.doseinfo-radar.com) where information on internal and external dose assessment is provided; most data are available directly from the site for immediate download. In this section, we describe the models chosen and methods employed by RADAR to calculate dose factors for use in internal dose assessment in nuclear medicine. The last publication of dose factors for radionuclides of interest in nuclear medicine in this format from the MIRD Committee was provided in 1975.[26] Tables of *S* values (rem/μCi-day) were also published as an ORNL document,[27] and selected *SEE* values were presented in ICRP Publication 30.[28] Dose factors, principally for nuclides of interest in nuclear medicine, were made available with the distribution of the MIRDOSE software,[29] but the factors themselves were never published. A new version of the software for internal dose assessment in nuclear medicine is under development by the RADAR group,[30] which will employ the dose factors described in this section. The compilation described in this work provides more radionuclides and models than have been previously available, and uses the RADAR service to make them immediately available to users through interactive processes.

Publication of the technical basis for these factors is important to users. With the advent of "electronic publishing" approaches, it is possible for the voluminous data tables to be distributed through peer reviewed journals, with the technical basis for the data being published in a few pages of journal text. RADAR has employed this method in a number of publications,[31,32] to document the basis for these dose factors and facilitate their distribution to users through an electronic format. In the RADAR system internal dose is calculated by the following simple equation:

$$D = N \times DF \tag{5}$$

where N is the number of nuclear transitions that occur in source region S, and DF is a "dose factor". The factor DF contains the various components shown in the formulas for S and SEE; basically it depends on combining decay data with absorbed fractions (AFs), which are derived generally using Monte Carlo simulation of radiation transport in models of the body and its internal structures (organs, tumours, etc.):

$$DF = \frac{k \sum_i n_i \, E_i \, \phi_i(T \leftarrow S)}{m} \tag{6}$$

2.2 Activity Quantification and Kinetic Analysis

The majority of the work in an internal dose assessment for a patient is in the gathering of good activity estimates and the kinetic analysis of the data. To design and execute a good kinetic study, one needs to collect the correct data, enough data, and express the data in the proper units. The information needed is the fraction (or percent) of administered activity in all important source organs and in excreta samples. It is very important, in either animal or human studies, to take enough samples to characterize the *distribution, retention*, and *excretion* of the radiopharmaceutical over the course of the study. The following criteria are essential:

- Catch the early peak uptake and rapid washout phase.
- Cover at least 3-5 effective half-lives (T_e) of the radiopharmaceutical. ICRU 67[33] suggests points typically be taken near 1/3, 2/3, 1.5, 3, and 5 times the T_e.
- Collect at least two time points per clearance phase.
- Account for 100% of the activity at all times.
- Account for all major paths of excretion (urine, faeces, exhalation, etc.).

Some knowledge of the expected kinetics of the pharmaceutical are needed for a good study design. For example, the spacing of the measurements and the time of the initial measurement will be greatly different if we are studying a ^{99m}Tc-labelled renal agent which is rapidly cleared from the body in a few hours or an ^{131}I-labelled antibody which clears over several weeks. A key point that may be overlooked is the characterization of excretion. Very often, the excretory organs (intestines, urinary bladder) are the organs which receive some of the highest absorbed doses in the body, as much of the activity eventually passes through one or both of these pathways. If excretion is not quantified, one may make the assumption that the compound is retained in the body and removed only by radioactive decay. For very short-lived radionuclides, this may not be a problem and in fact may be quite accurate. For moderately long-lived nuclides, this can cause an overestimate of the dose to most body organs and an underestimate of the dose to the excretory organs, perhaps significantly. The MIRD Committee has published a document which addresses these considerations.[34] The document outlines the proper techniques for data quantification and for appropriate temporal sampling. The authors show how to use the conjugate view technique to acquire quantitative data for dosimetry analyses, including proper choice of source and background regions, with corrections for overlapping source regions, background, and scatter. The use of SPECT and PET techniques also are discussed. Quantitative methods for analyzing

blood and excreta samples are described. On the subject of temporal sampling, they demonstrate that 2 or 3 time points per phase (either uptake or clearance) are needed to adequately describe the kinetics. They also show graphically the amount of error in $\tilde{A}$ (or t) that occurs from neglecting an organ's wash-in phase or not adequately assessing the wash-out phase.

Before beginning clinical trials with radiopharmaceuticals, biokinetic data are usually gathered in animal studies. Organ, uptake measurements are made either after sacrificing the animal (using autoradiography or necropsy techniques) or through use of quantitative imaging methods; these data should be supplemented with blood and excreta measurements. Extrapolation to human values of uptake and clearance is far from a definite science, but these preliminary data provide a basis for going forward with clinical trials if the results are generally favourable. In clinical studies, quantitative imaging methods are used to obtain activity in various organs as a function of time. Planar or Single-Photon Emission Computed Tomography (SPECT) data may be acquired.[34] Obtaining quantitative estimates of the activity per organ is an arduous task, requiring careful attention to details, including corrections for photon scatter, attenuation, septal penetration, and other effects in the gamma camera.[35] When this phase is satisfactorily completed, data may be fit to various retention functions (often sums of exponentials), in closed compartment models using the SAAM II software,[36] or with other tools. Time integrals of activity are calculated and these may be used with dose conversion factors (as discussed below) to obtain radiation dose estimates.

2.3 Available Models for Dose Factors

2.3.1 Body and Organ Models

The current generation of anthropomorphic phantoms began with the development of the Fisher-Snyder phantom.[37] This phantom used a combination of geometric shapes - spheres, cylinders, cones, etc. - to create a reasonably accurate anatomic representation of the body. This phantom was used with Monte Carlo computer programs which simulated the creation and transport of photons through these various structures in the body, whose atomic composition and density were based on data provided in the ICRP report on Reference Man.[38] This report provided various anatomical data assumed to represent the average working adult male in the Western hemisphere. Although this was most directly applicable to use with adult males, the phantom also contained regions representing organs specific to the adult female. Using this phantom, radiation doses were calculated for adults based on activity residing in any organ and irradiating any other organ. Absorbed fractions at discrete photon energies were calculated and published by the MIRD Committee.[39] In addition, dose conversion factors, called S values, which represented the dose to a target region per nuclear

transition in a source region (approximately 20 source and target regions were defined) for over 100 radionuclides, were also published.[26]

The development of the series of phantoms by Cristy and Eckerman[4] allowed dose calculations for individuals of different size and age. Six phantoms were developed, which were assumed to represent children and adults of both genders. Absorbed fractions for photons at discrete energies were published for these phantoms, which contained approximately 25 source and target regions. Tables of *S* values were never published, but ultimately were made available in the MIRDOSE computer software.[29] The 15-year-old phantom was sometimes used as a model of the adult female.

With the publication of phantoms for the adult female, both nonpregnant, and at 3 stages of pregnancy,[40] a more specific model for the adult female was made available, and generally replaced the Cristy/Eckerman 15-year-old phantom in dose calculations for women. These phantoms attempted to model the changes to the uterus, intestines, bladder, and other organs that occur during pregnancy, and included specific models for the foetus, foetal soft tissue, foetal skeleton, and placenta. *S* values for these phantoms were also made available through the MIRDOSE software.

Others have also proposed more detailed models of some organs and other structures, including the brain,[41,42] eye,[41,43] peritoneal cavity,[44] prostate gland,[45] bone,[46,47] rectum,[48] and small unit density spheres, possibly to represent tumours, organs in small animals, or other structures.[5,49]

2.3.2 Bone and marrow models

Spiers *et al.* at the University of Leeds[50] first established electron absorbed fractions (*AFs*) for bone and marrow in a healthy adult male, which were used in the dose factors (*DFs*), or *S* values, in MIRD Pamphlet No. 11.[26] Eckerman re-evaluated this work and extended the results to derive *DFs* for fifteen skeletal regions in six models representing individuals of various ages.[46] The results were used in the MIRDOSE 3 software[29] to provide mean marrow dose, regional marrow dose, and dose-volume histograms for different individuals. Bouchet *et al.*[47] used updated information on regional bone and marrow mass, and calculated new *AFs* using the EGS4 Monte Carlo code. Although the results of the latter two efforts are similar in most characteristics and reported results, the models differed in three areas: (1) Eckerman multiplied *AFs* in the marrow space (MS) by the fractional cellularity (cellularity factor, *CF*) of a given bone to obtain red marrow *AFs*, but Bouchet *et al.* did not, (2) Eckerman assumed that surface-distributed sources of bone-seeking nuclides reside on an infinitely thin layer adjacent to the bone mineral, while Bouchet *et al.* assumed that they were distributed by volume within the 10 micrometer layer of tissue representing the dividing cells on the surfaces of growing bone, and (3) electrons entering the 10 micrometer layer of tissue on bone surfaces were assumed by Eckerman to have entry angles distributed

randomly chosen between 0 and 180 degrees, while Bouchet *et al*. chose angles from a cosine distribution. These differences resulted in significant differences in the reported *AFs* and *DFs* in some cases. Recently, a revised model has been derived[51] which resolves these model differences in ways best supported by currently available data. In this model, marrow-to-marrow electron absorbed fractions were derived using the Bouchet *et al*. values at low energies and the Eckerman values at medium-to-high energies. The rationale for this is that at very low energies, when the electron range is comparable to the dimensions of the cells in the marrow, the electron absorbed fractions should approach 1.0, and not the *CF* in that bone, as in the Bouchet *et al*. model. As the electron energy increases, and the electrons traverse many cell diameters or even marrow cavities before depositing all of their energy, the fraction of total energy deposited in the active marrow should be the *CF* times the fraction deposited in the total marrow space (MS) (as in the Eckerman model), and not simply the absorbed fraction in the MS (as in the Bouchet *et al*. model). As discussed in the article by Stabin *et al*.[51], new data were only available to guide the modifications for the adult model.[52] Nonetheless, adjustments were made as well to the values of $AF(RM \leftarrow RM)$ for the paediatric models, assuming that the values approached 1.0 at low energies and followed similar trends with energy as were noted in the adult. A revision has also been implemented to the cortical bone *AFs* reported by Eckerman and Stabin.[46] Skeletal average *AFs* for all bone regions employed in the calculations in this study are given in Stabin reference above. Values of $AF(RM \leftarrow RM)$ and values of $AF(BS \leftarrow CB\text{-}S)$ and $AF(BS \leftarrow CB\text{-}V)$ differ from those originally published by Eckerman and Stabin[46] for the reasons discussed above.

2.4 Development of Image-based Dosimetry Approaches

As noted in previous chapters of this book, imaging of patients to obtain anatomical and physiological information has matured greatly in the last decade. Anatomic information obtained with Magnetic Resonance Imaging (MRI) or Computed Tomography (CT) is usually expressed in 3 dimensions (3D) in voxel format, with typical resolutions on the order of 1 mm. Similarly, SPECT and PET imaging systems can provide 3D representation of activity distributions within patients, also in voxel format, with typical resolutions of around 5-10 mm. The newest systems now combine CT with both PET and SPECT state of the art imaging systems on the same imaging gantry, so that patient anatomy and tracer distribution can be imaged during a single imaging session without the need to move the patient, thus greatly improving and facilitating image registration. The use of a well-supported radiation transport code such as MCNP or EGS4 with knowledge of patient anatomy will result in a significant improvement in the accuracy of dose calculations.

As a step in the direction of providing pre-calculated dose conversion factors for use with 3D voxel data, voxel source kernels were developed by

Williams *et al.*[53] and by the MIRD Committee.[54] These dose conversion factors in principle allow calculation of 3D dose information from patient-specific data, but only in limited spatial areas where there are no changes in material composition or density (these kernels were developed only for a soft tissue medium, and thus do not work in lung or bone, or where soft tissue/bone or soft tissue/lung interfaces occur). Several other established efforts have attempted to use 3D patient data including heterogeneous voxel distributions with radiation transport algorithms, including several well developed computer codes, as will be described below.

3. Calculational Tools

3.1 MIRDOSE/OLINDA

The MIRDOSE code series began with a Tektronix PC-based MIRDOSE 1 code,[55] then migrated to the MIRDOSE 2 code in the PC-DOS environment, and MIRDOSE 3 and 3.1 in the PC-Windows environment.[29] MIRDOSE has been widely used by the nuclear medicine community and cited in the literature and in presentations at scientific meetings as the basis for presented internal dose estimates. The MIRDOSE 2 and 3 codes implemented the use of whole body MIRD stylized mathematical phantoms representing adult males and females, children, and pregnant women. The codes automated the calculation of internal dose for a large number (>200) of radiopharmaceuticals in these phantoms, the rapid comparison of calculations for different cases, examination of dose contributions to different organs, and regional marrow dose calculations. The code was also widely used as a tool for teaching internal dosimetry in universities and professional training centres. Sample input and output screens from the code are shown in Figure 1. The left panel in Figure 1 illustrates a user interface where information such as the type of radionuclide and element, age, sex, and physiological parameters are specified. The right panel in Figure 1 illustrates the output of the total dose for a reference adult male who is injected Y-90. The output also include information about absorbed doses to various organs that are clinically important.

MIRDOSE is being updated to a new generation code, named OLINDA (Organ Level INternal Dose Assessment),[30] employing the Java programming language and the Java Development Kit environment. The entire code was rewritten, but all of the basic functions of the MIRDOSE code were retained, and others were extended. More individual organ phantoms were included, the number of radionuclides was significantly increased (to over 800, including many alpha emitters), and the ability to perform minor patient-specific adjustments to doses reported for the standard phantoms was made available. A sample output screen from the OLINDA code and a screen involving the implementation of patient-specific organ mass adjustments are shown in Figure 2.

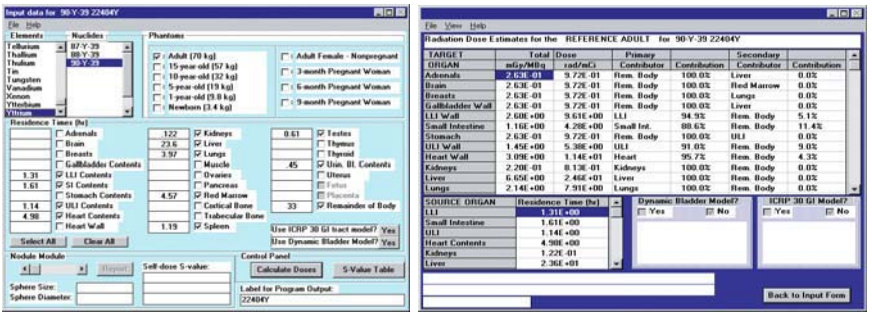


FIGURE 1. Sample input (left) and output (right) screens from the MIRDOSE 3.1 software.

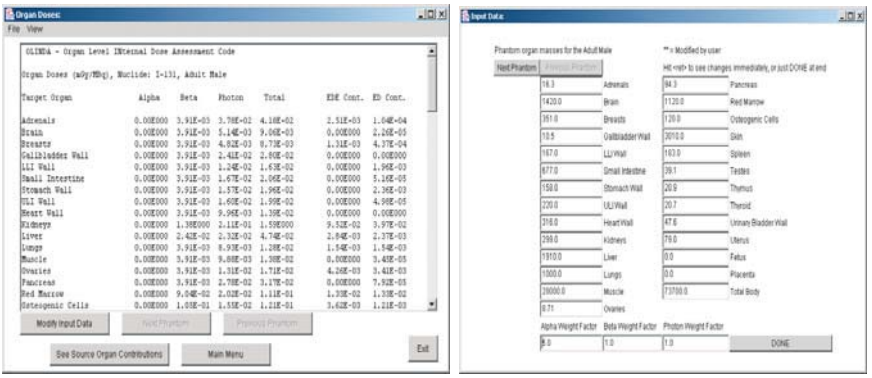


FIGURE 2. Sample output screens from the OLINDA software.

3.2 MABDose

The MABDOSE[56] PC code performs dose calculations; with the possible inclusion of tumour source regions, and integration of organ time-activity curves (the MIRDOSE software requires that the user perform these integrations separately).

3.3 Image-based Computational Tools

Several of the efforts to use image data to perform dose calculations, as described above, include the 3D-ID code from the Memorial Sloan-Kettering Cancer Center,[57,58] the SIMDOS code from the University of Lund,[59] the RTDS code at the City of Hope Medical Center,[60] the MABDose computer code,[56] and the DOSE3D code.[61] The code with the most clinical experience to date is the 3D-ID code. These codes either rely on the standard geometrical phantoms (MABDose and DOSE3D) or patient-specific voxel phantom data (3DID and SIMDOS) and various in-house written routines to perform photon transport. Neither has a particularly robust and well supported

electron transport code, such as is available in EGS or MCNP. The PENELOPE code[62] (PENetration and Energy LOss of Positrons and Electrons) performs Monte Carlo simulation of electron-photon showers in any kind of material. Their adopted scattering model treats radiation transport in the energy range from a few hundred eV to about 1GeV and generates electron-photon showers in complex material structures consisting of any number of distinct homogeneous regions (bodies) with different compositions.

The usual approach used in these codes is to assume that electron energy is absorbed wherever the electron is first produced. The development and support of electron transport methods is quite complex, as evidenced by ongoing intensive efforts by both the EGS4[63] and MCNP[64] computer code working groups. It is not reasonable to expect in-house written codes to deal effectively with electron transport. In areas of highly nonuniform activity distribution, such as an organ with multiple tumours evidencing enhanced uptake of an antibody, explicit transport of both photons and electrons is needed to characterize dose distributions adequately.

Investigators at Vanderbilt University have demonstrated the capability for performing radiation transport in voxel phantoms with the MCNP Monte Carlo radiation transport code for internal sources[65,66] in the voxel phantom provided by the group at Yale,[67] while investigators at Rensselaer Polytechnic Institute have demonstrated the capability using the EGS code for external sources[68,69] in the VIP man voxel phantom.[70] Jones[71] reported work performed at the NRPB, UK on an adult male model called NORMAN using MR images of 2-mm $\times$ 2-mm resolution and 10-mm slice thickness. Their model was used to estimate organ doses from external photon sources over a range of energies and irradiation geometries. When comparing their calculations to those which used a MIRD-type stylized model, differences in organ doses were found to range from a few percent to over 100% at photon energies between 10 and 100 keV. Petoussie-Henss *et al.*[72] reported a family of tomographic models developed from CT images of 2-mm $\times$ 2-mm resolution and 8-mm slice thickness. Dose coefficients from external irradiation with these phantoms were substantially different than values derived using the MIRD phantom, suggesting to these authors that the MIRD models do not represent well a large proportion of the population.

4. Design of Studies to Obtain Dosimetry Data

The key to a successful dosimetry study is in obtaining all of the correct data. The analysis phase is relatively straightforward, at least for the majority of work done today, which relies on the standardized anthropomorphic phantoms described above. In the future, as more and more approaches involve image-based methods with Monte Carlo calculations performed for each subject, additional concerns will become important, including issues regarding image registration, organ segmentation, statistical reliability of the

Monte Carlo results, presentation and evaluation of the 3D dose results, and other matters. For studies involving animal data or planar images from human subjects, the key issues are:

- Accurate quantification of data.
- Collecting enough data to make it possible to perform a good kinetic analysis (as described in some detail above).
- Characterizing both organ retention and systemic excretion.
- Developing a cohesive biokinetic model, including prudent assumptions about biokinetic behaviour after the time of the last datum (typical assumptions include assuming only radioactive decay, or assuming the same rate of biological removal as in the last observed phase(s) of clearance).
- Correct characterization of the time-integrals of the biokinetic functions, for use in the dose calculations.
- Choice of an appropriate anthropomorphic model for dosimetry.
- Possible patient-specific modifications to the calculated doses (see below).

5. Application of Methods and Tools in Clinical Trials and Clinical Practice

5.1 Balancing Need for More Data with Logistical Considerations

In either animal or human studies, there is always a tension between the desire for more data, which ultimately improves the understanding of the system and the accuracy of the reported results, vs. logistic considerations, including costs, time, and animal or patient comfort. As was noted above, an absolute minimum of two data points are needed per phase of organ uptake or clearance, so this dictates the minimum number of points needed to account for routine data fluctuations and uncertainties. For very short lived nuclides, as is the case with many positron emitters, there is rarely enough time for more than one phase of clearance to be observed, so two or three points is usually sufficient. For longer lived nuclides, the biological behaviour of the nuclide is usually of most importance. Biphasic clearance patterns are very common with radiopharmaceuticals, from individual organs or the whole body, so 4-6 points is usually the minimum possible (with the points suitably spaced to ensure obtaining two points in each phase. If a pharmaceutical's clearance is truly monophasic, then obtaining 5 or more data points is really not necessary; 3-4 spaced over two or more effective half-times will be sufficient. With animals imaged or sacrificed at each time, the main consideration is cost, of the animals and the staff time to manage the study. With patients or healthy volunteers, one must consider what is logistically possible for the staff and what is tolerable by the subjects.

5.2 Limitations on Accuracy of Data from Planar and Tomographic Images

Planar images provide a projection of activity as viewed by the collimated detector which has a varying resolution and sensitivity as a function of depth in the body. The best resolution is at the surface with poorer resolution at increasing depth. Sensitivity also falls with depth due to attenuation resulting from absorption and scatter. Further imperfections arise when there is superposition of activity from different overlapping source regions. Tomographic images are able to overcome these sources of distortion given knowledge of the physics of the device, and the interactions of the radionuclide in the body (scattering and absorption), and in the detector (collimator and imaging receptor). The geometric mean of the number of events sensed by opposing detectors is able to decrease the magnitude of variations of counts per emitted event coming from different depths, assuming a uniform distribution of activity in non overlapping structures. The accuracy of quantitative estimates from planar images is less than that achieved from PET or SPECT procedures, where accuracy of the order of $\pm 5\%$ can be achieved. With planar imaging, it is possible to achieve 10% accuracy, although this will vary as a function of source size (smaller objects will emit fewer counts, suffer from partial volume losses more than larger objects). Imaging of 140 keV emissions from Tc-99m is optimum for a gamma camera due to low collimator penetration, and good stopping power of the NaI(Tl) detector. Scatter distortion is worse at low energies, such as from Tl-201, and from other commonly used nuclides with multiple energy emissions (e.g Ga-67, and In-111). PET images have the benefit of better means of correcting for attenuation, but suffer from the blur in the images due to the range of the positron before it comes to a stop and results in the back to back emission of a pair of 511 keV photons, at approximately 180 degrees apart.

5.3 Relating Image Data to Anatomic Regions

The recent development of imaging instruments that combine high resolution CT with PET, SPECT, and planar imaging devices has greatly simplified the ability to relate the *in vivo* distribution sensed by the nuclear imaging devices to body regions. The use of transmission sources in the past provided reasonable accuracy in the reconstruction of activity that matched the spatial resolution achieved in the nuclear images. However, delineation of organ boundaries was very limited, and this has been improved remarkably now that high resolution clinical quality CT systems are in common use on what are now called hybrid cameras. There are two ways in which image registration is achieved. In the first, CT images taken before or after the nuclear scan achieve good positional registration since the patient is on the same imaging bed, and is not moved between procedures. Other methods have been developed where a higher energy gamma source is imaged at the same time as the

lower energy emissions from the labelled tracer are acquired. In this case, any internal or external movements that occur affect both imaging results, and although they may be more blurred, that is a better reflection of reality than when the images are performed sequentially. The relative merits of the two approaches will require time and experience to judge.

Given activity in different spatially defined regions, the next task is to relate those locations to anatomic regions in the body for all imaging sessions. For planar images, experienced investigators can usually identify most important anatomical structures, manually draw boundaries that encompass these structures, and extract the number of counts from the regions from images taken at different times. Relating counts to absolute activity may be done by relative means (using the number of counts known to represent 100% of the activity) or absolute methods (from previously performed attenuation and calibration studies). In tomographic imaging, the different anatomic images (CT/MRI) must be registered with each other and with the nuclear medicine images, and then the different anatomic structures in which the radioactivity is localized, and where absorbed dose is to be calculated, must be segmented. The task of registration of PET/CT and SPECT/CT has been simplified by the commercial development of machines embodying both techniques (SPECT or PET with CT in a single camera system). Many groups have studied methods for segmentation (e.g. Dawant *et al.*[73]) and registration (e.g. Maurer *et al.*[74]) of brain and spine images from whole body MRI imaging. Techniques include manual delineation and editing capabilities, basic image processing operations such as linear and non-linear filtering, morphological operations, multi-spectral classification (maximum-likelihood classifiers and neural networks), and deformable model algorithms (see chapters 9 and 10).

5.4 Making Patient-specific Modifications

Some patient-specific modifications to dose calculations made with standardized anthropomorphic phantoms are possible, and are certainly indicated in therapy applications. One can make the dose estimates calculated with standardized anthropomorphic phantoms more patient-specific through mass-based adjustments to the organ doses:[75]

- Specific absorbed fractions for electrons and alphas scale linearly with mass ($\Phi' = \Phi(m'/m)$).
- Absorbed fractions for photons scale with mass to the 1/3 power ($\phi' = \phi(m'/m)^{1/3}$).

One generally *cannot* accurately:

- Account for patient-specific differences in organ geometry.
- Account for patient-specific marrow characteristics.
- Calculate dose distributions within organs.

To perform accurate patient-specific dose calculations, one needs a patient-specific anatomical model to be used with patient-specific biokinetic data. A ''one-dose-fits-all'' approach to radiation therapy with internal emitter treatments, as is currently advocated by many, is not likely to be adequate (due to the narrow range between tumour ablation and bone marrow toxicity). Individual patients not only have significantly different uptake and retention half-times of activity of the radioactive agent, but also have significantly different physical characteristics and radiosensitivities. Many cancer patients have failed other treatments, and may enter radiotherapy with compromised marrow due to their previous treatments. Thus, their therapies should be optimised, taking into account effects of previous therapy as well as the other measured parameters as much as is possible.

5.5 Relating Dose to Effect

The goal of all forms of radiation therapy against cancer is to give lethal doses of ionizing radiation to malignant cells while not exceeding the radiation tolerance of involved normal tissues. Knowledge of the tolerance of normal tissues to radiation is thus essential to the successful design of therapy. Understanding of these tolerances has grown over the years since radiation therapy has become routine, from experience with many patients and sharing of data. A number of standard toxicity scoring systems have been developed, including those of the National Cancer Institute (NCI), Southwestern Oncology Group (SWOG), Radiation Therapy Oncology Group (RTOG), and World Health Organization (WHO). In these systems, severity of effects is separated into categories, or ''grades'', with scores ranging from 0 to 4 or 5, with higher grades implying more severe toxicity, even to the point of mortality. For hematological parameters, grades are generally associated with numerical values of the levels of various elements in peripheral blood (platelets, lymphocytes, etc.), whereas for organs the grades are established by less objective, phenomenological observations (inflammation, degradation of function, ulceration, etc.). It is important to remember that criteria for grading may be quite circumstance-dependent, e.g. the definition of ''normal'' blood levels in cancer patients (and thus the changes that suggest that a particular grade of toxicity is indicated) may vary from that for normal healthy adults. Some scales, e.g. the RTOG, use different criteria for acute (defined as up to 90 days post-treatment) versus late toxicity. Quality of Life (QOL) indicators may also be employed in clinical trials and other evaluations, including measures such as pain, use of analgesics, constipation, mood, fatigue, and other measures of performance status, e.g. physical activity.

New radiolabelled agents under investigation for possible use in treatment of cancer are evaluated in clinical trials in which safety and efficacy are studied carefully. Safety concerns include possible toxicity associated with radiation exposure of healthy tissues in the body. Generally, dosages

thought to be useful against the disease are given in graded steps, with careful observation of deleterious side effects. If a drug can be shown to have good efficacy while resulting in low, manageable rates of occurrence of normal tissue toxicity, it is likely to be approved for more widespread use. Patient cohort size varies, as do the definitions of maximum tolerable doses (MTDs), but generally at least 5 patients must be studied in Phase I trials, and the MTD is generally defined as the dose that results in moderate to severe toxicity in the majority of this cohort. Escalation of administered dosages (and thus radiation doses) is generally accomplished in fixed steps of 20-25%, while maintaining vigilance over undesired side effects of the treatment. In Phase II and III trials, more patients are evaluated and doses lower than MTDs are given, with an eye on efficacy and clinical feasibility. The use of MTD as the target dose is open to some debate as there circumstances in which different effects are induced at high doses that may block the beneficial effects observed at lower doses.

Organ tolerance data for external beam irradiation has been derived generally using treatment plans which deliver doses at high dose rates (HDR), perhaps 2 Gy in a single fraction at dose rates of around 1 Gy/min, with 5 fractions received per week for up to 8 weeks. Thus the actual dose rates for a single treatment are high, but 24 hours elapse between most fractions, with 72 hours between some. When radiation is delivered using radionuclides, the dose is delivered in a continuous but constantly decaying fashion over periods of days to weeks, depending on the effective half-life of the radiopharmaceutical in the tumour or organ. The dose rates experienced will be generally lower, but there are no time periods during which radiation dose is not being received. Thus the relationship to be expected between known organ tolerances for fractionated HDR external beam therapy and dose from continuous, LDR internal sources is not well established. In addition, profound differences may exist among radionuclides with respect to their radiotoxicity.

From both several decades of widely published data and from the multi-year focused effort funded by the NCI in the late 1980's, tolerances to external beam fractionated radiation for a number of normal tissues in the human have been well documented. Data are expressed in terms of NTCPs of 5/5 and 50/5 (i.e. a 5% or 50% complication probability in 5 years following exposure) for 1/3, 2/3, or 3/3 of an organ or tissue being irradiated.[76] By contrast, tolerance doses for radionuclides are only available from a relatively small number of patients. A recent overview article by Meredith,[77] summarizes current knowledge in this area. In her paper, Dr. Meredith points out several problems with the interpretation of these data, namely that (1) the internal dose data are generally not as accurate as external dose data, as they do not as carefully account for lack of radionuclide homogeneity within the tissues, tissue density changes which may have affected the image quantification, and other factors, (2) tracer studies used to establish the biokinetics (and thus the dosimetry) may have different

biokinetic patterns than when the full therapeutic dose is given, as has been documented in several studies, and (3) lack of uniformity of reporting of internal dose results, depending on which image quantification methods, computer programs, dose conversion factors, etc. were used. Interpretation of the findings is difficult due to the complexities in the format of the reported results as well. Nonetheless, the data may suggest that a higher tolerance has been observed for internally administered radionuclides, implying that the low dose rate effects cause a lower tissue response given the same level of absorbed dose (Gy). It is also, however, reasonable to expect that this effect will vary considerably with differences in radiopharmaceutical effective half-time in the tissue. Short-lived agents will deliver their dose in a shorter time, thus more approximating the dose rates of external beam radiation, while longer lived agents, whose dose is delivered over many days to weeks, will deliver the dose to the tissues at a considerably lower dose rate especially at later times.

Bone marrow toxicity is the major dose-limiting factor in radioimmunotherapy (RIT), but there is no consensus on how to calculate that dose accurately, or of individual patients' ability to tolerate the planned therapy.[78-85] Experience with external beam therapy has shown that there is a narrow margin between the delivered dose that kills a tumour and that which causes serious injury to the patient. Given that with modern methods the dose from external beam treatments is accurately known, the remaining major issues are tumour sensitivity, and patient tolerance. In nuclear medicine, in contrast to beam radiotherapy, marrow absorbed dose is not accurately evaluated, even when a treatment planning dosimetry study is performed.[86] Both external beam and internal emitter therapy share the same uncertainty regarding predicted patient tolerance. Another problem is that nuclear medicine physicians, in general, have less experience in the therapy regime and typically undertreat patients in order to minimize the risk of even low-grade toxicity. If treatment of bone cancer with radioisotopes and RIT are to become a primary means of treating cancer patients, we need better ways of determining the amount of activity to be administered to individual patients, which is based on dose needed to treat the tumour and patient tolerance.[87]

Hepatotoxicity with labelled ^{90}Y-labelled microspheres was suggested to be lower than would have been expected from knowledge of dose-related toxicity from external sources of radiation.[88] Behr *et al.* have extensively studied radiation nephrotoxicity following therapy with radiopeptides and antibody fragments, including the development of methods to prevent these side-effects in pre-clinical as well as clinical settings.[89-91] In a recent editorial,[91] they point out that "... no maximum tolerated renal dose (MTD) is known for internal emitters. The assumption of similarity between external beam irradiation (doses ~20-25 Gy may cause nephrotoxicity in 1%-5% of patients) ... and internal emitter tissue tolerance doses may be a conservative and thus a safe one, but it is not supported by any experimental or

clinical evidence'' and that ''. . . dose rate, penetration range and heterogeneity of microdosimetry differences suggest higher MTDs for internal emitters than for external beam radiation.'' Behr *et al.*[92] have suggested that the nephrotoxic potential of Auger electron emitters (such as ^{111}In, ^{67}Ga, or ^{140}Nd) is much lower than that of ^{177}Lu-labelled peptides. They suggest that perhaps as a consequence of their short path length, they selectively irradiate the tubular cells, thereby sparing the much more radiosensitive glomeruli. Auger electron radiation has demonstrated in some preclinical trials higher anti-tumour effects when the radionuclide is internalized into the tumour cells, as is the case for most peptide type radiopharmaceuticals. Auger electron emitters such as ^{195m}Pt may prove to be more effective than traditional beta emitters in therapy.[93,94]

6. Future Developments Needed/Anticipated and Their Impact on Dosimetry

Rapid ongoing developments in molecular biology, radiation sciences, and computing technology are having a major impact on many areas in biology and medicine. As one gains increasing knowledge of the normal and altered pathways that govern cellular processes, it becomes increasingly possible to design drugs that interact with and modulate specific selected targets. Advances in chemical synthesis, directed synthesis, as well as combinatorial chemistry are providing designer molecules for diagnosis and therapy. Linkage of these molecules to suitable carrier systems, is further aided by the addition of toxins such as radioactive nuclides that track the biokinetics, and depending on the goal (diagnosis vs. therapy), one uses radionuclides with different properties. Where the drug is delivered close to the target of interest, and therapy is desired, a short range high LET agent such as an alpha particle or Auger electron emitter can be the optimum choice. Where diagnosis is desired as part of a screening or treatment planning paradigm, then a low LET emitter with high photon yield is the optimum choice.

For diagnosis, given an ideal tracer, one that only goes to the region of interest and has energy emissions well suited for the available imaging devices; even a poor imaging device will provide excellent results. Given, a tracer with poor uptake and biodistribution characteristics, the converse is not true, and then a very high quality imaging device is needed to achieve acceptable results. Much improved high speed, low cost computing resources are now available and these have made it possible for industry to design, build and commercially deliver excellent X- and gamma-ray imaging devices for planar and tomographic imaging. The ability to image single photon emitters with gamma cameras has improved remarkably so that high spatial and temporal resolution can be achieved and displayed as 2-, 3-, and 4-D images and similar advances have been made in the imaging of

positron-emitting radionuclides. The development promises to continue using new radiation detector materials, and associated readout methods improving performance in terms of sensitivity, resolution, field of view, and possibly in the range of energies that can be imaged.

For a long time there has been an interest in verifying dose by direct measurement. Thermoluminescent (TLD) devices have a wide relatively flat response independent of energy, and a wide linear sensitivity to dose. These have been used mostly in inanimate phantoms, but have been used in animal studies, and have been proposed for some human applications. These are integrating devices and read out the dose absorbed between the time implanted and the time when retrieved from the subject. Small thermoluminescent dosimeters (TLDs) have been very useful in anthropomorphic phantoms in calibration of diagnostic and therapeutic external radiation sources, and attempts were made to place them into small animals, principally in tumours to measure accumulated radiation doses over time.[95,96] Calibration of the *in vivo* TLDs and interpretation of the data, however, proved problematic, most likely due in part to interactions of the dosimeter material with tissues. Glass-encapsulated MOSFET detectors have been successfully used for radiation dosimetry with IMRT, and RIT[97,98] and are under continued development. These devices show potential for use with high energy beta monitoring (such as with Y-90), but have difficulties due to attenuation in the glass wall of the device. Nonetheless, they may be of value in the validation of absorbed dose estimates from external beam and from internal emitter therapy, such as with Y-90 spheres being used in nuclear medicine therapy.

References

1. ICRU, International Commission on Radiation Units and Measurements; Report No. 33, 1980.
2. ICRP, International Commission on Radiological Protection; Publication 60, 1991.
3. Loevinger R., Budinger, T. F. and Watson, E. E., *MIRD primer for absorbed dose calculations, Revised*. The Society of Nuclear Medicine, New York, (1991).
4. Cristy M. and Eckerman, K. F., Oak Ridge National Laboratory; Report No. ORNL/TM-8381/V1 to ORNL/TM-8381/V7, 1987.
5. Stabin M. G. and Konijnenberg, M. W., Re-evaluation of absorbed fractions for photons and electrons in spheres of various sizes. *J Nucl Med* **41:** 149-160 (2000).
6. Lathrop K. A., Johnston, R. E., Blau, M. *et al*., MIRD Dose Estimate Report No. 1: selenium-75-L-selenomethionine. *J Nucl Med* **14:** 49-50 (1973).
7. Cloutier R. J., Watson, E. E., Hayes, R. L. *et al*., MIRD Dose Estimate Report No. 2: Summary of current radiation dose estimates to humans from 66Ga-, 68Ga-, and 72Ga-citrate. *J Nucl Med* **14:** 755-756 (1973).
8. Atkins H. L., Cloutier, R. J., Lathrop, K. A. *et al*., MIRD Dose Estimate Report No. 3: Technetium-99m-sulfur colloid in various liver conditions. *J Nucl Med* **16:** 108A-108B (1975).

9. Cloutier R. J., Freeman, L. M., Mc Afee, J. G. *et al.*, MIRD Dose Estimate Report No. 4: Au-198-Colloidal Gold in Various Liver Conditions. *J Nucl Med* **16:** 173-174 (1975).
10. Berman M., Braverman, L. E., Burke, J. *et al.*, MIRD Dose Estimate Report No. 5: I-123, I-124, I-125, I-126, I-130, I-131 and I-132 as Sodium Iodide. *J Nucl Med* **16:** 857-860 (1975).
11. Blau M., Mc Afee, J. G., Rohrer, W. S. *et al.*, MIRD Dose Estimate Report No. 6: Hg-197- and Hg-203-Labeled Chlormerodrin. *J Nucl Med* **16:** 1095-1098 (1975).
12. Freeman L. M., Patton, D. D., Rosenthall, L. *et al.*, MIRD Dose Estimate Report No. 7: Summary of current radiation dose estimates to humans from 123I, 124I, 126I, 130I, and 131I as sodium rose bengal. *J Nucl Med* **16:** 1214-1217 (1975).
13. Lathrop K. A., Atkins, H. L., Berman, M. *et al.*, MIRD Dose Estimate Report No. 8: Summary of current radiation dose estimates to normal humans from 99mTc as sodium pertechnetate. *J Nucl Med* **17:** 74-77 (1976).
14. Atkins H. L., Robertson, J. S., Croft, B. Y. *et al.*, MIRD Dose Estimate Report No. 9: Radioxenons in Lung Imaging. *J Nucl Med* **21:** 459-465 (1980).
15. Blau M., Wicks, R., Thomas, S. R. *et al.*, MIRD Dose Estimate Report No. 10: Albumin microspheres labeled with Tc-99m. *J Nucl Med* **23:** 915-917 (1982).
16. Robertson J. S., Price, R. R., Budinger, T. F. *et al.*, MIRD Dose Estimate Report No. 11: Fe-52, Fe-55, and Fe-59 Used to study Ferrokinetics. *J Nucl Med* **24:** 339-348 (1983).
17. Thomas S. R., Atkins, H. L., Mc Afee, J. G. *et al.*, MIRD Dose Estimate Report No. 12: Tc-99m-diethylenetriaminepentaacetic acid. *J Nucl Med* **25:** 503-505 (1984).
18. Weber D. A., Makler, P. T., Jr., Watson, E. E. *et al.*, MIRD Dose Estimate Report No. 13: Radiation absorbed dose from technetium-99m-labeled bone imaging agents. *J Nucl Med* **30:** 1117-1122 (1989).
19. Atkins H. L., Thomas, S. R., Buddemeyer, U. *et al.*, MIRD Dose Estimate Report No. 14: Radiation absorbed dose from technetium-99m-labeled red blood cells. *J Nucl Med* **31:** 378-380 (1990).
20. Robertson J. S., Ezekowitz, M. D., Dewanjee, M. K. *et al.*, MIRD Dose Estimate Report No. 15: Radiation absorbed dose estimates for radioindium-labeled autologous platelets. *J Nucl Med* **33:** 777-780 (1992).
21. Atkins H. L., Weber, D. A., Susskind, H. *et al.*, MIRD Dose Estimate Report No. 16: Radiation absorbed dose from technetium-99m-diethylenetriaminepentaacetic acid aerosol. *J Nucl Med* **33:** 1717-1719 (1992).
22. Atkins H. L., Robertson, J. S. and Akabani, G., MIRD Dose Estimate Report No. 17: Radiation absorbed dose estimates from inhaled krypton-81m gas in lung imaging. *J Nucl Med* **34:** 1382-1384 (1993).
23. Mardirossian G., Brill, A. B., Harwood, S. J. *et al.*, MIRD Dose Estimate Report No. 18: Radiation absorbed dose estimates for indium-111-labeled B72.3, an IgG antibody to ovarian and colorectal cancer. *J Nucl Med* **39:** 671-676 (1998).
24. Weber D. A., Eckerman, K. F., Dillman, L. T. *et al.*, *MIRD: Radionuclide data and decay schemes*, The Society of Nuclear Medicine, New York, (1989).
25. Stabin M. G., Siegel, J., Hunt, J. *et al.*, RADAR - the radiation dose assessment resource. An online source of dose information for nuclear medicine and occupational radiation safety. [abstract] *J Nucl Med* **42:** 243P (2001).

26. Snyder W. S., Ford, M. R., Warner, G. G. *et al*., *MIRD Pamphlet No. 11: 'S', Absorbed dose per unit cumulated activity for selected radionuclides and organs*, The Society of Nuclear Medicine, New York, (1975).
27. Snyder W., Ford, M., Warner, G. *et al*., Oak Ridge National Laboratory; Report No. ORNL-5000, 1975.
28. ICRP, *ICRP publication 30. Limits for intakes of radionuclides by workers*, Pergamon Press, New York, (1979).
29. Stabin M. G., MIRDOSE: personal computer software for internal dose assessment in nuclear medicine. *J Nucl Med* **37:** 538-546 (1996).
30. Stabin M. G. and Sparks, R. B., MIRDOSE4 does not exist. [abstract] *J Nucl Med Supplement* **40:** 309P (1999).
31. Stabin M. G. and da Luz, C. Q. P. L., New decay data for internal and external dose assessment. *Health Phys* **83:** 471-475 (2002).
32. Stabin M. G. and Siegel, J. A., Physical models and dose factors for use in internal dose assessment. *Health Phys* **85:** 294-310 (2003).
33. ICRU, International Commission on Radiation Units and Measurements; Report No. ICRU Report 67, 2002.
34. Siegel J. A., Thomas, S. R., Stubbs, J. B. *et al*., MIRD Pamphlet No. 16: Techniques for quantitative radiopharmaceutical biodistribution data acquisition and analysis for use in human radiation dose estimates. *J Nucl Med* **40:** 37S-61S (1999).
35. King M. A., Glick, S. J., Pretorius, P. H. *et al*., "Attenuation, scatter, and spatial resolution compensation in SPECT." in: *Emission Tomography: The Fundamentals of PET and SPECT*, edited by M Wernick and J Aarsvold Academic Press, San Diego, (2004), *in press*
36. Foster D. and Barrett, P., "Developing and testing integrated multicompartment models to describe a single-input multiple-output study using the SAAM II software system." Proc. Sixth International Radiopharmaceutical Dosimetry Symposium, Oak Ridge, TN: Oak Ridge Associated Universities, *pp* 577-599 (1999).
37. Snyder W. S., Ford, M. R., Warner, G. G. *et al*., MIRD Pamphlet No. 5: Estimates of absorbed fractions for monoenergetic photon sources uniformly distributed in various organs of a heterogeneous phantom. *J Nucl Med* **10:** Suppl 3: 5-52 (1969).
38. ICRP, *ICRP publication 23. Report of a task group on reference man*. Pergamon Press, New York, (1975).
39. Snyder W. S., Ford, M. R. and Warner, G. G., *MIRD Pamphlet No. 5 (Revised): Estimates of absorbed fractions for monoenergetic photon sources uniformly distributed in various organs of a heterogeneous phantom*., The Society of Nuclear Medicine, New York, (1978).
40. Stabin M. G., Watson, E. E., Cristy, M. *et al*., Oak Ridge National Laboratory; Report No. ORNL/TM-12907, 1995.
41. Eckerman K., Cristy, M. and Warner, G., "Dosimetric evaluation of brain scanning agents." Proceeding of Third International Radiopharmaceutical Dosimetry Symposium. HHS Publication FDA 81-8166, *pp* 527-540 (1981).
42. Holman B., Zimmerman, R., Shapiro, J. *et al*., Biodistribution and dosimetry of n-isoproyl p-123I iodoamphetamine in the primate. *J Nucl Med* **24:** 922-931 (1983).

43. Bouchet L. G. and Bolch, W. E., Five pediatric head and brain mathematical models for use in internal dosimetry. *J Nucl Med* **40:** 1327-1336 (1999).
44. Watson E. E., Stabin, M. G., Davis, J. L. *et al.*, A model of the peritoneal cavity for use in internal dosimetry. *J Nucl Med* **30:** 2002-2011 (1989).
45. Stabin M. G., A model of the prostate gland for use in internal dosimetry. *J Nucl Med* **35:** 516-520 (1994).
46. Eckerman K. F. and Stabin, M. G., Electron absorbed fractions and dose conversion factors for marrow and bone by skeletal regions. *Health Phys* **78:** 199-214 (2000).
47. Bouchet L. G., Bolch, W. E., Howell, R. W. *et al.*, S values for radionuclides localized within the skeleton. *J Nucl Med* **41:** 189-212 (2000).
48. Mardirossian G., Tagesson, M., Blanco, P. *et al.*, A new rectal model for dosimetry applications. *J Nucl Med* **40:** 1524-1531 (1999).
49. Siegel J. A. and Stabin, M. G., Absorbed fractions for electrons and beta particles in spheres of various sizes. *J Nucl Med* **35:** 152-156 (1994).
50. Spiers F. W., "Beta dosimetry in trabecular bone" in: *Delayed Effects of Bone-Seeking Radionuclides*., edited by C.W. Mays Univ of Utah Press, Salt Lake City, UT, (1969), *pp* 95-108.
51. Stabin M., Eckerman, K., Bolch, W. *et al.*, Evolution and status of bone and marrow dose models. *Cancer Biother Radiopharm* **17:** 427-433 (2002).
52. Jokisch D. W., Patton, P. W., Inglis, B. A. *et al.*, NMR microscopy of trabecular bone and its role in skeletal dosimetry. *Health Phys* **75:** 584-596 (1998).
53. Williams L. E., Liu, A., Raubitschek, A. A. *et al.*, A method for patient-specific absorbed dose estimation for internal beta emitters. *Clin Cancer Res* **5:** 3015s-3019s (1999).
54. Bolch W. E., Bouchet, L. G., Robertson, J. S. *et al.*, MIRD pamphlet No. 17: The dosimetry of nonuniform activity distributions. Radionuclide S values at the voxel level. *J Nucl Med* **40:** 11S-36S (1999).
55. Watson E. E. and Stabin, M. G., "BASIC alternative software package for internal dose calculations." Computer Applications in Health Physics, Proceedings of the 17th Midyear Topical Symposium of the Health Physics society, Pasco, WA, Feb 5-9, 1984, *pp* 79-86 (1984).
56. Johnson T. K., McClure, D. and McCourt, S., MABDOSE. I: Characterization of a general purpose dose estimation code. *Med Phys* **26:** 1389-1395 (1999).
57. Sgouros G., Chiu, S., Pentlow, K. S. *et al.*, Three-dimensional dosimetry for radioimmunotherapy treatment planning. *J Nucl Med* **34:** 1595-1601 (1993).
58. Kolbert K. S., Sgouros, G., Scott, A. M. *et al.*, Implementation and evaluation of patient-specific three-dimensional internal dosimetry. *J Nucl Med* **38:** 301-308 (1997).
59. Tagesson M., Ljungberg, M. and Strand, S. E., "The SIMDOS Monte Carlo code for conversion of activity distributions to absorbed dose and dose-rate distributions." Sixth International Radiopharmaceutical Dosimetry Symposium, Oak Ridge, TN: Oak Ridge Associated Universities, *pp* 425-440 (1999).
60. Liu A., Williams, L. E., Lopatin, G. *et al.*, A radionuclide therapy treatment planning and dose estimation system. *J Nucl Med* **40:** 1151-1153 (1999).
61. Clairand I., Ricard, M., Gouriou, J. *et al.*, DOSE3D: EGS4 Monte Carlo code-based software for internal radionuclide dosimetry. *J Nucl Med* **40:** 1517-1523 (1999).

62. Salvat F., Fernandez-Varea, J., Costa, E. *et al*., *PENELOPE - A Code System for Monte Carlo Simulation of Electron and Photon Transport*, Workshop Proceedings, Issy-les-Moulineaux, France, 5-7 November 2001, (2001).
63. Bielajew A. F. and Rogers, D. W. O., PRESTA: The Parameter reduced electron-step transport algorithm for electron Monte Carlo transport. *Nucl Instrum Methods Phys Res A* **B18:** 165-181 (1987).
64. Briesmeister J. F., Los Alamos National Laboratory, NM; Report No. LA-13709-M, 2000.
65. Yoriyaz H., Stabin, M. G. and dos Santos, A., Monte Carlo MCNP-4B-based absorbed dose distribution estimates for patient-specific dosimetry. *J Nucl Med* **42:** 662-669 (2001).
66. Stabin M. G. and Yoriyaz, H., Photon specific absorbed fractions calculated in the trunk of an adult male voxel-based phantom. *Health Phys* **82:** 21-44 (2002).
67. Zubal I. G., Harrell, C. R., Smith, E. O. *et al*., Computerized 3-dimensional segmented human anatomy. *Med Phys* **21:** 299-302 (1994).
68. Chao T. C., Bozkurt, A. and Xu, X. G., Organ dose conversion coefficients for 0.1-10 MeV electrons calculated for the VIP-Man tomographic model. *Health Phys* **81:** 203-214 (2001).
69. Bozkurt A., Xu, X. G. and Chao, T. C., Fluence-to-dose conversion coefficients from monoenergetic neutrons below 20 MeV based on the VIP-man anatomical model. *Phys Med Biol* **45:** 3059-3079 (2000).
70. Xu X. G., Chao, T. C. and Bozkurt, A., VIP-Man: an image-based whole-body adult male model constructed from color photographs of the Visible Human Project for multi-particle Monte Carlo calculations. *Health Phys* **78:** 476-486 (2000).
71. Jones D. G., A realistic anthropomorphic phantom for calculating specific absorbed fractions of energy deposited from internal gamma emitters. *Radiat Prot Dosimetry* **79:** 411-414 (1998).
72. Petoussi-Henss N., Zanki, M., Fill, U. *et al*., The GSF family of voxel phantoms. *Phys Med Biol* **47:** 89-106 (2002).
73. Dawant B. M., Zijdenbos, A. P. and Margolin, R. A., Correction of intensity variations in MR images for computer-aided tissue classification. *IEEE Trans Med Imaging* **12:** 770-781 (1993).
74. Maurer C. R., Fitzpatrick, J. M., Wang, M. Y. *et al*., Registration of head volume images using implantable fiducial markers. *IEEE Trans Med Imaging* **16:** 447-462 (1997).
75. Snyder W., ''Estimates of absorbed fraction of energy from photon sources in body organs.'' Medical Radionuclides: Radiation Dose and Effects. CONF-691212 AEC Symposium Series 20, *pp* 33-50 (1970).
76. Emami B., Lyman, J., Brown, A. *et al*., Tolerance of normal tissue to therapeutic irradiation. *Int J Radiat Oncol Biol Phys* **21:** 109-122 (1991).
77. Meredith R. F. and Knox, S., Radioimmunotherapy of B-cell NHL. *Curr Pharm Biotechnol* **2:** 327-339 (2001).
78. Siegel J. A., Wessels, B. W. and Watson, E. E., Bone marrow dosimetry and toxicity for radioimmunotherapy. *Antibody Immunoconj Radiopharm* **3:** 213-233 (1990).
79. Siegel J. A., Lee, R. E., Pawlyk, D. A. *et al*., Sacral scintigraphy for bone marrow dosimetry in radioimmunotherapy. *Int J Rad Appl Instrum* **16:** 553-559 (1989).

80. DeNardo D., DeNardo, G., O'Donnell, R. *et al*., Imaging for improved prediction of myelotoxicity after radioimmunotherapy. *Cancer* **15:** 2558-2566 (1997).
81. Lim S., DeNardo, G., DeNardo, D. *et al*., Prediction of myelotoxicity using radiation doses to marrow from body, blood and marrow sources. *J Nucl Med* **38:** 1374-1378 (1997).
82. Breitz H., Fisher, D. and Wessels, B., Marrow toxicity and radiation absorbed dose estimates from rhenium-186-labeled monoclonal antibody. *J Nucl Med* **39:** 1746-1751 (1998).
83. Eary J., Krohn, K., Press, O. *et al*., Importance of pre-treatment radiation absorbed dose estimation for radioimmunotherapy of non-Hodgkin's lymphoma. *Nucl Med Biol* **24:** 635-638 (1997).
84. Behr T. M., Sharkey, R. M., Juweid, M. E. *et al*., "Hematological Toxicity in the Radioimmunotherapy of Solid Cancers with 131I-Labeled Anti-CEA NP-4 IgG1: Dependence on Red Marrow Dosimetry and Pretreatment." Proceedings of Sixth International Radiopharmaceutical Dosimetry Symposium, Oak Ridge, TN: Oak Ridge Associated Universities, Vol. I; *pp* 113-125 (1999).
85. Juweid M., Zhang, C., Blumenthal, R. *et al*., Prediction of hematologic toxicity after radioimmunotherapy with (131)I-labeled anticarcinoembryonic antigen monoclonal antibodies. *J Nucl Med* **40:** 1609-1616 (1999).
86. Erdi A. K., Erdi, Y. E., Yorke, E. D. *et al*., Treatment planning for radioimmunotherapy. *Phys Med Biol* **41:** 2009-2026 (1996).
87. DeNardo G., DeNardo, S., Macey, D. *et al*., Overview of radiation myelotoxicity secondary to radioimmunotherapy using 131I-Lym-1 as a model. *Cancer* **73:** 1038-1048 (1994).
88. Gray B., Burton, M., Kelleher, D. *et al*., Tolerance of the liver to the effects of Yttrium-90 radiation. *Int J Radiat Oncol Biol Phys* **18:** 619-623 (1990).
89. Behr T. M., Sharkey, R. M., Sgouros, G. *et al*., Overcoming the nephrotoxicity of radiometal-labeled immunoconjugates: improved cancer therapy administered to a nude mouse model in relation to the internal radiation dosimetry. *Cancer* **80:** 2591-2610 (1997).
90. Behr T., Goldenberg, D. and Becker, W., Reducing the renal uptake of radiolabeled antibody fragments and peptides for diagnosis and therapy: present status, future prospects and limitations. *Eur J Nucl Med* **25:** 201-212 (1998).
91. Behr T., Béhé, M., Kluge, G. *et al*., Nephrotoxicity versus anti-tumour efficacy in radiopeptide therapy: facts and myths about the Scylla and Charybdis. *Eur J Nucl Med Mol Imaging* **29:** 277-279 (2002).
92. Behr T. M., Béhé, M., Angerstein, C. *et al*., Radiopeptide therapy with cholecystokinin (CCK)-B/gastrin receptor ligands:toxicity and therapeutic efficacy of Auger e-versus or emitters. [abstract] *J Nucl Med* **42:** 68P (2001).
93. Willins J. D. and Sgouros, G., Modeling analysis of platinum-195m for targeting individual blood-borne cells in adjuvant radioimmunotherapy. *J Nucl Med* **36:** 315-319 (1995).
94. Howell R. W., Kassis, A. I., Adelstein, S. J. *et al*., Radiotoxicity of platinum-195m-labeled trans-platinum (II) in mammalian cells. *Radiat Res* **140:** 55-62 (1994).
95. Demidecki A. J., Williams, L. E., Wong, J. Y. *et al*., Considerations on the calibration of small thermoluminescent dosimeters used for measurement of beta particle absorbed doses in liquid environments. *Med Phys* **20:** 1079-1087 (1993).

96. Yorke E. D., Williams, L. E., Demidecki, A. J. *et al*., Multicellular dosimetry for beta-emitting radionuclides: autoradiography, thermoluminescent dosimetry and three-dimensional dose calculations. *Med Phys* **20:** 543-550 (1993).
97. Gladstone D. J., Lu, X. Q., Humm, J. L. *et al*., A miniature MOSFET radiation dosimeter probe. *Med Phys* **21:** 1721-1728 (1994).
98. Gladstone D. J. and Chin, L. M., Real-time, in vivo measurement of radiation dose during radioimmunotherapy in mice using a miniature MOSFET dosimeter probe. *Radiat Res* **141:** 330-335 (1995).

18
Summary

H. Zaidi*

Nuclear medicine has a long tradition of incorporating quantitative analysis in its diagnostic and therapeutic procedures. Many of the clinical and research applications of emission tomography rely on a solid quantitative foundation, which is highly dependent on the performance characteristics of nuclear medicine instrumentation and the accuracy of image correction and reconstruction algorithms. In many situations, absolute quantification is desirable to obtain a truthful representation of the biological process or metabolic function being imaged. In quantitative functional imaging studies, there is always a tension between the desire for more data (e.g. MRI, CT, blood samples, ... etc), which ultimately improves the understanding of the process being studied and the accuracy of the reported results, vs. logistic considerations, including cost, time, imaging unit throughput and patients or healthy volunteers' comfort. One must consider what is logistically possible for the staff in a busy clinical nuclear medicine facility and what is tolerable by the patients and/or subjects. Likewise, accurate quantification requires extensive technical and organizational efforts that may be unaffordable for a small clinical department with limited scientific support. In the clinical setting, it has become standard practice to use simplified imaging protocols compared to the often complex methods developed for research using emission tomography.

Quantitative analysis has been used for several decades since the inception of nuclear medicine imaging. Over this time, its use has evolved from a tool available to only a few expert physicists and research biomedical scientists to one that may be integrated in commercial software that is now potentially available for routine use. This evolution has been driven by four fundamental and critical developments:

*PD Dr H. Zaidi, Geneva University Hospital, Division of Nuclear Medicine, CH-1211 Geneva, Switzerland

1. The ever-increasing sophistication of nuclear imaging instrumentation and associated hardware developments for image correction (e.g. transmission scanning for accurate attenuation correction);
2. The rapid evolution and widespread clinical acceptance of emission tomography (especially PET) and dual-modality imaging as effective diagnostic tools, requiring improved and accurate quantitative analysis;
3. The availability of open source libraries for simulation, image registration, reconstruction, processing and objective assessment of image quality methodologies; which spurred the development of more complex and ambitious computational models (e.g. anatomically-guided reconstruction and partial volume correction);
4. The near exponential increase in preclinical research studies using small-animal imaging devices; an area in which quantification is always essential.

These four interrelated developments have fostered the work examined in this book. It is worth emphasizing that quantification has been traditionally performed in the case of PET, which started mainly as a research tool where there was greater emphasis on accurate quantitative measurements, and more recently has been applied for SPECT. There are many historical and methodological reasons for that, which are addressed in this volume. Different strategies for image reconstruction and accurate attenuation, scatter and partial volume effect corrections have been proposed so far with various degrees of success. The development of newer, faster and more robust algorithms remains an open research field which requires further research and development efforts. In summary, quantitative analysis of nuclear medicine images is an area of considerable research interest and many research groups are very active in this field, leading the nuclear medicine community to forecast a promising progress during the next few years.

Biosketches

Habib Zaidi, Ph.D, P.D

Dr. Habib Zaidi is senior physicist and head of the PET Instrumentation & Neuroimaging Laboratory at Geneva University Hospital and is also visiting professor at the department of Physics of UAEU University. He received a Ph.D. in medical physics from Geneva University for a dissertation on Monte Carlo modelling and scatter correction in positron emission tomography. His academic accomplishments in the area of quantitative PET imaging have been well recognized by the medical faculty of Geneva University, which elected him to become Privat-Docent (PD). Dr. Zaidi is actively involved in developing imaging solutions for biomedical research and clinical diagnosis in addition to lecturing graduate and postgraduate courses on medical physics and medical imaging. His research centre on modelling nuclear medical imaging systems using the Monte Carlo method, dosimetry, image correction, reconstruction and quantification techniques in emission tomography as well as statistical image analysis in functional brain imaging, and more recently on novel design of dedicated high-resolution PET scanners in collaboration with CERN. He is an associate editor for Medical Physics, member of the editorial board of Computer Methods and Programs in Biomedicine and the International Journal of Nuclear Medicine, and scientific reviewer for leading journals in medical physics, nuclear medicine and scientific computing. He is also affiliated to several International medical physics and nuclear medicine organisations and member of the professional relations committee of the IOMP. He is involved in the evaluation of research proposals for European and International granting organisations and participates in the organisation of International symposia and conferences. He is the recipient of several awards and distinctions among which the prestigious 2003 Young Investigator Medical Imaging Science award given by the Nuclear Medical and Imaging Sciences Technical Committee of the IEEE to a young investigator in recognition of significant and innovative technical contributions to the field of medical imaging science and the 2004 Mark Tetalman Memorial Award given by the Society of Nuclear Medicine. Dr Zaidi has been an invited speaker of many keynote

lectures at an International level, has authored over 80 publications, including peer-reviewed journal articles, conference proceedings and book chapters and is the editor of this and another textbook on therapeutic applications of Monte Carlo calculations in nuclear medicine.
E-mail: habib.zaidi@hcuge.ch
Web: http://dmnu-pet5.hcuge.ch/

M'hamed Bentourkia, Ph.D
Dr M'hamed Bentourkia is a senior physicist and assistant professor at the department of nuclear medicine and radiobiology, faculty of medicine, University of Sherbrooke, Quebec, Canada. He received a B.Sc. at University of Quebec at Chicoutimi, a M.Sc. in applied nuclear physics at University of Montreal and a Ph.D. in radiobiology at University of Sherbrooke. His main research interests are in the field of PET imaging modality. Actual domains of research are: image reconstruction, image corrections and enhancement, kinetic modeling, Monte Carlo simulations and software development for image analysis. He is a peer-reviewing member of several international journals and a member of several scientific associations and committees.
E-mail: mhamed.bentourkia@USherbrooke.ca
Web site: http://www.mednuc.usherb.ca/

Michael Braun, Ph.D
Dr Michael Braun is a senior lecturer at the Department of Applied Physics, University of Technology, Sydney. Previously, he held academic appointments at the School of Electrical Engineering, University of Sydney, and School of Medicine, Flinders University of South Australia. His research interests encompass analysis and processing of medical images, including image registration, segmentation, tomographic reconstruction and image-based biomechanical modelling.
Email: michael.braun@uts.edu.au
Web: http://www.phys.uts.edu.au/

A Bertrand Brill, M.D, Ph.D
Randy Brill received his MD at the University of Utah in 1956, and his PhD at the University of California, Berkeley (Biophysics) in 1961. His dissertation advisor was Hardin Jones, and the topic was Radiation Leukemogenesis, a study based on the Hiroshima/Nagasaki A-Bomb Survivors. Randy started his career as a Public Health Service Officer assigned to the Atomic Bomb Casualty Comm. following which he returned to the University of California to complete his PhD. He was assigned to Johns Hopkins (Assistant Prof. Radiology), where he set up a radiation epidemiology study to investigate the possible relation between I-131 administered to people and leukemia, as a counterpart to the then known effects in the A-bomb survivors. Randy returned to academic medicine (Vanderbilt University) after 7

years in the PHS. He was an Assoc. Prof, then Prof. in Medicine/Physics/Radiology, co-Director, then Director of the Div. of Nuclear Medicine and Biophysics. After 15 years, he left for Brookhaven National Lab where he served as Nuclear Medicine Program Coordinator for 7 years. Thereafter, he served as Prof., and Research Director of the Nuclear Medicine Dept. at the Univ. of Mass Medical School for 10 years following, which he returned to Vanderbilt as Research Prof. of Radiology and Physics. His current work involves the development of new nuclear medicine imaging receptors, and radiation dosimetry, particularly oriented to radioimmunotherapy and bone marrow effects.
email: Aaron.brill@vanderbilt.edu

Abldel-Ouahab Boudraa, Ph.D
Dr Abdel-Ouahab Boudraa was born in Constantine, Algeria. He graduated with a B.Sc degree in Physics (Electronics Engineering) from Constantine Institute of Physics, University of Constantine, Algeria. He received the M.Sc in Biomedical Engineering from INSA, Lyon, a University degree in Nuclear Magnetic Resonance Imaging, a PhD degree in Image Processing and University degree in Statistics and Modeling in Positron Emission Tomography, all from the University of Claude Bernard, Lyon 1, France. He is currently Associate Professor of Electrical Engineering at Ecole Navale, Brest, France. His current research interests include computer vision, vector quantization, data structures and analysis, data fusion, time frequency analysis, hard and fuzzy pattern recognition and applications of fuzzy set theory to medical imaging. Dr. Boudraa is an Associate Member of IEEE Society.
Email: boudra@ecole-navale.fr
Web: http://www.ecole-navale.fr/

Eric C. Frey, Ph.D
Eric C. Frey has a Ph.D. in Physics from the University of North Carolina at Chapel Hill. He is currently Associate Professor of Radiology in the Division of Medical Imaging Physics of the Russell H. Morgan Department of Radiology and Radiological Sciences at Johns Hopkins University, in Baltimore, Maryland, USA. His research has been in the area of modeling and compensating for physical image degrading factors in SPECT to improve image quality and quantitative accuracy. He is the author of more than 50 journal papers.
Email: eric_frey@jhu.edu

Guido Germano, Ph.D
Dr Guido Germano is the Director of the Artificial Intelligence in Medicine (AIM) Program at Cedars-Sinai Medical Center, and a Professor of Medicine at the UCLA School of Medicine in Los Angeles. After receiving a B.S. in Electronic Engineering from the University of Naples, Italy, and an MBA

from the Italian government, he came to the United States on a Fulbright scholarship and received a M.S. and a PhD degree in Biomedical Physics from the University of California in Los Angeles. His current research interests include the automatic processing and quantification of nuclear cardiology SPECT and PET images in the context of myocardial perfusion and function assessment; the merging of clinical and quantitative parameters into automated, internally consistent reporting modules; the development of segmentation, display and quantification software for cardiac MRI and CT; intra- and inter-modality applications of 3-D image registration and fusion; and, teleradiology. Dr Germano is a Fellow of the American College of Cardiology and the American Heart Association, a member of the editorial boards of numerous national and international journals, a frequent lecturer in the field of nuclear cardiology, and the author of 140 book chapters and peer-reviewed papers.
Email: germano@csmc.edu

Bruce H. Hasegawa, Ph.D, Professor
Dr Bruce Hasegawa has undergraduate degrees in physics and mathematics, with a master's degree in mathematics from California State University, Fresno, a second master's degree in Radiological Sciences from the University of Colorado Health Science Center, and has pursued graduate work in physiology and biophysics at the University of Washington, Seattle. Dr. Hasegawa completed his doctoral degree in Medical Physics at the University of Wisconsin at Madison under the mentorship of Professor Charles Mistretta, where he developed exposure compensation techniques to improve image quality in chest radiography and digital angiography. Dr. Hasegawa currently is a Professor of Radiology at UCSF where he has been a faculty member since 1985. He also is a faculty member of the Joint Bioengineering Graduate Group at UCSF and UC-Berkeley, and holds a joint faculty appointment in the Department of Nuclear Engineering at UC Berkeley. Dr. Hasegawa has written numerous papers in medical x-ray and radionuclide imaging, and is the author of a book on medical imaging entitled ''The Physics of Medical X-Ray Imaging'', and subtitled ''The Photon and Me: How I Saw the Light''. He is the senior faculty member at the UCSF Physics Research Laboratory and is recognized as a pioneer of dual-modality imaging, especially SPECT/CT. His research interests involve the development of medical imaging techniques and instrumentation for radiography and nuclear medicine. His specific on-going research projects involve the development of dual-modality instrumentation for combined anatomical and physiological assessments of human disease and for imaging structure and function in small animals for biological research.
Email: hasegawa@radiology.ucsf.edu
Web: http://www.radiology.ucsf.edu/

Brian Hutton, Ph.D, Professor
Professor Brian Hutton graduated with B.Sc (Hons) in Natural Philosophy in 1973 and MSc in Medical Physics in 1974 from the University of Aberdeen. He worked at Royal Prince Alfred Hospital, Sydney from 1975-1995, leading the Physics Group in Nuclear Medicine from 1981. He played a central role in the establishment of Australia's first PET facilities around 1990 and spent 6 months at the MRC Cyclotron Unit at Hammersmith Hospital at that time. In 1995 he was appointed Head of Diagnostic Physics at Westmead Hospital, Sydney. He completed a PhD at University of Technology, Sydney in 1999 and became a Visiting Professorial Fellow at the Centre of Medical Radiation Physics, University of Wollongong in 2001. In 2004 he was appointed to the Chair of Medical Physics in Nuclear Medicine and Molecular Imaging Science in the Institute of Nuclear Medicine at University College London. His research interests focus on image and data processing in Nuclear Medicine, in particular reconstruction and quantification in emission tomography (SPECT and PET) and more recently image registration in Nuclear Medicine. He has published widely in these areas and has supervised many postgraduate research projects. He also has a reputation for teaching of Nuclear Medicine Physics and in particular has conducted many workshops on behalf of the IAEA. He has coordinated a major international IAEA educational scheme based on distance assisted training for nuclear medicine technologists that has been introduced to countries in Asia, Africa and Latin America. He has recently been responsible for development of accreditation and training of Medical Physicists in Nuclear Medicine in Australia and NZ.
Email: b.hutton@nucmed.ucl.ac.uk
Web: http://www.ucl.ac.uk/nuclear-medicine

Kenneth F. Koral, Ph.D
Ken Koral received his B.S. degree in physics from Case Institute of Technology and his Ph.D. in nuclear physics from its successor, Case Western Reserve University, Cleveland, Ohio. His dissertation advisor was Philip R. Bevington, author of "Data Reduction and Error Analysis for the Physical Sciences." Ken started his research career in medical physics during a year's postdoc in the United Kingdom. William H. Beierwaltes recruited him to the University of Michigan Medical Center where his current title is Senior Research Professor. Long-time collaborators and friends are Les Rogers and Neal Clinthorne. Ken's main research interests have been and are nuclear medicine scatter correction, activity quantification in SPECT, and tumor dosimetry. Assisted by Jeff Fessler and Yuni Dewaraja, he is the principal investigator of a U.S. National Cancer Institute grant entitled "Techniques for calculating tumor dosimetry from imaging". Ken belongs to a family consisting of himself, his wife Mary and their three adopted children, one each from Viet Nam, India and South Korea. Those children have children: Maekong, Dante and Meagen Ann. They all enjoy movies,

reading and hiking. Ken is a member of the American Association of Physicists in Medicine, the Society of Nuclear Medicine and the Institute of Electronic and Electrical Engineers.
Email: kenkoral@umich.edu
Web: http://www.rad.med.umich.edu/

David Mankoff, MD, Ph.D
David Mankoff is Associate Professor of Radiology and Medicine at the University of Washington in Seattle. He received a B.Sc in physics from Yale University, an MD from the University of Pennsylvania, and PhD in Bioengineering also from the University of Pennsylvania. He has served as the Director of Engineering for UGM Medical Systems and completed internal medicine and nuclear medicine residencies at the University of Washington. His research focuses on the application of radiotracer imaging to cancer, focusing on PET and breast cancer, and on the quantitative analysis of PET images. He is a member of the Society of Nuclear Medicine and the IEEE Nuclear Sciences Section. He is a reviewer for a number of journals and is on the editorial board of Nuclear Medicine and Biology.
Email: dam@u.washington.edu

Mark Muzi, M.Sc
Mark Muzi is Senior Staff Scientist and Director of Image Analysis and Modeling at the University of Washington in Seattle. He received a B.Sc in Zoology from the University of Wisconsin in Madison and an M.Sc in Biology from Eastern Washington University conducting research utilizing labeled hexoses to identify neural activation centers in homing salmon. He has many years of research experience on human brain cancer metabolism with labeled hexoses and a variety of other PET radiotracers for both clinical research and animal validation studies. His current research interests include model development for quantification of novel PET radiotracers in cancer detection and progression. He is a member of the Society of Nuclear Medicine.
Email: muzi@u.washington.edu

Johan Nuyts, Ph.D
Johan Nuyts obtained a masters degree in electronic engineering (1982) and in biomedical engineering (1983), and a PhD in applied sciences (1991) at K.U.Leuven. He did research in industrial image processing hardware at ESAT, K.U.Leuven until 1987. Since then, his research concentrates on image analysis and reconstruction in emission tomography. Since 1991 he is with the department of Nuclear Medicine (as a professor of the Faculty of Medicine, K.U.Leuven, Belgium).
Email: Johan.Nuyts@uz.kuleuven.ac.be
Web: http://www.kuleuven.ac.be/nucmed/

Olivier Rousset, Ph.D
Dr Olivier Rousset is currently a post-doctoral fellow in the department of Radiology at Johns Hopkins Medical Institutions, Baltimore, Maryland (USA). Although a Ph.D. graduate from the Claude Bernard University in Lyon (France), Dr Rousset did most of his graduate research work while at the Montreal Neurological Institute (Canada). Dr Rousset is particularly involved in developing algorithms for improving the accuracy and precision of nuclear medicine images. His interests include the development of simulation algorithms for validation of image processing methods used for the quantitative assessment of physiological parameters obtained with emission tomography.
Email: olivier@indigo.jhoc1.jhmi.edu

Piotr Slomka, Ph.D
Dr Piotr Slomka is a research scientist with the Artificial Intelligence in Medicine Program at Cedars-Sinai Medical Center in Los Angeles. He is also an Associate Professor of Medicine at the University of California, Los Angeles (UCLA) David Geffen School of Medicine. Dr. Slomka received his Ph.D in Medical Biophysics from the University of Western Ontario, London, Ontario, Canada and his M.A.Sc in Medical Instrumentation from the Warsaw University of Technology, Warsaw, Poland. He is widely recognized as one of the leading contributors in the world to the area of research in software algorithms for medical image analysis. His current focus is in the development of automated computer algorithms for myocardial perfusion quantification and multimodality image registration. Dr. Slomka has published or co-published numerous papers for peer-reviewed journals. He has presented lectures, both nationally and internationally and has received numerous awards from a variety of professional organizations and is a scientific reviewer for several leading journals.
Email: Piotr.Slomka@cshs.org

Michael G Stabin, Ph.D
Dr. Stabin's areas of specialty are the dosimetry of internal emitters and radiation protection. He worked for 15 years at the Radiation Internal Dose Information Center in Oak Ridge, TN, and currently works at Vanderbilt University in Nashville, TN. The emphasis of his current research is on radiation dosimetry for nuclear medicine patients. He works as an associate editor for several journals and is active in the Health Physics Society and the Society of Nuclear Medicine.
several: michael.g.stabin@vanderbilt.edu
Web: http://www.doseinfo-radar.com/

Benjamin W Tsui, Ph.D, Professor
Professor Benjamin M. W. Tsui graduated with B.Sc. in Physics in 1970 from Chung Chi College, Chinese University of Hong Kong, A.M. in

Physics in 1972 and Ph.D. in Medical Physics in 1977 from the University of Chicago. He worked in the Department of Radiology at the University of Chicago from 1977-1982. He joined the Department of Radiology and the Department of Biomedical Engineering at the University of North Carolina at Chapel Hill in 1982 and established a Medical Imaging Research Laboratory which focused its research in medical imaging including nuclear medicine and magnetic resonance imaging with emphasis in single-photon emission computed tomography (SPECT). Also, he directed and taught in a graduate level training program in medical imaging. In 2002, he and his entire laboratory joined the Russell H. Morgan Department of Radiology and Radiological Science at the Johns Hopkins University and established a new Division of Medical Imaging Physics. He was appointed the Director of the new division which focused in research and education in medical imaging physics especially in the areas of x-ray and nuclear medicine imaging. Dr. Tsui's research interests are in medical imaging theory, quantitative image reconstruction methods, Monte Carlo simulation, computer phantom development, image evaluation and clinical applications in cardiac and oncological SPECT. He is the principal investigator of several NIH research grants and industrial research contracts and has trained over 30 M.S. and Ph.D. students and postdoctoral fellows. He is a Fellow of the Institute of Physics (IOP) and the American Institute for Medical and Biological Engineering (AIMBE), and a senior member of the IEEE. He is active in several professional societies including the Society of Nuclear Medicine (SNM), Institute of Electrical and Electronics Engineers (IEEE), American Association of Nuclear Cardiology (ASNC) and American Association of Physicists in Medicine (AAPM), and Editorial Boards including the Physics in Medicine and Biology (PMB) and Journal of Nuclear Cardiology (JNC).
Email: btsui1@jhmi.edu

Koen Van Laere, M.D, Ph.D, D.Sc
Koen Van Laere is staff member at the Division of Nuclear Medicine of the University Hospital in Leuven, Belgium and associate professor at the Leuven University. He graduated in Experimental Physics (M.Sc) at Gent University (Belgium) in 1988. He finished a Ph.D in radiation physics at the same university in 1992, on the topic of Monte Carlo modelling of radiation transport in electron and photon beams and dosimetry using alanine. He received his medical degree (M.D) in 1997, and another Ph.D. in medical sciences in 2001 on quantification of perfusion and receptor brain imaging. His current research focuses on analysis methodology and preclinical/clinical PET and SPECT applications of brain imaging for neurodegeneration (movement disorders and dementia), brain tumours and epilepsy. He is scientific reviewer for the European Journal of Nuclear Medicine and the Journal of Nuclear Medicine. He is affiliated to several international nuclear medicine organisations and secretary of the Belgian Society for Nuclear Medicine Task Group on Brain Imaging. He is author of several

peer-reviewed publications, book chapters and conference proceedings in brain imaging, radiation dosimetry and general clinical nuclear medicine.
Email: koen.vanlaere@uz.kuleuven.ac.be
Web: http://www.kuleuven.ac.be/nucmed/

Index