QUANTITATIVE NUCLEAR MEDICINE IMAGING USING ADVANCED IMAGE RECONSTRUCTION AND RADIOMICS

by

Saeed Ashrafinia

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

March 2019

Abstract

Our aim is to help put nuclear medicine at the forefront of quantitation on the path to realization of personalized medicine. We propose and evaluate (Part I) advanced image reconstruction and (Part II) robust radiomics (large-scale data-oriented study of radiological images). The goal is to attain significantly improved diagnostic, prognostic and treatment-response assessment capabilities.

Part I presents a new paradigm in point-spread function (PSF)-modeling, a partial volume correction method in PET imaging where resolution-degrading phenomena are modeled within the reconstruction framework. PSF-modeling improves resolution and enhances contrast, but significantly alters noise properties and induces edge-overshoots. Past efforts involve a dichotomy of PSF vs. no-PSF modeling; by contrast, we focus on a wide-spectrum of PSF models, including under- and over-estimation of the true PSF, for the potential of enhanced quantitation in standardized uptake values (SUVs).

We show for the standard range of iterations employed in clinic (not excessive), edge enhancement due to overestimation actually lower SUV bias in small regions, while inter-voxel correlations suppress image roughness and enhance uniformity. An overestimated PSF yields improved contrast and limited edge-overshoot effects at lower iterations, enabling enhanced SUV quantitation. Overall, our framework provides an effective venue for quantitative task-based optimization.

Part II proposes robust and reproducible radiomics methods. Radiomics workflows are complex, generating hundreds of features, which can lead to high variability and overfitting, and ultimately

hampering performance. We developed and released a Standardized Environment for Radiomics Analysis (SERA) solution to enable robust radiomics analyses. We conduct studies on two unique imaging datasets – renal cell carcinoma SPECT and prostate cancer PET – identifying robust and reproducible radiomic features.

In addition, we evaluate a novel hypothesis that radiomic features extracted from clinically normal (non-ischemic) myocardial perfusion SPECT (MPS) can predict coronary artery calcification (CAC; as extracted from CT). This has important implications, since CAC assessment is not commonly-performed nor reimbursed in wide community settings. SERA-derived radiomic features were utilized in a multi-step feature selection framework, followed by the application of machine learning to radiomic features. Our results show the potential to predict CAC from normal MPS, suggesting added usage and value for routine standard MPS.

Thesis Committee:

Prof. Arman Rahmim
Department of Electrical and Computer Engineering,
Department of Radiology and Radiological Sciences (advisor, primary reader)

Prof. Yong Du
Department of Radiology (secondary reader)

Prof. Jerry L. Prince
Department of Electrical and Computer Engineering

Prof. Jin Kang
Department of Electrical and Computer Engineering

Acknowledgments

He who does not express gratitude to creatures, has not expressed gratitude to The Creator.

-Prophet Mohammad

Many individuals have contributed to the successful completion of my work during the Ph.D. program. I would like to thank my advisor, Dr. Arman Rahmim, for his consistent support and faith in me from the very beginning while providing me with extremely valuable guidance throughout these years. His decent mentorship and patience were a great motivation for my learning. He taught me how to be a world-class researcher that is respected by peers from both academia and industry across the world. Our detailed discussions and making sure that I have truly understood the topics investigated enabled me to develop excellent attention to details. Things I have learned from Dr. Rahmim are beyond those just related to my research. He was keen to educate us to be more productive and have goals. He was extremely humble when interacting with other people regardless of their education or wealth or position in society.

I would like to thank Dr. Martin Pomper, the Director of the Division of Nuclear Medicine and Molecular Imaging at Johns Hopkins University, Department of Radiology, for providing funding and support during my Ph.D. degree. His open vision of stepping beyond typical clinical projects in the division has paved the way for novel research involving artificial intelligence and machine learning/deep learning. Also, I would like to especially thank him for his trust, motivation, and providing me the opportunity to be involved in writing a major research grant.

I have learned from many individuals from the Division of Nuclear Medicine and Molecular Imaging, including faculty and residents. I would like to thank Dr. Yong Du, for kindly agreeing to be in my dissertation committee as the reader, and for being a humble friend and providing best advice on many different fronts to me. I would like to thank Dr. Steven Rowe for allowing me to engage in many different state-of-the-art clinical research projects in nuclear medicine. Many of my radiomics projects were co-advised by Dr. Rowe. I would also like to thank Dr. Martin Lodge for many fruitful discussions during all these years. Furthermore, I would like to thank Dr. Dean Wong and the neuropharmacology lab, including Drs. James Brasic and Anil Mahur. I would also like to express gratitude to our great collaborators Drs. Som Javadi and Chahid Civelek, as well as Andrew Crabb and Jeff Leal from our division.

A very special thanks to many radiology and nuclear medicine residents for their help. My special thanks go to Dr. Pejman Dalaie, a dear friend of mine with whom I spent a number of months collecting the cardiac dataset, and who was greatly dedicated to the project. He has taught me a lot about nuclear cardiology. Other great residents who helped us with our project and provided a better perspective on the clinical aspect of our research projects, include Drs. Charles Marcus, Payam Ghazi, Mehdi Taghipour, Esther Mena, Pouya Aghajari, Michael DiGianvittorio, Mohammad Salehi Sadaghiani, Rima Tulbah, Krystyna Jones, and Yafu Yin.

Other special thanks go to Dr. Jerry Prince from my primary department, electrical and computer engineering. Dr. Prince is without a doubt an excellent example of a great mentor and teacher. His medical imaging systems course was my most favorite course and the only course that I later TA'ed. The course and the teaching assistantship were both very challenging, but through his smooth guidance became great learning experiences. Also, as one of my Ph.D. qualifying exam examiners, he managed to turn a two-hour-long oral exam that I thought would be extremely difficult to one of my best exams ever, as Dr. Prince turned it to an extremely great learning experience that I eventually learned something from. I would also like to thank Dr. Jin Kang for kindly agreeing to be in my

dissertation proposal and defense committees. I would like to thank Drs. Jeffery Siewerdson and Jeffery Web Stayman for the great X-ray and CT imaging course that I learned a lot, and for kindly agreeing to be in my graduate board oral exam committee. I want to thank Dr. Michael Gorin from the department of urology for including me in very exciting research projects on prostate cancer and renal cancer.

I had the great opportunity to work besides and learn from some of the world's top researchers in nuclear medicine, some of which used to be a part of Johns Hopkins but moved elsewhere. A special thanks to Dr. Thomas Schindler for his great guidance and extremely fretful discussions, especially regarding the cardiac SPECT project, that won a fellowship and a young investigator award recognition. I would like to thank Dr. Rathan Subramaniam for allowing me to be a part of multiple oncologic projects during his presence at Johns Hopkins. Many thanks to Dr. Abhinav Jha for great discussions. I would like to also thank Drs. Richard Wahl and Yon Zhou. Other external faculty to whom I owe acknowledgments include Drs. Adam Alessio, Dan Kadrmas, Mathieu Hatt, Usama Mawlawi, Alex Zwanenburg, and Martin Vallières.

I wish to dedicate a paragraph to the amazing people in our Quantitative Tomography lab led by Dr. Rahmim, being some of the greatest and most supportive people. I would like to start by thanking Dr. Nicolas Karakatsanis for his excellent mentorship and great support during my earlier days at our lab. I also thank Dr. Hassan Mohy-ud-Din for very productive conversations. A special thanks to my friend Kevin H. Leung for being very contributing and dedicated to some of my projects during the past two years, and during our classes. I wish to thank Isaac Shiri for stepping forward and including me in great collaboration projects. I would like to also thank Drs. Madjid Soltani, Rongkai Yan, Yuanyuan Gao, as well as Sima Fotouhi, Yansong Zhu, and Ryan VanDenBurg.

I had the unique opportunity to collaborate with amazing individuals from industry. I would like to acknowledge and thank the support provided by Siemens Healthineers during the first two years of

my program, and then later to providing me the chance to spend time as an intern at the company. A very special thanks to the director of Research and Clinical Collaborations, Dr. Darrell Burckhardt, for providing me with the opportunity to try different areas of the research and development at Siemens molecular imaging (MI), and for his great mentorship. I would like to also thanks Drs. Michael Casey, Vladimir Pannin, Ramya Rajaram, and many others from Research and Development of Siemens MI. I also thank MIM software for providing us with the free research license used in many of our projects. Moreover, I thank the support provided to me by the society of nuclear medicine and molecular imaging (SNMMI), American association of physicists in medicine (AAPM), and international electrical and electronic engineers (IEEE) for providing me support in the forms of awards and fellowships.

A very sincere and special thanks goes to my best friends, Drs. Shahabeddin Torabian and Safoora Seddigh, and their lovely baby Aula for their limitless support, love, and friendship. Outside of work, we were just like a family, or even closer. They were there for me during all these years in Baltimore in times of happiness, sickness, and pressure. During my overloaded first semester that I was arriving home after 2AM, hot dinner was ready for me. I cannot thank them enough. Also, many thanks to our amazing community in Baltimore with awesome people that made some of my best memories. Special thanks to Drs. Seyyed Alireza Mortazavi, Mehdi Taghipoor, Elaheh Bordbar, Hossein Hezaveh, Pegah Ghahremani, Heydar Davoodi, Mohammed G. Saleh, Mehdi Hamzei, Narges Hamzei, Reza Kaimasi, Sajjad Salehi, Fereshteh Alamifard, Reza Seyfabadi, Ehsan Variani, and Ehsan Hamidi, as well as Zeinab Vessal, Zeinab Ameri, Ebrahim RajabTabar, Haniyeh Ataei, Farshid AlamBeigi, Hamid Foroughi, Somayyeh Gharaee, Mr. Azami and family, and Mr. Hosseini and family.

Last but definitely not least, my outmost appreciation goes to my beloved family, for their decades of uncompromising support and love. They were the main motivation behind my success from early days at school until today, making sure I am at the best place and school at every step of my education. May God reward them with health and prosperity.

Dedication

To my beloved family for their consistent and uncompromising love and support.

Table of Contents

List of Tables

List of Figures

1. Introduction

Nuclear medicine imaging technology, coupled with image generation techniques, has dramatically improved compared to a few decades ago. At the same time, the apparent complexity and/or limited accuracy of quantitation methods remain bottlenecks that motivate developmental research. Our aim in this dissertation is to help put nuclear medicine at the forefront of quantitation on our path to the realization of personalized medicine. Our work proposes, implements and evaluates advanced image reconstruction as well as image processing techniques in nuclear medicine imaging, aiming to attain significantly improved diagnostic, prognostic, and treatment response assessment capabilities.

1.1. Medical Imaging in Today's Medicine

Clinical practice has commonly consisted of diagnostic, prognostic, and treatment-related tasks [1]. Medical imaging has been a staple of clinical diagnosis workflow across many different specialties for decades and has substituted the need for some invasive procedures, particularly explorative surgeries.

At the same time, applications of medical imaging have also gone beyond diagnosis to other clinical territories. Many studies use interpretations from medical images to improve prognosis in oncology [2-6], cardiology [7-12], and neurology [13-17]. Furthermore, medical imaging and in particular functional imaging modalities that capture the functionality of tissue cells, such as positron emission tomography (PET), are widely utilized in treatment planning [18-20], treatment monitoring [21-24], and treatment response assessment [25-30]. Medical imaging has also stepped up to enable newly-

developed frontiers in medicine. An example is computer-aided diagnosis (CAD), which has been increasingly investigated and applied to assess diverse types of abnormalities as imaged via different modalities [31-39]. CAD-based assessment of breast cancer from mammograms is now a part of routine clinical work at many hospitals [31, 40-43], and there is significant research on different fronts.

Another emerging area that has benefited significantly from medical images is theranostics. Theranostics, a combination of the terms therapeutics and diagnostics, describes the integration of targeted therapy and targeted diagnostic tests; e.g. use of a radioactive drug to image (diagnose) and another radioactive drug to deliver therapy (treat) cancerous tissue [44, 45]. The idea is "to treat what we see and to see what we treat". Specific multifunctional materials can be employed for theranostics including polymers, magnetic, and inorganic nanoparticles that deliver therapeutic use, in addition to the ability to be fully functionalized with imaging agents. The latter enables one or more diagnostic imaging techniques such as MRI, nuclear medicine imaging (PET, single photon emission computed tomography (SPECT)), and even optical and fluorescence imaging [44, 46-52].

Overall, advancements in imaging techniques have made medical imaging an essential component in most episodes of care. An important branch of medical imaging is nuclear medicine imaging, which presently enjoys numerous clinical applications and wide clinical usage. Our dissertation focuses on the development and validation of image reconstruction and processing techniques towards improved quantitation in nuclear medicine imaging. In what follows, we provide a brief introduction to nuclear medicine imaging. Subsequently, we discuss medical image quantitation and the concept of biomarkers in order to clarify the aims of our efforts. Specifically, we elaborate on the link between image quantitation and personalized medicine. We then introduce the field of radiomics, an "image processing" technique to characterize phenotypes of regions of interests in medical images. Following these introductions, we re-capture and summarize our motivations for this research and a list of publications in the course of our work.

1.2. Nuclear Medicine and Tomographic Imaging

The field of nuclear medicine involves the administration of a radioactive-labeled pharmaceutical, or radiopharmaceutical, with the aim of providing diagnostic and/or therapeutic information in a wide range of disease states. Nuclear medicine imaging involves injecting a patient with a compound labeled with a gamma-emitting or positron-emitting radiopharmaceutical. These radiopharmaceuticals are sometimes referred to as radiotracers, or simply tracers. The tracer results in the emission of high-energy photons, some of which can exit the body. A set of detectors can then detect these exiting high-energy photons. The resulting data can be subsequently reconstructed to generated images of the distribution of the radiotracer inside the subject [53].

Unlike X-ray and CT imaging where X-rays are emitted from an external source, transmitting through the body. In nuclear medicine imaging, photons are emitted from inside the body. Hence, X-ray and CT images are produced through *transmission*, while nuclear medicine images are obtained through *emission* of photons. Moreover, in contrast to X-ray, CT and MRI modalities that often produce images of anatomical structures (hence performing *structural* imaging), in nuclear medicine, the explicit aim is always to capture the biological behavior of a substance in the body [54]. The biodistribution of a given radiotracer is determined by physiological and biochemical functioning of the body that may not be apparent via structural imaging; thus, nuclear medicine is considered a *functional* imaging modality.

Nuclear medicine imaging mainly consists of two broad classes: single photon emission, and positron annihilation photon emission. The former is the basis of *2D planar imaging* (also known as scintigraphy or gamma scan), as well as *single photon emission computed tomography* or SPECT imaging, and the latter forms *positron emission tomography* or PET imaging. In what follows, we first discuss single photon imaging including basics of Anger cameras and SPECT imaging, followed by a

discussion of positron emission and PET imaging. Subsequently, we discuss tomographic image reconstruction, and specifically, statistical image reconstruction. We then elaborate on the causes of degradation in quality and quantitative accuracy for nuclear medicine images.

1.2.1. Single photon emission

1.2.1.1. Anger camera

Single photon imaging involves the use of radiotracers that emit gamma-rays upon decay. It enables detection of a two-dimensional (2D) projection of the three-dimensional (3D) biodistribution of the radiotracer in the body. The resulting planar image, though 2D (with no explicit depth resolution) can still be diagnostically useful. Once the radiotracer is administered to the patient through injection, inhaling, etc., it gets circulated around the body and gets accumulated at certain tissue cells, e.g. bones, based on the type of the pharmaceutical. Subsequently, a considerable number of gamma-rays will be emitted, especially originating at these target sites. On the outside, a gamma camera can be used to detect these incoming gamma-rays, convert the absorbed energy from the incoming gamma-rays to the detectors to electrical signals, and then form or reconstruct the image using reconstruction algorithms. This process happens in the so-called Anger *scintillation camera* or the *gamma camera*, which was invented in the late 1950s by Hal Anger of the Donner Lab at the University of California at Berkeley [54].

The Anger camera is the most commonly-used imaging instrument in nuclear medicine today [54]. A simple diagram of its components is shown in Figure 1-1. The first component the gamma-rays encounter is the multi-hole lead collimator with a narrow gap that allows only the incoming photons at nearly normal angels to pass through. The gamma-rays originating from the body are emitted randomly in all direction. As such, the job of a collimator is to ensure that the detector only captures photons from a see-through angle of the collimator and prevent off-angled incoming photons to be

detected. In the case of commonly-used parallel-hole collimators (like the one shown in Figure 1-1), they only allow incoming gamma rays from nearly normal angles.

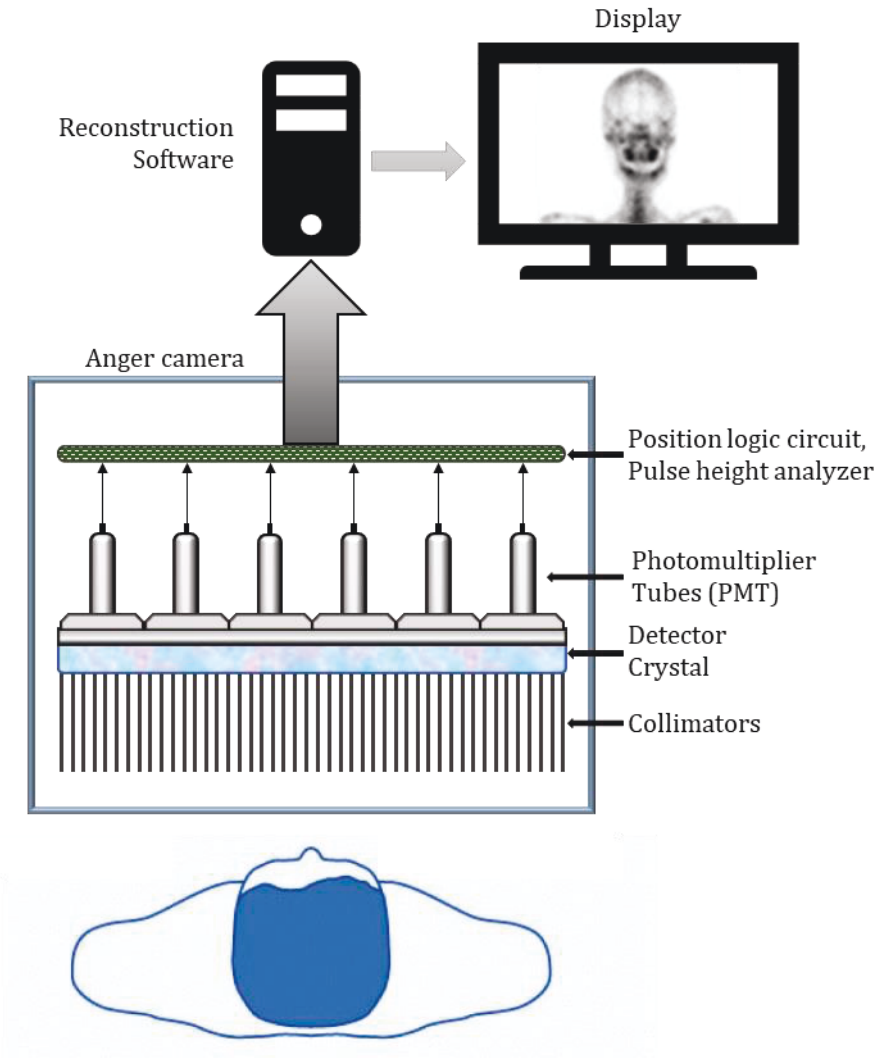

Figure 1-1. Components of a typical Anger camera

The next component is the scintillation crystal that can be a 10-to-25 inch circular, square, or rectangular block of thallium-doped sodium iodide (NaI [Tl]) or more recently, cadmium zinc telluride (CZT) where the latter provides significantly improved image sharpness and contrast [55]. This type of detector crystals emits light photons ('scintillation') after deposition of energy in the crystal by ionizing radiation. A gamma photon that is absorbed by the scintillation crystal produces

a burst of light in the crystal comprising thousands of light (scintillation) photons [54]. This light is then channeled out of the back of the crystal where it arrives at an array of photomultiplier tubes (PMTs). PMTs receive photons at their front photocathode. For every photon received at the photocathode, an electron is released on the inside of the PMT (via the photoelectric effect). The electron subsequently goes through a series of electrodes called dynodes and is amplified. Each dynode release 3-6 electrons upon receiving one electron, and thus, through having 10-14 successive dynodes, one electron that arrives at the photocathode results in 10^6 to 10^8 electrons at the other end of the PMT (the anode). The output of the anode arrives at a positioning logic circuit, which determines the location of the events occurring on the face of the crystal and the combined output of all the PMTs [53, 56].

Another component of the camera is the pulse height analyzer. The collimator only allows photons traveling in a predetermined direction to pass through and arrive seamlessly at the crystal; thus, a line originated at the scintillation event at the crystal through the collimator hole is assumed to arrive at the site of the origin of the photon, which is where the gamma-ray has left the radiotracer in the patient's body. However, it is possible for the gamma-ray to undergo Compton scattering during this path. In Compton scattering, an incident photon ejects a valence electron from the outer shell of an atom. The incident photon then loses some of its own energy and changes its direction. A photon is said to be attenuated when it goes through one or more scatterings and gets distracted from its original path or completely loses its energy. Now, if a photon undergoing this phenomenon passes through the collimators and reaches to the detectors, then the line drawn from the scintillation passing through the collimator hole may not arrive at the location where the gamma-ray was actually emitted from. Ideally, we do not want to include such scattered photon in the formation of the image. One solution to exclude such photons is by using the pulse height analyzer. Photons that have gone through Compton scattering have lost some of their energy relative to a typical photon that arrives at the detector without being scattered. The pulse height analyzer only accepts photons with their

energy in a certain range called acceptance interval or acceptance window and rejects many scattered photons that have been detected. The output of the Anger camera is interfaced to a computer where an image is being generated.

1.2.1.2. Single Photon Emission Computed Tomography (SPECT)

While the Anger camera is readily used for planar nuclear medicine scanning, it can be used in a slightly modified setting to provide a 3D image that contains the biodistribution of the radiotracer with depth information inside the body. The process of generating a 3D image from multiple 2D projections is referred to as "tomography". Tomography consists of two Greek words: "tomos" meaning 'slice' or 'section', and "graphō" meaning 'to write'. Since this is performed digitally, it is referred to as computed tomography. Therefore, single photon emission computed tomography or SPECT is the process of imaging a sample using the single photon emission technique by acquiring projection data multiple times at different orientations, ultimately yielding a 3D image of the sample structure using reconstruction algorithms. As such, tomographic reconstruction enables looking at slices through the investigated object without physically cutting it. Modern SPECT cameras consist of two or three camera heads, like the one in Figure 1-1, mounted on a single rotating gantry, that can move around the patient and cover various angles to detect a larger fraction of emitted photon. This yields an increase in signal-to-noise ratio and produces a 3D tomographic image [53, 54, 56].

1.2.1.3. SPECT/CT and its applications in nuclear medicine

Nowadays, SPECT systems increasingly include a CT component. Such a system, the SPECT/CT, allows the use of CT imaging to also estimate and compensate for attenuation of emitted photons [53, 54]. The CT image is also used to create fused SPECT/CT images in the three orthogonal planes for anatomical localization. The addition of CT component, in general, leads to improved sensitivity and specificity of single photon emission imaging, due to attenuation correction, as well as anatomical

localization, and morphological characterization rendered by the CT, while also facilitating contrast-enhanced CT scanning as part of the examination [57].

The significant role of SPECT/CT in nuclear medicine has been reviewed extensively [58-61]. An example clinical applications of SPECT/CT include characterization and localization of solitary pulmonary nodules and lung cancers, brain tumors, lymphoma, prostate cancer, neuroendocrine and endocrine tumors, as well as malignant and benign bone lesions [58]. SPECT/CT is also applied to thyroid and parathyroid imaging [62], breast cancer [63, 64], radio-guided biopsy [65] and cerebral masses [66]. Other non-oncological applications include infection and inflammation, and to precisely localize infectious foci that can be problematic [67, 68], gastrointestinal (GI) diseases such as upper/lower GI bleeding [69, 70], pulmonary embolism [71] and neurodegenerative disorders [72]. Another major application of SPECT/CT is in patients diagnosed or suspected of coronary artery disease. At the same time, myocardial perfusion SPECT (MPS) imaging maintains a dominant and unique place for evaluation of the physiological significance of coronary artery problems and atherosclerosis [73]. While SPECT has many oncological and neurological applications, most SPECT scans are performed in cardiology [74]. One of our main projects targets MPS imaging and will be discussed in chapter 4 of this dissertation.

1.2.2. Positron Emission

Positron-emitting radionuclides can also be used to label pharmaceutical compounds. Once the radiotracer is administrated intravenously to the patient and distributed in tissues in a manner determined by its biochemical properties, they decay by ejecting a positron from the nucleus. Positron-emitters include ^{11}C, ^{18}F, ^{13}N, ^{15}O, etc. that occur naturally in biological molecules, allowing them to be more readily incorporated into a wide variety of useful radiotracers, which is an advantage over commonly-used SPECT radionuclides such as ^{99m}TC.

A neutron-deficient nucleus can become stable by either capturing a new electron that is subsequently captured by a proton thus transforming into a neutron or by emitting positron that happens from a proton p decay. In this case, the following transmission takes place:

$$p \rightarrow n + \beta^+ + \nu_e \qquad 1.1$$

where n is the neutron, β^+ is the positron, and ν_e is another subatomic particle called neutrino. For a typical radionuclide ${}^{M}_{Z}X$ with mass number M and atomic number Z, the following process occurs:

$${}^{M}_{Z}X \rightarrow {}_{Z-1}{}^{M}X + \beta^+ + \nu_e \qquad 1.2$$

Positrons are emitted with a continuous range of energies up to a maximum, due to the presence of neutrino. The freed positrons give up their enormous kinetic energy mostly through Compton scattering with electrons within the surrounding medium, until they reach thermal energies. At this stage, they start to interact by electrons, mostly leading to positron annihilation[1]. Higher energy positrons require to traverse a larger distance (on average) in the medium before reaching thermal energies and interacting with electrons. This distance is referred to as the *positron range*. Properties of commonly-utilized positron-emitting isotopes are more elaborated in [75].

Table 1-1. Properties of commonly-utilized positron-emitting isotopes [76, 77]

	$T_{1/2}$ (min)	E_{avg}(MeV)	E_{max} (MeV)	Mean range in water (mm)
C-11	20.3	0.39	0.96	1.1
N-13	9.97	0.49	1.19	1.5
O-15	2.03	0.73	1.79	2.5
F-18	109.8	0.24	0.64	0.6
Cu-64	762	0.28	0.66	6.07
Ga-68	4,086	0.83	1.9	2.9
Rb-82	1.25	1.52	3.38	5.9
I-124	6,019	0.69	2.15	3.46

[1] A second possibility for positron-electron interaction is the formation of a hydrogen-like orbiting couple called positronium. In this case they may either decay by self-annihilation and generating two 511 keV photons (for an ortho-positronium) or self-annihilate by emission of three photons (for a para-positronium) which is very rare.

When a positron and an electron annihilate, based on the mass-energy equivalence principle, they emit two gamma-ray photons. Given $E = mc^2$, inserting the masses of an electron and a positron, yields an energy of 511 keV for each of the two emitted photons. These two photons travel almost anti-parallel at a 180° angle. The two 511 keV photons can be detected via detectors connected to a coincidence-detection circuit that only accepts two incoming photons at two detectors within a short timing-window at the order of nanoseconds, just enough to capture photons emitted from the same annihilation. This scenario is depicted in Figure 1-2.

1.2.2.1. Positron emission tomography (PET)

Imaging *in vivo* using positron emission makes use of radiopharmaceuticals that decay by emitting positrons. The resulting anti-parallel 511 keV photons are subsequently detected by a 360° ring of detectors. The data is collected in many angles and is processed with a tomographic reconstruction algorithm to produce an image of radiotracer distribution in the body. Such a procedure is referred to as positron emission tomography or PET imaging. Although these gamma-ray photons can also be detected using SPECT system that operates in a single-photon-counting mode, they are not optimally designed for the relatively high energy of 511 keV photons and have low detection efficiency at this energy range. More importantly, PET replaces physical collimation that comes with a cost of reduced sensitivity with coincidence detection circuit that acts as electronic collimation, increasing sensitivity by at least one order of magnitude [54, 56]. PET imaging has gained widespread clinical application and is commonly utilized along with other imaging modalities such as CT, MRI, and SPECT.

1.2.2.2. Image acquisition in PET

In single photon imaging, as shown in Figure 1-1, collimators restrict the acceptance angle of incoming photons, and without them, a detected photon could have come from *any* region in the field of view (FoV). In positron emission imaging, however, much higher sensitivity is achieved by noting that collimation can be performed *electronically* as depicted in Figure 1-2. The coincidence detection

circuit can determine if two events arriving at two detectors within a certain coincidence time window are dually-emitted gamma-rays from a single positron emission. As such, a line-of-response (LOR) can be drawn between the two detectors (dashed line in Figure 1-2).

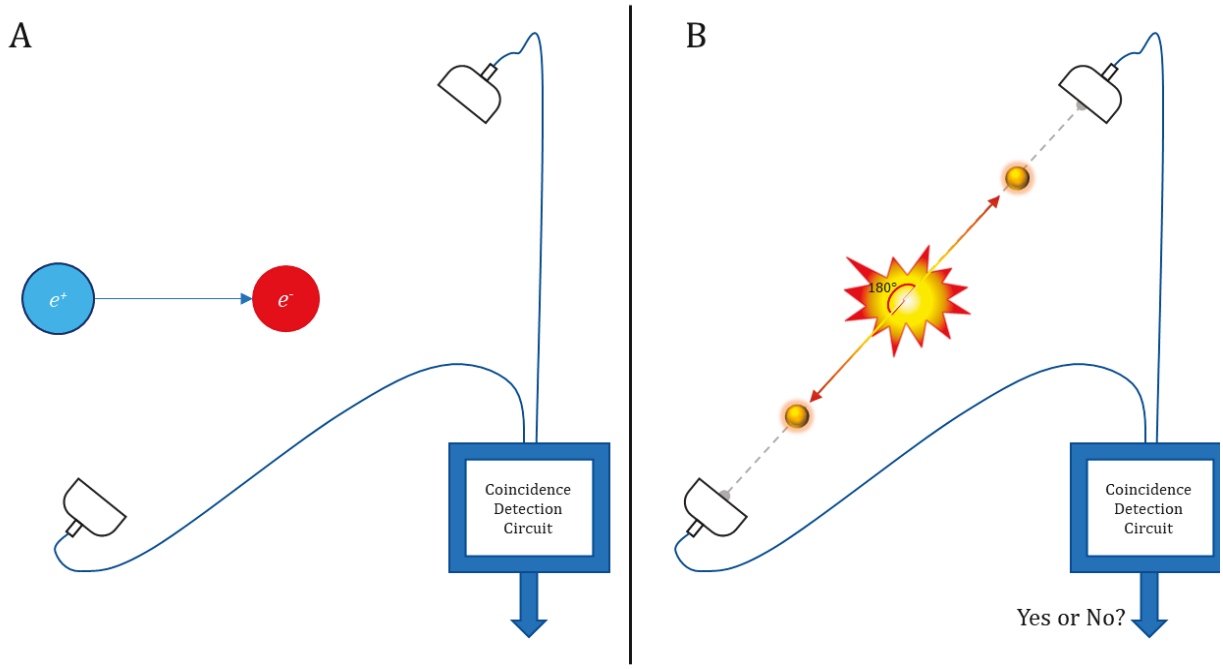

Figure 1-2. Positron annihilation. A) positron reaches thermal energy and meets an electron. B) they annihilate, producing two 511 keV gamma-ray photons travel 180° apart. A coincidence detection circuit can then determine whether they originate from a single annihilation event.

PET scintillation crystals

A PET scanner is commonly comprised of a 360° arrangement of crystal detectors. The incoming radiation excites electrons in the scintillation crystals through Compton scattering (see section 1.2.1.1) or photoelectric absorption, and cause the scintillation crystal to "glow". The other end of the crystal is attached to the photocathode of a PMT (see section 1.2.1.1) in order to amplify the detected signal, as measured at the anode of the PMTs. Various types of PET-specific scintillation crystals exist. Properties of these materials are characterized based on several criteria such as:

1- Linear attenuation coefficient

2- Decay time—the time for the excited electrons in the crystal to return to their ground state

3- Spectral distribution—wavelength of the output light from the crystal that should be efficiently detected by PMTs
4- Linearity—proportionality of the amount of light produced to the energy deposited by radiation,
5- Conversion efficiency—a fraction of the radiation energy converted to detectable scintillator light.

Table 1-2 provides a list of such properties for some of the commonly-used scintillation crystals. In the past, bismuth germanate (BGO) was commonly used in PET scanners, whereas, nowadays, LSO (Lutetium Oxyorthosilicate), or its yttrium-doped version, LYSO, are used more often due to the improved decay time, high conversion efficiency, and energy resolution.

Table 1-2. Table of commonly-used scintillation materials. Bismuth germanate (BGO), lutetium oxyorthosilicate (LSO), sodium iodide (NaI), and gadolinium oxyorthosilicate (GSO).

	NaI[Tl]	**BGO**	**LSO**	**GSO**
Atomic No. (Z)	50	73	66	59
Lin. atten. coef. (cm^{-1})	0.34	0.92	0.87	0.62
Index of refraction	1.85	2.15	1.82	1.85
Light yield [%NaI:Tl]	100	15	75	41
Peak wavelength (nm)	410	480	420	430
Decay const. (nanosec.)	230	300	40	56
Fragile	Yes	No	No	No
Hygroscopic	Yes	No	No	No

PMTs and photo-detectors

Originally, PET scanners involved the one-to-one coupling of detector crystals and PMTs. In 1985, a new technology was developed to couple one PMT to a block of crystals [78], which continues to this day. In this scheme, a block of crystals (e.g. 8x8) is coupled to four PMTs, so that one crystal can disperse light in all four PMTs based on their distance and the crystal pattern. The crystal position and thus the location of the incident photon can be determined by linear averaging of the light output in all four PMTs.

In recent years, there has been an increased interest to replace PMTs with light-sensitive semiconductor detectors such as silicon photomultipliers (SiPMs). These detectors consist of an array of microscopic, parallel connected avalanche photodiodes that operate in limited Geiger-mode. In this mode of operation, each cell is biased enough just above the breakdown threshold, resulting in a cumulative avalanche breakdown within the diode depletion-region that results in an excessive current upon encountering a light photon [79, 80]. They exhibit a higher quantum efficiency (fraction of incident photons absorbed in the photosensitive area, QE) compared to PMTs, yet essentially do not amplify the signal internally like PMTs. The Silicon photodiodes have an internal gain that enables the detection of low light levels with a good signal-to-noise ratio in a very small size of only a few millimeters thick [81, 82]. Table 1-3 provides a comparison between PMTs and SiPMs. Some of the key advantages of SiPM over PMT is as follows:

- Solid-state (SiPM) vs. vacuum-tube technology
- Higher quantum efficiency for SiPM (25% to 40% vs. up to 80%)
- Lower operation voltage (20-40 V vs. 1-3 kV)
- Insensitivity to the magnetic field (SiPM)
- Miniaturization (SiPM)
- No damage in bright light (SiPM)

Table 1-3. Comparison of two types of photo-detectors typically used in PET: PMTs and SiPMs [83]

	PMT	**SiPM**
Gain	10^6	$\sim 10^6$
Rise time (ns)	~1	~1
QE at 420 nm (%)	~25	~25 – 75
Bias (V)	>1,000	30 – 80
Temperature sensitivity (%/° C)	< 1	1 – 8
Magnetic field sensitivity	Yes	No
Price/channel ($)	> 200	~ 50

Time-of-flight (ToF) PET

Theoretically, it is possible to precisely determine the location along the LOR between two detectors at which the annihilation photons originated. This can be done using the exact difference in the time at which the dual photons arrived at the pair of detectors defining the LOR. As such, for a system with superb timing resolution, there would be no need for reconstruction algorithms, as the point of the original can be identified. However, such technology is presently far-fetched. With advances in scintillation and detector technology, detector response time has improved as short as a few hundred nanoseconds; however, this is enough to only narrow down the probability of annihilation to an interval of a few centimeters along the LOR. This technology that utilizes the differences between the two detection times to somewhat localize the annihilation is called time-of-flight or ToF. Let us denote the difference in arrival times of the two photons by Δt. In that case, the location of the annihilation even with respect to the mid-point between the two detectors, Δd, is given by:

$$\Delta d = \frac{\Delta t \times c}{2} \tag{1.3}$$

where c is the speed of light at $\sim 3 \times 10^{10}$ cm/sec. Based on this equation, to achieve a 1 cm resolution requires a timing resolution of about 66 picoseconds. Nowadays, superfast electronics can handle such speeds, but the bottleneck is the rise time of light outputs from scintillators currently available for PET imaging that are too slow to provide this level of timing resolution. The fastest clinical ToF technology up-to-date is 214 picoseconds with the Siemens Biograph Vision [84], which narrows down the localization to ~3.2 cm along the LOR. Images acquired at such timing resolution, though not improving on the resolution, still have a significantly higher signal-to-noise ratios images reconstructed with no ToF information. This is because individual events can now be constrained within a smaller interval in the image reconstruction process, thus limiting propagation of noise in the projection operations within image reconstruction.

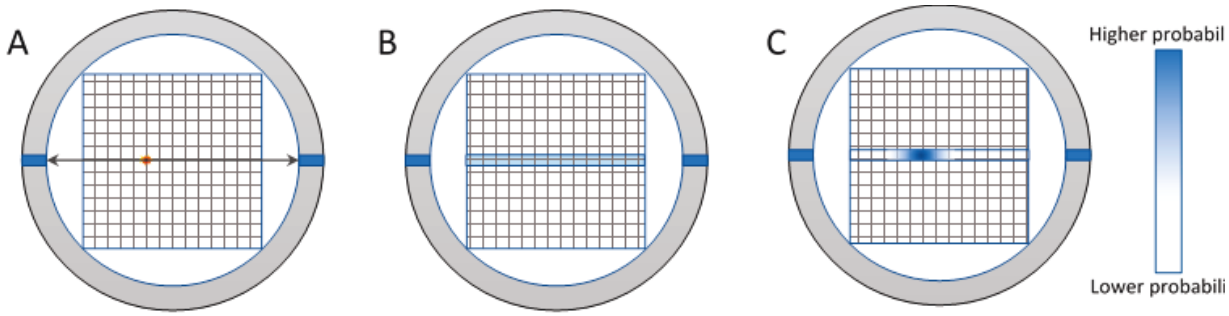

Figure 1-3. ToF PET. A) a pair of photons from an annihilation is emitted and detected by opposing detectors. B) without ToF, there is no information regarding the whereabouts of the source along the LOR. During the backprojection, the event is backprojected with a uniform probability of originating anywhere along the LOR. C) In the presence of ToF information, some degree of localization is possible.

1.2.2.3. Hybrid PET/CT and PET/MRI

PET systems are most commonly manufactured as combined PET/CT systems or sometimes as PET/MRI systems. PET/CT systems and some PET/MRI systems are configured in tandem or back-to-back arrangement. The power of such hybrid systems is that the two scans are acquired in quick succession, and thus both the PET and CT/MRI images can be naturally and conveniently fused, providing important underlying anatomical information to contextualize the PET images. The CT or MRI can also be used to provide a map of tissue attenuation values to compensate for photon attenuation as well as scatter. Such multi-modality imaging is also convenient in terms of cost and scheduling perspectives as both images are acquired in a single imaging session. Since 2003, standalone PET systems have not been manufactured (with few exceptions) because of the convenience and advantage of PET/CT imaging whose value was increasingly recognized. In contrast to CT, the MRI component can provide further soft-tissue contrast information that is particularly useful in brain studies. At the same time, estimating attenuating information from MRI images is much less straightforward than from CT images [85]. However, ongoing developments in this area have created viable solutions [86-88]. At the same time, the significantly higher costs associated with PET/MRI systems, and lack of definitive clinical 'killer apps' so far, have limited widespread utility of PET/MRI systems compared to PET/CT systems.

It should be noted that since most PET systems have an axial FoV ~15-26 cm, they often image the whole-body using a motorized bed that can operate in the so-called step-and-shoot mode. In this method, the patient body is scanned for some minutes at every bed position and then is moved to the subsequent bed position. Newer generation PET scanners (e.g. improved crystal/electronics technology including improved ToF) have enabled high-quality images with scan times of ~2-4 min depending on the application and patient size [89]. Some recent PET scanners can also run under continuous bed motion acquisition mode, where the bed is continuously moving, also allowing with variable scan times for different parts of the body [90]. The advantage of such systems is the ability to customize the speed to specifically reduce the speed during the acquisition over more important organs to capture more incoming photons, thus improving signal-to-noise ratios.

1.2.2.4. Applications of PET imaging

PET is considered both research and clinical tool. Its main application is in clinical oncology but also includes applications in other clinical areas such as cardiology and neurology, in addition to supporting drug development and pre-clinical studies.

One important component of a PET scan procedure is the specific radiotracer administrated to the patient. Some examples are summarized in Table 1-4. Next, we briefly discuss some areas of applications.

Table 1-4. Examples of positron-emitting radiotracers used in PET imaging, including targets and clinical applications.

Isotope	*Compound*	*Target*	*Clinical application*
Oncology			
F-18	FDG	Glucose metabolism	Tumor imaging
F-18	FMISO	Hypoxic cell tracer	Hypoxic tumor imaging
F-18	FET	Amino acids	Glioma
F-18	DCFPyL	Prostate specific membrane antigen (PSMA)	Prostate cancer tumors
C-11	Acetate	Intracellular phosphatidylcholine membrane	Tumor imaging
C-11	Choline	Phospholipids synthesis	Cancer cell proliferation
Ga-68	PSMA-11	PSMA	Prostate cancer tumors
Cardiology			
O-15	Water	Blood flow	Myocardial perfusion
N-13	Ammonia	Blood flow	Myocardial perfusion
Rb-82	Chloride	Myocardiocytes	Myocardial perfusion
C-11	Acetate	Myocardial oxidative metabolism	Myocardial blood flow
Neurology			
F-18	FDOPA	Pre-synaptic dopaminergic	Nigrostriatal dopaminergic pathway
F-18	Amyvid	β-amyloid neurotic plaque	Diagnosis of Alzheimer's disease
C-11	Raclopride	Dopamine D2 receptor	Movement disorders
C-11	SCH23390	Dopamine D1 receptor	Schizophrenia
C-11	NMSP	Dopamine, serotonin	Neural stem cells transplant
O-15	Oxygen	Oxygen metabolism	Cerebral blood volume

PET in oncology

The most widely-used radiotracer in PET imaging is fluorine-18 fluorodeoxyglucose, or [^{18}F]-FDG. FDG is a glucose analog taken up by the tumor cells, then phosphorylated to FDG-6 phosphate by hexokinase, and then trapped in the cell in almost all tissues because further downstream glycolysis is not possible [91]. Tumor cells require more energy than other normal cells, resulting in significantly higher glucose metabolism in malignant cells. As this tracer is trapped in the cell, the attached F-18 radionuclide decays by emitting positrons that are detected by the PET scan. [^{18}F]-FDG is the staple of initial cancer staging, restaging, recurrence, and monitoring response to treatment in

many types of cancers and are covered by the Centers for Medicare and Medicaid Services (CMS) [92-95].

A range of other radiotracers is used in oncological clinical practice and research settings. Here we mention one very active research area in PET oncology, namely that of imaging prostate cancer patients using radiotracers targeting the prostate-specific membrane antigen (PSMA) molecule, which is a transmembrane protein that is considerably overexpressed in most prostate cancer cells [96]. PSMA imaging is shown to have increased sensitivity and specificity compared with current standard imaging with CT, MRI and bone scintigraphy in patients with primary intermediate or high risk of prostate cancer [97]. Just as PSMA has gained interest in molecular imaging with PET, it has also gained interest in targeted radioligand therapy [98]. Furthermore, new concepts have emerged on PSMA targeted theranostics using PSMA tagged by Ga-68 for PET imaging and by Lu-177 for therapy and SPECT imaging [99].

PET in cardiology

Although SPECT imaging dominates nuclear cardiology imaging, PET is also used as a clinical imaging tool for quantitative assessment of myocardial perfusion and characterization of tissue viability in patients with coronary artery disease [100]. Cardiac PET can accurately identify and assess coronary artery stenosis severity as a basis for choosing and following effects of interventions. PET imaging is not only a reliable tool for managing coronary artery disease in traditional cardiology practice based on symptoms but also facilitates management of asymptomatic coronary atherosclerosis non-invasively. Such accurate evaluations can be used to identify patients who need cardiac catheterization and avoid unnecessary invasive procedures in patients with mild/no coronary artery disease [101]. Evaluating blood pool activity and perfusion defect severity, comparing myocardial

metabolism to perfusion, and acquiring PET imaging in conjunction with electrocardiogram (ECG) gating are among other PET imaging applications in nuclear cardiology [102].

PET in neurology

PET has numerous applications in neurology and brain imaging. It can provide information to pinpoint and evaluate a range of brain pathologies. Certain radiotracers like O-15 oxygen and [^{18}F]-FDG can be used to measure brain oxygenation and metabolism, respectively, which significantly decrease in patients with Alzheimer's disease [103]. Unlike surface electroencephalogram (EEG), depth electrocorticography, interoperative corticography, and structured MRI that lack high specificity in determining epilepsy, PET image can provide information about information about the regions of the brain that is causing seizure [104]. PET is also useful in evaluating neurodegenerative disease such as Huntington's disease [105, 106], Parkinson's disease [107, 108], Alzheimer's disease [109-112], and multiple sclerosis [113, 114]. It is also used in neuropsychology and cognitive neuroscience to explore links between specific psychological processes or disorders, and the brain activity [115, 116].

Dynamic PET

Current clinical PET protocols use patterns established for traditional nuclear medicine, where they are optimized for qualitative as opposed to quantitative assessment. The radiotracer is administered to the patient who then waits for a period prior to image acquisition, which in the case of [^{18}F]-FDG is ~60 minutes. This period allows the radiotracer to accumulate in the organs of interest and to washout from surrounding organs [117]. Then the patient is imaged on the PET scanner based on protocol (e.g. 20 minutes depending on center, scanner, and patient), as a result, a single *static* image is acquired during this acquisition time. However, radiotracer distribution in the body is a dynamic

process that is essentially different for normal organs, different tumors, and among different patients [118, 119]. An alternative approach involves the *dynamic* acquisition of temporal images that enables a more complete measurement of tracer kinetics between different physically or chemically-distinct states or compartments. The exchange of tracer between compartments can be modeled using ordinary differential equations, whose coefficients are kinetic parameters. These resulting kinetic parameters and models have been validated to produce reliable quantitative measurements of various clinically important physiological processes [120-122].

An example of a widely-popular compartmental method is the Patlak model [123]. This graphical analysis method evaluates sequential data such as tissue and blood concentrations over time. It uses linear regression to analyze pharmacokinetics of tracers such as [^{18}F]-FDG [124]. This method assumes that the tracer can be modeled as having by a nearly irreversible compartment where the radiotracer enters and gets trapped during the measurement [125]. Patlak linear graphical analysis directly estimates the tracer influx or uptake rate constant K_i and blood distribution volume V, which when quantified at the individual voxel-level enable parametric imaging. Assuming reversible dynamics only, alternative non-Patlak graphical methods can be used (e.g. Logan method [126]). Dynamic imaging and kinetic modeling can provide more accurate quantitation of the tracer relative to conventional standard uptake value (SUV) in PET imaging, as discussed in 1.3.2.1. Figure 1-4 shows an example co ofnventional SUV image vs. parametric images obtained from kinetic modeling of dynamic images of a patient scanned with [^{18}F]-FDG PET. The Patlak K_i image has significantly reduced background uptake compared to conventional SUV, while high background PET signals are observed in SUV and Patlak V images.

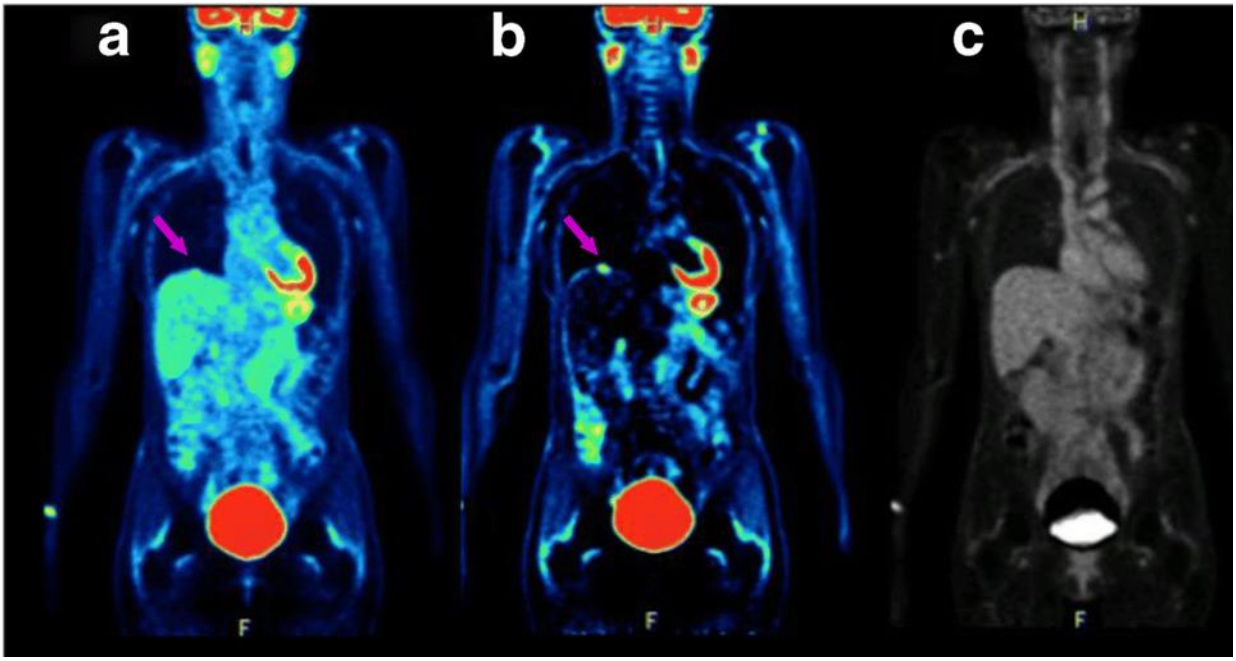

Figure 1-4. An illustration of static and dynamic PET. A) static conventional SUV imaging (70-90 minutes post-FDG injection). Dynamic (parametric) images (0-90 minutes) of b) K_i, and c)V, generated from a dynamic whole-body PET scan using Patlak plot with an image-derived input function and linear regression. Note the tumor uptake marked with a purple arrow on top of the liver that is notably visible in "b" but has dissolved in the background uptake in "a".

Parametric kinetic uptake imaging has implications in various sectors of clinical practice. Dynamic cardiac PET followed by kinetic modeling has significant applications in the clinic [127]. In oncologic PET, dynamic imaging and trace uptake quantification based on compartmental modeling has been shown to improve tumor characterization and treatment response monitoring [128-133]. In the past, dynamic imaging was mostly limited to a single bed position over the organ of the interest, but recently whole-body parametric imaging protocols have been introduced [134, 135], which refers to coverage of the body using multiple bed positions [136] or using the continuous bed motion technique [137], both involving multiple passes. The patient is sometimes injected with the radiotracer right on the scanner bed to capture the blood input function over the heart and can be scanned up to 95 minutes based on different imaging protocols [138]. Alternatively, population-based input functions can be used, and scaled for the specific patient. Dynamic whole-body PET is an emerging technology with promising clinical potential, specifically for improved quantification and assessment of systemic disease, including cancer, inflammation, and infection[138]. It combines

visualization and tracer uptake quantitation across the whole body, while minimizing dependence on SUV activity. It removes background uptake that allows small and less FDG avid tumors to be identified. In addition, conventional SUV image can be readily obtained by summation of multiple passes through the patient [138]. For more clinical applications of dynamic whole-body imaging, its advantages and challenges please refer to [138].

1.2.3. Tomographic image reconstruction

1.2.3.1. Image reconstruction basics

In this subsection, we discuss the basics of tomographic image reconstruction. Our focus is on PET imaging, but application to SPECT is very straightforward. First, we elaborate on the line-integral model. Without considering some degradation effects such as attenuation, scatter and randoms, detector deficiency, etc. that will be discussed in section 1.2.4, the total number of detected coincidences is related to the total amount of tracer contained in the volume or line of response (LOR), a tube that joins two face-to-face detectors. In what follows, we will discuss the 2D tomographic reconstruction. For more information regarding 3D PET reconstruction please consult [53, 54, 139].

2D PET imaging only considers LORs lying within a specific imaging plane from a 2D object $f(x, y)$ as indicated in Figure 1-5. Values of $p(s, \phi)$ are the line integral of the object across the LOR for a fixed ϕ. The collection of all projections for $0 \leq \phi \leq 2\pi$ creates a 2D map of s and ϕ, referred to as a *sinogram* (Figure 1-5). The name "sinogram" comes from the fact that if there is a point-source in the object, it will create a sinusoidal trend on the sinogram.

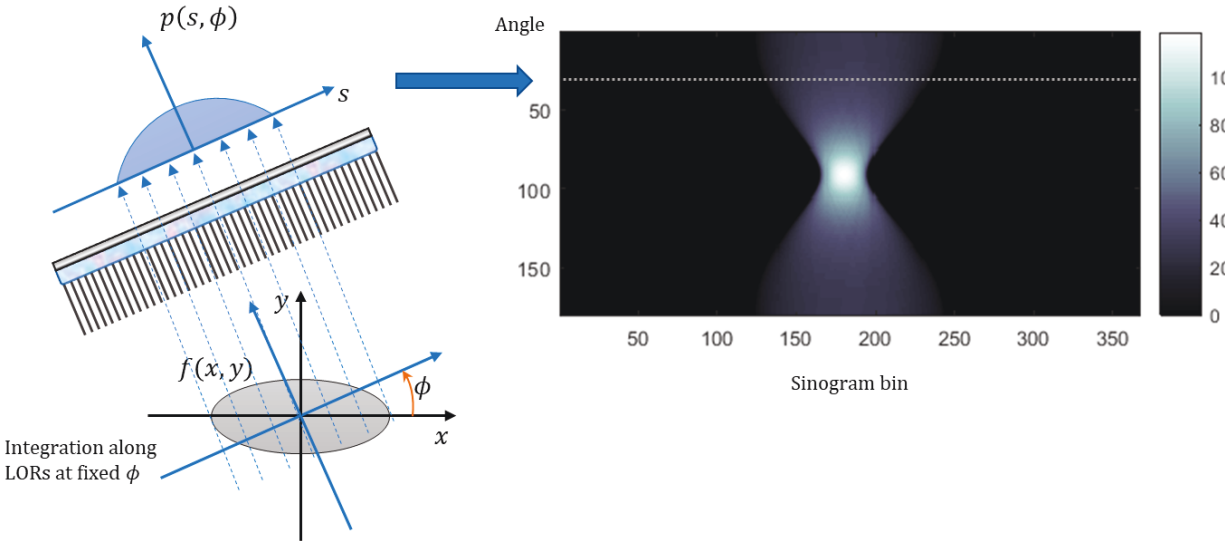

Figure 1-5. A projection $p(s,\phi)$ is generated from integration along all parallel LORs at an angle ϕ. The projections from all angles are stacked on top of each other to generate a sinogram.

From Figure 1-5, we can define a line in the plane by

$$L(s,\phi) = \{(x,y) | \, x\cos\phi + y\sin\phi = s\} \tag{1.4}$$

where s is the lateral position of the line and ϕ is the angle normal to the line. The line integral of the object $f(x,y)$ is given by

$$g(s,\phi) = \iint_{-\infty}^{\infty} f(x,y)\delta(x\cos\phi + y\sin\phi - s)\, dx\, dy \tag{1.5}$$

Here, the delta function is used to make the integrand to be zero everywhere except on the LOR. For a fixed ϕ, equation 1.5 is called a *projection*, and for all s,ϕ, then $g(s,\phi)$ is called the 2D *Radon transform*. The Radon transform collects line integrals, also called projection data, at multiple views from the underlying object, where the line is $s = x\cos\phi + y\sin\phi$. For the forward model, the Radon transform is called the *forward projection operator*, and for the inverse model, the adjoint of the Radon transform is referred to as the *backprojection operator*. The forward projection operator generates the projection data, while the backprojection operator reconstructs the image from the projected data.

Images of radiotracer activity concentration can be reconstructed from the projection data by solving the inverse of equation 1.5. Image reconstruction algorithms can be categorized into 1) analytical and 2) statistical reconstruction methods. The focus in this dissertation is on statistical image reconstruction which we discuss next; for details on analytical reconstruction methods such as the Fourier slice theorem and filtered back-projection, readers are referred to other references [53, 54].

1.2.3.2. Statistical image reconstruction

Maximum-likelihood expectation-minimization (ML-EM)

Statistical image reconstruction has demonstrated superior performance over analytical methods [140]. Statistical methods allow (i) accurate modeling of statistical noise, (ii) complex detector geometrics, as well as (iii) the ability to include corrections for various image degradation effects (section 1.2.4) [141, 142]. Here we briefly discuss the mathematics of image reconstruction in PET [143].

Let $Y_i, i = \{1, \dots, M\}$ denote the projection data as elements of the sinogram matrix, where M is the number of sinogram bins. Each Y_i can then be modeled as an independent Poisson random variable:

$$Y_i \sim \text{Poisson}\{\overline{y_i}\} \tag{1.6}$$

Let $\boldsymbol{y} = [y_1 \quad y_2 \quad \cdots \quad y_M]^T$ and $\overline{\boldsymbol{y}} = [\overline{y_1} \quad \overline{y_2} \quad \cdots \quad \overline{y_M}]^T$ denote column vectors of the measured and expected counts, respectively. The expected counts are related to the unknown activity distribution $\boldsymbol{x} \in \mathbb{R}^N$ by an affine transform:

$$\overline{\boldsymbol{y}} = \boldsymbol{P}\boldsymbol{x} + \boldsymbol{r} \tag{1.7}$$

where $\boldsymbol{P} \in \mathbb{R}^{M \times N}$ denotes the detection probability matrix, or the forward projection operator with each element (i, j) referring to the probability of detecting an event from the j^{th} voxel at the i^{th} detector pair, and $\boldsymbol{r} \in \mathbb{R}^M$ is the contributions of scatter and random events in the projection data (see section 1.2.4). Now, the log-likelihood function for the Poisson distributed projection data can be written as:

$$L(\boldsymbol{y}|\boldsymbol{x}) = \sum_{i=1}^{M} y_i \log \bar{y}_i - \log \bar{y}_i - \log \bar{y}_i ! \quad 1.8$$

The image reconstruction task can be framed as an optimization problem and can be solved using the maximum-likelihood expectation-minimization (ML-EM) [144, 145]. The ML-EM has been proven to converge, and in addition, it incorporates a non-negativity constraint such that if the initial estimate is non-negative, image estimates at every iteration are non-negative. The inverse problem or the maximum likelihood estimate can be then written as:

$$\hat{\boldsymbol{x}} = \arg\max L(y|x) \quad 1.9$$

$$\boldsymbol{x}^{n+1} = \frac{\boldsymbol{x}^n}{\boldsymbol{P}^T I_M} \boldsymbol{P}^T \frac{\boldsymbol{y}}{\boldsymbol{P}\boldsymbol{x}^n} \quad 1.10$$

where $\boldsymbol{x}^n$ is the image estimate at the n^{th} iteration, $\boldsymbol{P}^T$ is the backprojection operator, and I_M is a column vector of ones.

An accelerated version of ML-EM algorithm is called the ordinary-subset expectation-minimization (OS-EM) algorithm [146]. It divides the projection data into subsets and uses those subsets to iteratively update the estimation image. The number of OS-EM subsets are approximately equal to its speed-up factor. However, unlike ML-EM, OS-EM is not necessarily convergent, though some subsidized algorithms have been proposed that can converge [147, 148].

Noise in statistical reconstruction

Tomographic reconstruction is an ill-posed problem, and the ML estimates are extremely noisy. The convergence property of ML-EM is for its noise-free implementation, which is quite unrealistic. In practice, noise contributes to the detected counts. Moreover, as the number of ML-EM (or OS-EM) iterations increase, the impact of noise on the reconstructed images are more pronounced. That is why in practical implementations of ML-EM, the number of iterations is fixed somewhere around 60 to 80 updates, which for instance, correspond to 3 to 4 iterations with ~20 OS-EM subsets. Furthermore, the noise will be further reduced by post-smoothing the image with a Gaussian filter.

Other methods have been proposed to regularize noise to obtain more accurate estimates of activity distributions, such as using a penalized likelihood function [149, 150], where instead of maximizing the log-likelihood function, a penalized log-likelihood function is maximized that encourages smooth solutions with a penalty function and a tuning parameter (hyperparameter) β to control the contribution of the penalty function and the resolution-noise properties of image estimates [151-153]. Although penalized (also known as Bayesian) reconstruction has recently become available on commercial PET scans [154], the hyperparameter β can pose a challenge since it may have to be finely tuned for different patients/radiotracer/reconstruction settings. The penalty function is obtained by incorporating *a priori* distribution on the estimated image; i.e. $p(x) = \frac{1}{Z} e^{-\beta R(x)}$, where Z is a normalizing constant, and $R(x)$ is a penalty function. The penalized log-likelihood is also called *maximum a posteriori* or MAP estimate [155-158].

1.2.4. Causes of image degradation and quantitative inaccuracy in nuclear medicine

The relationship between the object and the projection space can be defined by a system matrix, which is the detection probability matrix $\boldsymbol{P}$ introduced in equation 1.7. Statistical reconstruction methods allow modeling of physical degradation factors which will be presented in the next section in the reconstruction framework. The system matrix can be factorized as [159]:

$$\boldsymbol{P} = \boldsymbol{P}_{det.sens}\boldsymbol{P}_{res.blur}\boldsymbol{P}_{attn}\boldsymbol{P}_{geom}\boldsymbol{P}_{positron} \quad 1.11$$

where $\boldsymbol{P}_{positron}$ models positron range in the image space, $\boldsymbol{P}_{geom}$ is the detection probability matrix, $\boldsymbol{P}_{attn}$ is a diagonal matrix of attenuation correction factors, $\boldsymbol{P}_{res.blur}$ accounts for phenomena that result in resolution blur, and finally $\boldsymbol{P}_{det.sens}$ is a diagonal matrix containing detector block normalization factors. We elaborate more on these degradation factors in the subsequent section.

PET imaging is impacted by several physical phenomena that result in degradation in image quality and quantitation. It is important to understand these phenomena to develop correction methods to restore the quality of the image to a better extent. Some of the most important factors are:

1) Attenuation of annihilation photons
2) Detection of scattered events
3) Detection of random events
4) Positron range effect
5) Photon non-collinearity
6) Detection deadtime
7) Detector blurring
8) Variations in detector sensitivity
9) Decay of radioactivity
10) Geometric correction
11) Patient motion

From the above list, items 2, 4, 5, 10, and 11 are specific to PET, while others are common between PET and SPECT. Corrections for detector blurring, positron range, and photon non-collinearity have been traditionally ignored because of the inherent low resolution of the PET camera. However, with the advent of higher-resolution scanners, it has become more important to account for these factors in the reconstruction task. Compared to SPECT imaging, attenuation fractions are larger in PET, but their corrections are easier. Meanwhile, detection of scattered evens is much more prevalent in 3D PET and complicates quantitative reconstruction using PET.

Before we discuss these phenomena, we need a clear understanding of a *true* event. This is similar in both PET and SPECT, though we explain it from the perspective of PET. As mentioned in section 1.2.2, once a positron annihilates, it produces two anti-parallel gamma-ray photons. In a degradation-free

system, these photons will travel in straight lines (along a LOR) towards the detectors and fall within the coincidence time window. Such an annihilation event that has not been diverted from its LOR and has arrived within the coincidence time window is called a true event. In the following subsections, we discuss the various phenomena and briefly review some of the related correction techniques. Later in this dissertation, we model a number of these phenomena and present a novel method for their correction in chapter 2: Adaptive Point-Spread Function (PSF) Modeling for Enhanced Quantitation in PET Image Reconstruction.

1.2.4.1. Attenuation

Gamma-rays interact with matter as they travel through a medium, through photoelectric absorption, or Compton scattering. Photoelectric absorption is the dominant photon-matter interaction in tissue for photon energies below 100 keV. On the other hand, Compton scattering is more dominant for photon energies between 100 keV and 2 MeV [53], and is the only important interaction process at 511 keV within subjects. As discussed in section 1.2.1, in a Compton interaction, the photon interacts with a free or outer-shell electron, causing a decrease in the energy of the photon and a change in its direction. The energy of the scattered photon is then given by the following equation:

$$E_{scatter} = \frac{E_0}{1 + \frac{E_0}{0.511}(1 - \cos(\theta))} \qquad 1.12$$

where E_0 is the energy of the gamma photon before interaction and θ is the angle of scattering. This phenomenon can cause the scattered photon to be deflected out of the FoV so as (i) it is not detected, or (ii) it arrives at another detector and is detected by it. Either of these cases results in a loss of the true LOR. This phenomenon is referred to as *attenuation*. Attenuation correction should be incorporated in the image reconstruction process to avoid underestimation of the radiotracer

distribution. Defining the survival probability A_a as the probability of a photon not interacting as it propagates along a path a, the *Beer-Lambert law*, or simply Beer's law says:

$$A_a = e^{-\int_a \mu(x)\,dx} \tag{1.13}$$

where $\mu(x)$ is the *linear attenuation coefficient* of the medium photon is traveling in. Linear attenuation coefficient provides insights on how effective a given material is in promoting photon interactions at position x per unit thickness. It increases with higher matter densities and decreases with higher photon energies. Subsequently, the probability of attenuation often called *attenuation factor* (AF) is given by $1 - A_a$, which is the probability that a photon travels through a medium without interacting with matter.

The survival probability of photons along a LOR is the product of the probabilities of each photon not interacting as they propagate along their paths a and b:

$$A_L = e^{-\int_a \mu(x)\,dx} \times e^{-\int_b \mu(x)\,dx} = e^{-\int_L \mu(x)\,dx} \tag{1.14}$$

where L refers to the union of a and b, or the entire LOR. In PET imaging, more than 60% of all emitted photons interact with tissue; however, equation 1.14 shows that attenuation along a LOR in PET is independent of the position of the annihilation event along the LOR. This contrasts with SPECT where attenuation is depth-dependent on the distance to the detector. This key observation result in more straightforward attenuation correction in PET compared to SPECT.

Nowadays the most widely-used method of attenuation correction involves the use of the CT image, a key motivation behind PET/CT and SPECT/CT scanners. The CT component of the system can quickly acquire an image of the body, with its voxel intensities corresponding to the attenuation coefficients of that location in the body. Historically, PET-only scanners used a transmission scan for attenuation correction, which would rotate a radioactive source around the patient inside the detector gantry and acquire the image without and with the object inside the FoV. Then the ratio

between these two scans along each LOR was a measure of the probability that the generated annihilation photons are not attenuated along each LOR. This method was much more time consuming than CT, and its quality was inferior to a CT image in terms of statistics. Examples non-attenuation-corrected (NAC) attenuation-corrected (AC) PET images are depicted in Figure 1-6. The CT image in (C) is used to correct the PET image shown in part B, and the image is usually displayed as a fused form of CT and AC PET, as shown in (D), to be able to anatomically localize the PET uptake using CT. Part A of this figure demonstrates some common artifacts encountered because of attenuation: unusually-high uptake can be seen on the skin, and the general uptake decreases as we track the image from the skin towards the inner parts of the body. By contrast, in the AC PET image, high skin uptake is gone, and inner-body uptake is more uniform.

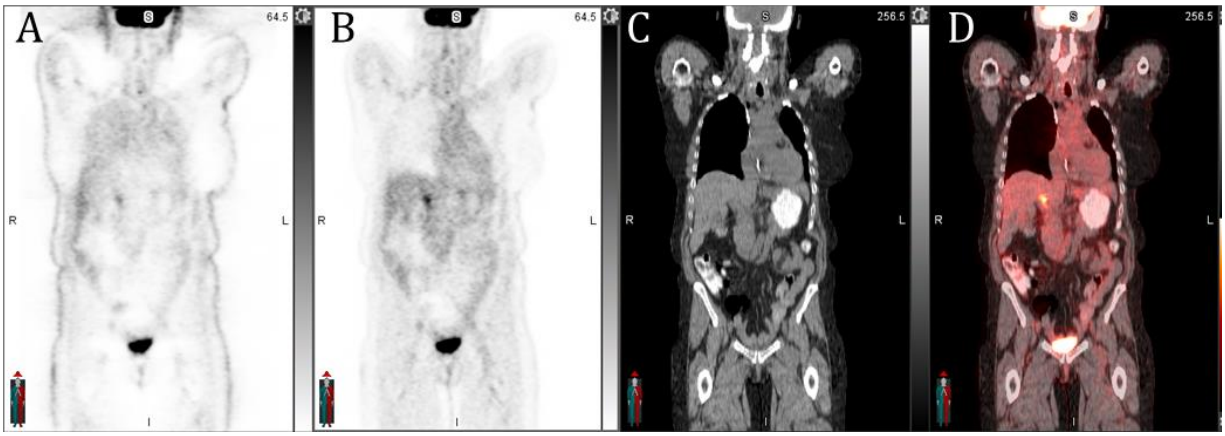

Figure 1-6. Effect of attenuation correction in PET imaging. A) Non-attenuation-corrected (NAC) PET image, B) Attenuation-corrected (AC) image using C) CT image acquired with the PET/CT. D) A fused attenuation-corrected and CT image together.

One point to remember is that the energy level of the photons generated by the x-ray tube of the CT component is of the order of 40~140 keV, which is less than 511 keV of the annihilation gamma-rays [160]. As previously mentioned, the linear attenuation coefficient is dependent on the energy level of the photon in the medium. Therefore, some corrections are required to convert the linear attenuation correction of the image from the CT scan to those that would be obtained at 511 keV [160, 161], posing some challenges. More discussion about attenuation correction in PET/CT and PET/MRI systems was provided in section 1.2.2.3.

1.2.4.2. Scattered events

A scattered event is a term that refers to a specific type of attenuation. When an annihilation photon is scattered due to Compton scattering yet is still detected and considered as a coincidence, it is referred to as a scattered event. Scattering can result in an assignment of an incorrect LOR to the incoming photons, as depicted in Figure 1-7.

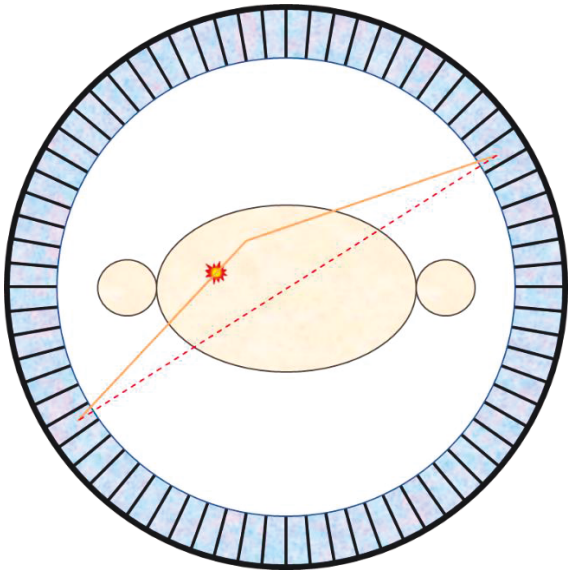

Figure 1-7. Example of a scattered event. One of the annihilation photons got scattered, but both are detected under an incorrect LOR

Mispositioning of the LOR degrades the spatial resolution of the underlying activity distribution. Scatter fractions range from 10% to 20% in 2D PET, were septa are used — thick layers of lead inserted between crystal rings axially to ensure that the incoming photons enter the crystal only from the transaxial slice which the crystal ring circumambulates [162, 163]. This technology resulted in poor sensitivity of the scanner. In more advanced 3D PET systems were the septa are removed, thus significantly improving sensitivity, the added challenge is that scattered events now account for 35% to 50% of all detected coincident events and scatter correction is more importantly required than in SPECT [164]. Scatter correction can be performed using convolution subtraction methods [165],

comparisons of 2D and 3D distributions [166], or direct calculation of scattering distribution using the Klein-Nishina equation [167, 168], the latter being the most commonly invoked.

1.2.4.3. Random events

A random event occurs when two gamma-rays originating from two distinct annihilation events are detected within the coincident time window. To understand this, note that only a 'single event' can be detected from two anti-parallel photons corresponding to a given annihilation. As such, a random event occurs when two independent single events occur within a coincidence time window. Random events cause localization of an annihilation event along an incorrect LOR. Single events occur because of one of the following reasons:

1) One gamma-ray photon is attenuated (due to photoelectric absorption or scattering) and is not detected
2) One photon passes through the detectors without being detected
3) One photon does not arrive at the detector at all due to the location and direction of the originally emitted gamma-rays

These are depicted in Figure 1-8. The rate of random events along a LOR connecting detectors i and j is given by:

$$R_\tau = 2\tau S_i S_j \qquad 1.15$$

where τ is the coincidence time windows, and S_i and S_j refer to the single event rates at the two detectors. From the above equation, one may infer that 1) reducing the coincidence time windows results in reduced random coincidences, and 2) since single rates are proportional to the activity in the subject, random rates are quadratically proportional to the activity. This implies that higher

injected activity to the patient can result in much higher random rates, and thus can be a limiting factor to determine the maximum optimal injected activity.

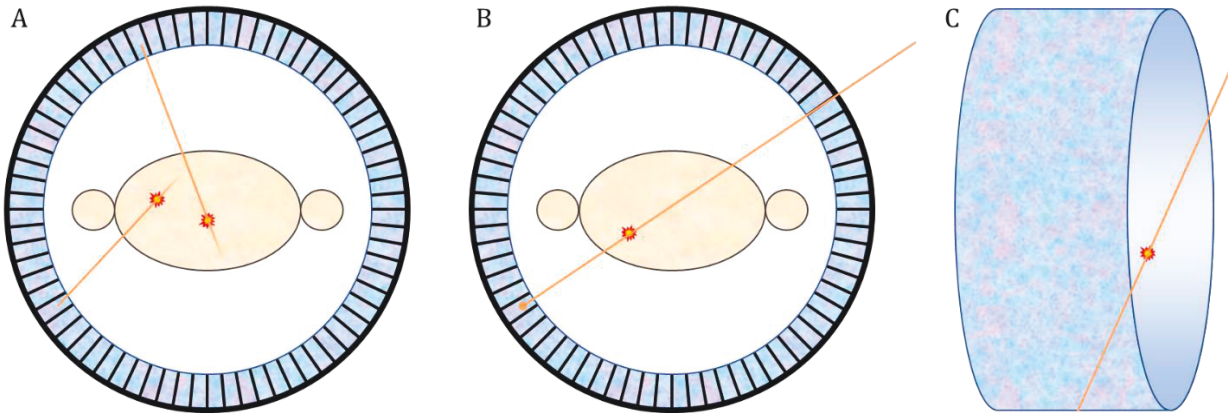

Figure 1-8. Examples of single events that may contribute to random events. A) one photon never gets to the detectors due to photoelectric absorption or scattering. B) one photon passes through detectors without being detected. C) One photon does not meet detectors due to the orientation of the annihilation. The detection of the two single events, like in "A", that happens within the coincidence time window, results in a random event.

Correcting for random events can be achieved using image-based convolution-subtraction [168], or estimating random events using (i) delayed-time window methods [169], or (ii) by invoking above equation 1.15 [170], the latter most common. The estimated random events can then be directly subtracted from the measured data, or incorporated within image reconstruction, the latter is more meaningful as it preserves statistical considerations in image reconstruction. The introduction of ToF technologies has significantly reduced random rates due to the smaller time-windows [171].

1.2.4.4. Positron range effect

Once an unstable radionuclide decays and a positron is emitted, it has a very high kinetic energy that is reduced by a sequence of Coulomb interactions and getting to a threshold of thermal energy that is suitable to interact with an electron and annihilate. The distance positron travels to reach thermal energies and to interact with an electron is called the positron range. Positron range depends mainly

on the radionuclide and the kinetic energy of the released positrons. The positron range of some of the most widely-used PET radionuclides was mentioned in Table 1-1.

Positron range induces spatial blurring in the reconstructed image due to misplacing of the LOR. The detected LOR is slightly shifted from the true LOR because of the distance that positron has traveled (Figure 1-9).

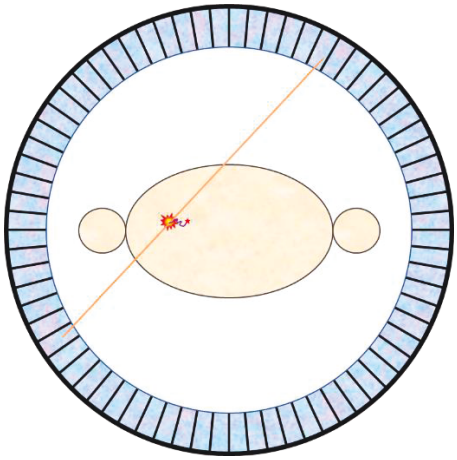

Figure 1-9. Positron range effect. The emitted positron travels a distance and reduces its kinetic energy to thermal energy and annihilate upon meeting an electron. The detected LOR does not necessarily pass through the location where positron was emitted.

As the positron range effect induces a spatial blur in the image, its impact on image resolution depends on the intrinsic resolution of the scanner. As an example, F-18 has a 0.6 mm mean positron range which is quite small compared to the ~4 mm resolution of present-day PET scanners, and thus is negligible. However, contributions for radionuclides with higher positron ranges such as Rb-82 or Ga-68 are higher and may be corrected to reduce blurring from the reconstructed images. This effect is usually modeled as an image-space blurring kernel that can be incorporated within the image reconstruction framework or can be corrected using deconvolution methods [76, 172-175].

1.2.4.5. Photon non-collinearity

As a positron encounters an electron with which it annihilates, neither one of them may have a zero kinetic energy prior to annihilation. As such, when the two particles annihilate, conservation of momentum can make the emitted dual photons slightly deviate from a 180° angle (Figure 1-10).

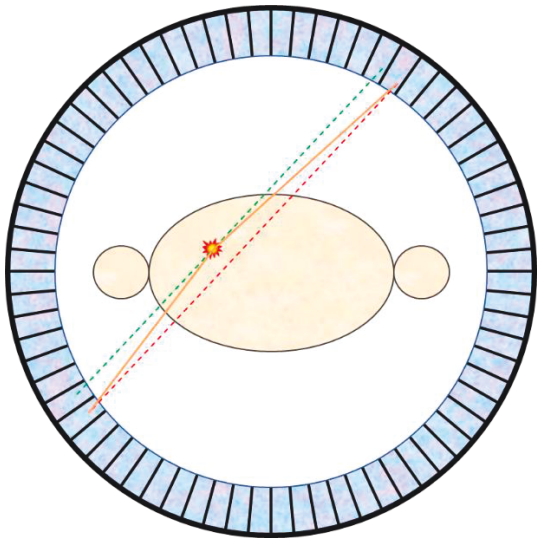

Figure 1-10. Photon non-collinearity effect. A slight deviation of two gamma-rays from 180° results in detecting the incidence from an incorrect LOR (dashed red) instead of the true LOR (dashed green).

This deviation is around 0.4° [176]. The resolution loss due to this effect can be modeled as a function of the diameter of the FoV and can be written as:

$$\text{FWHM} = 0.0022\, D \qquad 1.16$$

where D is the diameter of the FoV. This equation implies that the larger the scanner diameter, the more contribution of photon non-collinearity to resolution degradation. As such, small animal PET scanners are less impacted by this effect relative to a clinical PET scanner. Photon non-collinearity can be modeled as a sinogram blurring kernel in the image reconstruction framework [177].

1.2.4.6. Detection deadtime

Deadtime refers to the time it takes to process an event, and it limits the counting rate of the scanner. As the incoming count rate increases, e.g. due to the higher dose of radioactivity injected to the subject, a larger portion of the incoming counts are lost due to deadtime effects. As such, this is another condition that limits the injected radioactivity to a patient for optimal imaging. The bottlenecks that contribute to deadtime can be the scintillator decay time, crystal identification, energy discrimination, overall coincidence detection, or delays in electronics. Therefore, deadtime effects can occur at the level of the crystal, block or the subsequent electronic circuitry. Deadtime can be corrected by scaling the measured counts by deadtime correction factors measured from the single rates [53].

1.2.4.7. Detector blurring

An incoming photon to a crystal excites electrons in the crystal by Compton scattering and/or photoelectric absorption. If all these processes occur in a single crystal where the phone had entered, the event and subsequently the LOR will be properly positioned. However, three complications may happen:

Inter-crystal penetration

The first issue is when the incident photon enters a crystal at an oblique angle, passing through the crystal undetected and only detected in the adjacent crystal. This is referred to as *inter-crystal penetration*, and results in a mispositioned LOR. Subsequently, as the annihilation event originates at a distance from the center of the FoV and closer to the edges of the scanner, photons enter the detector at a higher angle, with a higher likelihood of penetration. This results in the so-called depth of interaction effect (elaborated more later) and manifests itself in the reconstructed images as the

so-called parallax effect: degrading spatial resolutions away from the center of the FoV. The stopping power and density of the scintillation crystal directly determine the extent of this effect.

Inter-crystal scattering

The second issue is when the photon arriving at a crystal gets scattered inside the crystal block and gets detected in a neighboring crystal. This effect is called *inter-crystal scattering*, and it also causes misplacement of a LOR. This effect also worsens as the incoming photons to the detectors arrive at a more oblique angle.

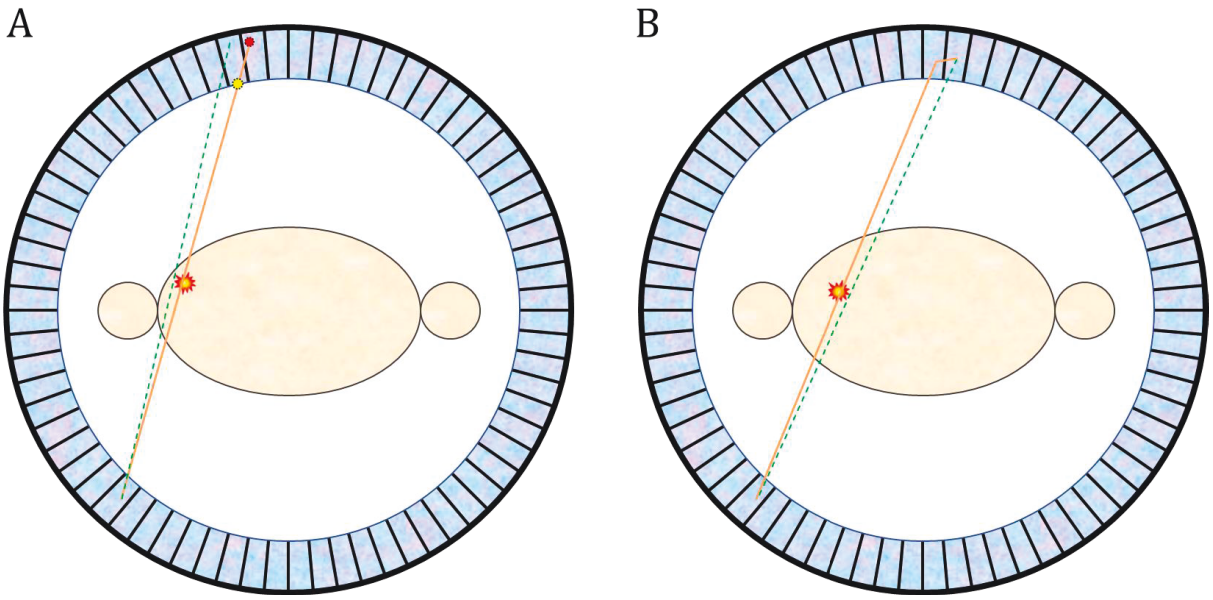

Figure 1-11. Detector blurring effects. A) inter-crystal penetration, where the photon penetrates the adjacent crystal where it gets detected and causes a mispositioned LOR. B) inter-crystal scattering, where the scattering scintillation light gets detected on the other end of the crystal at an adjacent detector, causing a mispositioned LOR.

These two effects are corrected in SPECT using resolution recovery [178], and in PET using point-spread function (PSF) modeling. The latter is the main topic of chapter 2: Adaptive Point-Spread Function (PSF) Modeling for Enhanced Quantitation in PET Image Reconstruction, and will be extensively discussed.

Depth of interaction

A scintillation crystal requires notable thickness to stop 511 keV photons. Gamma cameras used in SPECT systems have often used NaI crystals with a thickness of 1.25 cm or less. PET crystals such as BGO and LSO are usually thicker (2 to 3 cm) with greater stopping power to stop the higher energy 511 keV photons. When the source of emission is placed away from the center of the FoV, increasingly, inter-crystal penetration (discussed before) results in the so-called *depth of interaction* (DOI) effect. Figure 1-12 illustrates this.

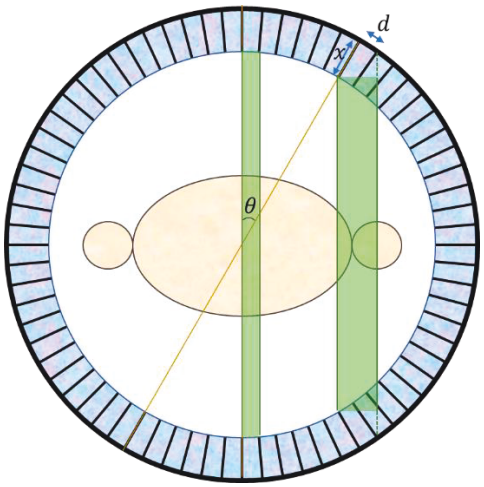

Figure 1-12. The depth of interaction effect. It lowers the resolution for incidents occurs farther from the center of the FoV.

When a source is located near the center of the FoV, spatial resolution is determined by the width of the detector front $R_{det} = d/2$, where d =width of the detector. However, if the source is located away from the FoV, the apparent width of the detector becomes

$$d' = d\cos\theta + x\sin\theta \qquad 1.17$$

where d, x, and θ are indicated in Figure 1-12. This apparent change is because the more angulated detector lacks knowledge about the depth at which an interaction has occurred within the scintillation crystal. The spatial resolution then becomes $R'_{det} = d'/2$ as follows:

$$R'_{det} = \frac{d}{2} \times \left[\cos\theta + \frac{x}{d}\sin\theta\right] \qquad 1.18$$

$$= R_{det} \times \left[\cos\theta + \frac{x}{d}\sin\theta\right]$$

This equation shows that the depth of interaction is described by a multiplicative factor applied to detector resolution at a midpoint between a pair of directly opposing detectors. For a typical PET scanner with x around 2 to 3 cm, d around 0.3 to 0.6 cm, and a diameter of 80 cm, the DOI effect causes around 40% less resolution degradation at 10 cm away from the center of the FoV [53].

Some PET scanners such as the high-resolution research tomograph (HRRT) [179] handle DOI effect through a method called DOI-encoding. In DOI-encoding used in HRRT scanners, the crystal block comprises of a double scintillator crystal layer that provides some discrete level of DOI information and minimizes the misplacement of LORs [180].

1.2.4.8. Variations in detector sensitivity

The various detectors in a PET (or SPECT) camera may have variations in their efficiencies, which are determined by several factors including (i) the cross-section or effective surface area of the crystal, (ii) the change in the effective depth-of-interaction of crystals, (iii) crystal material imperfections, (iv) light guide (i.e. isolation between crystals) variations and imperfections, (v) PMT gain inconsistencies, and (vi) further inconsistencies in electronics to detect PMT signals. To correct for detector variations, usually a positron-emitting germanium cylindrical phantom with uniform activity is scanned by the PET scanner for a long period during the quality check (QC) process. In this case, all detectors presumably should have uniform inputs. This scan is called a *normalization scan* that produces a *normalization sinogram*, which provides coefficients that are proportional to the reciprocal of the number of counts obtained for each LOR. More details are provided in other references [181-184]. An example of a normalization sinogram is provided in Figure 1-13.

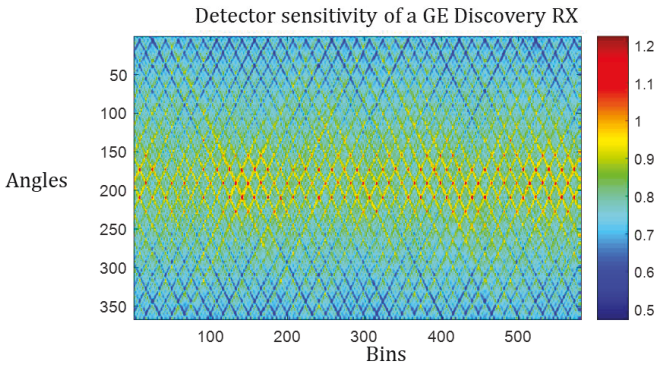

Figure 1-13. An example of a normalization sinogram for a 2D transaxial slice of a GE discovery PET scanner.

1.2.4.9. Decay of radioactivity

Radioactive nuclides undergo radioactive decay, and as time goes on, the rate of decay itself decreases thus having a lower number of disintegrations per second. Decay can be corrected in a straightforward manner by rescaling the image counts using decay correction factors that are specific to the radionuclide used in the radiotracer. Examples of decay half-lives of some commonly-used radionuclides were presented in Table 1-1.

1.2.4.10. Geometric correction

The 360° ring arrangement of the detectors in PET images results in a non-uniform sampling of the LOR in the projection profile. As such, LORs would be more narrowly-spaces (denser) as one steps away from the center of the FoV. This non-uniform sampling can be resolved by applying a simple geometric correction factor

$$\Delta r = \Delta d \sqrt{1 - \frac{2r}{D}} \qquad 1.19$$

where Δd is the detector width, r is the LOR position, D is the diameter of the scanner, and Δr is the correction factor [185].

1.2.4.11. Patient motion

Today's PET scanners have a spatial resolution of 3.5 to 6 mm FWHM. With such improvements in spatial resolution, patient movement during the scan can become more apparent and contribute more to image quality degradation. In a very general sense, patient motion can be categorized into: motion due to the movements of the lungs (respiratory motion), movements due to heart functioning (cardiac motion), and other unintentional or intentional patient body movements (bulk motion). Motion correction in PET has been a large field of research for decades and is beyond the scope of our work. For more information please refer to review articles [186-193]

1.3. Medical Imaging Quantitation and Biomarkers

Modern medical imaging modalities have experienced tremendous progress and can provide an unprecedented level of spatial details and cellular/functional information [194]. At the same time, there have been advancements in our understanding of the molecular underpinnings of disease and the rise of more statistical and evidence-based approaches to diagnosis and treatment. This has paved the way to leverage quantitative techniques in medical imaging for supporting a diverse set of clinical and research goals [195, 196]. In this section, we explain the basics of image quantitation, following by quantitative imaging biomarkers, and we briefly discuss one of the commonly-used quantitative imaging biomarkers in PET.

1.3.1. Image quantitation

Quantitative imaging refers to the extraction and utilization of quantifiable features from medical images for assessment of normality or the severity, degree of change, and status of a disease relative to normal conditions [197, 198]. Research in quantitative imaging includes the development, optimization, standardization, and application of anatomical, molecular and functional imaging acquisition protocols, structured reports, data analyses, and validation against other clinical data

[198, 199]. The concept of medical imaging quantitation is relatively close to the definition of biomarkers, defined as a characteristic objectively-measured and evaluated, which indicates a normal biologic or pathologic process, or response to a therapy [200]. Subsequently, a quantitative imaging biomarker can be defined as an objectively-measured characteristic derived from a medical image and can be correlated with relevant physiological and anatomical parameters such as disease presence, characterization, and severity, as well as its prognosis (predicted disease course with or without treatment), and treatment response monitoring and assessment [197]. More relevant definitions to quantitative imaging biomarker development are provided in Table 1.

Table 1-5. Definitions related to quantitative imaging and biomarkers [197, 198, 201]

Term	Definition
biomarker	A characteristic that is objectively measured and evaluated as an indicator of normal biologic or pathogenic processes, or response to a therapeutic intervention
Predictive biomarker	A biomarker that is used to forecast the efficacy of a therapy/therapies
Diagnostic biomarker	A biomarker that improves the accuracy of patient diagnosis
Prognosis biomarker	A biomarker that improves the accuracy of patient prognosis
Response biomarker	A biomarker whose change after treatment predicts if treatment would lead to a beneficial outcome
Monitoring biomarker	A biomarker that (usually) is regularly measured to detect relapse or emergence of toxicity
Quantitative imaging	The extraction and utilization of various quantitative features, such as numerical and statistical, from medical images
Quantitative imaging biomarker	An objectively-measured characteristic derived from an *in vivo* image as an indicator of a normal biological or pathogenic process, or a response to a treatment
Repeatability	The amount of agreement between subsequent measurements measured under the same condition
Reproducibility	The amount of agreement between subsequent measurements performed under varying conditions

Quantitative imaging has significantly contributed to the value of diagnostic testing and become more prominent in preclinical studies, clinical practice and clinical research [202-206]. The main advantage of quantitative imaging is its potential for standardization and higher precision of image

interpretation. It further helps with increased diagnostic accuracy, improved reproducibility across multiple devices/centers, decreased variability and subjectivity, more structured reporting, and more robust association of image findings with clinical and biological data [207]. Evidence-based medicine is also another driver for developing quantitative imaging as diagnoses in many clinical fronts can be reinforced with quantitative imaging data and biomarkers [208, 209]. Clinical trials are another demanding area for quantitative imaging, in which quantitative measurements of tumor response are measured and reported to assess the efficacy of investigational therapy. Another long-term stipulated utilization of quantitative data has been in guidelines for image-based response assessment such as response evaluation criteria in solid tumors (RECIST) [210], that is based on the measurement of tumor size and frequently used for response assessment in oncology, and have been widely used for many years and successfully validated against long-term patient outcomes [211, 212].

1.3.2. Quantitative imaging biomarkers

Biomarkers are useful only if they provide added value in predicting a clinical outcome [198]. These biomarkers can be derived via methods as simple as using calipers (e.g. to measure the length), or a complex measurement of a functional parameter associated with a dynamic relationship of image measurements to an external stimulus. Quantitative imaging biomarkers can be generally classified as structural, morphological, textural, functional or physical [213]. Nevertheless, there are factors that can inherently affect the measuring of quantitative imaging biomarkers and thus reducing their reliability, repeatability, and reproducibility, which need to be properly addressed and investigated. More information regarding necessary attributes of biomarkers, as well as methods for assessing these attributes (e.g. reproducibility) are discussed later (section 3.2.1). As an example, the next

subsection 1.3.2.1 discusses a very important and widely popular quantitative image biomarker in nuclear medicine imaging. Another set of biomarkers will be introduced in subsequent section 1.4.

1.3.2.1. Standard Uptake Value (SUV)—an important quantitative biomarker

A well-known example of a quantitative imaging biomarker is the standard uptake value (SUV) in PET imaging with 2-deoxy-2-(^{18}F)fluoro-D-glucose or fludeoxyglucose F-18, also known as [^{18}F]-FDG, for oncologic imaging. Increased accumulation of FDG in tumors relative to normal tissue is shown to be a useful marker for detection and staging in many cancers [214]. Moreover, since changes in FDG accumulation have been shown to provide useful information for assessing response to a therapy, SUV is also being used for monitoring individual treatment response as well as an evaluation tool for new drugs and therapies [215]. In PET scanners, the *in vivo* radioactivity concentration is quantified (e.g. in kBq/ml or mCi/cc). The tissue uptake varies between patients due to different injected doses and patient sizes. The SUV is thus used as a relative measure of FDG uptake. It is defined as

$$\text{SUV} = \frac{c(t)}{I/BW} \tag{1.1}$$

where $c(t)$ is the radioactivity activity concentration [kBq/ml] measured by PET for a region of interest (ROI) at time t and is decay-corrected to $t = 0$, I is the injected dose [kBq] to the patient at time $t = 0$, and BW is the body weight [g] of the patient at the time of the imaging. If all injected SUV is retained in the body and evenly distributed, the SUV everywhere in the body would be 1g/ml regardless of patient size and injected dose. If we assume 1 ml of tissue weighs 1 gram, then SUV would be dimensionless.

Many parameters can affect the numerator of the SUV equation such as the definition of the ROI and its size, image resolution, reconstruction settings, and uptake period [216]. Concerns regarding its normalization factors in the denominator are mainly regarding using BW vs. body surface area vs.

lean body mass [119, 217], and BW is sometimes substituted by the other two. This is because if SUV is only corrected for BW, it does not consider the relatively lower FDG accumulation in fatty tissues in the fasting state. Subsequently, SUV corrected for lean body mass would be a more effective quantitative biomarker than BW or body surface area especially for obese patients [119, 218]. There are different formulas in the literature for estimating a patient's lean body mass that usually differs for men vs. women [219], as well as international guidelines for standardizing the entire process of image acquisition in SUV parameters [220]. Following RECIST criteria, a PET-based assessment method was introduced, namely the PET Response Criteria in Solid Tumors (PERCIST), which can include quantification of maximum SUV in an ROI called SUV_{max}, or the average SUV in a 1cc sphere moved around the tumor until it is maximized, referred to as SUV_{peak} [204]. In any case, from this discussion, we can clearly identify SUV as an important quantitative imaging biomarker in oncologic PET imaging.

1.4. Towards personalized medicine with radiomics

This section introduces the topic of personalized medicine, following the emerging topic of radiomics and its applications. Personalized medicine aims at tailoring therapy to each individual for the best response and highest safety margin to ensure better patient care (Figure 1-14) [221]. It paves the way for each patient to receive an earlier diagnosis and optimal treatment customized to one's specific clinical/molecular/genomic profile that ultimately improves healthcare and lowers costs. Personalized medicine is an innovative approach that takes into account a vast spectrum of data – from pharmacogenetic and pharmacogenomics information and protein-based biological markers, to molecular and anatomical diagnosis and response and targeted therapies, to family history, environment, and lifestyle – in order to achieve tailored decisions for individual patients [222-224].

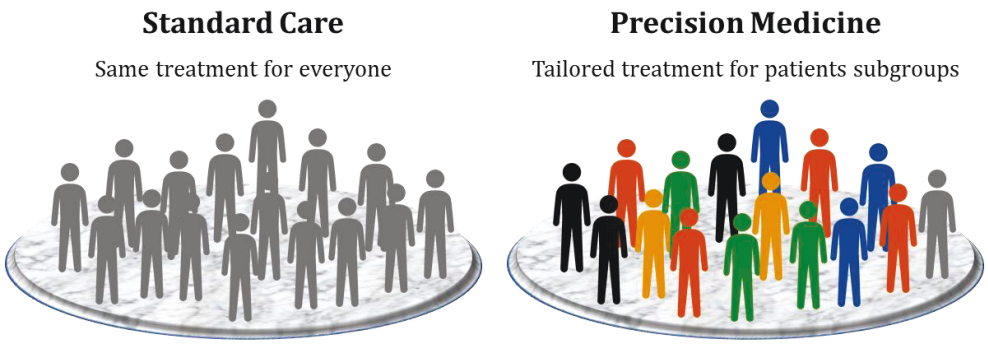

Figure 1-14. An illustration of standard care vs. precision medicine. In the former approach, all patients undergo the same process/treatment, whereas in precision medicine, a subgroup of patients receive tailored care optimized to their clinical/molecular/genomics profile.

Traditionally, the focus of biology has been more towards *describing* biological phenomena and mechanisms. But recently, it has become an increasingly data-rich subject [224]. As such, many fields of biological studies adopt new terminology for this novel approach to data-oriented studies. They add the suffix "-omics" on previously-used terms. More specifically, adding the suffix "-omics" refers to the data-oriented study of that field in biology (e.g. genomics that is a large-scale study of genes), and adding "-ome" addresses the object of study of that field (e.g. genome is the collection of all genes of an organism). Researchers nowadays extensively study different levels of biological "-omics", such as: genomics (large-scale data-oriented study of genes), proteomics (large-scale data-oriented study of proteins) [225], metabolomics (large-scale data-oriented study of metabolites) [226], pathomics (large-scale data-oriented study of pathological data) [227], and more recently, radiomics (large-scale data-oriented study of radiological images) [228] (Figure 1-15). They have further worked on combinations of two or more of these "-omics" research areas to yield interesting results, such as a combination of pathomics and radiomics [229], or many studies combining radiomics and genomics that has emerged as a new era of radiogenomics [230, 231].

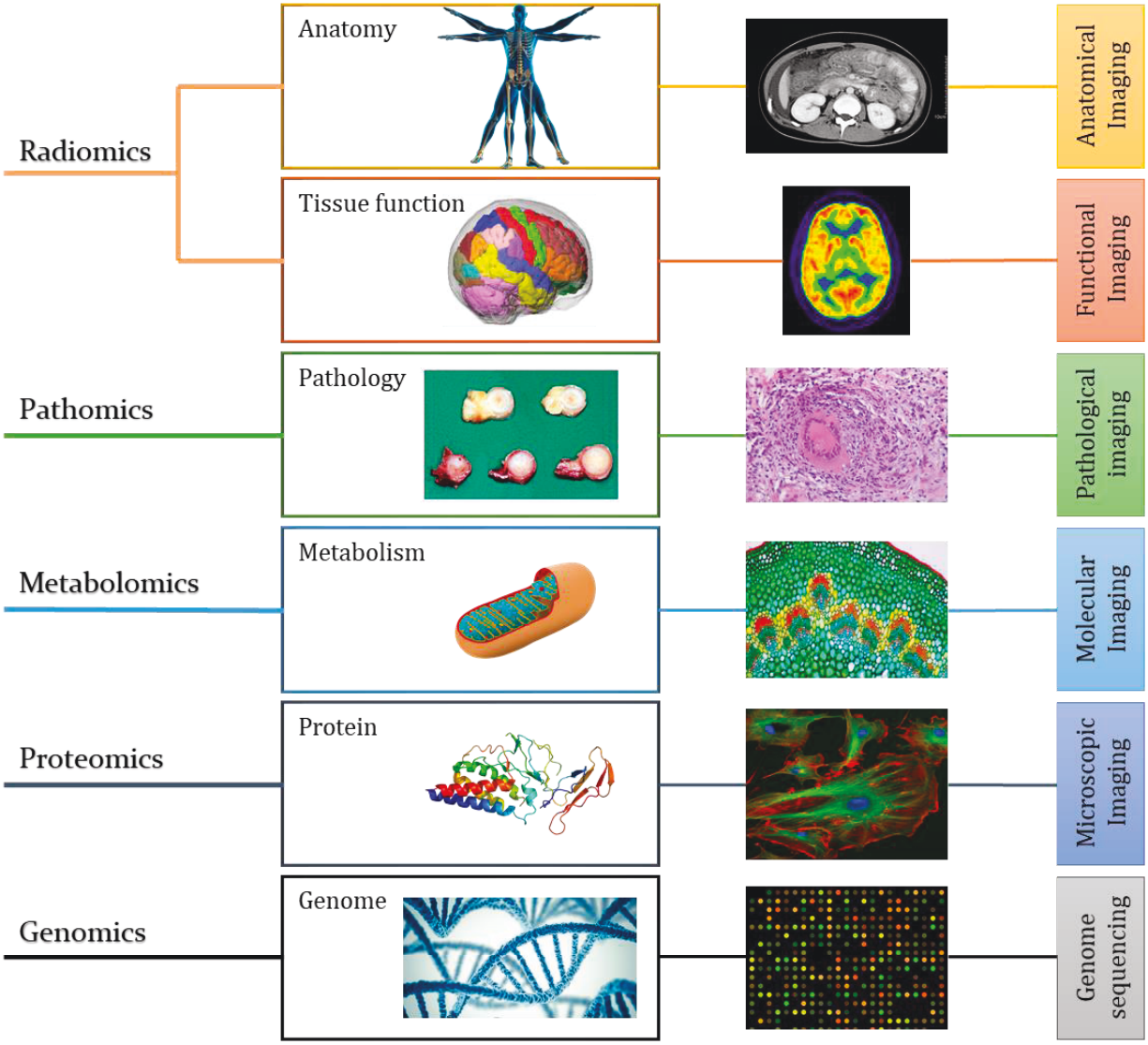

Figure 1-15. Imaging and 'omics' in various levels of biological studies

1.4.1. Personalized medicine in oncology

It is estimated that over 18.1 million new cases and 9.6 million cancer death occurred worldwide in 2018 [232]. Cancer was second to cardiovascular diseases in the annual number of deaths by cause [233]. The trend in global cancer burden is expected to increase due to several factors including population growth, aging of the population, lifestyle adoptions, unhealthy diet, smoking, physical inactivity and reproductive changes [234]. Nonetheless, it is possible to mitigate a substantial amount of cancer cases globally with advancements in cancer control and preventive knowledge, early detection, appropriate treatments and palliative care [235]. With modern diagnostic modalities

becoming more ubiquitous, it is possible to collect, process and analyze large volumes of high-dimensional patient-specific data that eventually propels the progress in personalized medicine in oncology that aims to improve cancer prevention and prognosis.

Personalized medicine in oncology aims to customize cancer care, such that cancer can be detected in very early stages, or the success of preventive and therapeutic interventions maximizes with minimum side-effects. Most research in this area focuses on lower-level personalized medicine using genomics and proteomics approaches to characterize tumors, which is invasive and requires biopsies. Such studies have enabled the identification and validation of many genes that are cancer-drivers that can cause malignancy in a model system [236]. Despite all these successful efforts, the proper realization of personalized medicine in oncology is still limited. One contributing factor is that tumors exhibit diverse spatial and temporal heterogeneity, both within (intra-) and between (inter-) tumors, causing intra- and inter-tumor heterogeneity [237]. Such inter- and intra-tumor heterogeneities are likely to cause phenotypic variations that ultimately affect treatment response and resistance. Moreover, heterogeneity can be observed over the course of cancer progression in different pattern changes, from initiation to metastases, and even relapse after surgery or other therapies [238]. For example, Hatt *et al.* studied PET and CT images of patients with lung cancer and showed that tumors with more non-uniformity, i.e. more heterogeneous tumors, are more resilient to therapy compared to more uniform tumors [239] (Figure 1-16). These spatial and temporal (longitudinal) intra-tumor heterogeneity assessments require multiple biopsies, which adds to the patient's burden due to their invasiveness. As such, this necessitates the development of non-invasive approaches for profiling tumor phenotypes.

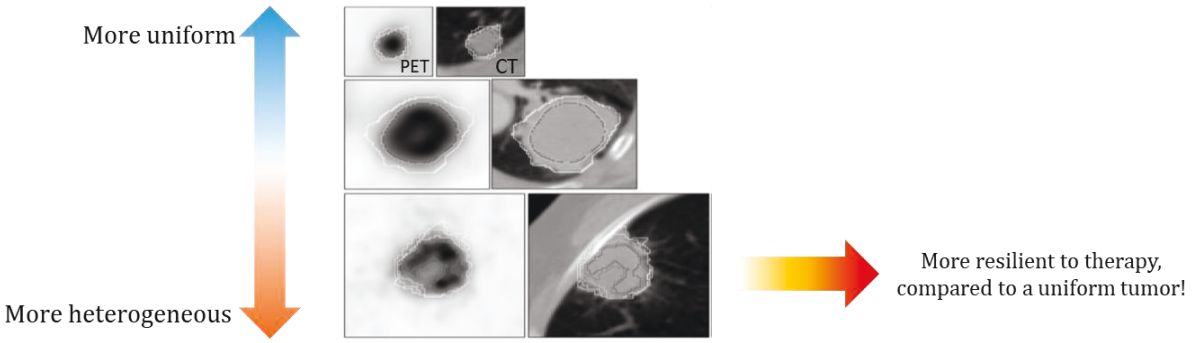

Figure 1-16. Heterogeneous tumors are more resilient to therapy compared to uniform tumors [239].

Medical imaging has unprecedented potential for non-invasive tumor characterization. Today's higher resolution and advanced image acquisition systems can facilitate the extraction of a wide array of tumor phenotype characteristics [228]. So how should we step up from typical visual interpretation of tumors to a more quantitative data-oriented approach that better enables personalized medicine?

In a typical radiology setup, tumor appearances (i.e. imaging features) are typically described visually, and subsequently interpreted by radiologists who then may select suspicious lesions for further clinical investigation such as biopsy [240] (Figure 1-17). However, this process can be enhanced by a combination of both quantification and visual assessment of images. Quantitation methods benefit from image processing techniques that can describe a diverse set of properties that quantify the image from different perspectives. They can describe size and shape, intensity-based textural properties, and other more advanced morphological and textual features with or without additional filtering in a quantitative scheme. These imaging features can then be combined with statistical modeling techniques to find a specific pattern or predict a certain clinical outcome. This field of research is referred to as "Radiomics" as discussed in the next subsection.

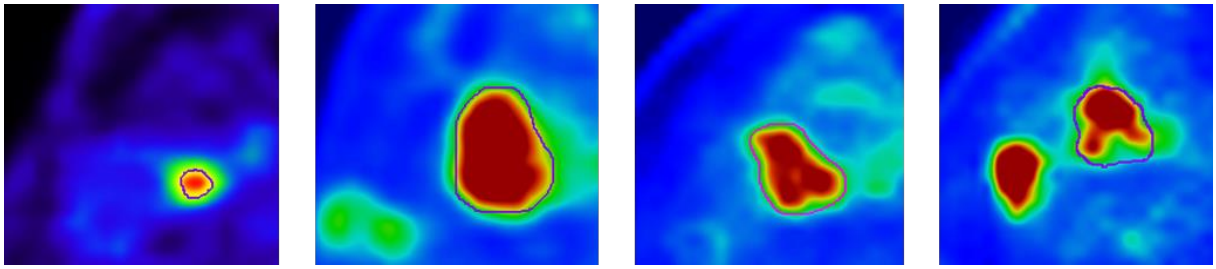

Figure 1-17. PET images of four patients with head and neck cancer with their primary tumors delineated by a nuclear medicine physician (maroon line). Based on conventional quantitative imaging, all these tumors have almost the same SUV_{max}. the right three ROIs have almost the same volume, too. But does it mean that all these four patients have the same diagnosis and need the same therapy, or can we derive more information about the tumor phenotypes from these images? That's where radiomics comes into play.

1.4.2. Radiomics

Radiomics transforms digitally encrypted medical images that contain information regarding tumor pathophysiology into mineable high-dimensional data [228, 241]. In other words, it hypothesizes that different phenotypic characteristics of tumors such as intra- and inter-tumor heterogeneity can be quantified as features called "radiomic features" through advanced image processing and computer vision techniques [228]. The information is harnessed through image processing and quantitative image analyses [242] and can be leveraged via clinical decision support systems to improve decision making and personalized medicine [243]. Radiomics can extract a massive amount of data from medical images of different modalities (CT, MRI, PET, etc.) to uncover advanced underlying features that non-invasively characterize tumor through data analysis. It aims to identify new reproducible and repeatable quantitative image biomarkers for disease staging, predicting tumor response, therapy assessment, and understanding tumor evaluation and its intrinsic biology. Ultimately, radiomics may pave the way for personalized medicine in different areas of clinical practice and provide clinicians with crucial information to guide their clinical decisions.

1.4.3. Applications of Radiomics

The term "radiomics" was introduced in 2011 (though related research was pursued prior to it, the term and field were better crystallized later). Radiomics has witnessed significant interest from

researchers from all around the world in recent years. The number of published articles in PubMed has seen an exponentially increase (Figure 1-18). Several well-written reviews are also published [244-248]. Nowadays radiomics is not only a term used by medical physicist and those at the technical side of the clinical world. Clinicians and clinical research faculty are becoming more familiar with the term. Some radiomic features, especially first-order families of radiomic features, are understandable, communicable and easy to explain, and if properly explained to clinicians and radiologists, they themselves may be able to translate these approaches to read clinical images to assess radiomic dimensions of tumors, similar to the example in Figure 1-16 where more heterogeneous tumors are found to be more resilient to therapy.

More recently with the fast growth of deep learning applications, there is some interest in the comparison of methods utilizing explicit (radiomics) vs. implicit (deep learning) image patterns [249-251]. Such comparisons and efforts are beyond the scope of our present work, which focuses on explicitly defined radiomic features. But generally, deep learning-based methods require larger datasets, which is a limitation. Moreover, unlike radiomic features that may be interpretable, deep learning is like a black-box approach that impacts its translation to the clinic. In this subsection, we provide a brief overview of some major area of radiomics applications.

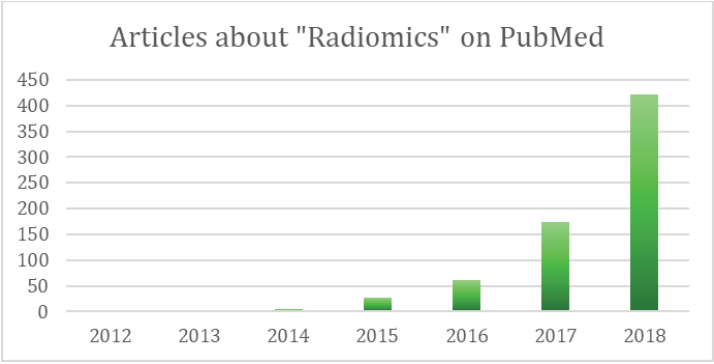

Figure 1-18. The number of published articles about Radiomics shows an increase in interest

1.4.3.1. Oncology

The majority of radiomics research has been performed for oncology. Radiomics is able (to some extent) to predict tumor characteristics such as histology [252] and genetic footprint [253-255], in addition to response to therapy in the form of pathological response from primary tumor [256, 257], lymphadenopathy [258], response to chemotherapy [259], response to chemo-radiotherapy [260], recurrence [261, 262], lymph node involvement [263], distant metastasis [264-266], and survival [267, 268] for a spectrum of pathologies. Many studies have demonstrated strong prognostic powers of radiomics in CT [265, 269-271], MRI [272-275], PET [247, 276, 277], as well as multiparametric imaging [278, 279]. In terms of clinical outcome, some studies have demonstrated the discriminating capability of radiomics for stratification of tumor stages [280], tumor histology [281], and other clinical outcomes [282].

1.4.3.2. Non-oncological applications

Application of radiomics is not limited to oncologic imaging. In neurology, radiomics has been used in Alzheimer's disease [283], [284-287], multiple sclerosis [288, 289], autism [290], as well as some publications from our group in using radiomics for correlating brain DaTSCAN SPECT with motor function in Parkinson's disease patients [291, 292]. We have also developed radiomics analysis of cardiac SPECT imaging [293, 294] which this dissertation especially elaborates. Radiomics has also been recently studied for immunotherapy [295]. It has also been utilized in assessing radiation injury in patients post-radiotherapy [296-298].

1.4.3.3. Combination of radiomics and other "-omics"

Following the emergence of "omics" research and the promising results, researchers have started to look beyond a particular modality and combine "omics" for more accurate outcome analysis and new applications. In light of this interdisciplinary research, radiomics has been mostly studied in combination with genomics and termed "radiogenomics" [230]. It is expected that such combinations

may notably contribute to the realization of personalized medicine [299]. Lung cancer has been a prominent area in radiogenomics research, due to its prevalence, an abundance of imaging data especially from CT, as well as publicly available datasets such as the national lung screening trial [300]. Many groups have studied using radiogenomics for diagnosis, prognosis, and predicting optimal therapy in lung cancer (see review article [248]). We have also participated in radiogenomics research with external collaborators where we used radiogenomics to predict EGFR and KRAS mutation of patients with lung cancer [301, 302].

1.5. Our Motivation and Overview of Efforts

The main motivation behind this dissertation research was to **enhance quantitation in nuclear medicine imaging**, especially in PET and SPECT. Above, I elaborated on the significance of quantitation for improved patient care and towards the realization of personalized medicine. My multiple research studies can be generally classified under two themes (both of which aim to enhance quantitation in nuclear medicine): methods in "**image reconstruction**" and in "**image processing**". In the former theme, I specifically studied the impact of point-spread function (PSF) modeling in PET image reconstruction, and introduced the concept of generalized PSF modeling, demonstrating its superior performance compared to existing PSF modeling techniques. In the latter theme, I focused on the radiomics analysis of nuclear medicine images (both PET and SPECT) in oncologic and cardiologic applications. I developed a standardized radiomic framework for calculating standardized and reproducible quantitative biomarkers (radiomic features), followed by specific clinical applications. I developed various extensions and workflows, and have released two software packages: (i) PET Simulation and Image Reconstruction; and (ii) Standardized Environment for Radiomics Analysis (SERA) that are available for scientific purposes to the imaging community

The work and research in the course of my Ph.D. studies resulted in 4 journal and 29 conference publications [293, 294, 301-331], of which 11 were as first author, as well as 3 journal submissions under review [332-334]. Specifically, we have 1 journal paper [303] and 4 conference works [304-307] on aspects of point-spread function (PSF) modeling in PET image reconstruction, which chapter 2 elaborates. We also have 3 journals papers [308-310] and 25 conference works [293, 294, 301, 302, 310-331],, on image processing and/or radiomics in different settings. These include my standardized radiomics efforts which are elaborated in chapters 3 and 4, including collaboration with the image biomarker standardization initiative (IBSI). Some publications also involved collaborative usage of our radiomics package (SERA) or other tools developed, such as extensions and workflows developed to extract clinical images and quantitative measures via MIM and Matlab software packages.

In what follows, Chapter 2 presents our new approach to PSF modeling—adaptive PSF modeling for enhanced reconstruction. Chapter 3 elaborates concepts behind standardized radiomics workflows, following by our IBSI collaborative efforts, and subsequently, two studies on reproducibility analysis of radiomic features in two distinct nuclear medicine clinical contexts. Chapter 4 contains a comprehensive study of using radiomics in myocardial perfusion SPECT stress tests to predict coronary artery calcifications in patients with a normal scan. Finally, concluding remarks and opinions have been expressed in chapter 5.

Part I: Enhanced Quantitation Using Advanced Image Reconstruction

2. Adaptive Point-Spread Function (PSF) Modeling for Enhanced Quantitation in PET Image Reconstruction

In this chapter, we fist discuss the partial volume effect (PVE) in PET imaging, shedding light on how it impacts images. Subsequently, we discuss methods of partial volume correction (PVC) to compensate for PVE. Specifically, we focus on a well-known and widely-used PVC technique referred to as point-spread function (PSF) modeling. Since PSF modeling has pros and cons, there is a debate as to whether or not enable it during PET reconstruction. We subsequently propose an alternative approach to the dichotomy of using or not to use PSF modeling—namely, an adaptive, contextualized PSF modeling framework for enhanced quantitation in PET. For evaluation, we first model our idea using a simple framework of deblurring a Gaussian-filtered image. Next, after demonstrating promising results from this simplistic model, we move forward to introducing and implementing our adaptive PSF modeling in a realistic PET reconstruction framework, and demonstrate how it outperforms the abovementioned dichotomy modeling in terms of image quantitation. The first part was presented as a conference proceeding in SPIE annual meeting [304], while the main contribution of this project was published in the journal *Physics in Medicine and Biology* [306].

2.1. Partial Volume Effect in PET

PET imaging continues to be affected by several resolution degradation factors, some of which were discussed earlier in section 1.2.4, resulting in image blurring, also referred to as the partial volume

effect (PVE). PVE technically refers to two distinct phenomena that ultimately make the PET voxel values different than what they ideally should be [335]. The first effect is the 3D image blurring caused by the finite spatial resolution of the PET system. This limitation is caused by the extent of the detectors, as well as degradation factors mentioned earlier including positron range (section 1.2.4.4), detector blurring (section 1.2.4.7), photon non-collinearity (section 1.2.4.5), depth of interaction (section 1.2.4.7), as well as patient motion (section 1.2.4.11). These degradations result in blurred images that subsequently cause spillover between regions. Due to this spillover, a region with higher-than-background uptake will appear larger, but dimmer. This effect can be mathematically described by a 3D convolution operation, where the image is generated by convolving the true image with the 3D image response or point-spread function (PSF) of the imaging system.

To illustrate this first effect (the spillover effect) due to PVE, we use a digital simulation of a national electrical manufacturing association (NEMA) NU-2 image quality phantom that is commonly used for image quality calibrations for PET scanners [336]. A phantom is a specially-designed container that includes some internal structures (e.g. rods, spheres, etc.) with known measurements that can be filled with radioactivity, and when imaged with a scanner, can demonstrate how accurately its image is reconstructed. The NEMA NU-2 phantom consists of a large semi-cylinder container that includes six spheres with inner diameters of 10, 13, 17, 22, 28 and 37 mm (Figure 2-1). The container and the spheres are filled with radioactive solutions with known activities and then imaged using a PET scan. The reconstructed image is used to calculate measurements to determine the quality of the scanner and the reconstruction protocol.

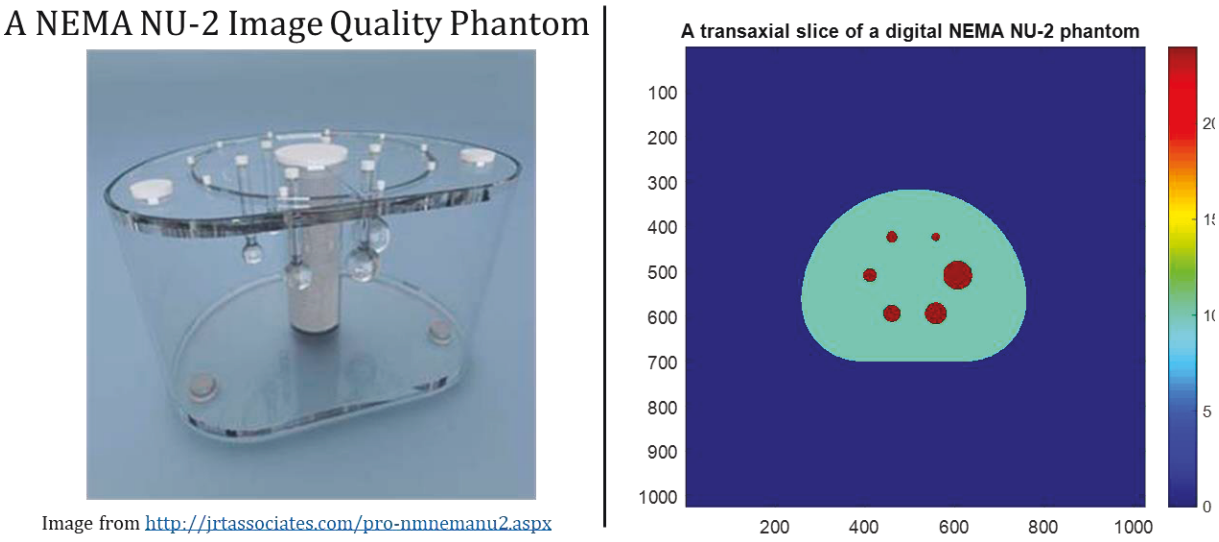

Figure 2-1. NEMA NU-2 image quality phantom. Left: an actual phantom design by a manufacturer. Six spheres inside the container are filled with known radioactivity. Right: a transaxial slice of the digitally-simulated NEMA NU-2 phantom that passes through the center of all the spheres.

Figure 2-2 shows the effect of PVE in action. The left image shows a 3D illustration of a 2D transaxial slice of a digital simulation of a NEMA NU-2 phantom (a 3D illustration of the right figure from Figure 2-1). The right image in Figure 2-2 shows a noise-free reconstruction of the left image given to our in-house developed realistic PET simulation and reconstruction framework, which was forward-projected (no noise added) and then reconstructed using OS-EM algorithm with 2 iterations and 14 subsets. It can be observed from the peak of the six spheres in the right image that they are lower than the actual value in the true image in the right. Also, the edges of the regions are blurred. Moreover, the bottom of the peaks is spread over more voxels compared to the true image. These are all signs of PVE.

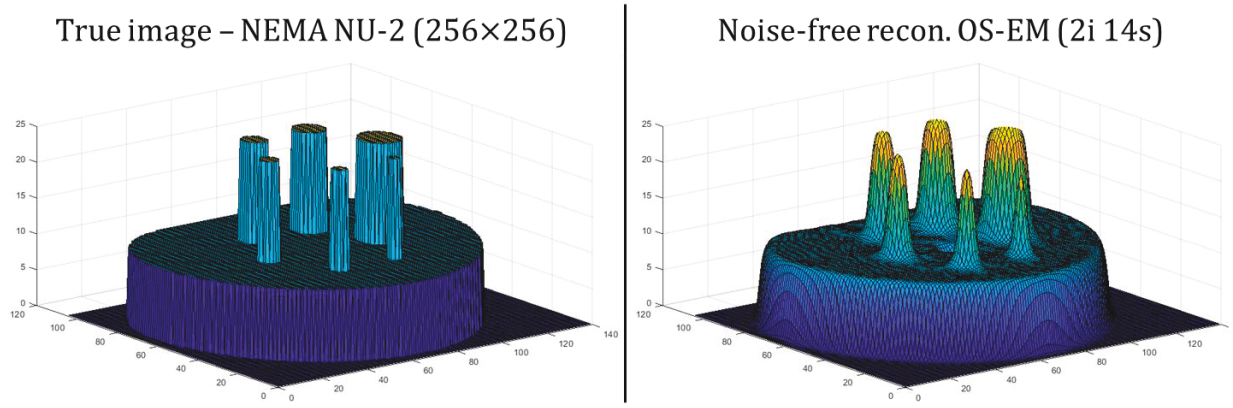

Figure 2-2. A 3D illustration of the partial volume effect (PVE). PVE results in spillover to nearby voxels and blurring. Left: 3D illustration of the true image of a 2D transaxial slice of NEMA NU-2 image quality phantom. Right: 3D illustration of the noise-free reconstruction of the slice on the left with OS-EM algorithm with 2 iterations and 14 subsets.

At the same time, this simulation cannot perfectly illustrate the impact of PVE. The reason is an actual PET scanner images an essentially-continuous object of a human body or phantom, and imaging the object and reconstructing it to an image will result in pixilation of the true object that inherently has a lower resolution. This brings us to the second phenomenon causing PVE, which is due to the finite sampling of the radiotracer uptake on a voxel grid. The boundaries of regions and contours of voxels in a reconstructed image do not match the actual contours of the tracer distribution. As a result, a single voxel may contain radiotracer uptake distribution from multiple tissues.

PVE can severely impact reconstructed images—both qualitatively and quantitatively. For instance, the signal from a small-sized high-uptake (so-called "hot") lesion over a lower uptake (so-called "cold") background is spread out due to PVE. Thus, the maximum value inside the reconstructed hot region will be less than the actual maximum value (Figure 2-2-right). Moreover, PVE can impact the apparent tumor size. This issue becomes more concerning when PET is used to assist radiotherapy planning. In this case, the reconstructed region of a tumor may be larger than the true metabolically-active part of the tumor due to the image degradation phenomena as well as the limited spatial resolution of PET reconstructed images that are around 4 mm. The CT image captured by a PET/CT scan can be helpful in determining the contour more accurately; however, CT images do not necessarily capture the metabolically-active regions of an organ, which is the main reason we use

PET imaging. Therefore, the quest for correcting PVE and producing higher resolution PET images remains highly relevant.

2.1.1. Methods for Partial Volume Effect Correction

The first medical applications of PET imaging were focused on the brain, and earlier PET prototypes were developed specifically to image functionality of the brain [337]. Similarly, the first methods for partial volume correction (PVC) were also developed for neurologic PET procedures to enhance their quantitative capabilities [338, 339]. MRI images were used in PVC strategies relying on an adjunct co-registered structural image, due to the higher gray and white matter contrast compared with CT [340, 341]. Nowadays, the anatomical information provided by the CT component of the hybrid PET/CT scanners (section 1.2.2.3) also may enable PVC in other body organs including cardiovascular [342], atherosclerosis [343], and whole-body oncologic imaging applications [335, 344].

PVC methods can be tackled via reconstruction-based or post-reconstruction-based techniques that are broadly categorized into ROI-based and voxel-based methods [345]. The aim of ROI-based methods is to produce improved estimates of mean ROI uptake [339, 346, 347]. This is often achieved by using anatomical information from MRI images; however, these techniques usually involve an assumption of homogeneous PET uptake distribution in the anatomical regions, and more importantly, do not produce images. Another technique is the use of recovery coefficient, where the uptake in a region is multiplied by a correction factor, which is pre-calculated for an object whose size and shapes are similar to those in the ROI [346]. Voxel-based methods, on the other hand, produce images. Some examples of voxels-based PVC techniques include multiresolution [348-350] and partition-based methods [347], both of which typically include simplifying assumptions. Another example is the use of iterative deconvolution [351], which is often used for image restoration—i.e. to

recover spatial resolution. Deconvolution can severely enhance noise levels, but promising performances can be achieved when utilizing regularization [352] and denoising [353]. Some reconstruction-based PVC methods involve incorporating anatomical information provided from the CT or MRI components in the context of Bayesian maximum *a posteriori* (MAP) PET reconstruction. In this method, anatomical regions are often segmented and used as prior images wherein intervoxel PET intensity variations are penalized while allowing large intervoxel variations across the boundaries. A distinct approach to voxel-based PVC is point-spread function (PSF)-modeling, also referred to as resolution modeling (RM), which we elaborate in the next section.

2.1.2. PSF Modeling in PET

PSF-modeling aims to capture object-domain and/or detection-domain resolution degrading effects and facilitate more accurate modeling of the measurement [354]. As a result, it can reduce image degradations caused by model-mismatch and yield improvements in reconstructed image quality as it compensates for some partial volume effects. As such, PSF-modeling has attracted considerable interest in PET over the past decade [177, 354], and has been adopted by major PET vendors in their state-of-the-art PET scanners [84, 355-359].

2.1.2.1. Types of PSF modeling

PSF-modeling implementations are commonly divided into (i) image-space, and (ii) projection-space methods. Theoretical analyses of these two approaches have been provided [177, 360], and a few studies have performed preliminary comparisons between the two [361, 362].

Imaged-based PSF modeling

Image-based methods attempt to incorporate resolution blurring effects entirely in the image-space, based on the idea that the reconstructed image can be considered as the blurring of the true image by a point-spread function. This can be performed within the image reconstruction (PSF modeling in

the image space component of the system matrix, thus applied before forward projection and after back projection operations), or performed post-reconstruction in the form of iterative deconvolution [351, 354, 357, 360, 361, 363-365]. These methods are straightforward to implement, do not impose a significant computational burden, and produce images with high quality so that certain vendor PET scanners have already provided this option [358].

Projection-based PSF modeling

On the other hand, projection-space methods perform modeling of degradation effects within the projection-space component of the system matrix. Iriarte et al. published a review on system models for statistical reconstruction of PET data; they indicate that the most popular approach of combining models of physical degradation factors is to factor the system matrix as a product of independent matrices, each one describing one or a collection of effects [354]. This arrangement is a well-established approach that has led to high-quality efficient reconstructions and yields substantial enhancements in storage preserving and computational time. Also, some studies on image-based PSF modeling considered anisotropy [357, 362, 364, 365], using more complex functions, e.g. mixtures of Gaussians and exponential [360, 362, 365], and PSF estimation methods based on measurements of or arrays of point sources [361, 362, 366].

The projection-based methods can be categorized into: (a) empirical methods utilizing measured data points [139, 367], (b) Monte Carlo simulations [159, 354, 368-370], (c) analytical models [172, 371-374] including additional modeling for positron range [174, 375-378], and (d) hybrid approaches incorporating combinations of these methodologies; e.g. starting with a simple geometrical calculation, and then imposing additional effects [373, 379-381]. These studies focus on the benefits of different methodologies and exploit their synergies to compute the system matrix [354]. In this study, we focus on the analytical models and will elaborate more in section 2.3.2.

2.1.2.2. Pros and Cons of PSF modeling

Advantages of PSF modeling

PSF-modeling provides a number of advantages: (i) improved spatial resolution and contrast recovery [357, 361-363, 377, 382]; (ii) reduced spatial noise or image roughness (IR), resulting in a visually smoother image [383], and (iii) improved focal lesion detectability performance [384-386].

Disadvantages of PSF modeling

At the same time, PSF-modeling poses two concerns [387]: First, it impacts noise characteristics of the reconstructed images as they appear smoother. Moreover, the resulting noise power spectrum (NPS) of the PSF modeled reconstructed image is seen to be amplified in the mid-frequency domain while exhibiting smaller values at higher frequencies [388]. Some efforts have been devoted to the analysis of the resulting noise properties that have important implications for quantitation and lesion detectability performance in PET imaging. These studies performed an experimental evaluation of noise characteristics on real data sets [388-391], or through Monte Carlo simulations [392, 393], and subsequently analyzed the impact of reconstruction parameters by adopting a variety of figures-of-merit (FOMs). PSF-modeling has different effects on different noise metrics [384, 388]. Rahmim *et al.* used analytic models of noise propagation [394, 395] to investigate the impact with and without resolution modeling. In a PSF modeled system matrix, more lines-of-response (LORs) contribute to a single voxel, as each voxel is related to more measurement locations compared to a non-PSF modeled system. This results in a more ill-conditioned inverse problem that suffers from slow convergence [396]. Moreover, at matched iterations, voxels in a PSF reconstruction depict lower voxel variance and higher inter-voxel correlations vs. no-PSF [384, 397]. As a result, PSF modeling noticeably alters the noise texture. Tong *et al.* derived analytical expressions relating image roughness and ensemble noise to voxel variance and inter-voxel correlations [384]. Due to PSF modeling, images become smoother, but the ensemble standard deviation of ROI mean uptake (a measure of reproducibility)

may remain unchanged [384] or even be amplified for smaller ROIs [398, 399]. We elaborate more on how this phenomenon impacts different noise metrics in section 2.

The second issue concerning PSF-modeling is its susceptibility to produce edge overshoot effect – a reminiscence of Gibbs ringing patterns at the edges of a region [396, 400, 401], manifesting as overshoots in smaller regions. This issue may compromise the accuracy of signal quantitation in small regions [177]. Snyder suggested using a less blurred (i.e. underestimated) version of the true PSF in the reconstruction [400, 401]. This method was shown to be effective at suppressing the edge overshoot effect [396, 399]. Snyder [400] also suggested that a possible reason for the appearance of edge overshoots is the mismatch between estimated PSF and true PSF and that the small mismatch can be amplified due to the instability of deconvolution process. Nonetheless, it was shown [307, 363, 396] – as we also demonstrate in this work – that even reconstruction with the true PSF results in the edge overshoot effect. Furthermore, for specific detection or quantitation tasks, it is plausible that such an effect may even enhance task performance, as we show in this work for quantitation.

2.1.3. Motivation

Clinical tasks vary from pure detection-related tasks (e.g. diagnosis and staging) to quantitation-related tasks (e.g. therapy response assessment and prognostication). The features that make a PET image suitable for detection task differ from those that make an image effective for tumor quantitation. In fact, noise contributes differently to these two general tasks, and PSF-induced noise propagation can result in improved quantitation but may reduce lesion detectability performance, or vice versa [388]. However, such analyses and optimizations have only been performed for the dichotomy case of no-PSF vs. PSF modeling. **It is possible to generalize PSF modeling to include overestimated and underestimated PSF kernels. This provides a wider range of options to study the impact of PSF modeling on image quantitation tasks**, thus facilitating task-based

optimization for quantitation. Furthermore, given the challenges with reproducibility as well as edge overshoots in PSF modeling, **it is plausible that generalized PSF modeling may provide adaptive kernels that perform optimally, properly balancing different effects**. The present work pursues such a generalized adaptive PSF framework in the context of tumor quantitation, which in future efforts can be thoroughly evaluated for lesion detectability tasks as well.

In the following sections, first, we present the adaptive PSF modeling idea utilizing a simplistic image blurring-deblurring framework that comprises an image-space deconvolution-based PSF modeling. We used Gaussian blurring with a predetermined filter size to blur the image and then deblurred that image with a range of filter sizes to assess the plausibility of the idea that generalized PSF modeling enhances image quantitation. After showing some encouraging preliminary results, we present our practical implementation of this idea using an in-house developed realistic PET scanner simulation and reconstruction framework. We adopted an "analytically-modeled" "projection-based" PSF modeling (section 2.1.2.1). It is important to show the feasibility of these findings in a practical setup, therefore, in this step, we intended to design our study as close to a real-world scenario as possible. We used analytically modeled PET degradation phenomena and ultimately used them in our simulation and reconstruction framework that models PET scanner with its various degradation phenomena. We also used a digital anthropomorphic phantom using realistic uptake values by modeling tracer kinetic uptake using patient-derived kinetic modeling parameters. We simulated a wide variety of scenarios for different tumor sizes with and without noise and introduced a wide set of figures of merit to assess the performance of adaptive PSF modeling from different perspectives. As mentioned earlier in this chapter, what appears in section 2.2 is published as a conference proceeding in [304], and the subsequent section 2.3 is published in the journal of physics and medicine and biology [306].

2.2. Adaptive Image-Based PSF modeling

2.2.1. Introduction

In this section we propose a generalized PSF modeling framework, including extensive task-based optimization, wherein we continualize the conventionally discrete framework of PSF modeling vs. no PSF modeling, to include varying degrees of PSF modeling. The proposed framework has the advantage of providing a trade-off between the enhanced contrast recovery by PSF modeling and the reduced inter-voxel correlations in the absence of PSF modeling, and to enable improved task performance.

2.2.1.1. Aims of the study

The concerns regarding PSF modeling in PET discussed earlier have prompted us to propose an extensive task-based assessment of a generalized PSF modeling framework, wherein we continualize the conventionally discrete framework of PSF modeling vs. no PSF modeling, to include varying degrees of PSF modeling, including overestimation and underestimation of the 'true PSF'. Such a generalized scheme allows consideration of a much wider array of images, which are subsequently analyzed in the context of different imaging tasks. We elaborate upon these next.

2.2.2. Methods

2.2.2.1. Generalized image-based PSF modeling

The investigated context was that of oncologic FDG PET imaging. In this initial assessment, we focused on lung tumor imaging, with SUV images at 45min, simulated based on kinetic parameters extracted from the literature ([402], Table II). We used a single slice of an XCAT anthropomorphic

phantom (Figure 2-3) [403] and placed a squared-shape lung tumor in the center of the right lung (Figure 2-4).

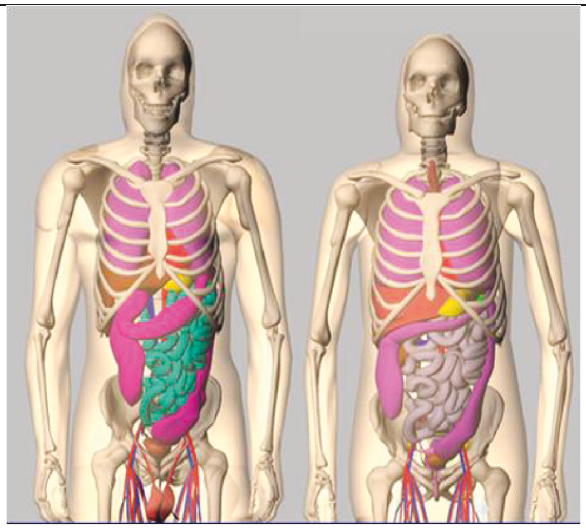

Figure 2-3. XCAT digital anthropomorphic phantom, capable of realistically modeling the human body and widely used in imaging research.

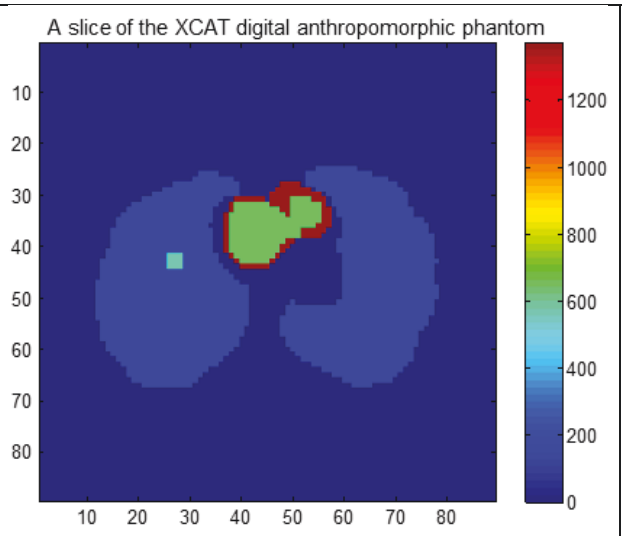

Figure 2-4. A transaxial slice of the XCAT phantom with a simulated lung tumor. The slice includes regions from lung, myocardium and blood pool.

Blurring was simulated in the image-space domain (being generalized in subsequent work to incorporate projection-space blurring), using a Gaussian blurring kernel with FWHM h, which denotes the "true PSF" (Figure 2-5). We subsequently add Poisson noise to the blurred image to quantify the impact of correlated noise as induced by a PET scanner (Figure 2-6).

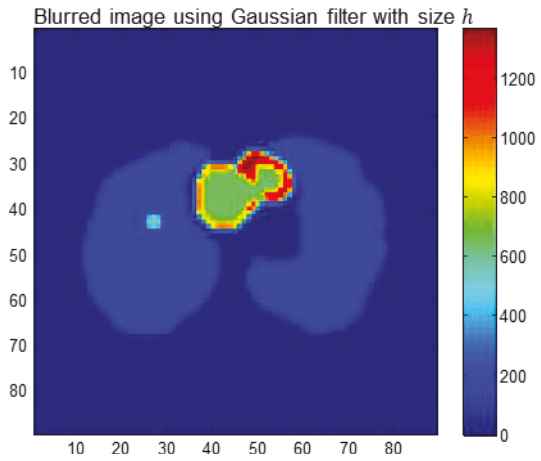

Figure 2-5. Blurring the true image using a Gaussian filter with size h, which models the image degradation.

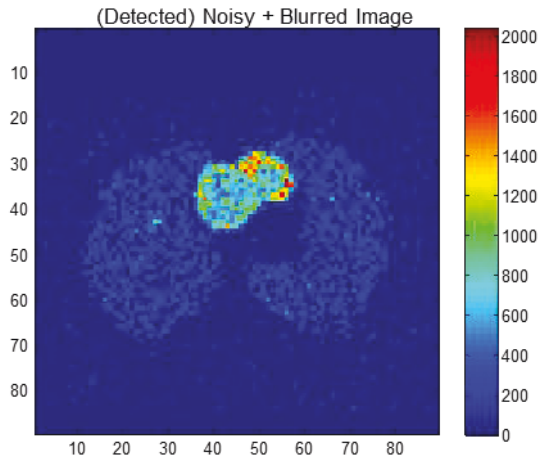

Figure 2-6. Poisson noise added on top of the already-blurred image to model the effect of noise.

At this step, performed iterative EM including PSF modeling using a wide range of generalized PSF kernels with FWHM denoted by $\hat{h}$. In our simulations, we set $h = 6$ mm (true PSF kernel), and varied the reconstruction PSF kernel size $\hat{h}$ from 0 (no PSF modeling) to 12 mm, thus performing both underestimation and overestimation of the true PSF. An example of a reconstructed image with an overestimated PSF kernel (Gaussian filter) size $\hat{h} = 9$ mm is shown in Figure 2-7. We then focus on quantification task assessment and seek to assess and optimize the performance of the generalized PSF modeling framework as we discuss next.

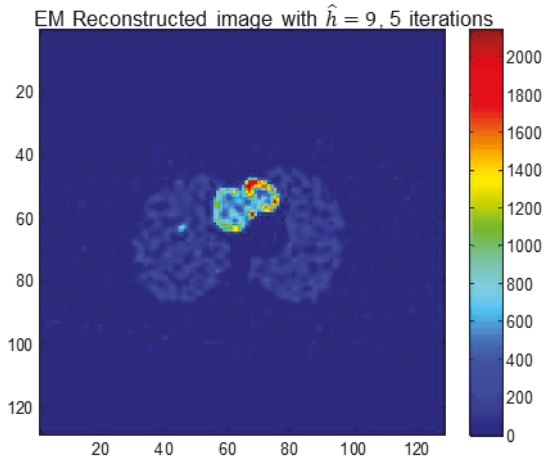

Figure 2-7. An example of a reconstructed image with $\hat{h} = 9$ and iteration 5.

2.2.2.2. Quantification task performance

Analysis of noise-bias pattern is a popular method to assess quantitative performance. We generated noise vs. bias trade-off curves, as generated with increasing iterations into the various generalized PSF modeling algorithms. We also performed convergence analysis, in which an algorithm was quantified as converged when the last 10 iterations altered the bias no more than a certain threshold. We then computed the coefficient-of-variability (COV), $\sigma_{ensemble}$ defined below, compare to the mean uptake. We further include the contrast vs. noise trade-off. These two noise performance metrics were specifically defined as:

i) Spatial variance ${\sigma^2}_{spatial}$, calculated for an image at a given noise realization (in the case of multiple noise realization measurements, this expression is subsequently averaged) via the following:

$$\sigma^2_{spatial} = \frac{1}{N-1}\sum_{i=1}^{N}(s_i - m)^2 \qquad 2.1$$

where s_i denotes the image values at any voxel i within a given ROI (e.g. tumor) consisting of N voxels and having a mean m. In the case of multiple noise realization measurements, this expression is subsequently averaged.

ii) Ensemble variance of ROI mean uptake m_r across multiple noise realizations $r = 1, \dots, R$, with the average ROI mean uptake $\bar{m}$:

$$\sigma^2_{ensemble} = \frac{1}{R-1}\sum_{r=1}^{R}(m_r - \bar{m})^2 \qquad 2.2$$

2.2.3. Results

Figure 2-8 and Figure 2-9 depict SUV_{mean} COV and SUV_{max} COV vs. SUV_{mean} bias; respectively. In these two figures, neither of no PSF modeling $\hat{h} = 0$ nor full PSF modeling ($\hat{h} = h = 6$) achieve the best noise-bias trade-off; yet, Figure 2-8 demonstrates the best performance achieved through $\hat{h} = 4$. The higher SUV_{max} COV in Figure 2-9 compared to SUV_{mean} COV in Figure 2-8 is due to the greater Gibbs ringing artifacts for higher $\hat{h}$. Figure 2-8 also shows that besides the underestimated filter $\hat{h} = 4$ mm, slightly-overestimated kernels such as $\hat{h} = 8$ mm also demonstrates superior performance to both no PSF and full PSF modeling. We also observe from Figure 2-9 that slight overestimation results in less SUV_{mean} bias and less SUV_{mean} noise at matched iterations compared to both no PSF and full PSF.

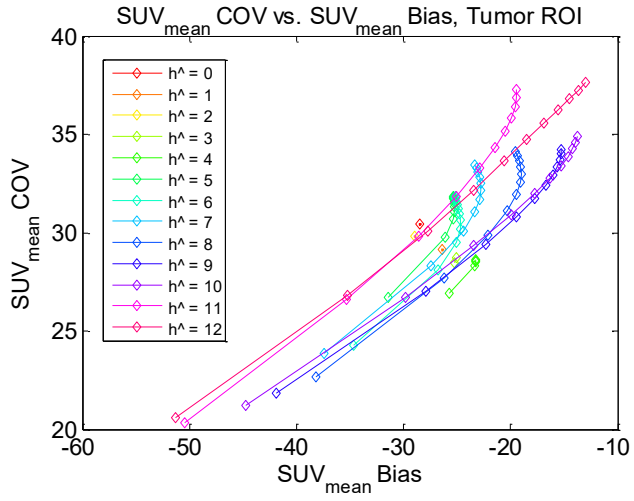

Figure 2-8. SUV_{mean} COV vs. SUV_{mean} bias trade-off for iterations 1-20

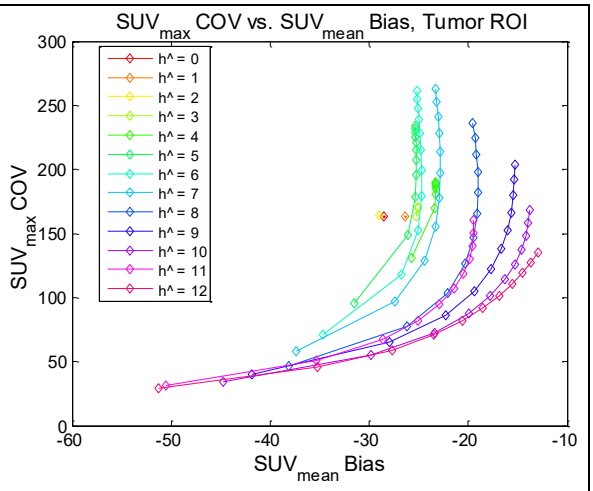

Figure 2-9. SUV_{max} COV vs. SUV_{mean} bias trade-off for iterations 1-20

Figure 2-11 contains noise-contrast trade-off plots and shows higher contrast for higher $\hat{h}$ because of greater overshoot of the edges. This effect can be observed more in detail in Figure 2-11, where the CRC vs. iterations is plotted. This plot shows contrast converges faster for smaller filter sizes, without a significant increase in its value. On the other hand, overestimating filters with higher $\hat{h}$ require more iterations to converge and their CRC value converges to a value higher than 1. The

reason for this unusual CRC of over 100% is the presence of the edge artifact, and the fact that edges become so pronounced that they grow further than the truth. It is interesting to see from this image that PSF modeling with the overestimated size of $\hat{h} \geq 8$ do not seem to have converged during the first 10 iterations plotted in this figure and need more iteration to reach convergence, which yields an even higher CRC at their convergence. However, if we can stop the EM algorithm during earlier iterations of reconstructing with slightly-overestimated kernels, it is possible that we can actually a CRC value closer to 100%—an achievement that neither true PSF nor no PSF modeling were able to achieve, which seems very interesting and promising to be applied to the realistic PSF modeling scheme.

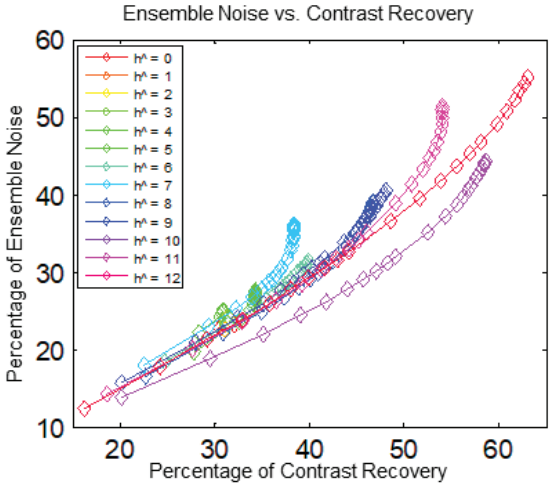

Figure 2-10. Ensemble noise percent of added noise vs. contrast percent trade-off for iterations 1~100

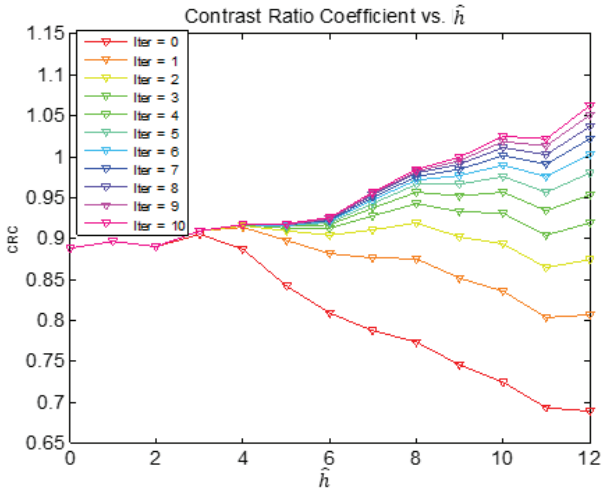

Figure 2-11. CRC trend for fixed iterations for different reconstruction filters. Filters with higher spread converge later than smaller filters.

The converged ensemble noise values based on the criteria described in section 2.2.2.2 is shown in Table 2-1. In this table, we see that $\hat{h} = 4$ mm results in a relatively-low increase in COV with respect to no PSF modeling, while the impact is considerably higher for $\hat{h} = 6$ mm and $\hat{h} = 7$ mm. Overall comparison of the above results reveals that underestimation of the true PSF can result in enhanced quantitation capabilities, while considerably lowering COV compared to full PSF modeling.

Table 2-1. Ensemble Noise Values at convergence.

$\hat{h}$ FWHM	0	1	2	3	4	5	6	7	8
COV	3.44	3.42	3.47	3.54	3.58	3.93	4.12	4.75	5.01
%change COV	0	-0.74	0.14	2.77	4.08	14.23	19.58	37.73	45.36

2.2.4. Discussion

As the first step towards assessing generalized PSF modeling scheme, we proposed a generalized PSF modeling framework with extensive task-based optimization to continualize the conventionally-dichotomized framework of PSF modeling vs. no PSF modeling and applied various degrees of PSF modeling for reconstruction. We showed how this framework provides a trade-off between the enhanced contrast recovery and the reduced inter-voxel correlations for no PSF modeling and improves task performance. We assumed a blurring kernel with FWHM of h and performed iterative EM including PSF modeling with varying widths $\hat{h}$, to demonstrate underestimated and overestimated resolution blurring kernels $\hat{h}$ enhance task performance in terms of lower SUV_{mean} noise and bias. Slightly-overestimated kernels showed the possibility to reach a contrast recovery of 100%. Overall, the results reveal that generalized PSF can result in enhanced quantitation capabilities, while lowering COV compared to full PSF modeling, thus providing an attractive solution for both diagnostic and treatment response monitoring applications.

With this promising preliminary result, we now move on to implement generalized adaptive PSF modeling using a more realistic scheme.

2.3. Adaptive projection-based PSF modeling

2.3.1. Introduction

Following the promising preliminary study of generalized PSF modeling for a simplistic image-based blur-deblur scheme, in this section, we present the study of adaptive generalized PSF modeling in projection-space.

Frameworks explored in the past involve a dichotomy of PSF vs. no-PSF modeling. By contrast, the present work focuses on quantitative performance evaluation of standard uptake value (SUV) PET images, while incorporating a wide spectrum of PSF models, including those that under- and over-estimate the true PSF, for the potential of enhanced quantitation of standardized uptake values (SUVs). The developed framework first analytically models the true PSF, considering a range of resolution degradation phenomena (including photon non-collinearity, inter-crystal penetration, and scattering) as present in data acquisitions with modern commercial PET systems.

In the following subsections, first, we elaborate on analytically modeling on the PET system and various image-degradation phenomena. Then, in the methods section, we describe our realistic PET simulation and reconstruction framework. We subsequently present our results and discussion.

2.3.2. Modeling a PET system in analytical PSF modeling

In this section, we explain a systematic approach to modeling PET imaging and reconstruction, following by analytical modeling of some of the PET image degradation phenomena. These models were subsequently be used to realistically model PET forward projection and reconstruction and then were implemented in our in-house developed PET simulation and reconstruction software. These models were later used to generate a spectrum of PSF kernels to assess our proposed generalized PSF modeling scheme.

Suppose $p(i,j)$ is the element of the detection probability matrix $P \in \mathbb{R}^{M\times N}$ that represents the probability of detecting an emission from pixel j $(j = 1, \dots, N)$, at detector pair i $(i = 1, \dots, M)$. Currently factorized schemes for the system matrix are based on the proposed works of Mumcuoglu *et al.* [159] and Qi *et al.* [368]for 2D and 3D acquisitions respectively:

$$P = P_{det.sens}P_{det.blur}P_{atten}P_{geom}P_{im.blur} \quad 2.3$$

$P_{im.blur}$ accounts for image-based blurring effects, particularly the positron range. A detailed discussion on analytically modeling positron range effects in statistical image reconstruction can be found in [368]. As the current work focuses on [^{18}F]-FDG scanning known to exhibit short mean positron range (0.64 mm), its effect can be safely ignored in the PSF model.

$P_{geom}$is the geometric projection matrix where the $(i.j)$ element defines the probability that a photon pair produced in voxel j reaches the front faces of the LOR i in the absence of attenuation and assuming perfect photon-pair collinearity. We used the built-in Radon transform command in Matlab® to perform the geometric projection. The number of projection bins of this function is sufficient to compute the projection at unit intervals, even along the diagonal.

$P_{det.blur}$ accounts for blurring in the sinogram space and includes photon pair non-collinearity, inter-crystal scattering, and crystal penetration [53]. Technically, modeling radial, angular and inter-sinogram blurring requires a 3D blurring scheme. However, in the present work, we assume a small axial acceptance angle, and these blurring effects are confined to a single sinogram using a 2D blurring model. Each of these effects can be analytically modelled and eventually combined. We briefly discuss modeling of resolution degrading effects in sinogram space and later we show how to exploit these analytical expressions to create adaptive generalized PSF kernels.

Photon non-collinearity has to be technically modeled in the geometric component P_{geom} in equation 1. However, as an approximation to considerably simplify the system matrix computation, one can

assume photon non-collinearity is depth independent and model it in the projection-space component $P_{det.blur}$ of the system matrix [159, 388]. The FWHM of the Gaussian approximation that models the effect of blurring on the spatial resolution using simple geometric calculations can be given as:

$$\text{FWHM} \cong \left(0.25 \times \frac{\pi}{180}\right)\frac{L}{2} = 0.0022 \times L \qquad 2.4$$

where L is the detector separation and can be related to the scanner diameter D for different angles of incidence θ via $L = 2D\cos(\theta)$. Therefore, equation (2) becomes:

$$\text{FWHM} \cong 0.0022 \times 2D\cos(\theta) \qquad 2.5$$

This results in an angular dependent Gaussian blurring kernel $D^{\theta}_{non-col}(x, z)$, which models blurring due to photon non-collinearity along the (x, z) in radial and axial directions of the sinogram, respectively.

Although the remaining two effects – inter-crystal scattering and penetration – are often not distinguished from one another, it would be very beneficial to conceptually separate them for proper modeling: inter-crystal penetration occurs when a photon penetrates the incident detector element and is detected in the adjacent crystal; whereas inter-crystal scattering of photons can occur even when the angle of incidence is 90 degrees.

The penetration effect can be modeled using our knowledge of the 511 keV attenuation coefficient of crystals, μ, as well as the angle of incidence θ. If we model the individual detectors' penetration by $p^{\theta}(x)$, then we can calculate the resulting penetration distribution $D^{\theta}_{penet}(x)$ for the coincident pair as [371, 372] :

$$D^{\theta}_{penet} = \int_{x} p^{\theta}(X)p^{\theta}(2x - X)dX \qquad 2.6$$

where the 1D attenuation distribution $p^{\theta}(x)$ can be described by the angular-dependent exponential function:

$$p^{\theta}(x) = e^{-\mu x'/\sin(\theta)} \tag{2.7}$$

where x' denotes the tangential direction along the detector surface, such that $x = x'\cos(\theta)$.

For the scattering component, which can be effectively decoupled from the penetration effect [172], we used real measured projection data of a point source at the center of the field-of-view (FOV) and determined the average radial and axial scatter blurring in the sinogram space. After correcting for the non-collinearity effect through subtraction in squares, we arrive at an estimate for $D_{scatter}(x,z)$. An important point is that, since the non-collinearity effect is independent of crystal blurring, it can be convolved with the corresponding crystal blurring kernels once they are created, as we briefly discuss next. Consequently, combining the above analytical models, we arrive at the overall projection-space blurring kernel $D^{\theta}_{proj}(x,z)$ through the following convolution expression:

$$D^{\theta}_{proj}(x,z) = D^{\theta}_{non-col}(x,z) * D^{\theta}_{penet}(x) * D_{scatter}(x,z) \tag{2.8}$$

P_{atten} in Eq. 1 is a diagonal matrix containing the attenuation coefficients. The attenuation image in current work was derived by forward projecting the attenuation map created by the XCAT anthropomorphic phantom for a typical 80kVp CT scan. The sinogram was then corrected for 511 keV γ-rays using a three-step correction method proposed by Abella in [372]. Finally, $P_{det.sens}$ is also a diagonal matrix that contains the detector efficiencies for normalization. Here, we obtained a 3D normalization sinogram of the GE scanner and incorporated it into our model.

2.3.3. Methods

In this section, we describe our proposed approach to incorporate and assess generalized PSF kernels in the image reconstruction framework. First, we describe a simulation configuration, followed by the image reconstruction method incorporating the true PSF kernel. We then explain the methodology of generating a spectrum of PSF kernels from the true PSF. Finally, we define the FOMs for assessing and analyzing the results.

2.3.3.1. Simulation and phantom configuration

We used the 4D anthropomorphic XCAT phantom [403] to generate dynamic FDG-PET images of different tumors, as well as the corresponding attenuation map. In this study, we chose to implement six liver tumors of different diameters (10, 13, 17, 22, 28 and 37 mm), which was in agreement with the NEMA NU-2 image quality phantom [336]. The position of the center of the tumor spheres is fixed across all six images for consistency. Figure 1 depicts two of these six reference images. The original reference image has a transaxial dimension of 1024×1024 and a voxel size of 0.5856 mm, and 2D OS-EM reconstructions with seven subsets are performed into 256×256 images with voxel dimensions 3.47 × 3.47 × 3.27 mm^3. Starting from a higher resolution image is more realistic to better capture the spatial continuity of the actual object (patient) being scanned, albeit its contribution to the reconstruction time.

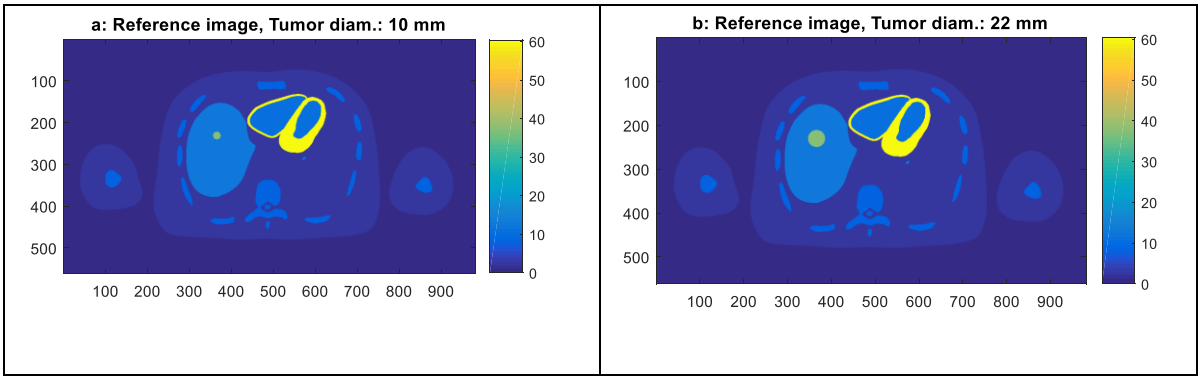

Figure 2-12. XCAT generated phantom as reference images with liver tumor sizes of (left) 10mm, and (right) 22mm.

We simulated a set of 60-min post-injection SUV PET images for a scan duration of 3 minutes. The FDG tracer kinetics were modeled based on a patient-derived input function [404], a set of realistic kinetic parameters reported in the literature (table 1), and the standard two-tissue compartment kinetic model for FDG. Thus, a respective set of time-activity curves (TACs) were generated for each tissue and tumor to allow calculation of the activity concentration levels at 60-min post-injection. Lesion spheres were also modeled based on rate parameters in the liver region. Additionally, based on an evaluation of multiple-patient [^{18}F]-FDG PET scans, we found out that the activity concentration values of the soft tissue background outside the liver were about 21% of the corresponding value in the liver. We set the liver rate constants to derive the background activity TAC accordingly. The dynamic acquisition protocol consisted of 9 passes (from 30 to 90 minutes, 45 seconds for each bed position). The uptake activities were then calculated by temporal integration for the duration of the scan.

2.3.3.2. Image reconstruction

We performed the simulations were performed using an in-house validated reconstruction software [388]. First, noise-free emission images were generated by assigning the modeled values to the respective regions of the voxelized XCAT human torso digital phantom. Then, forward projection of the emission images was performed [405, 406] based on the geometry of a GE Discovery RX PET/CT [407]. The generated sinograms were subsequently attenuated according to the XCAT attenuation factors and scaled based on the reported sensitivity of the scanner (normalization).

Our reconstruction software performs an analytic OS-EM projection-space based PSF-modeling reconstruction and models positron range, geometric projection, photon non-collinearity, inter-crystal scattering, crystal penetration, and corrects for attenuation and detector deficiencies. A detailed modeling of our analytic reconstruction is provided in Appendix A.

Analytic simulations were performed for the images reconstructed using 367 radial bins (60 cm radial field of view) and 581 angular samples covering 180° with. Poisson noise was subsequently simulated to generate 200 independent noise-realizations. Finally, the generated PET projection data were reconstructed, using the proposed methods to produce PET images, as described in the next subsection.

Artifacts in reconstructed PET images are location specific; so ideally the location of the masked ROI has to be the same for both the tumor and the background to perform more precise quantitative analysis. Therefore, we ran our simulation once with tumors present (each of the six tumors) and once with the tumor absent, and then use the mask from the tumor-present reconstructed image to mask out the background region in the tumor-absent reconstructed image. To add more accuracy to our quantitative analysis, we also assure that for every PSF-kernel and every iteration of the noise-free reconstructed images, the ROI location in the tumor-absent matches the actual location in the tumor-present.

Table 2-2. Kinetic parameters used in the simulation of the anthropomorphic phantom for [^{18}F]-FDG tracer. References: myocardium and normal lung [408], normal liver [409], liver tumor [409] and bone [410].

Tissue Compartment	K_1 (mL min^{-1}g^{-1})	k_2 (min^{-1})	k_3 (min^{-1})	k_4 (min^{-1})	V_b
Lung	0.301	0.864	0.097	0.001	0.168
Liver	0.864	0.981	0.005	0.016	0.00
Bone	0.091	0.469	0.0023	0.067	0.00
Myocardium	0.6	1.2	0.1	0.001	0.00
Liver tumor	0.243	0.78	0.1	0.00	0.00
Background activity	0.183	0.981	0.005	0.016	0.00

2.3.3.3. Generalized PSF-modeling

In this section, we describe how to generate a spectrum of PSF (generalized PSF) kernels from the true PSF kernel. We propose an analytical approach to generate a wide spectrum of PSF kernels that portray both *underestimations* and *overestimations* of the true PSF, in addition to no-PSF and true

PSF. The “no-PSF” kernel assumes the incoming rays are solely detected at their incident detector, whereas the true PSF kernel matches exactly with the forward-projector based on scanner parameters and mathematical models of blurring as we explain below.

In order to implement image reconstruction via a spectrum of PSF kernels that has a smooth transition from the no-PSF kernel (identity matrix) to the analytically-modeled “true PSF” and beyond, we observed that we have to simultaneously vary the outputs of three equations that model photon non-collinearity, inter-crystal scattering, and penetration. We constructed a series of generalized PSF kernels that included under- and overestimation of the true PSF by applying a line-space of incremental scaling factors to these three modeled terms. More specifically, three sequences of numbers are multiplied by (i) the mass attenuation coefficient for the crystal (LYSO in this case) in Eq. 2.7 that models inter-crystal penetration, and (ii, iii) the FWHMs of $D_{non-col}^{\theta}$ (Eq. 2.6) and $D_{scatter}$ that model non-collinearity and inter-crystal scattering, respectively.

Table 2-3 contains the three scaling factors multiplied by the attenuation coefficient of crystals and FWHM of non-collinearity and crystal scattering used in generating 20 generalized PSF modeled kernels presented in this study. The column on the right contains the factor we multiplied by the attenuation coefficient of LYSO crystal that is 0.087. The second to the right column includes the factor we used to rescale the FWHM of photon non-collinearity in Eq. 2.5. The third column contains the FWHM of inter-crystal scattering effect being used in Eq. 2.6.

Table 2-3. List of scaling factors used to generate 20 PSF modeled kernel

	Kernel index	FWHM of inter-crystal scattering	Factors multiplied by the FWHM of photon non-collinearity	Factors multiplied by the attenuation coefficient of crystals
No-PSF	1	0.04	0	111.1
Underestimated PSF	2	0.392	0.091	44.534
	3	0.785	0.182	18.746
	4	1.438	0.409	6.983
	5	1.962	0.545	2.826
	6	2.354	0.682	1.319
	7	2.616	0.773	1.240
	8	2.877	0.818	1.156
	9	3.139	0.909	1.080
True PSF	10	3.27	1	1
Overestimated PSF	11	3.662	1.091	0.919
	12	3.924	1.182	0.840
	13	4.185	1.227	0.76
	14	4.447	1.318	0.68
	15	4.709	1.409	0.6
	16	4.970	1.5	0.519
	17	5.232	1.545	0.440
	18	5.493	1.636	0.360
	19	5.755	1.727	0.280
	20	5.886	1.773	0.240

We defined and used multiple metrics for quantitative task-performance analysis. This included four types of noise (image roughness (IR), SUV_{mean} coefficient of variation (CoV), SUV_{max} CoV, and average max-min difference) and two types of bias (SUV_{mean} bias and SUV_{max} bias). In addition, mean-squared error (MSE) of each voxel and MSE of SUV_{mean} were computed. Definitions of these metrics are provided in detail in the next section.

2.3.3.4. Signal and noise figures of merits for quantitative analysis

We used several figures of merit (FOM) to quantify the performance of generalized PSF modeling kernels in our study. Below we define these metrics and elaborate on why each FOM is required for studying different aspects of image quantitation.

We denote the total number of independent noise realizations by R, the total number of image voxels by N, and the total number of voxels in a (tumor) region-of-interest (ROI) by M. If $\hat{v}_i$ and v_i^r refer to the true and reconstructed uptake values of the i^{th} voxel ($i = 1, \dots, N$) at the r^{th} noise realization, respectively, then: (i) $\bar{v}^r = \text{SUV}_{\text{mean}}^r = \frac{1}{M}\sum_{i=1}^{M} v_i^r$ represents the spatial mean of voxels across an ROI at noise realization r, (ii) $\bar{v}_i = \frac{1}{R}\sum_{r=1}^{R} v_i^r$ represents the mean of a given voxel i, averaged temporally across all noise realizations, and (iii) $\bar{v} = \frac{1}{R}\sum_{r=1}^{R} \bar{v}^r$ is the ensemble mean of all $\text{SUV}_{\text{mean}}^r$ values across R noise realizations. Since each ROI is set to have a uniform uptake in the true image, we denote $\hat{v}_i = \hat{v}$ to represent the value of all the voxels inside that region. Subsequently, the following noise and bias metrics can be derived:

Contrast recovery curves (CRC)

We define two types of CRC: SUV_{mean} CRC and SUV_{max} CRC. The former is defined as the ratio between the contrasts of the reconstructed image and the true object in terms of the ROI average. The latter would be the same ratio but in terms of the ROI maximum voxel.

Bias in SUV_{mean} and SUV_{max}

Biases in activity uptake quantitation were defined in terms of SUV_{mean} and SUV_{max}, calculated as follows [384, 388]:

$$\text{SUV}_{\text{mean}}\ \text{Bias} = \frac{1}{R}\sum_{r=1}^{R}(\text{SUV}_{\text{mean}}^r - \hat{v}) \qquad 2.9$$

$$\text{SUV}_{\text{max}}\ \text{Bias} = \frac{1}{R}\sum_{r=1}^{R}(\text{SUV}_{\text{max}}^{r} - \hat{v}) \qquad 2.10$$

where $\text{SUV}_{\text{max}}^{r} = \max_{i\in\{1\ldots M\}}\{v_i^r\}$ represents the voxel with the maximum uptake value inside the tumor ROI of realization r.

Equations 2.9 and 2.10 show that SUV_{mean} and SUV_{max} biases are basically defined as the deviation of the reconstructed image from the *true* value of the object. Both biases are then normalized to the true value for plots in the results section. It is worth noting that studies of quantitative task performance using clinical patient data would not have this level of rigor in determining the bias (given typical lack of access to the true value of every voxel).

The coefficient of variability (CoV) of SUV_{mean} and SUV_{max}

We characterized the CoV for both SUV_{mean} and SUV_{max}. The former was defined as follows:

$$\text{SUV}_{\text{mean}}\ \text{CoV} = \left(\frac{1}{R-1}\sum_{r=1}^{R}(\text{SUV}_{\text{mean}}^{r} - \bar{v})^2\right)^{\frac{1}{2}} \qquad 2.11$$

where $\bar{v}$ was defined at the beginning of this section. Similarly, the maximum uptake CoV defined as the variability of the maximum voxel of selected ROI across all realizations, and can be calculated as:

$$\text{SUV}_{\text{max}}\ \text{CoV} = \left(\frac{1}{R-1}\sum_{r=1}^{R}(\text{SUV}_{\text{max}}^{r} - \overline{\text{SUV}_{\text{max}}})^2\right)^{\frac{1}{2}} \qquad 2.12$$

where $\overline{\text{SUV}_{\text{max}}}$ denotes the SUV_{max} values averaged across the noise realizations. Both equations will be normalized to $\bar{v}$ in plots of the results section 2.3.4.3for more proper comparison.

Mean-squared error (MSE)

MSE is a more general metric that combines noise and bias of voxels within a single quantity. It can be averaged over all voxels within an ROI to represent the MSE of the ROI, and then for the R realizations, the mean of all MSEs would be calculated:

$$\text{MSE} = \frac{1}{RM}\sum_{r=1}^{R}\sum_{i=1}^{M}(v_i^r - \hat{v})^2 \qquad 2.13$$

Due to the importance of both the accuracy and precision of SUV_{mean} value in clinical practices and to better study the effect of different reconstruction kernels to the mean uptake, we defined the next FOM as the MSE of the mean uptake. MSE can also be calculated by summing the squared noise and squared bias [411]. The MSE of SUV_{mean} was defined as:

$$\text{MSE of SUV}_{\text{mean}} = \text{SUV}_{\text{mean}}\,\text{Bias}^2 + \text{SUV}_{\text{mean}}\,\text{CoV}^2 \qquad 2.14$$

Image roughness (spatial noise)

Image roughness (IR) measures the voxel by voxel variability in the image and can be calculated even for a single realization. Within a given ROI containing M voxels, image roughness was defined as the variability of the voxel values with respect to SUV_{mean}. This was then averaged over R noise realizations:

$$\sigma_{spatial} = \frac{1}{R}\sum_{r=1}^{R}\left(\frac{1}{M-1}\sum_{i=1}^{M}(v_i^r - \text{SUV}_{\text{mean}}^r)^2\right)^{\frac{1}{2}} \qquad 2.15$$

where $\bar{v}^r$ is the mean of all voxels, v_i^rs, inside the given ROI of realization r. The noise values plotted in the results section were normalized to $\bar{v}^r$.

Voxel variation (σ_0)

This metric provides a measure of the variability of individual voxels over multiple noise realizations:

$$\sigma_0 = \frac{1}{M}\sum_{i=1}^{M}\left(\frac{1}{R-1}\sum_{r=1}^{R}(v_i^r - \bar{v}_i)^2\right)^{\frac{1}{2}} \quad 2.16$$

Voxel variation impacts both image roughness and SUV_{mean} CoV [384]. Assume a uniform region consisting of M voxels with voxel variance σ_0^2 for each voxel and inter-voxel covariance $\mathrm{cov}(i,j)$ between two voxels i and j. Tong et al. showed that the expectations of image roughness (Eq. 2.15) and SUV_{mean} CoV (Eq. 2.11) is given by:

$$E\left[\sigma_{spatial}^2\right] = \sigma_0^2 - \frac{1}{(M-1)M}\sum_{i\neq j}\mathrm{cov}(i,j) = \sigma_0^2 - \frac{2}{(M-1)M}\sum_{i>j}\mathrm{cov}(i,j) \quad 2.17$$

and

$$E[\{SUV_{mean}\mathrm{CoV}\}^2] = \frac{\sigma_0^2}{M} + \frac{1}{M^2}\sum_{i\neq j}\mathrm{cov}(i,j) = \frac{\sigma_0^2}{M} + \frac{2}{M^2}\sum_{i>j}\mathrm{cov}(i,j) \quad 2.18$$

In Eq. 2.17, it is seen that the reduced voxel variance and increased inter-voxel covariance due to PSF modeling result in overall reduction in image roughness. In Eq. 2.18 however, these two works against one another and the increasing (positive) covariance contributes positively to SUV_{mean} CoV.

Averaged differences of max and min uptake

To better quantify edge effects, we assessed the range of uptake within the ROI after reconstruction. For a total of R realizations, the average max-min difference was calculated as follows:

$$\text{Averaged Max-Min difference} = \frac{1}{R}\sum_{r=1}^{R}(\mathrm{SUV}_{\max}^r - \mathrm{SUV}_{\min}^r) \quad 2.19$$

where $\mathrm{SUV}_{\min}^r = \min_m\{v_r\}$, referring to the lowest uptake values within the ROI of the r^{th} realization. The result was then averaged over all realizations. For plotting purposes in Results section, we normalized this measure to $\bar{v}$. This subtraction of the minimum undershoot from the maximum

overshoot of every reconstructed ROI evaluates the deviation from the true flat uptake due to noise or false texture (including edge effects). Then, averaging its value over all noise realizations essentially provides a FOM for *shape*.

2.3.4. Results

depicts a selection of eight out of 20 generalized PSF isocontours, including "true PSF" and "no-PSF" plots, in addition to three underestimated and three overestimated kernels. Here we follow the convention of Alessio et al. to demonstrate these kernels by plotting isocontours of their radial profiles [382]. The dashed horizontal line represents the LOR perpendicular to the detector element, in which the maximum probability of detection happens.

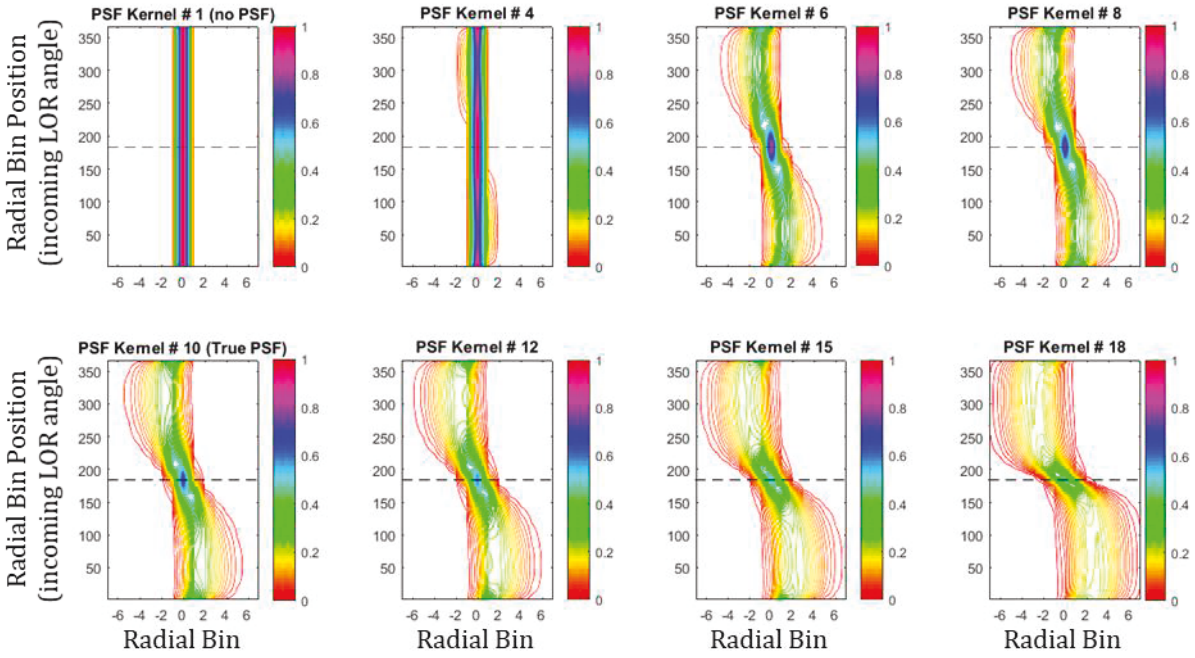

Figure 2-13. Isocontours of selected PSF modeled radial profiles: radial bins positions vs. radial bins. The intensity of contours is the probability of an incoming radial bin (LOR) from different angles (vertical axis) to a particular bin and its seven neighbor bins (zero for the centred bin and ±7 bins in the horizontal axis). The dashed line represents the LOR perpendicular to the detector element. Kernels 4, 6 and 8 are examples of underestimated and kernels 12, 15 and 18 are examples of overestimated PSF kernels.

It is worth breaking down how each of the three degradations phenomena affects the PSF kernel. Inter-crystal scattering symmetrically blurs the neighboring crystals of the incident detector. Equation 2.5 addressing photon non-collinearity also yields a symmetric blur. However, in inter-crystal penetration, photons penetrate the neighboring crystals and cause the parallax effect. This skews the PSF with respect to the true LOR, thereby inducing symmetry.

The generalized PSF-modeling kernels presented here has an advantage over the underestimated PSFs performed in image-space in previous studies that characterize PSF kernels by varying the FWHM of the measured PSF [396, 412]. Those approaches overlooked two issues with the realistic PSF kernels that we can observe in Figure 2-13. First, realistic PSF kernels are anisotropic, so their FWHM varies with the angle of LOR. Second, under- and overestimating the true PSF not only changes its FWHM but also shifts its peak location that is angular dependent. This can be observed in Figure 2-13, where the peak of radial bins corresponding to LORs entering with an oblique angle (radial bins 1 to 150 in Figure 2-13) drifts from 1 to 4 as we increase the PSF kernels width.

2.3.4.1. Reconstructed images

Noise-free reconstruction.

Figure 2-14 shows images of the noise-free reconstruction with 10 iterations and 7 subsets. PSF-modeling is known to improve resolution and enhance contrast. This can be observed by comparing the no-PSF reconstructed images in the left column with the columns representing kernel #7 (slight underestimation) and beyond. The two major drawbacks of PSF modeling can also be addressed here, as we point out in some observations from this figure:

(i) Following a few iterations, edge ringing phenomenon – a staple aftermath of PSF-modeling – starts to appear from kernel 6 (not shown in this figure – an intermediate underestimation of the true PSF – in all tumor sizes) and intensifies as the PSF kernel index – i.e. its deblurring effect – increases. This

observation challenges the idea of using underestimated PSF kernels as a solution to eliminate edge overshoot effects [400, 401].

(ii) The other prominent issue with PSF modeling is that it causes higher inter-correlation between voxels. The background (normal liver) regions in the first two kernels (no-PSF and kernel #3) are observed to have a different texture in comparison to the subsequent kernels, as the noise loses some of its high-frequency content. This difference becomes more conspicuous when comparing the noise texture of the noisy reconstructed images.

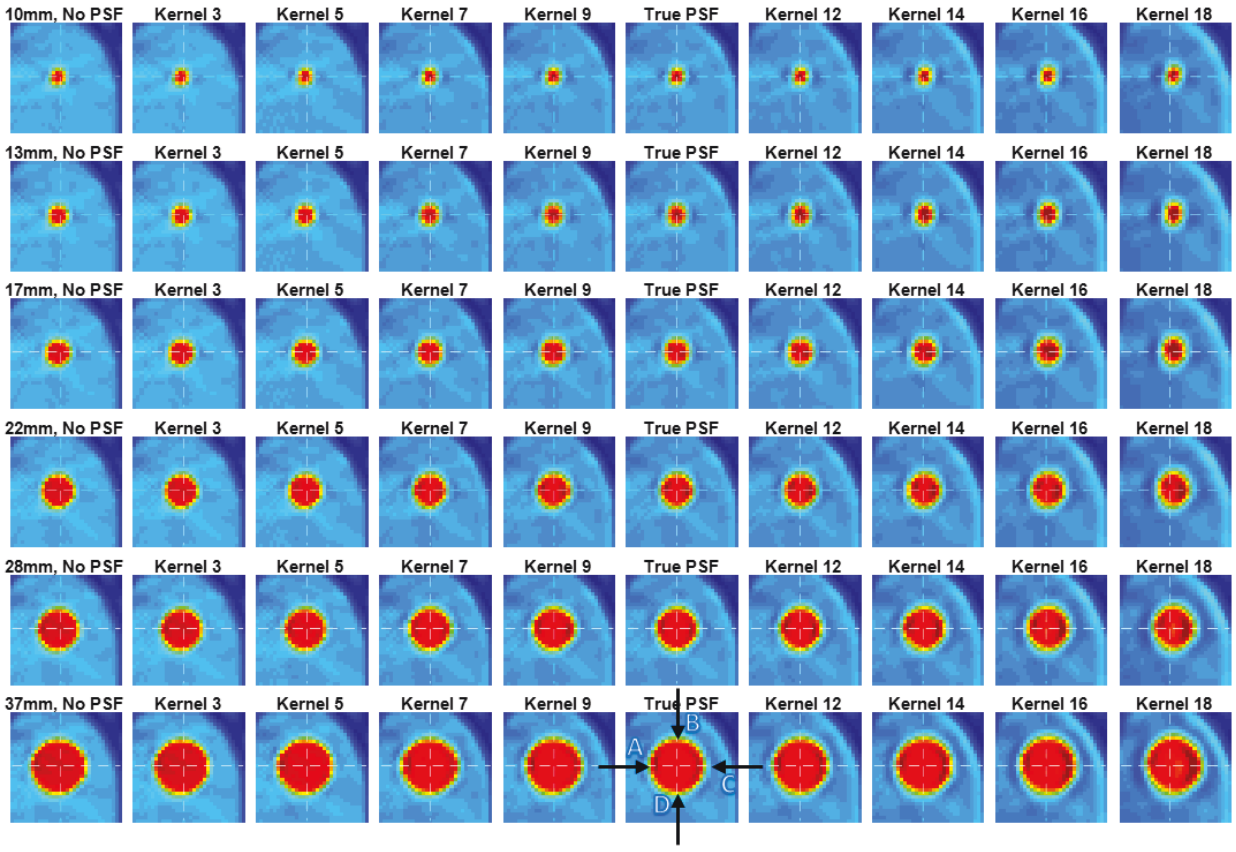

Figure 2-14. Noise-free reconstruction images of liver tumor and background (cropped to include liver tumor and its background tissue) after 10 iterations and 7 subsets. Rows represent different tumor sizes. Columns starting from the left indicate no-PSF reconstruction, four under estimating PSF kernels (#3, #5, #7 and #9), true PSF, and four overestimating PSF kernels (#12, #14, #16 and #18). The intersection of white dashed lines indicates the center of the tumor in the true object. The center of the FOV is located at the left-hand side of the tumor, and hence the tumor edges in its left and right sides pointed at by A and C arrows are more pronounced than top and bottom indicated by B and D.

There are three more interesting observations in Figure 2-14:

(iii) The edge overshoot in PSF reconstructed images of tumors larger than 17mm is not uniform across its ring; i.e. the edge is more pronounced in the left and right, compare to the top and bottom. This can be observed in bottom-middle reconstructed image in Figure 2-14 by comparing the regions pointed to by "A" and "C" having a darker red color with "B" and "D". The reason is closely related to the parallax effect. Photons from annihilation events away from the center of the FOV may experience significant inter-crystal penetration. Thus, the apparent LOR may not exactly match the true LOR and would be closer to the center of the FOV. In no-PSF modeling reconstruction, this LOR mismatch resulting in skewed lesions towards the center of the FOV will not be "deblurred", whereas it will be deblurred by incorporating a true PSF-modeling kernel. The edge overshoot appears as an aftermath of this deblurring. The overshoot would be more pronounced in the direction of the parallax effect that skews the regions towards the center. In this figure the center of the FOV is located approximately on the left side of the tumor, so the left and right edges of tumor undergo more deblurring compared to the top and the bottom ("A" and "C" directions compare to "B" and "D"), thus exhibiting more edge overshoot.

(iv) Furthermore, the overshoot on the right side of the ring (pointer "C") is longer than the one on the left side (pointer "A"). The reason is the partial ring section in the right is farther with respect to the center of the FOV than the left. Therefore, the amount of deblurring and edge overshoot is larger, and subsequently, an asymmetric edge overshoot will appear on the left and the right of the region.

(v) The final observation is that the apparent tumor location manifested in the reconstructed image drifts *away* from the center of the FOV as we apply higher kernels. This movement can be tracked using the white dashed lines representing the center of the tumor in each image. The reconstructed ROI with the 10th kernel (true PSF) is in a perfect position; while it slightly shifts towards the center of the FOV for underestimated kernels including no-PSF, and slightly shifts away from the center for overestimated ones. These effects result from under-/over-correcting for the parallax effect by various PSF kernels. By applying the underestimated kernels, the full correction (i.e. deblurring) is

not yet accomplished, thus the apparent position of the ROI is not back in its initial location; whereas the overestimated the kernels are actually over-correcting (over-deblurring) the region in the reconstruction.

Noisy reconstruction.

Figure 2-15 shows noisy reconstructed images. These images display the edge overshoot in the reconstructed ROIs of kernels 7 and above, in addition to its asymmetry, as explained in the last section. However, they also demonstrate another principal of PSF modeling: modified noise texture. With more blurring kernels, images look smoother, voxel variance reduces in both the tumor and the background, and the noise becomes more correlated and blobby. The inter-voxel correlation increases as we apply wider PSF kernels, thus the images look smoother with a blobby noise-texture.

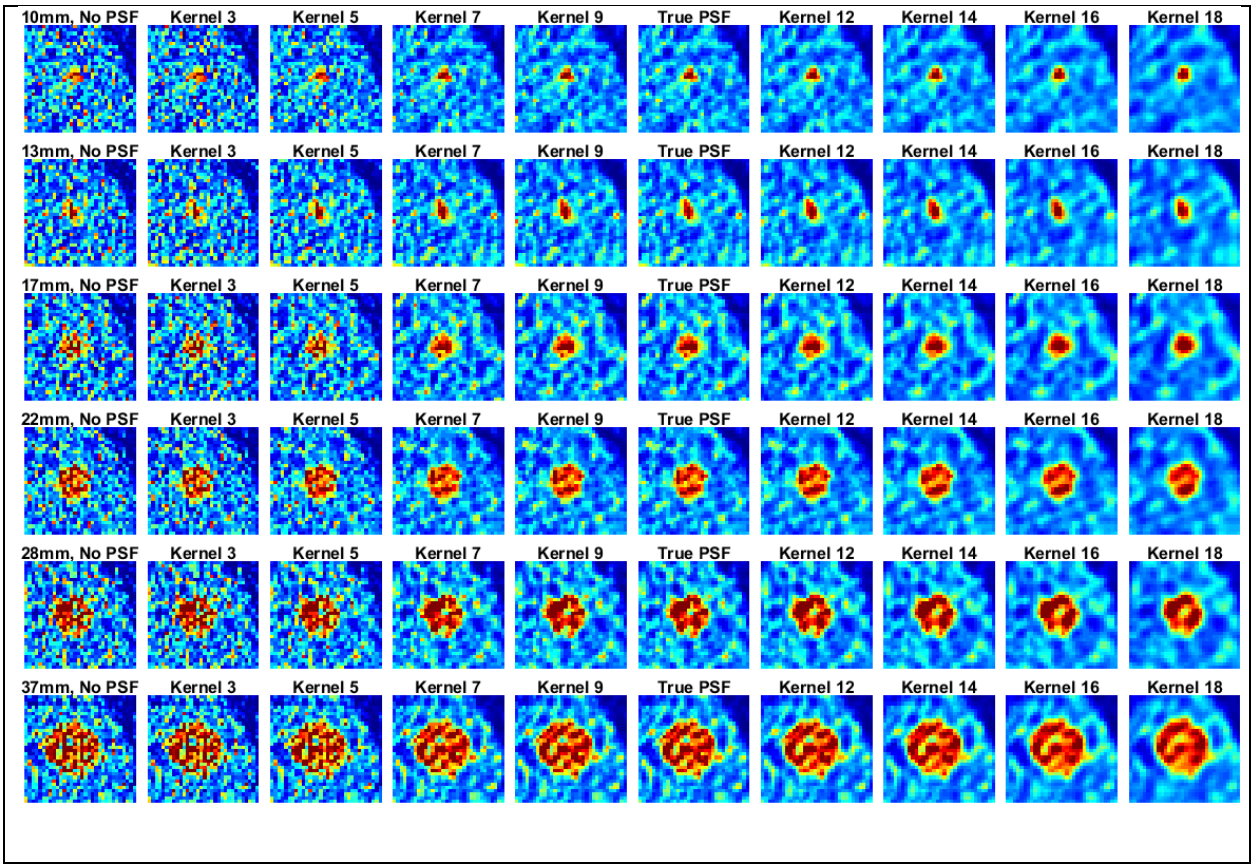

Figure 2-15. Noisy reconstruction images of liver tumor and background (cropped to include liver tumor) for iteration #10 iterations with 7 OS-EM subsets and no post-smoothing. Rows represent different tumor sizes. Columns starting from the left indicate no-PSF reconstruction four under estimating PSF kernels (#3, #5, #7 and #9), true PSF, and four overestimating PSF kernels (#12, #14, #16 and #18)

2.3.4.2. Contrast recovery analysis

Figure 2-16 shows plots of contrast recovery for SUV_{mean} and SUV_{max} (CRC_{mean} and CRC_{max}, respectively) of the tumor reconstructed with 20 PSF kernels. The first six images show that neither PSF nor no-PSF kernels can yield a CRC of one. PSF overestimation, however, yields a CRC value closer to one. Yet in most cases, extreme overestimation, i.e. kernels 15 and above, results in CRC_{mean} higher than one, which is as undesirable as CRC<1 for underestimated kernels (mostly due to PVE at the edges) and induces an overestimation bias in the reconstructed region that alters quantification.

The CRC_{mean} curves have a smooth and monotonic transition with respect to increasing PSF kernel width. PSF modeling corrects for PVE and thus reduces blurring at the edges, therefore no-PSF modeling yields the maximum PVE and lowest CRC_{mean}.

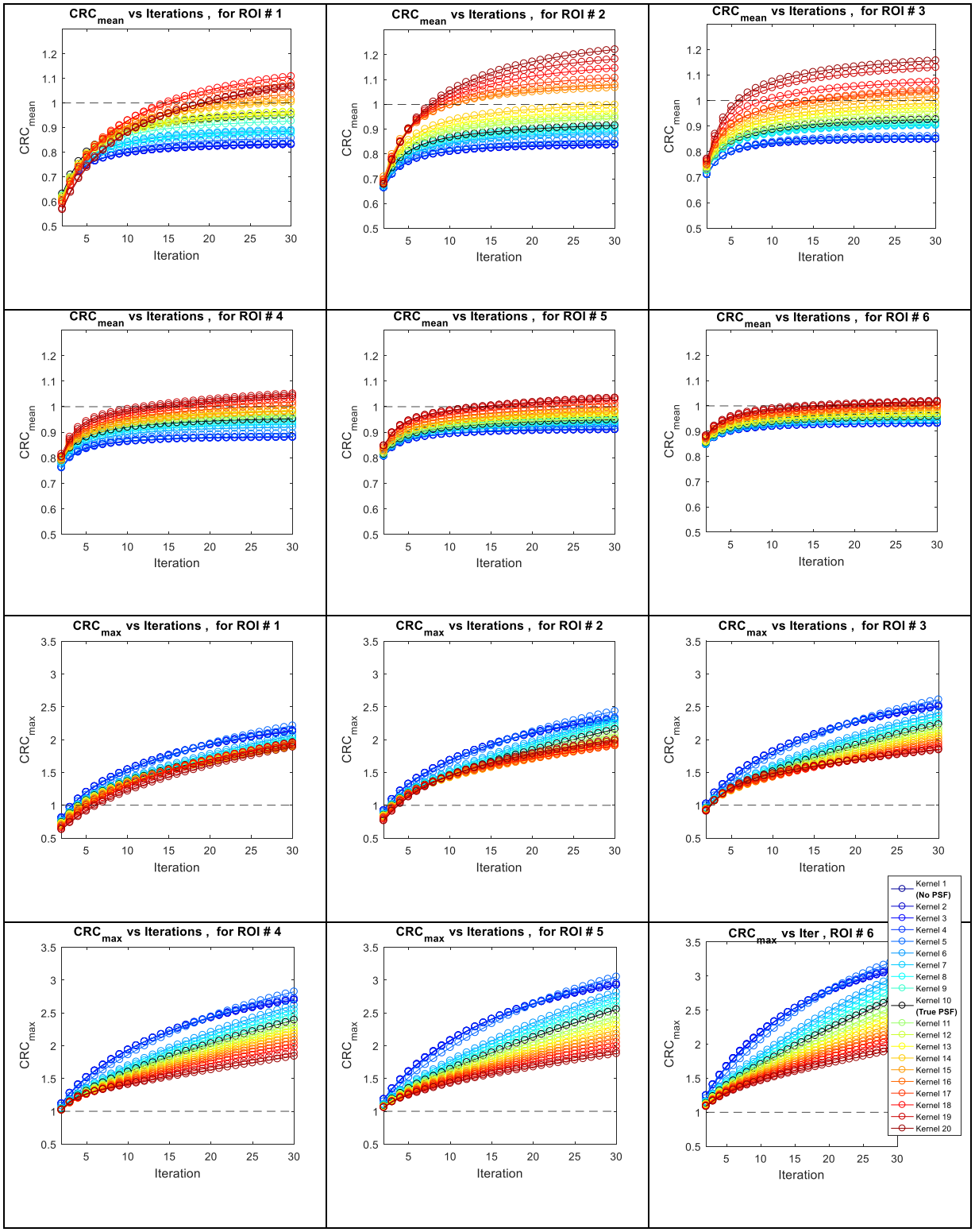

Figure 2-16. Averaged CRC of SUV_{mean} and averaged CRC of SUV_{max} vs OS-EM iterations for six tumors over 200 noise realizations. The dashed line highlights CRC=1. CRC_{mean} plots have a fixed vertical axis range of [0.5, 1.3], and the range for CRC_{max} plots is fixed to [0.5, 3.5].

Another observation from this figure is that the true PSF reconstruction yields CRC_{mean} less than one. The reason is that the overshoot at the edge of the ROI does not actually involve the outmost set of voxels of the region. The set of voxels undergoing overshoot are encircled by another ring(s) of voxels that (i) contain more voxels than the overshoot ring, and (ii) have less uptake than the reference truth. This is because the edge has not been completely recovered until after 40-60 (total) iteration updates, and the algorithm has not yet perfectly converged. These surrounding voxels at the very edge of the ROI decrease CRC_{mean} to less than one—even in the presence of the overshoot edge—and consequently induce negative bias, as can be seen in the results of section 2.3.4.3. The EM algorithm is known to improve with every iteration towards convergence that eventually reconstructs edges perfectly after a massive number of iterations. However, this is impractical in PET reconstruction due to the presence of noise and its severe amplification. Therefore, the PVE at the edges of the region impacts CRC_{mean} and causes it to be suboptimal. Thus, the observed underperformance is an attribute of the EM algorithm in PET reconstruction, and it disturbs image reconstruction with any degree of PSF modeling. But it can be observed from Figure 2-16 that overestimated PSF modeling kernels tend to mitigate this deficiency.

CRC_{max} plots interestingly follow a reverse pattern, where, in contrary to CRC_{mean}, underestimated and no-PSF attain higher CRC_{max} values and the curves decline as kernels' widths increase. We observe that most of the curves lay above one, which is due to (i) the presence of noise and (ii) not performing any post-smoothing on the images that are shown to reduce CRC [396]. Moreover, iterative reconstruction algorithms, including OS-EM, are known to intensify the noise as they iterate. Therefore, this produces a monotonic increase regardless of the generalized PSF kernel. However, PSF reconstruction with wider kernels yields more correlation between the voxels. As a result, voxels cannot oscillate freely in the presence of the noise and their fluctuation decrease as the PSF kernel width increases. This inter-voxel correlation not only contains the oscillation of each voxel due to

noise but also limits the overshoot from rising too much in reconstruction with overestimated PSF kernels. The combination of these two effects contributes to the reduction of SUV_{max} thus CRC_{max}.

The CRC_{max} curves also show that the smaller tumors (first three ROIs) have smaller CRC values; even the first few iterations may generate a CRC_{max} of one. The reason is again that the region is not fully recovered within 30 iterations. It is because the few voxels of the region mostly have values less than the reference truth due to the PVE, even when reconstructed with wide PSF kernels. The noise will then be added on top of this PVE and cause the CRC_{max} to become closer or even exceed one. More importantly, edge overshoot has not yet developed in early iterations of smaller tumors because the very few voxels across the ROI have not created enough extent for the overshoot to rise. Larger tumors, on the other hand, have an ample amount of space for multiple overshoot- and undershoot rings to appear. Therefore, as the diameter of the region grows, CRC_{max} increases, but it decreases with wider PSF kernels.

2.3.4.3. Noise-bias performance comparison

We defined three measures of bias, six measures of noise and two types of MSE in section 2.3.3.4, and are presenting their plots in this section. Note that the curves representing reconstruction with PSF kernels in all plots follow the same legend as Figure 2-17. In all figures, the starting iteration for plotting is two.

Image roughness (IR) vs. SUV_{mean} bias.

Figure 2-17 shows image roughness vs. SUV_{mean} bias for six tumors. The range on all six plots is fixed for a better comparison unless otherwise stated.

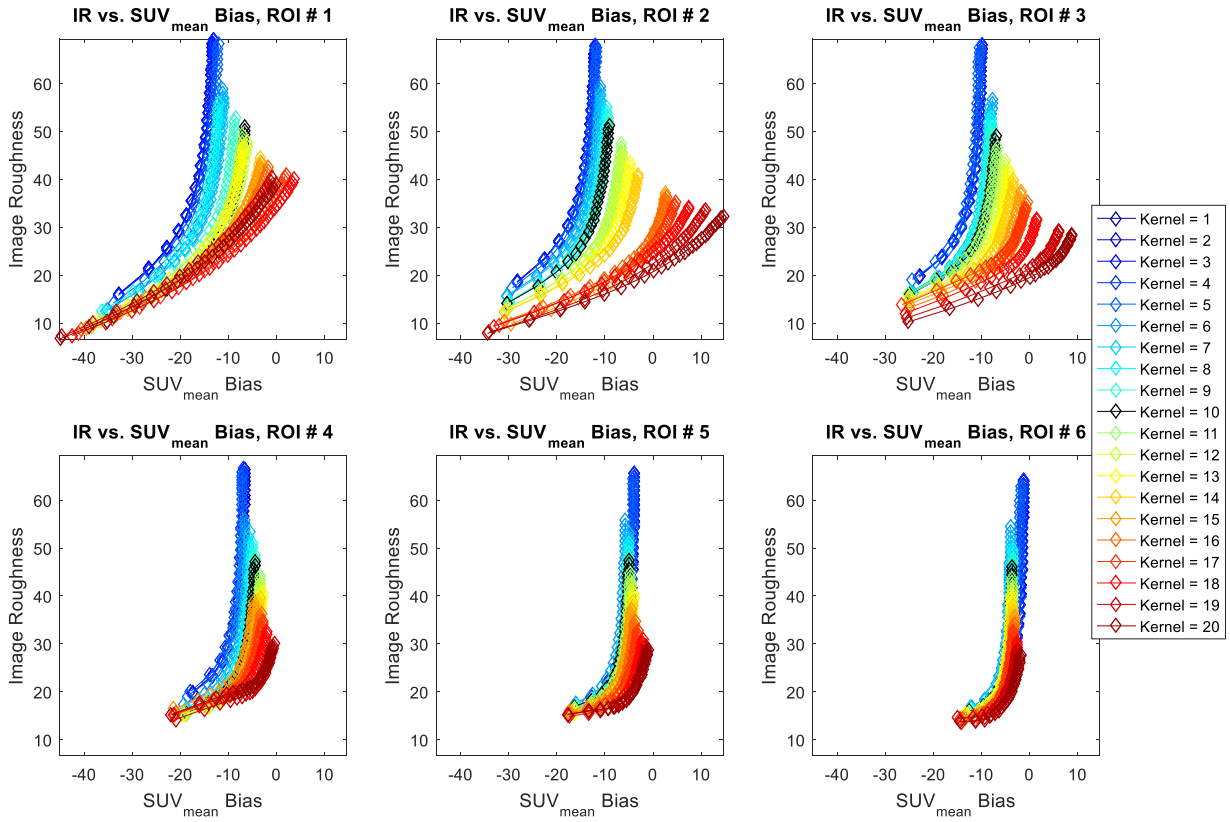

Figure 2-17. Image roughness vs. SUV_{mean} bias for six tumors. Each point in the curves represents the results for a single OSEM iteration.

IR decreases as the kernel index increases, which is consistent with PSF-modeling reducing spatial voxel variation and yielding a smoother image. Comparing no-PSF with true PSF (kernel #10) and a medium overestimated PSF kernel #15 (orange curve) at matched iterations shows a range of 25%~35% and 38%~45% less noise for kernels #10 and #15 within all six tumors, respectively. At matched noise, SUV_{mean} biases of these two kernels compared to no-PSF degrade significantly for first four tumors (25%~45% less bias for kernel #10 and 60%~94% for kernel #15), while its variation with respect to different kernels drops for two larger tumors (-16%~2% difference in bias for kernel #10 and 6%-16% less bias for overestimated kernel #15. Excessive overestimation, such as for kernel indices over 16, usually leads to a positive bias.

The behavior of IR curves can be explained by Eq. 2.17. No-PSF and underestimated kernels have higher σ_0 and lower covariance value; both of which contribute to amplify the spatial noise. As the kernels approach, true PSF and its overestimation, σ_0 degrades and voxels exhibit more covariance, and both yield lower image roughness. Moreover, IR has a small increase for larger ROIs. These regions consist of more voxels that result in less weight of the second term on the right-hand side of Eq. 2.17; which also results in a lower range of IR for larger tumors.

SUV_{mean} CoV vs. SUV_{mean} Bias

The plot of SUV_{mean} noise vs. bias is shown in Figure 2-18. Unlike the other figures displayed earlier, each of the six plots has different axis ranges for more clear differentiation of the curves.

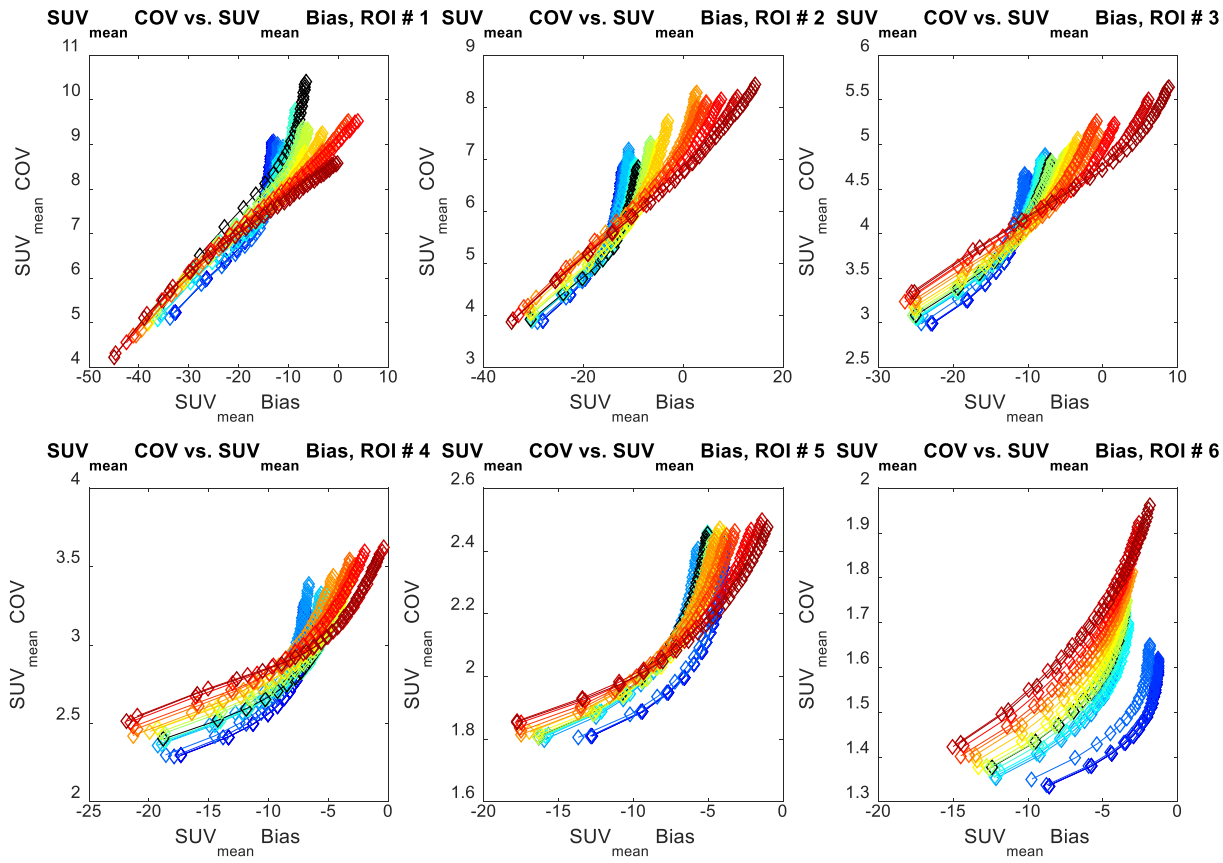

Figure 2-18. SUV_{mean} CoV vs. SUV_{mean} bias for six tumors. Note that axes ranges are not the same for each plot in this figure.

Compared to the previous noise vs. bias performance curves with noise extending over 60%, the three plots for larger tumors in Figure 2-18 have much lower values and a smaller range of SUV_{mean} CoV. Observing lower values for SUV_{mean} is predictable, because not only it is an averaging process, but also the PVE at the edges contributes largely to the negative bias. As the PSF kernel's width increases, some undershoots also may appear that increase the impact of overshoots and contribute to maintaining lower SUV_{mean} values. For matched iterations (7th iteration; same for earlier two figures), no-PSF reconstruction demonstrates 0~5% and -1%~10% less noise compared to kernels #10 and #15, respectively. One immediate reason for such a small SUV_{mean} CoV level for these larger ROIs is a higher number of voxels. Nonetheless, in the first three smaller ROIs, the SUV_{mean} CoV values and ranges are slightly higher: no-PSF shows (mostly) improved noise performance compared to

kernels #10 and #15 by -2%~15% and 7%-11%, respectively. Some relevant discussions are provided in the next section explaining the observed patterns.

SUV_{max} CoV vs. SUV_{max} bias

The plot is shown in Figure 2-19 displays an approximately linear relationship between noise and bias.

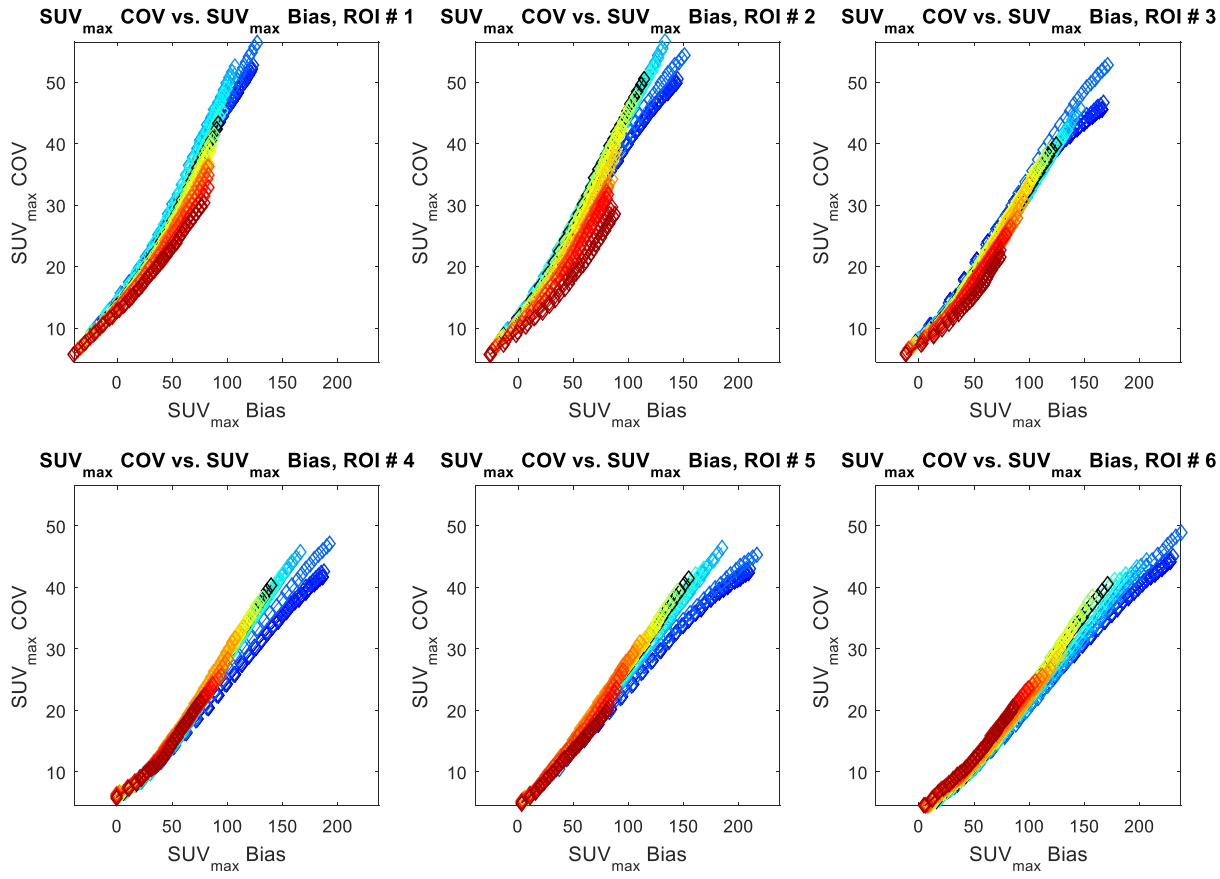

Figure 2-19. SUV_{max} CoV vs. SUV_{max} bias for all six tumors.

SUV_{max} noise vs. bias curves of the overestimated PSF kernels typically demonstrate lowered noise, lowered bias and thus a higher quantitation performance compared to other earlier PSF kernels. In terms of SUV_{max} CoV, kernels #10 (true PSF) and #15 show 43%~57% and 58%~72% less noise,

respectively, as compared to no-PSF. In terms of SUV_{max} bias, the numbers are 18%~42% and 43%~55%, respectively. This is mainly due to increased inter-voxel correlation explained for CRC_{max} in section 2.3.4.2. The first three curves in this figure also show that smaller ROIs exhibit a negative SUV_{max} bias for the first few iterations (maximum of 7) for the overestimated PSF curves, whereas the three larger ROIs do not show a negative bias, due to reasons explained in 2.3.4.1 about the smaller ROIs greatly suffering from PVE.

Average max-min difference vs. SUV_{mean} bias

The noise vs. bias plot is presented in Figure 2-20. It was mentioned in section Averaged differences of max and min uptake that this noise metric essentially is a measure of shape and quantifies regions' non-uniformities due to PVE and the edge overshoot effect.

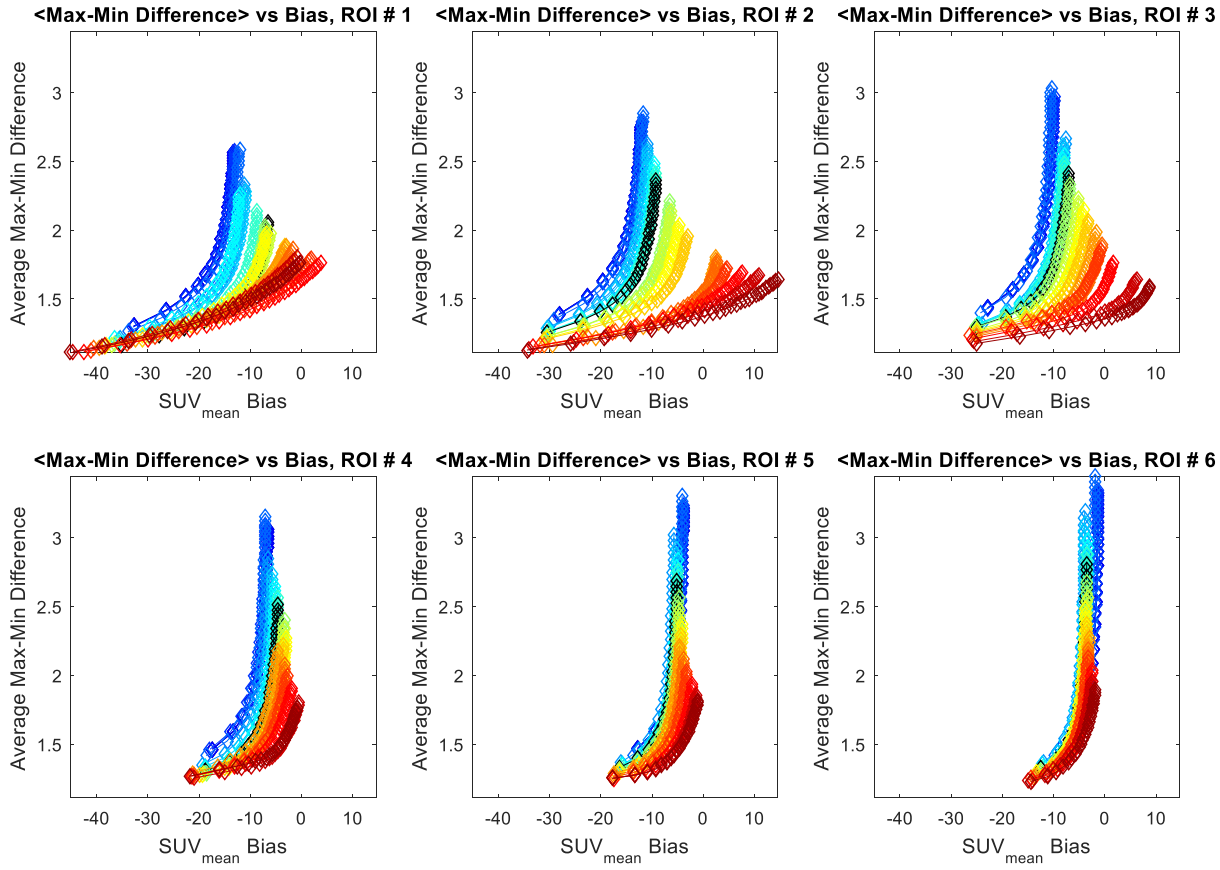

Figure 2-20. Averaged max-min difference vs. SUV_{mean} bias for each of the six tumors studied.

The significance of this metric is its ability to assess edge overshoot effect and PVE. Comparing this shape metric for no-PSF with kernels #10 and #15 with matched iterations shows 15%~20% and 21%~28% less shape variability, respectively. Underestimated kernels are more prone to noise and post-smoothing was not performed here, so most of the higher variabilities measured in the smaller blurring kernels are due to the noise. Nonetheless, it is interesting to observe the monotonic decrease of this shape variability that mostly depicts the effect of higher correlation because of more deblurring with wider PSF kernels.

MSE vs. PSF modeled kernels indices

Figure 2-21 depicts plots of MSE vs. PSF kernels, where MSE effectively combines noise and bias within a single metric. Every line in the plot corresponds to an OS-EM iteration, and the results are shown for different ROIs. Interesting observations can be made. Except for the first few iterations in the smallest ROI, we see generally decreasing MSE values with increasing PSF widths for each given iteration. Also, plotting MSE and performing minimization while allowing for iteration number to vary – i.e. looking at the bottom of each plot – we see that least MSE is obtained, in all six ROIs, for overestimated kernels.

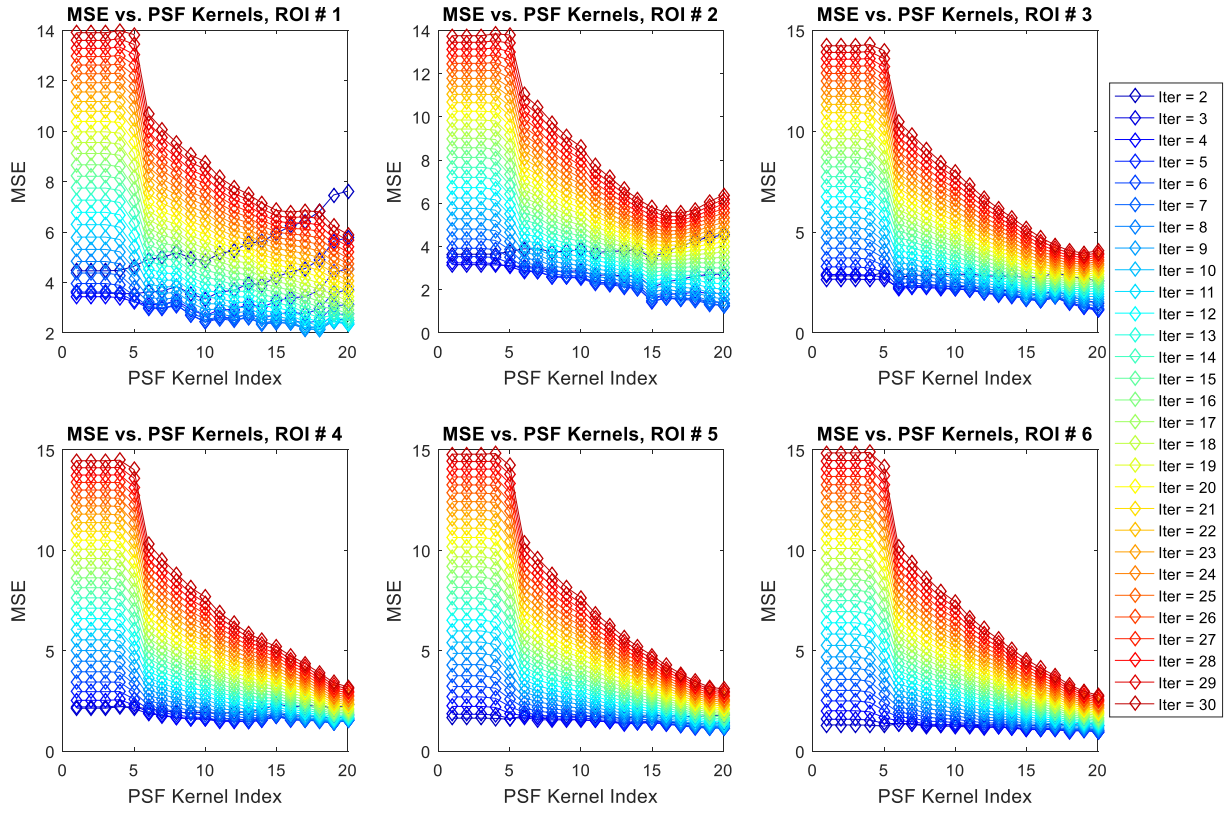

Figure 2-21. MSE vs. PSF kernels. Each color indicates an OS-EM iteration.

MSE of SUV_{mean} vs. PSF modeled kernels indices

We calculated the MSE values for SUV_{mean} using equation 12 and plotted them vs PSF kernel indices, as shown in Figure 2-22 (note that each plot has a different horizontal axis range). We indicated that the MSE of SUV_{mean} captures the effects of both its noise and bias, thus providing us with an indicator of the overall performance of SUV_{mean}, which itself is robust to spatial noise and can be used to quantify PVE. In these figures, following the trends of increasing iterations implies wider PSF kernels require more iterations to converge than narrower PSFs. Furthermore, we observe that the best performance, in terms of the minimum MSE of SUV_{mean}, was obtained for overestimated PSF kernels. More specifically, it is seen that for small tumors, minimum overall MSE is obtained with medium PSF overestimation, whereas for the two largest regions, the slight decrease in SUV_{mean} CoV (higher SUV_{mean} reproducibility, as explained in SUV_{max} CoV vs. SUV_{max} bias) boosts the performance of underestimated PSF kernels. However, the MSE improvement between underestimated PSFs and true PSF for the two largest regions are 0.10 and 0.11, while the improvements for the overestimated PSFs (kernel 14 for instance) vs. the true PSF are 0.56, 0.61, 0.36 and 0.08 for ROIs 1 to 4, respectively. Therefore, overestimated PSF achieve more significant improvement.

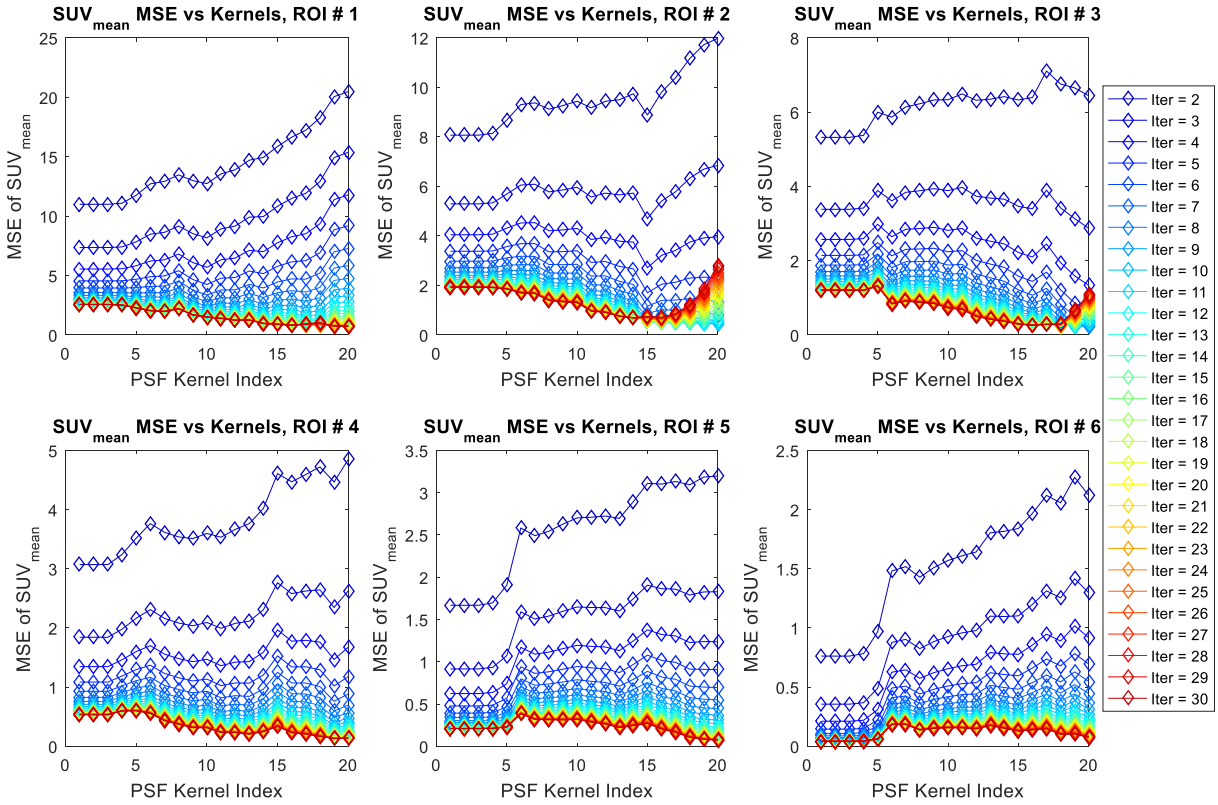

Figure 2-22. MSE of SUV_{mean} vs. PSF kernels. Note that each plot has a different scale.

Plots vs. sphere diameters

In the following Figure 2-23, six of the FOMs are plotted vs. ROI sphere diameters. The iteration is fixed at 10 for plotting:

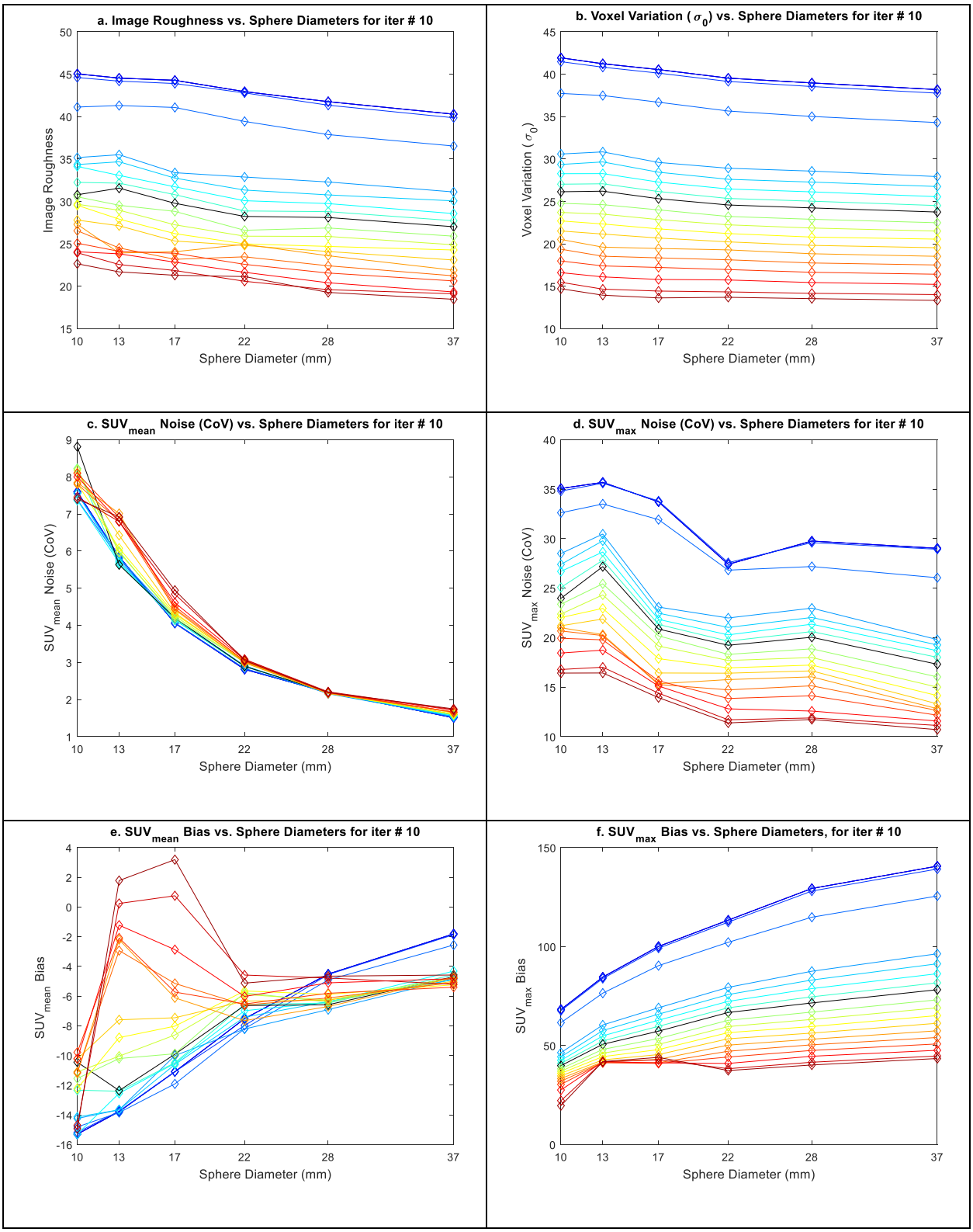

Figure 2-23. Various quantification metrics vs. ROI diameters (in mm) at iteration 10 from different kernel sizes: (a) image roughness, (b) voxel variation, (c) SUV_{mean} CoV, (d) SUV_{max} CoV, (e) SUV_{mean} bias, and (f) SUV_{max} bias. Legends are the same as Figure 2-17.

These six plots provide better intuition about the impact of PSF kernels for ROIs with different sizes and help make interesting observations. Figure 2-23.a shows a small reduction in average voxel variation as the diameter increases. This is related to the fact that with larger ROIs, a smaller fraction of voxels is impacted by edge overshoot effect, which itself amplifies voxel variability [388]. Also, as seen in Figure 2-23.b, image roughness decreases with increasing ROI size. This is also related to the above effect, as well as the fact that second term in Eq. 2.17 decreases with increasing voxels (M). Figure 2-23.c shows an important trend. As mentioned in section SUV_{mean} CoV vs. SUV_{mean} Bias, the only plots that show overestimated PSF kernels have inferior performance were for SUV_{mean} CoV. But as we explained, both Figure 2-8 and Figure 2-23.b indicate that this inferior reproducibility is not substantial; even in its worst case the CoV for the 10mm and 13mm ROIs, they vary within a range of 1.5% from minimum to maximum SUV_{mean} CoV. This figure also shows a considerable decline in CoV with an increase in a number of voxels in an ROI, as expected, as the mean measure becomes substantially more robust.

An analogous trend observed in Figure 2-23.d, Figure 2-23.e and Figure 2-23.f is a peak in SUV_{mean} CoV, SUV_{mean} bias and SUV_{max} bias curves, respectively for 13-17mm tumors, especially with PSF overestimation. This pattern is mostly generated because of edge overshoot effect. We explained in 2.3.4.2 that although this aftermath is referred to as *edge* overshoot, it does not involve voxels exactly at the edge of the ROI. The outermost ring of voxels always undergoes PVE (even at extreme PSF overestimation) where they have not yet recovered their true value. Excluding this outermost ring, four rings of voxels can contribute to the first overshoot ring as can be observed in large ROIs in the last row of figure 3 for kernel #10 (true PSF), and even more rings in the case of extreme overestimation. In larger ROIs the distance between one end of the region to the other is much larger than 10 voxels, therefore a complete doughnut can arise in the edge of the region. The center of this doughnut accommodates smaller undershoots and overshoots in mid-sized ROIs, as well as a flat region in larger ROIs. However, in smaller lesions, such as the 13mm and 17mm tumors, the extent

from one edge to the other is less than what is needed for these two overshoots to remain separate (thus generate a doughnut). Hence these two overshoots *merge* together and create a single overshoot with a value higher than each of the individual edges. This results in higher SUV_{max} FOM values, including higher SUV_{max} noise (Figure 2-23.d) and bias (Figure 2-23.f). Due to this phenomenon in smaller regions, most of the voxels inside the reconstructed ROI undergo an overshoot and exhibit a positive bias, resulting in a more positive SUV_{mean} bias. This is the main reason for the peak in 13mm and 17mm regions in Figure 2-23.e. In larger tumors, the outermost ring of voxels contains more voxels than the overshoot rings exhibiting overshoot. Moreover, the negative bias that these outermost voxels experience due to PVE is much larger than the positive bias that inner ring of voxels exhibits due to the edge overshoot, which eventually causes the SUV_{mean} bias to become negative.

2.3.5. Discussion

2.3.5.1. Noise vs. bias analysis

In the current work, we plotted various noise vs. bias curves to assess the quantitation performance of true PSF, no-PSF and generalized PSF-modeling kernels. We performed a comprehensive analysis of generalized PSF modeling reconstruction for assessing quantitative task-performance, including noise vs. bias analysis between four types of noise (IR, SUV_{mean} CoV, SUV_{max} CoV, and average max-min difference) and two biases (SUV_{mean} bias and SUV_{max} bias). Past efforts have focused on quantitation performance comparison between PSF vs. no-PSF-modeling reconstructions. These included bias (and/or contrast) vs. noise trade-off curves, commonly illustrating outperformance of PSF when defining noise as IR [357, 361, 363, 413], or SUV_{mean} CoV [360, 362, 370, 382]. Our present study shows improvement in IR, SUV_{max} CoV, and the average max-min difference for the true PSF kernel vs. no PSF, complying with previous reports and demonstrates medium overestimated PSF

kernels outperforming the true PSF. At the same time, reproducibility in terms of SUV_{mean} CoV between true PSF vs. no PSF shows less significant improvement (for small tumors with the same number of OS-EM iterations) or even a slight degradation (for small regions for PSF vs. no PSF or which complies with the previous reports; both of which complies with previous reports [384, 398, 399]. The same behavior is observed for overestimated PSF vs. true PSF. However, since SUV_{mean} metric is involved averaging voxels yielding a smoother value, its coefficient of variability is generally very low. Therefore, in practice, reproducibility degradation of overestimated PSF vs. no-PSF is negligible (<10% variability for a CoV of 3%).

We also notice that more thorough analysis of reproducibility vs. bias can be performed for many other metrics (beyond SUV_{mean} and SUV_{max}) in the emerging area of radiomics and heterogeneity quantification [247, 306, 309, 414-428]. This requires a distinct effort which can be pursued in the context of varying PSF kernels.

2.3.5.2. Important factors in quantitation analysis

In the present work, it was seen that two essential and determining parameters need to be carefully tuned for task-performance optimization: (i) number of iterations and (ii) segmentation of the target region or thresholding the ROI.

The number of iterations in reconstruction.

It can be seen from the plots in sections 2.3.4.2 and 2.3.4.3 that matched iterations cause different levels of contrast, noise or bias in reconstructions with different PSF kernels. This is because the degree of convergence in true PSF versus no-PSF reconstructions differs at the same number of iterations. This is very important when assessing and analyzing quantitation task performance. As an example, consider the four plots of SUV_{mean} bias vs. sphere diameters in Figure 2-24 for iterations 5, 7, 11 and 13. We observe that these plots follow a comparable pattern, where curves peak at 13mm

and 17mm ROIs. However, the values are considerably changing. No curve demonstrates a positive bias in iterations 5 and 7, whereas in iterations 11 and 13 the maximum value of the plot increases, creating positive bias. As such, merely comparing images reconstructed with different PSF kernels at the same number of iterations for PSF and no-PSF would not be sufficient (although this practice commonly appears in the literature [354, 412]). Therefore, it is necessary to observe the trend of quantification metrics with increasing iterations to obtain an accurate and meaningful comparison.

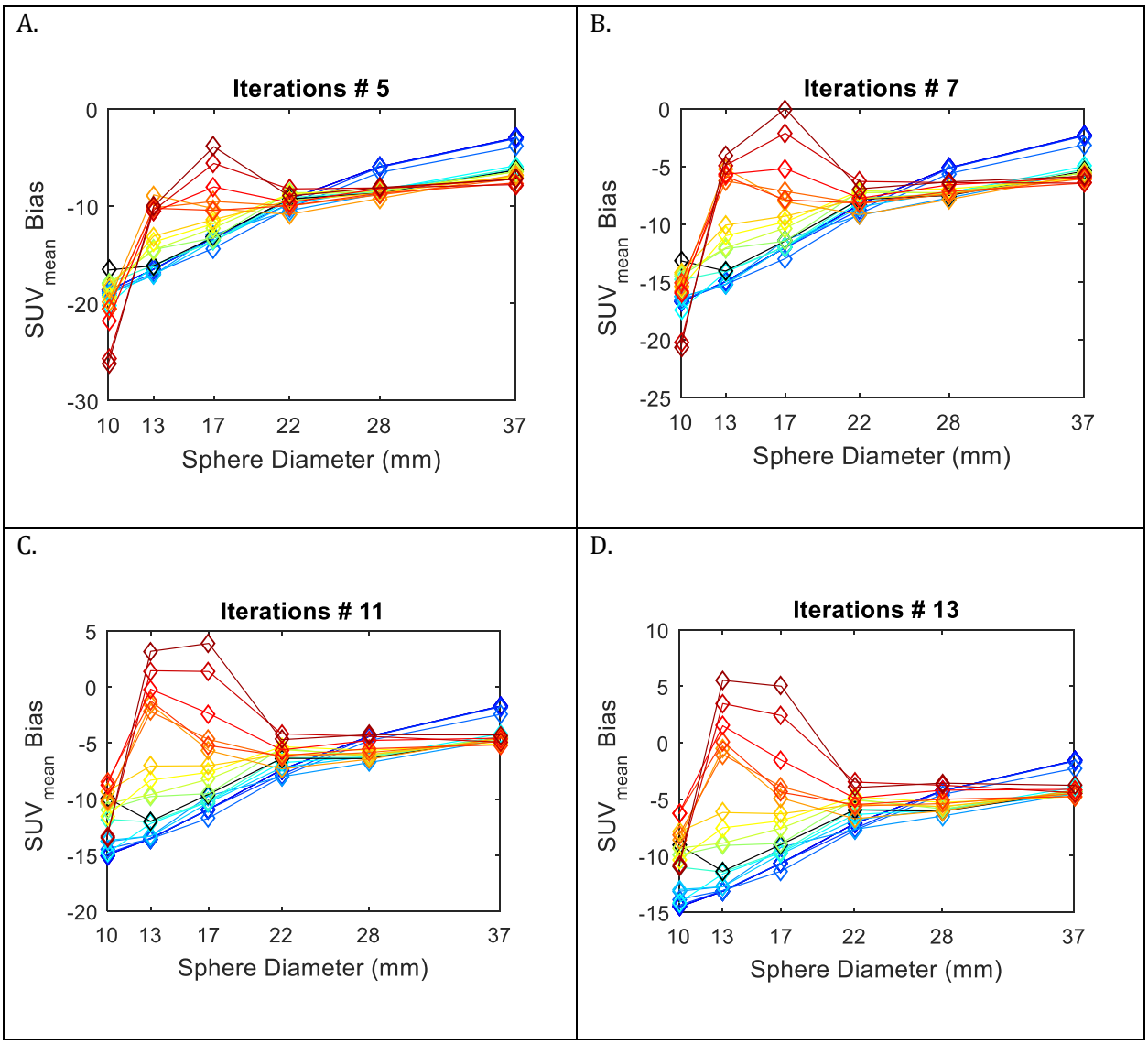

Figure 2-24. SUV_{mean} bias vs. sphere diameters for iterations number (a) 5, (b) 7, (c) 11, and (d) 13. Legend is the same as Figure 2-17.

ROI segmentation.

ROI segmentation was performed using thresholding, and the level was set to 55% in the present work, which we explored and assured visually to correspond very closely to the reference truth region. In any case, we observed that ROI segmentation has a significant impact on quantitation. This is because the uptake values of the outermost voxels at the boundaries of a reconstructed ROI suffer from PVE and hence have relatively lower values compared to the reference truth, which eventually impacts the noise and bias metrics. Correspondingly, the maximum voxel in a noise-free reconstructed ROI may represent either the overshoot due to the edge overshoot effect in mild underestimated, full, and overestimated PSF-modeling, or the true value of the ROI in no- and heavily underestimated PSF-modeling. On the other hand, the minimum voxel in such an ROI may represent a voxel in an undershoot of the ringing effect in mild underestimated, full, and overestimated PSF-modeling, while it may also refer to a low uptake in the very edge of the region due to PVE. In this case, the thresholding should be defined properly so that the masked ROI excludes low uptakes at the edges. Otherwise, in the case of mild underestimated, full, and overestimated PSF-modeling, it would be hard to determine whether the minimum uptake is due to an undershoot of the edge effect that occurs inside the first overshoot ring, or it is a blurring due to PVE at the border of the ROI. To prevent this confusion, we optimized the thresholding ratio to preserve the region shape, while excluding blurred voxels due to PVE in noise-free reconstruction, and then apply it to the noisy reconstructed images.

2.3.5.3. Noise metrics

From the analysis provided in section Voxel variation (σ_0) of Section 2.3.3.4, it can be deduced that noise, when measured spatially, can be significantly reduced by PSF modeling (given the same iteration number) whereas SUV_{mean} CoV actually could be unchanged [384], increase [399], even

multi-fold [398]. This potential increase in SUV_{mean} CoV means a reduction in SUV_{mean} reproducibility. The present work explored how generalized PSF modeling (under or over-estimated kernel estimation) impacts quantitative performance given these different metrics, and the effect of iterations number was also especially considered.

We observed in section SUV_{mean} CoV vs. SUV_{mean} Bias that SUV_{mean} CoV vs. bias curves of underestimated PSF kernels demonstrated slightly better performance for small regions. Visual assessment of these plots reveals that the underestimated PSF kernels outperform true and overestimated PSF kernels only within the first few iterations. However, if iterations exceed 40 (not depicted), the noise vs. bias curves of underestimated PSF kernels for the largest ROI in Figure 2-18 would eventually follow the pattern observed in Figure 2-17, Figure 2-19 and Figure 2-20—i.e. overestimated PSF curves outperforms underestimated and true PSF. The reason for this behavior can be partially explained by equations 2.17 and 2.18 in section 2.3.3.4; especially Eq. 2.18, in which the two terms of σ_0 and covariance act against each other. Although more iterations increase σ_0 and decrease covariance, the effect of a $1/M$ factor in the first term of Eq. 2.18 further diminishes the effect of σ_0 compared to the first term in Eq. 2.17, thus the first term cannot impact SUV_{mean} CoV in Eq. 2.18 as much as it impacts IR in Eq. 2.17. Furthermore, increased covariance values in earlier iterations contribute to SUV_{mean} CoV, increasing it and thus decreasing reproducibility of the higher PSF modelled kernel indices. Nonetheless, this degradation of reproducibility for overestimated kernels, particularly the kernel #15, was lower (<10%) compared to improvements (reductions) in IR (~50%) and σ_0 (30%~50%), while increasing CRC (a CRC_{mean} of 0.95 for kernel #15 vs. 0.85 for kernel #10 and 0.80 for no-PSF). In the present work, we did not include any noise-suppression or control in the current analysis, but surely studies should be pursued to analyse methods such as post-reconstruction smoothing, penalized EM, and other credible methods and study their effects along with various degrees of PSF modeling in quantitation. Also, the present work was focusing on a single tumor contrast derived from clinically derived kinetic parameters for a liver tumor. Future studies

need to be performed that includes tumors with higher contrasts. The present work proposes and expands on a particularly neglected dimension of imaging, namely on the extent of PSF kernels. But in future studies it s,hould be integrated along with various filtering regimes and number TOF kernels to find the optimal combination for optimizing the overall PET quantification.

2.3.5.4. Comparison with past efforts

A preliminary study of projection-space generalized PSF-modeling was performed involving a spectrum of underestimated PSF kernels [307]. In that study, we quantitatively analyzed projection-space reconstructions with a spectrum of PSF kernels generated from the true PSF using a convex combination approach; i.e. $(1-\alpha)\times$ "no-PSF kernel" + $\alpha\times$ true "PSF kernel", $\alpha \in [0,1]$, where the "no-PSF" kernel is a delta function that assumes the incoming rays are solely detected at their incident detector. Those intermediate kernels are considered *underestimations* of the true PSF kernels. The convex combination method used to generate PSF kernels for that study cannot be extended to generate overestimated kernels, thus we chose to rescale reconstruction parameters as we explained in Appendix B. Therefore, although some of the underestimated PSF kernels on that study do not perfectly match the PSF kernels used in this work, the results follow the same pattern: the SUV_{mean} CoV of slightly underestimated kernels showed a 12% increase compared to true PSF.

The present work pursues such a generalized PSF framework in the context of quantitation [305], which in future efforts can be thoroughly evaluated for detection tasks. Additional work is being pursued in this area on the front of heterogeneity analysis, including studies on the effect of PSF modeling on heterogeneity quantification task performance, with the aim of adopting more reproducible and robust shape and textural features and optimizing them for enhanced prediction and prognostic tasks [306].

Positron range is another PET degradation that induces more blur to the system matrix, as it increases the FWHM of the PSF. Some vendor PET scanners use Ge-68 point sources to characterize the PSF as modeled within PSF-reconstruction [139]. However, in comparison, the most popular isotope, F-18, has a relatively small positron range. Therefore, utilizing a PSF kernel obtained from Ge-68 point-sources that has a significant positron range to reconstruct F-18 PET data is approximately equivalent to an overestimated PSF kernel in the reconstruction, since both increase the FWHM of the PSF. It is interesting to note, based on our observations with overestimated PSF kernels, that this may not be a problem in fact, and may effectively lead to improvements in quantitative performance, though further analysis is required to implement the exact model of radiotracers with higher positron range, that is left for future study.

2.3.6. Conclusion

PSF-modeling is an increasingly employed partial volume correction method. We studied the impact of an array of projection-space-based PSF models on PET reconstructed images for optimized quantitative task performance. The system PSF was constructed using models of photon non-collinearity, inter-crystal scattering, and inter-crystal penetration, as well as the patient attenuation map and scanner normalization sinogram. Using these models, we generated 20 generalized PSF-modeling kernels – including no and true PSF, as well as 8 underestimated and 10 overestimated PSF kernels. We used an XCAT anthropomorphic phantom with 6 different liver tumor sizes and kinetically derived [^{18}F]-FDG time-activity curves to reconstruct noise-free, as well as 200 noisy images using the OS-EM algorithm. The quantitative figures of merit included contrast recovery, mean-squared error, and various noise metrics (image roughness, voxel variation, SUV_{mean} and SUV_{max} coefficient of variability (CoV), averaged max-min difference) and biases (SUV_{mean}, SUV_{max}). We evaluated these metrics for different tumor sizes/iterations/PSF kernels/noise realization.

The results of our main study follow those of our preliminary image-based PSF modeling presented in section 2.2 earlier in this chapter. Our results show that **for the standard range of iterations employed in the clinic (not excessive), edge enhancement due to overestimation counter-intuitively lowered SUV bias in small tumors, while inter-voxel correlations suppressed image roughness and enhanced uniformity in all tumors**, only slightly degrading SUV_{mean} reproducibility in the smallest tumors. One may at first imagine that overestimating the PSF would lead to higher overshoots at the edges. However, we only observed this at higher iterations. In fact, using an overestimated PSF resulted in improved contrast and limited edge overshoot effect at lower iterations, in turn enabling enhanced SUV quantitation. **Overall, our work suggests that one ought *not* necessarily to utilize an exactly matched system PSF for enhanced image reconstruction performance, and that slightly overestimated PSF modeling can improve PET image quantitation**. Our proposed framework can as such be pursued as a powerful and viable approach in quantitative task-based optimization including in prognostication and treatment response assessment.

Part II: Enhanced Quantitation using Advanced Image Processing and Radiomics

3. Standardized Radiomics in Nuclear Medicine Imaging

3.1. Introduction

"Images are more than pictures, they are data" [244]. Radiomics refers to the mapping of images to data and then mining that data towards improved clinical decision support. Initially, radiomics was utilized for oncologic studies, but it has expanded to other clinical areas. Also, while first efforts on radiomics were mostly performed on high-resolution images from CT and MRI, starting in 2009 radiomics-type approaches were also studied in PET imaging, and later on, our group applied radiomics to SPECT imaging [291]. In the field of radiomics, image features (also known as radiomic features) are extracted from regions of interest (ROIs). An ROI can consist of a functional region; e.g. a tumor (possibly also including its immediate background), or an anatomic region (e.g. putamen in brain studies, left anterior descending segment of the heart, or other tissues of interest). However, this relatively new field faces substantial challenges on its path to routine clinical usage [244].

Chapter 1 provided an introduction to radiomics and some of its applications (section 1.4.3). In this chapter, we elaborate on the radiomics workflow. Then, we discuss the importance of standardized and reproducible radiomics, following by our collaborative efforts with the image biomarker standardization initiative (IBS) to standardize radiomic features and processes. Finally, we present two of our studies applying standardized radiomics to PET and SPECT clinical datasets, where we perform radiomics feature selection, a first and important step towards reliable radiomics analysis.

3.1.1. Motivation

As we explain in detail in section 3.2, radiomics workflows suffer from different sources of variability, and it is crucial for a systematic and reliable radiomics study to account for and present a method to correct for all these variabilities. In this chapter, we describe radiomics standardization efforts and our contribution to a global initiative aiming at standardizing radiomics studies (section 3.2.3) that have been published in the form a global collaboration effort [311, 325, 333, 429]. We have also released our work as a software package entitled Standardized Environment for Radiomics Analysis (SERA). Furthermore, in section 3.3, we present our studies on reproducibility of radiomic features in two distinct nuclear medicine projects, some of which we have previously presented to the community [314, 317].

3.1.2. Radiomics workflow

Figure 3-1 shows a sample radiomics workflow. Following the acquisition of high quality and standardized image, the ROI is first defined and segmented, since most radiomics analyses are defined for ROIs, not the entire image. The segmentation task can be performed manually by an imaging expert such as a radiologist or radiation oncologist for a CT or MRI image, or a nuclear medicine physician for a PET image. It can also be performed semi-automatically or be fully automated, but it is recommended for these efforts to be supervised by an imaging expert. The segmented ROI may then need to be resampled to a cubic voxel size, and this resampling also involves interpolation.

In the next step, radiomic features are calculated automatically from the segmented ROI. Radiomic features can be from different feature families (or classes) based on the property that they characterize from the ROI. Features families are namely divided into first-order, and higher-order features. The first order features are those that are directly calculated from the ROI. Examples include

morphological feature family such as volume, surface area, etc. of an ROI. Higher-order feature families capture more complex characteristics of a region, and they are calculated first by mapping the ROI into a specific matrix, and then calculate features for that matrix. An example of such feature families is a grey-level zone-size matrix (GLSZM) that is a matrix that records the counts of zones with the same values, and then several features are calculated for this matrix, such as small-zone emphasis, large-zone emphasis, and so on. Some 1st-order radiomic features and all higher-order radiomic features require preprocessing of the ROI prior to feature calculation. Such preprocessing procedures, as well as more details about radiomic feature classes, are discussed in section 3.1.3.

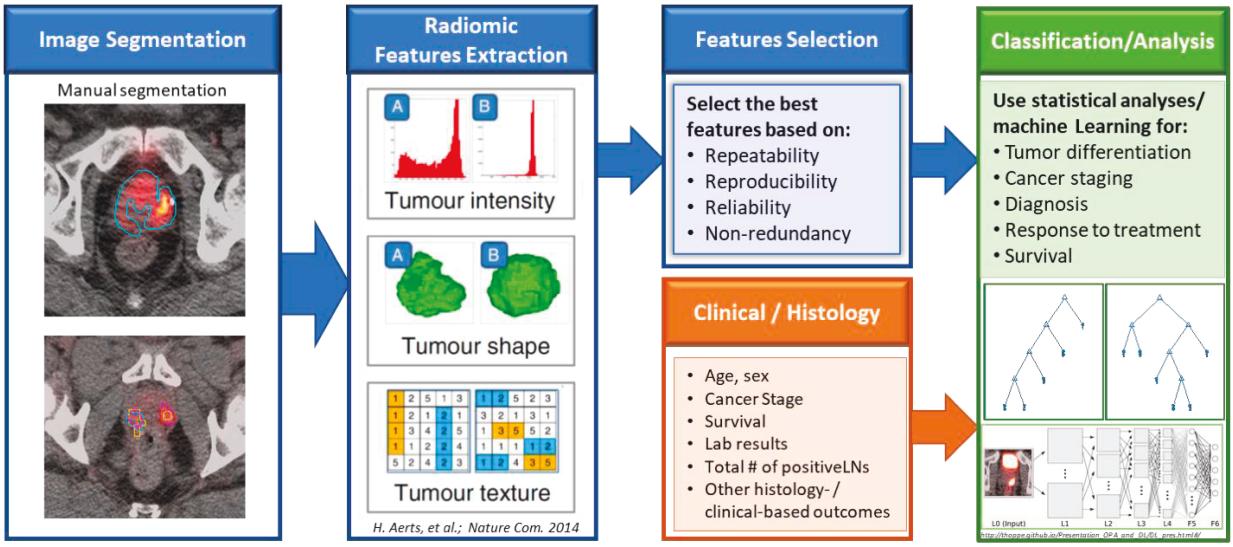

Figure 3-1. A typical radiomics workflow. Following image acquisition, the ROIs are segmented, then radiomics features are calculated from the segmented ROI. Features are narrowed down in the feature selection step.

Following feature calculation, a statistical model is used to select "useful" features and to eliminate features that are redundant, identical, highly correlated, or useless in any other sense. This is an important phase to prevent overfitting of data. In the last step, different statistical techniques can be used to: classify the features, or correlate them with an outcome, or predict that outcome, or enhance patients' stratification. Some other clinically-relevant data such as patient clinical history, risk factors, lab test results, etc., may also be incorporated along with radiomic features.

3.1.3. A brief introduction to radiomic features calculations

The detailed procedures at every step of the radiomic workflow, including the definitions of over 100 essential radiomic features, are not provided in this dissertation, and the reader can refer to the IBSI documentation for these details [429]. In this section, we briefly describe some of the preprocessing procedures prior to feature calculation following by the main radiomic feature classes, and the name selected features from each class.

3.1.3.1. Image preprocessing

Below we briefly describe some image preprocessing steps prior to feature calculations.

Segmentation

Segmentation can be performed manually by radiology or nuclear medicine expert, or semi-automatically using methods such as gradient-based approaches, thresholding based on maximum voxel (e.g. SUV_{max}), thresholding based on deviation from reference (liver) uptake, or performed automatically using available deterministic techniques and/or deep learning. The segmentation mask would be a 3D binary matrix of the same size as the image following a similar grid system.

Interpolation

Texture features may require resampling to have isotropic (cubic) voxel sizes to be rotationally invariant. It also enables a comparison between analyses from different samples and cohorts. Resampling requires voxel interpolation, and it affects image feature values since they are sensitive to variations in the voxel size [430-432].

Interpolation algorithms map voxel intensities from the original image grid to an interpolation grid, where voxels are spatially represented by their centers. Commonly-used interpolation algorithms include:

- Nearest neighbors
- Linear
- Cubic convolution
- Cubic Spline

More details on interpolation processing including grid orientation and grid translation can be found in the IBSI document [429].

Re-segmentation

Re-segmentation is a process following image segmentation and interpolation in which certain voxels may be removed from the ROI if they are outside of a certain range. One example that demonstrates the benefit of re-segmentation is excluding air voxels from an ROI of a lung tumor. Two main methods to perform re-segmentation are:

Range re-segmentation

This method excludes voxel intensities beyond a predefined interval (e.g. exclude voxels with intensities outside of $[-500,400]$ HU in a lung tumor radiomics study.

Intensity outlier filtering

In this re-segmentation method voxel with intensities that are considered outliers may be removed from the ROI. One method is to calculate the mean μ and the standard deviation σ of the grey-levels (defined below) inside the ROI and then exclude voxels outside of the range $[\mu - 3\sigma\ , \mu + 3\sigma]$ [264].

The ROI is called **the morphological mask** before performing re-segmentation and after re-segmentation is referred to as **the intensity mask**.

Intensity discretization

ROI voxel intensity discretization or quantization is often required for calculation of higher-order texture and heterogeneity features to make them tractable and suppress noise [245]. During the discretization process, voxel intensities are discretized into new values called **grey-levels (GL)**. They are two commonly-used approaches to ROI discretization:

Fixed bin-number discretization

In this method, voxel intensities $X_i, i = 1, \dots, N_{ROI}$, where N_{ROI} is the total number of voxels in the ROI, are discretized into a fixed number of N_g bins as follows:

$$X_i^D = \begin{cases} \left\lfloor N_g \dfrac{x_i - X_i^{\min}}{X_i^{\max} - X_i^{\min}} \right\rfloor & x_i < X_i^{\max} \\ N_g & x_i = X_i^{\max} \end{cases} \qquad 3.1$$

where $X_i^{\min}$ and $X_i^{\max}$ are the minimum and maximum values of all $x_i, i = 1, \dots, N_{ROI}$.

Fixed bin-number discretization imposes some attributes to the ROI. First, it removes the relationship between image intensity and the physiological underlying matter (if any). Second, it has a normalizing effect which is an advantage in case of imaging modalities with arbitrary units such as MRI and SPECT. Third, certain classes of radiomic features are highly dependent on the number of GLs, and fixed bin-number discretization enables a direct comparison of feature values across multiple ROIs or multiple patients.

Fixed bin-size (or fixed bin-width) discretization

This discretization type is a simple concept where a new bin is assigned for every intensity interval with a width w_b. In this case, w_b is the *bin width*, and it starts at a minimum fixed GL referred to as

X_{gl}^{min}. This parameter has a user-set value and can be set to zero, another fixed value, or to be the same as the lower-bound of the re-segmentation interval $X_i^{\min}$ as defined in the previous subsection. Here is the formulation for fixed bin-size discretization:

$$X_i^D = \left\lfloor \frac{x_i - X_i^{\min}}{w_b} \right\rfloor + 1 \qquad 3.2$$

It is highly recommended to set a fixed X_{gl}^{min} across all ROIs in the dataset to main consistency between samples. This can be 0 SUV for PET images and -500 HU for CT images. Fixed bin-size discretization is not recommended for use in case of arbitrary-unit modalities such as MRI and SPECT. Unlike fixed bin-number, the current method has the advantage of maintaining a direct relationship with the original units of the image. This can become helpful in case of functional imaging modalities such as PET. Examples of fixed bin-size bin width w_b is 25 HU for CT images and 0.5 SUV for PET images. An example of a discretized and resample ROI is given in Figure 3-2.

3.1.3.2. Radiomic features

Radiomic features are divided into several feature classes or feature families. Table 3-1 outlines some popular feature classes and the number of features per class, as pursued by the IBSI [429]. We briefly describe the feature classes and only name a few features inside each class. For a complete definition of features and their formulations, please refer to the IBSI document [429].

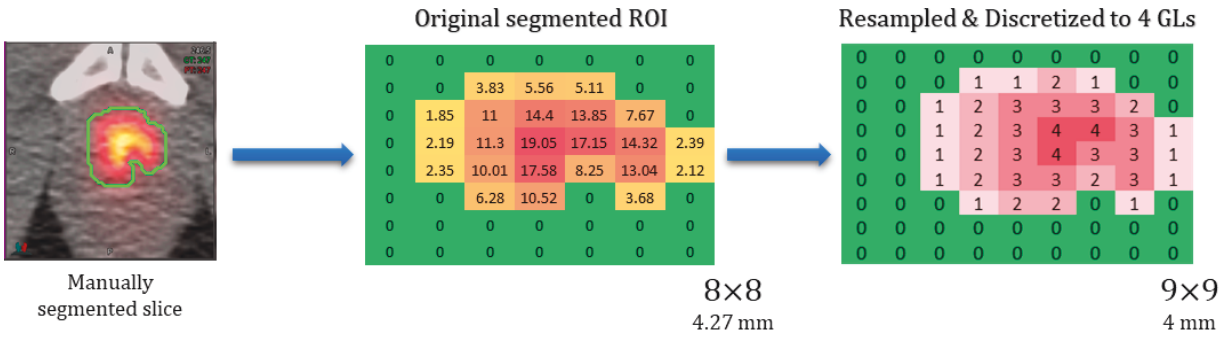

Figure 3-2. A 2D ROI over a prostate tumor on a PET image. The original segmented ROI has SUV values (middle matrix) with a certain size. The ROI is then resampled to a finer resolution and discretized with a fixed bin-size of $w_b = 5$ SUV.

Table 3-1. List of radiomic features in compliance with IBSI guidelines. For details on subtypes, refer to section Grey-level co-occurrence matrix (GLCM) on page 128.

	Feature Family	Subtypes	Number of Features
First-order features	Morphology	-	29
	Local Intensity	-	2
	Intensity-based Statistics	-	18
	Intensity Histogram	-	23
	Intensity-Volume Histogram	-	7
Higher-order features	Gray Level Co-occurrence Matrix (GLCM)	2D Averaged 2D Slice-Merged 2.5D Direction Merged 2.5 D All Merged 3D Averaged 3D Merged	25 25 25 25 25 25
	Gray Level Run Length Matrix (GLRLM)	2D Averaged 2D Slice-Merged 2.5D Direction Merged 2.5 D All Merged 3D Averaged 3D Merged	16 16 16 16 16 16
	Gray Level Size Zone Matrix (GLZSM)	2D 2.5 D 3D	16 16 16
	Gray Level Distance Zone Matrix (GLDZM)	2D 2.5 D 3D	16 16 16
	Neighborhood Grey Tone Difference Matrix (NGTDM)	2D 2.5 D 3D	5 5 5
	Neighboring Grey Level Dependence Matrix (NGLDM)	2D 2.5 D 3D	17 17 17
	Total		**487**

Morphological features

Morphological features or shape features describe geometric aspects of an ROI. These features are based on ROI voxel representations of the volume. Voxel representations of morphological features can be set as coordinates of the voxel centers, or a surface mesh representation of the ROI. The latter

can be performed by different algorithms including Marching Cubes [433]. These different definitions can impact the calculated values. For instance, the definition of "volume" based on the mesh representation calculates the volume of the space contained by the triangulate mesh surrounding the ROI, whereas the volume based on the voxel representation would be simply the number of voxels times the volume of a single voxel. The IBSI guideline refers to the first approach as "volume" and the second approach as "approximated volume" since the mesh representation provides a value nearer to the truth. Below is the list of some morphological features.

1- Volume
2- Approximate volume
3- Surface area
4- Surface to volume ratio
5- Compactness 1
6- Compactness 2
7- Spherical disproportion
8- Sphericity
9- Asphericity
10- Centre of mass shift
11- Maximum 3D diameter
12- Major axis length
13- Minor axis length
14- Least axis length
15- Elongation
16- Flatness
17- Volume density-AABB
18- Area density-AABB

19- Volume density-OMBB

20- Area density-OMBB

21- Volume density-AEE

22- Area density-AEE

23- Volume density-MVEE

24- Area density-MVEE

25- Volume density-convex hull

26- Area density-convex hull

27- Integrated intensity

28- Moran's I index

29- Geary's C index

Most of the above features solely depend on the morphological ROI (as defined earlier), and are independent of voxel values. For the definition and formulation of these features please refer to the IBSI documentation [429].

Local Intensity features

Local intensity refers to two specific features that are computed based on the voxel intensities within a defined neighborhood around a centered voxel. The centered voxel should belong to the intensity ROI, but the neighboring voxels may be from outside of the intensity ROI if the centered voxel is at the edge of the ROI. The two local intensity features are:

1- Local intensity peak

2- Global intensity peak

The former is the maximum mean intensity of a $1cm^3$ sphere, that is, voxels found in a set of voxels within a radius of $r = \left(\frac{3}{4\pi}\right)^{1/3} \approx 0.62$ cm of the centered voxel. The difference between the two is that in the local intensity peak, the center voxel should belong to the intensity ROI, whereas in the global peak, the centered voxel can be from anywhere on the image. The former is close to the definition of SUV_{peak} in PET imaging.

Intensity-based statistical features

These features describe how voxel values within the ROI are distributed. These voxel-values do not need discretization. They intensity-based features would not be meaningful in the case of arbitrary-unit modalities such as MRI and SPECT. Below is a list of these features. For their formulation please refer to the IBSI document [429].

1- Mean
2- Variance
3- Skewness
4- Kurtosis
5- Median
6- Minimum
7- Percentile 10
8- Percentile 90
9- Maximum
10- Interquartile range
11- Range
12- Mean absolute deviation
13- Robust mean absolute deviation

14- Median absolute deviation

15- Coefficient of variation

16- Quartile coefficient of dispersion

17- Energy

18- Root mean square

Intensity histogram features

To calculate these features, first, we have to generate an intensity histogram by discretizing the original set of voxel intensities x_i into grey-level bins based on one of the approaches described in section 3.1.3.1. Subsequently, a range of features can be calculated from the histogram:

1- Mean

2- Variance

3- Skewness

4- Kurtosis

5- Median

6- Minimum

7- Percentile 10

8- Percentile 90

9- Maximum

10- Mode

11- Interquartile range

12- Range

13- Mean absolute deviation

14- Robust mean absolute deviation

15- Median absolute deviation

16- Coefficient of variation

17- Quartile coefficient of dispersion

18- Entropy

19- Uniformity

20- Maximum gradient

21- Maximum gradient grey level

22- Minimum gradient

23- Minimum gradient grey level

Intensity-volume histogram (or cumulative histogram) features

The (cumulative) intensity volume histogram (IVH) is generated by producing a cumulative histogram from the distribution of discretized voxels into bins. It describes the relationship between a GL and the fraction of the volume of the histogram containing at least that particular GL. IVH is the only feature class that may have its own discretization type. For consistency, it is recommended to set the range of the IVH to be the same as the re-segmentation range. The following features can be calculated from IVH:

1- Volume fraction at 0.10 intensity

2- Volume fraction at 0.90 intensity

3- Intensity at 0.10 volume

4- Intensity at 0.90 volume

5- Difference volume fraction at 0.10 and 0.90 intensity

6- Difference intensity at 0.10 and 0.90 volume

7- The area under the IVH curve

Below we continue with higher-order feature classes. For brevity, we only explain three higher-order classes of features through the rest of this section. Readers can refer to the IBSI guideline for definitions of other higher-order feature classes [429] which we have implemented in our SERA software.

Grey-level co-occurrence matrix (GLCM)-based features

These features were originally developed to assess the texture of a surface in 2D images, but they can be extended 3D objects as well. Voxel intensities are usually discretized prior to calculation of the texture features with methods that were explained in section 3.1.3.1.

The grey-level co-occurrence matrix (GLCM) is a matrix that expresses how two voxels "co-occur" with respect to one another; in other words, it expresses how combinations of discretized GLs of *neighboring* voxels are distributed along a certain direction and with a certain distance [434]. GLCM with a neighboring distance of 1 has a 26-connected neighborhood in 3D and 8-connected neighborhood in 2D, yielding 13 unique directions in 3D and four in 2D. Thus, for a 3D approach with a distance of one, an ROI has 13 unique GLCMs for every 13 direction.

GLCM calculation across a certain direction is as follows. Let M_m be the $N_g \times N_g$ GLCM matrix with N_g number of GLs present in an ROI intensity mask, and m is a certain direction from one of the possible choices we presented in the previous paragraph. The element (i, j) of the GLCM matrix M_m contains the frequency at which combinations of GL i and GL j co-occur in neighboring voxels along the direction m. The GLCM, consequently, is a symmetric matrix. An example of GLCM with some highlights to help understand the matrix evaluation is presented in Figure 3-3.

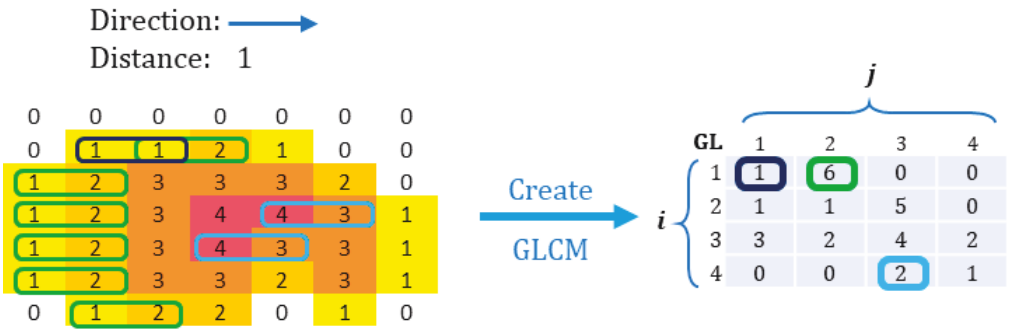

Figure 3-3. A sample GLCM 2D matrix for direction (1,0) and distance 1. GLCM is calculated on the discretized intensity ROI. Each element of the GLCM matrix is the frequency of neighboring voxels i and j in the given direction and the given distance. For example, the GLs in this image are between 1 to 4. they are 6 co-occurrence of GL "1" with GL "2" in the "right" direction with distance one in the figure above as marked by green color.

GLCM features are then derived from the GLCM matrix. Below is a list of GLCM features. The formulation of GLCM features is presented in [429].

1- Joint maximum

2- Joint average

3- Joint variance

4- Joint entropy

5- Difference average

6- Difference variance

7- Difference entropy

8- Sum average

9- Sum variance

10- Sum entropy

11- Angular second moment

12- Contrast

13- Dissimilarity

14- Inverse difference

15- Inverse difference normalized

16- Inverse difference moment

17- Inverse difference moment normalized

18- Inverse variance

19- Correlation

20- Autocorrelation

21- Cluster tendency

22- Cluster shade

23- Cluster prominence

24- Information correlation 1

25- Information correlation 2

Each of the above features is derived from GLCM and returns a single value; i.e. each of the above features would have 13 values for each of the 13 GLCMs for 13 directions. GLCM can be calculated both in 2D with 4 different directions and in 3D with 13 directions, yielding 4 2D and 13 3D features, respectively. But only one value is ultimately reported as the final value, which means either GLCMs or features should be somehow combined. According to IBSI guidelines, there are six different methods to aggregate GLCMs and arrive at a single feature value, which results in six different feature calculation options as mentioned in Table 3-1. These methods are:

1- **2D, averaged:** Features are first computed from each 2D directional matrix. Then average feature over all directions and all slices.
2- **2D, merged:** First merge all 2D directional GLCMs over all directions per slice, then calculate features for this single merged GLCM, then average calculated features over all slices.
3- **2.5D, averaged:** First merge all 2D directional GLCMs over all slices per direction, then calculate features for this single merged GLCM, then average calculated features over all directions.

4- **2.5D, merged:** First merge all 2D directional GLCMs over all slices and all directions, then calculate the feature based on this single merged GLCM.

5- **3D, averaged:** Compute feature for every 3D directional GLCM, then average over all calculated features.

6- **3D, merged:** Merge all 3D GLCMs, then calculate a single value from this merged GLCM.

This means that the single class of GLCM can result in $25 \times 6 = 150$ different features. This emphasizes the high complexity and significant potential variability between studies reporting feature performance. Such variability in feature computations poses a major challenge towards reproducible radiomics studies.

Grey-level run length matrix (GLRLM)-based features

The grey-level run length matrix (GLRLM) captures another aspect of image feature [435]. Similar to the GLCM, GLRLM also requires a discretized ROI to start. But unlike GLCM that assesses co-occurrence of GLs within a neighborhood of voxels, GLRLM assesses run length or frequencies. A run length is the length of a consecutive sequence of voxels with the same GL along a certain direction. GLRLM directions are similar to GLCM directions. As such, just like GLCM had six different methods to average and merge matrices and slices over 2D and 3D, GLRLM has the same approach in computing its features. Figure 3-4 shows an example of a GLRLM matrix.

Figure 3-4. An example of GLRLM generation. The generated GLRLM is based on the (1,0) direction. Each element of GLRLM refers to the number of instances that a certain GL (rows) has appeared in the ROI with a certain run-length (columns). For example, there are only 2 instances of GL "3" with run-length of "2" as marked by green color.

Below is a list of IBSI features for GLRLM:

1- Short runs emphasis
2- Long runs emphasis
3- Low grey level run emphasis
4- High grey level run emphasis
5- Short run low grey level emphasis
6- Short run high grey level emphasis
7- Long run low grey level emphasis
8- Long run high grey level emphasis
9- Grey level non-uniformity
10- Grey level non-uniformity normalized
11- Run length non-uniformity
12- Run length non-uniformity normalized
13- Run percentage
14- Grey level variance
15- Run length variance
16- Run entropy

Elaborating each feature is outside of the scope of this thesis, yet, we briefly describe an interesting pattern for generation of some GLRLM features that make them more intuitive to understand and occurs in other higher-order classes such as GLSZM, GLDZM and NGLDM. The first eight GLRLM features with an *emphasis* on high or low GLs and/or run-lengths are calculated by distributing weights in different patterns on the GLRLM matrix and then performing a weighted average based on these patterns to arrive at a single number. Figure 3-5 can be used to explain this. In the figure, darker colors represent higher weights, and lighter color represent lower weights. We observe that e.g. for short runs emphasis (SRE), more weight is given to the left side of the GLRLM where it

represents the elements with shorter run-length. As another example, the long run low grey-level emphasis (LRLGLE), gives more weight both to (i) "long runs", i.e. towards the right of GLRLM, and (ii) to "low grey-levels", i.e. towards the upper side of the matrix. Therefore, the pattern is to give more weights towards the upper-right side of the GLRLM.

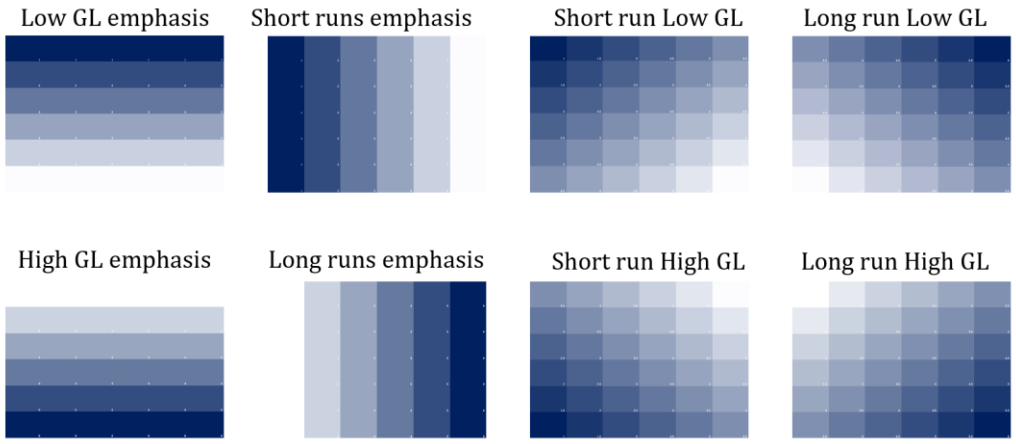

Figure 3-5. The first 8 GLRLM features can be visualized by these weighting patterns for a 6×6 GLRLM. For example, short-run low GL, puts more weight towards "short run-lengths" that is towards the left side of the matrix, and "low GLs" that is towards the upper side of the matrix, creating a weighting pattern towards the upper-left side of the matrix, and calculating the feature based on this weighted-average.

Grey-level size zone matrix (GLZSM)-based features

The grey-level size-one matrix (GLSZM) counts the number of zones, or connected components, of linked voxels. Voxels are considered in the same zone if their neighboring voxel has the same GL. As such, GLSZM, like other higher-level feature classes, require discretization of the ROI. However, unlike GLCM and GLRLM, GLSZM is not directionally-dependent.

Let M be the $N_g \times N_z$ GLSZM with N_g representing the maximum GL present in the ROI intensity mask, and N_z is the maximum zone size of any connected group of voxels. Then, an element of this matrix $s_{i,j}$ represents the number of zones with GL i and size j. Figure 3-6 shows an example of GLSZM generation.

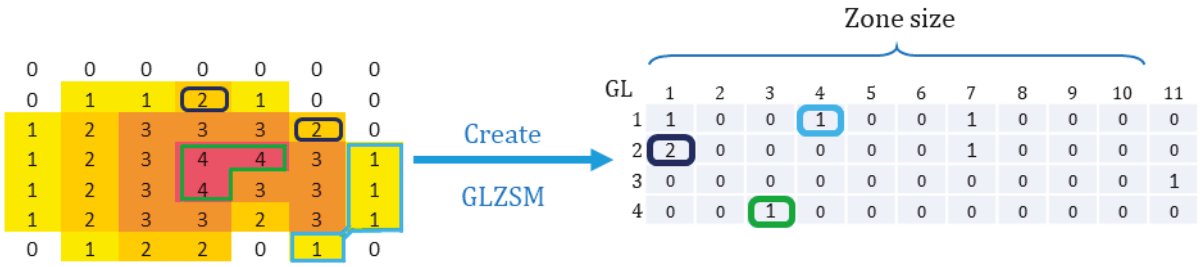

Figure 3-6. An example of GLSZM. Each element (i, j) represents the number of zones of connected voxels with grey level i (rows) and zone size j (columns). For instance, they are two instances of grey level 2 with size 1, as it marked with a black color.

Then a number of features can be generated from the GLSZM [429]:

1- Small zone emphasis

2- Large zone emphasis

3- Low grey level emphasis

4- High grey level emphasis

5- Small zone low grey level emphasis

6- Small zone high grey level emphasis

7- Large zone low grey level emphasis

8- Large zone high grey level emphasis

9- Grey level non-uniformity

10- Grey level non-uniformity normalized

11- Zone size non-uniformity

12- Zone size non-uniformity normalized

13- Zone percentage

14- Grey level variance

15- Zone size variance

16- Zone size entropy

GLSZM can be calculated both for 2D and 3D ROIs. In the case of 2D calculation, once the feature has been determined for every slice, the overall value of a feature is calculated by averaging that value

across all slices. The first 8 GLSZM features have the same intuition behind as the first 8 features of GLRLM as discussed in the previous page. They are based on weighted-average of GLSZM with a certain pattern described by the feature name, as shown in Figure 3-7.

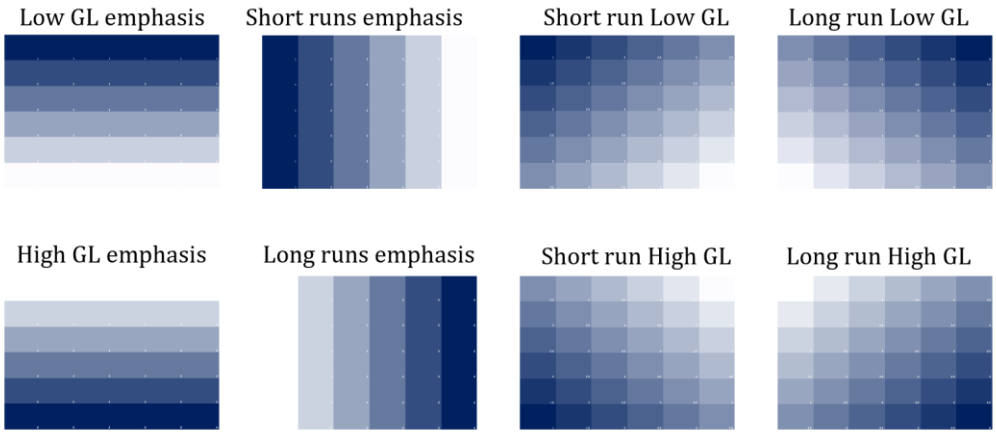

Figure 3-7. Weighing patterns of the first 8 GLSZM features. A similar approach to Figure 3-5.

Other higher-order feature classes

As mentioned before, there are other more complex higher-order radiomic feature classes that we do not elaborate here, and readers can refer to the IBSI guideline for detailed elaboration [429]. In a nutshell, these feature classes are as follows:

Grey level distance-zone matrix (GLDZM)-based features

GLDZM counts the number of groups of connected voxels with a specific GL value and distance to ROI edge. It contains an extra-level of information compared to GLSZM: it captures the relation between the distance from the edge, zone-size, and GL. Sixteen feature are derived from GLDZM.

Neighboring grey level dependence matrix (NGLDM)-based features

NGLDM captures the *coarseness* of the overall texture. Seventeen features are derived from the NGLDM.

Neighborhood grey tone difference matrix (NGTDM)-based features

NGTDM contains the sum of GL differences of voxels with a specific GL and the average discretized GL of neighboring voxels within a distance. Five features are computed from NGTDM.

Overall, in our SERA product, we defined 487 unique radiomic features in 11 classes. These features can be calculated with different discretization levels (either fixed bin-size or fixed bin-number). Under the assumption that only two discretization levels are considered, the number of features in our parameters space is easily increased to 1000. At the same time, most radiomics studies hardly include more than a few hundred data points. As such, the important question is: Can we use these hundreds or thousands of features to correlate with or predict clinical outcome?

3.2. Properties of Responsible and Reproducible Radiomics Research

Although radiomics is a relatively young discipline and has experienced relatively fast growth (Figure 1-18), it has not been yet been translated to routine clinical practice. This may be due to the *low reproducibility* of most current studies [436]. Radiomics has a complex workflow involving many steps that often suffers from incomplete reporting of methodologic information. Consequently, few radiomics studies available in the current literature are readily reproducible from start to the end [436]. Another major issue is the relatively small number of images in radiomics research datasets that may induce overfitting and high false-positive rates. This further worsens with the tendency to report overly-optimistic results [436].

Guidelines and protocols are available for quality control measures in nuclear medicine imaging to standardize patient preparation, dose production and administration, image acquisition, image reconstruction, SUV normalization, etc., such that the absolute SUV values are interchangeable in multicenter studies [216]. To perform radiomics in PET, such standardized SUV measurements are crucial; yet, the methodology to prepare the image and calculate radiomic features is also subject to

variability, showing a crucial need for standardization [247, 437, 438]. Several studies have shown the importance of robust and standardized protocols to enable reliable quantification of heterogeneity with textural features. They demonstrated an important need to standardize the computation methods due to the complexity of the radiomics workflow [216, 247, 436, 439]. In what follows, first, we briefly mention the causes of variability in radiomics workflow, following by an introduction to the IBSI effort, and then elaborate on the properties for a responsible radiomics research.

3.2.1. Causes of variability in the radiomics workflow

A serious challenge with the calculation of radiomic features is the very large number of features that can theoretically be calculated and different ways to calculate them. In addition, acquisition protocol, scanner variations, quantitative corrections, reconstruction algorithm settings (e.g. PSF vs. no PSF, number of iterations, etc.), post-reconstruction processing (Gaussian post-smoothing), as well as user-induced variations such as ROI definitions are all other sources of variability that increase the complexity and hamper the reproducibility. Figure 3-8 shows a flowchart of a typical radiomics workflow and summarizes a few of the choices that the user confronts during each step of the process.

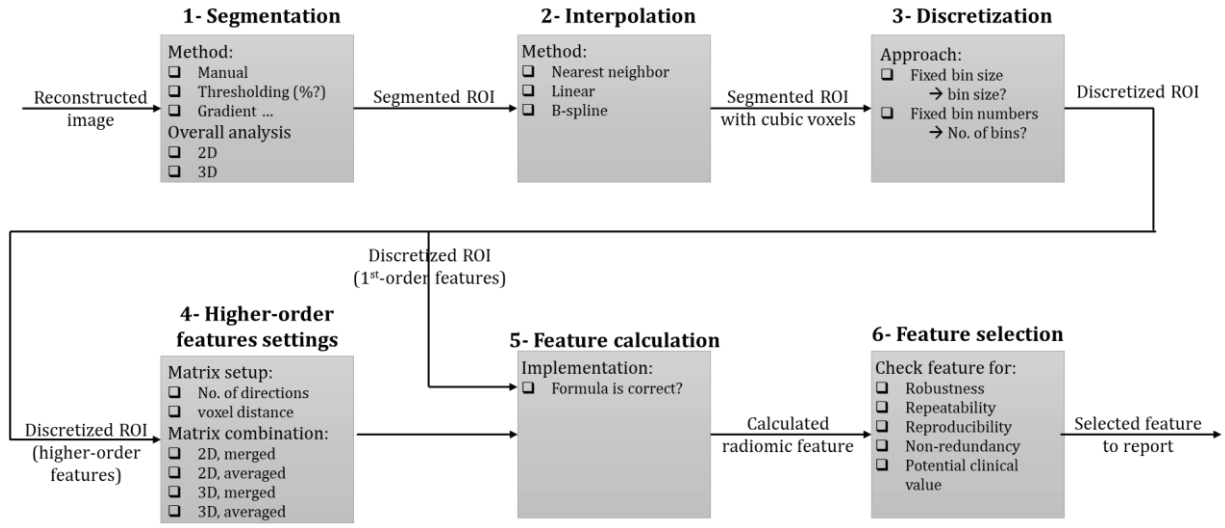

Figure 3-8. Points of variability in a radiomics workflow aiming at radiomic feature calculation and selection.

3.2.2. The image biomarker standardization initiative (IBSI)

Since September 2016, an initiative comprising over 50 researchers from 21 top universities and cancer centers from 8 countries have participated in the image biomarker standardization initiative (IBSI). IBSI aims at standardizing feature computation and image preprocessing phases of the radiomics workflow.

For this effort, first, a simple 5×4×4 digital phantom with 6 GLs was shared to standardize the computation of 172 features from 11 categories [429, 440]. Following that step, a CT image of a patient with lung cancer with a single large lung tumor plus an ROI defining the tumor was shared between all centers. Then, five different configurations of the features were provided (Table 3-2), and centers had to calculate all 487 feature combinations and report them.

Table 3-2. Different configurations of the IBSI standardization effort

	Config. A	**Config. B**	**Config. C**	**Config. D**	**Config. E**
Approach	2D	2D	3D	3D	3D
Interpolation	-	Yes	Yes	Yes	Yes
Voxel dimension	-	2×2	2×2×2	2×2×2	2×2×2
Interpolation Method	-	Linear	Linear	Linear	Cubic
Range re-segmentation	[-500,400]	[-500,400]	[-1000,400]	-	[-1000,400]
Outliers re-segmentation	-	-	-	3σ	3σ
Discretization	FBS	FBN	FBS	FBN	FBN
Bin size	25 HU	32 bins	25 HU	32 bins	32 bins

3.2.3. Standardized Environment for Radiomics Analysis (SERA)

Using the comprehensive guidelines from the IBSI effort, we developed an entirely in-house Matlab®-based framework to perform radiomics analysis, named as Standardized Environment for Radiomics Analysis (SERA). SERA is able to process images from different clinical imaging modalities such as CT, MRI, PET, and SPECT. Radiomic features calculated with SERA are standardized and in compliance with IBSI, which ensures their reproducibility, and have been adopted for a number of published research studies [293, 294, 310, 312, 325-328, 333].

SERA calculates 487 IBSI-standardized features, including 79 first-order features (Morphology, Statistical, Histogram, and Intensity Histogram features), 272 higher-order 2D features, and 136 3D features. In addition to 487 IBSI features, it also calculates 10 Moment Invariant features, that are not included in IBSI. The detailed list of the number of features in each class is the same as Table 3-1.

3.2.3.1. SERA configurations

SERA has options to set and modify all parameters that have been defined or used in the IBSI guideline. The image preparation configuration includes the following:

Resampling and interpolation:

SERA can resample to 2D and 3D isotropic voxel sizes; interpolation algorithm used in resampling image and ROI (nearest/linear/cubic); partial volume threshold (mostly used for CT HU).

Discretization:

Parameters to change include bin size, discretization type (fixed bin-size/fixed bin-numbers), and separate discretization for IVH features.

Other settings

Other settings including grey-level rounding, image re-segmentation (range re-segmentation, outliers).

3.2.4. Results of our IBSI standardization effort

From the 21 participating centers, SERA was one of the only three packages that included 100% of the IBSI features. It took several trials for some of the features to be fine-tuned, where we fixed our definitions and code by receiving feedback from IBSI organizers. Table 3-3 shows the results of our most recent feature calculation on the five categories defined in Table 3-2. It includes a number of features in each category, the number of features SERA matched with IBSI benchmarks, and the overall accuracy of each category.

Table 3-3. Results of our IBSI standardization effort

	Config. A	**Config. B**	**Config. C**	**Config. D**	**Config. E**
No. of features for config	351	351	215	215	215
No. of matched features	320	315	207	209	20
No. of no matched	18	16	3	1	5
No. of partial matched	0	0	0	0	0
No. of features with no consensus	14	20	5	5	191
Accuracy	94.7%	95.1%	98.6%	99.5%	80%

3.2.4.1. Observations

One very interesting and important observation from our IBSI effort is that even with different centers using the same detailed workflow configuration and feature formulation, the reported feature values were highly variable when using the exact same image and the same ROI! This further emphasizes the variability issue in radiomics research, though it also demonstrates that inter-center agreements can increase over time with coordinated efforts. Another observation was more variability in configurations D and E in Table 3-2 where more sophisticated interpolations were used.

3.2.5. Properties of an effective radiomic feature

Aside from standardization of radiomic features, further steps should be taken within a radiomics workflow to reduce the feature-space dimension and to prevent overfitting. Below we discuss some important properties of an ideal radiomics feature, as well as some recommended techniques to assess them [247].

3.2.5.1. Repeatability

Repeatability is related to the variability of a given radiomic feature when obtained under the same conditions; i.e. from two separate scans of the same subject performed close together in time. To assess repeatability, some studies have suggested comparing metrics calculated on test-retest PET images using e.g. the Bland Altman method for comparison. [441, 442].

3.2.5.2. Reproducibility

Reproducibility is related to the variability of a given radiomic feature when obtained under varying conditions; e.g. by varying processing parameters to compute the features (Table 1-5). Section 3.2.3 outlined several causes that may induce variability before, during, or after data acquisition. As an example, different configuration settings within the radiomics workflow itself, such as segmentation

method, discretization type, discretization bin size, etc. can be sources of variability. To assess the reproducibility of radiomic features, they can be calculated through differing analysis pipelines [442]. Features can be computed for different parameter settings, followed by analysis to identify features with substantial variability (i.e. low reproducibility) when varying certain parameter settings. Examples of these analyses will be provided later in section 3.3.

3.2.5.3. Redundancy

Features can be quantified through statistical and rank correlation analyses in order to find and eliminate redundant features [314, 317, 438, 443]. Certain machine learning techniques can be used to select features or combine them [444].

3.2.5.4. Offers value with regards to a given clinical endpoint

Features that survive the above refinement steps should also be analyzed with respect to the tasks or outcomes of interest, e.g. response to treatment, diagnosis, prognosis, etc., to select only those that offer value. A number of robust machine learning techniques are available to use for these regression or classification techniques [445].

3.3. Feature Selection and Reproducibility of Radiomic Features in Nuclear Medicine

In this section, we present two studies on feature selection and reproducibility of radiomic features in SPECT and PET. Each study was performed on a unique nuclear medicine dataset. We aimed to investigate the feature-space to eliminate non-reproducible, or redundant features. Some of these efforts have been presented to the community [314, 317]. At the same time, we did not perform repeatability analysis as it requires test-retest imaging datasets which are areas of ongoing efforts.

3.3.1. Reproducibility of Cold Uptake Radiomics in ^{99m}Tc-Sestamibi SPECT Imaging of Renal Cell Carcinoma

^{99m}Tc-Sestamibi SPECT/CT imaging of renal cell carcinoma (RCC) has recently shown significant promise to distinguish benign oncocytomas from malignant RCC, where the former appears as high uptake in SPECT images and the latter as cold uptake. We aim towards *radiomics analysis of cold uptake* in SPECT images towards another significant yet more daunting task of discriminating between RCC subtypes.

3.3.1.1. Introduction

Recent studies on ^{99m}Tc-sestamibi single-photon emission computed tomography / computed tomography (SPECT/CT) imaging of renal tumors have shown promising results for distinguishing renal cell carcinoma (RCC) from benign oncocytomas [446]. While normal renal tissue has positive radiotracer uptake on ^{99m}Tc-Sestamibi SPECT, these studies reported positive (hot) uptake for benign cases and negative (cold) uptake for RCC. RCC has several histologic subtypes including clear cell, papillary, and chromophobe variants (Figure 3-9). It is an open question as to whether cold ^{99m}Tc-sestamibi uptake has the potential to discriminate between these subtypes of RCC.

Radiomics analysis, via its combination of effective metrics that quantify shape and texture/ heterogeneity, holds significant promise towards identifying patients with higher risk and to help realize personalized medicine. It is been widely studied in oncologic MRI, CT, and PET [229, 291, 424, 447-449], and recently also was explored in brain SPECT imaging by our group [292]. This work presents another novel exploration of radiomics in SPECT analysis. In addition to, it is, to our knowledge, the first application of radiomics to cold uptake imaging. Overall, we aim to preform radiomics analysis of ^{99m}Tc-sestamibi SPECT for the potential to discriminate different RCC types. The purpose of the present specific study is to perform feature selection, by assessing the reproducibility and reliability of radiomic features in ^{99m}Tc-Sestamibi SPECT images, and to derive

robust features for correlation with tumor histology. Radiomics analysis usually involves evaluating hundreds of radiomic features for correlating with clinical data. The present reproducibility analysis aims to exclude radiomic features that are non-reproducible, non-repeatable, and redundant, in order to avoid data overfitting and to enable proper classification with clinical endpoints [247, 314, 317, 442, 445, 450-452].

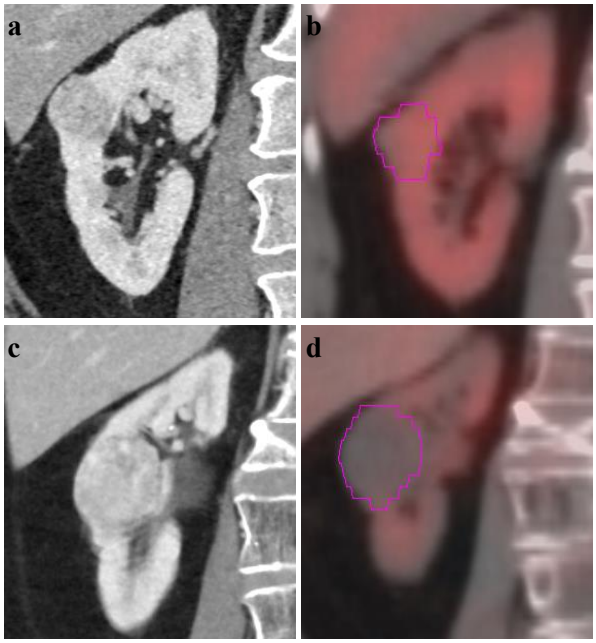

Figure 3-9. a) postcontrast CT image of a patient with Oncocytoma, b) ^{99m}Tc-sestamibi SPECT/CT scan of the patient shows high uptake, c) postcontrast CT image of clear cell RCC (CCRCC), d) ^{99m}Tc-sestamibi SPECT/CT scan of the same patient shows cold uptake.

3.3.1.2. Methods

50 patients presenting with a solid solitary clinical T1 renal mass were imaged with ^{99m}Tc-sestamibi SPECT/CT as part of a prospective study evaluating the diagnostic performance of this imaging test. All patients also had a contrast-enhanced CT or MRI scan. A trained radiologist used contrast-enhanced CT or MRI scans along with co-registered SPECT/CT images to manually segment a region of interest (ROI) around the renal mass on the SPECT image. Segmenting the tumor with a cold uptake on the SPECT image is rather challenging due to three reasons. First, the need for accurately co-

registering two image sets: 1) CT of SPECT/CT and 2) a contrast-enhanced CT or MRI to anatomically localize the mass. Then, both CT images should be co-registered with the SPECT. Second, low-resolution SPECT images and their large voxel sizes require more meticulous contouring to minimize exclusion of tumor voxels. Third, the inclusion of normal renal tissue with high activity right next to the tumor conspicuously disrupts the cold uptake ROI and should be also minimized. At the same time, the segmented cold-uptake ROI should be free from any voxels from the cold-uptake background, too. These reasons evidently show the importance of studying the reproducibility of segmentation for ^{99m}Tc-sestamibi SPECT RCC images.

To study the impact of segmentation on reproducibility, we generated three more ROIs from the manually segmented ROIs, with removing one (shrinkage) and adding one and two voxel-layers (enlargement) to the ROI in all three dimensions. These ROIs were created by thresholding a Gaussian-blurred version of the binary ROIs with different values. Images were then uniformly quantized using the fixed bin-number discretization into eight different gray-levels (GLs) (2^2, 2^3, …, 2^9). We used SERA (section 3.2.3) with all features except 2.5D higher-order features as presented in Table 3-1. Table 1 contains the details about the radiomic features classes.

For the four segmentations, Spearman's rank correlation was calculated between each quantization pair to find relatively consistent quantization levels. The intra-class correlation (ICC) between remaining GLs across all patients was used to adopt robust features for segmentation. Furthermore, features were studied with regards to their characteristics to propose a more practical and reproducible set of features.

3.3.1.3. Results

Reconstructed RCC SPECT images are not expressed in the form of quantitative metrics such as standard uptake value (SUV) in PET imaging. Instead, voxel intensities denote the number of counts,

which might be further different from the true counts due to specific scaling by some scanners. In the current dataset, the range of the maximum voxel intensities was 600~8000. For such a non-normalized dataset, radiomic features assuming normalized voxel intensities are meaningless. Thus, only a subgroup of the non-normalized statistical features will be practical. These features include variance, skewness, kurtosis, the coefficient of variation and quartile coefficient of dispersion. Moreover, a fixed number-of-bins quantization should be performed, as fixed bin-width quantization is impractical due to the same reason of non-normalized images.

We evaluated the intra-class correlation (ICC) of features vs. segmentation to assess their reproducibility to segmentation. We defined two sets of ROIs: all four ROIs, including the manually segmented plus the three resized ROIs denoted by G1, and all ROIs except the shrunk one denoted by G2. We subsequently evaluated the ICC of features to segmentations in G1 and G2 separately. We also evaluated the ICC for different GLs, too. Figure 3-10 shows four plots depicting ICC of features vs. segmentation (both G1 and G2) and for two GLs (32 and 512). The comparison of these ICC values had led us to several interesting observations. The general trend shows a higher ICC for moderate GLs (e.g. 32) as can be seen in Figure 3-10.c and Figure 3-10.d. A large group of features evaluated with higher GLs exhibits lower ICC. This is due to the presence of outlier voxels from the nearby high-uptake tissue that becomes more pronounced in higher GLs causing more variability. Moreover, we observe that ICC generally decreases as we include the shrunk ROI in the segmentation set (i.e. G2 vs. G1) for most features without altering the trend (Figure 3-10). This indicates a high sensitivity of most of the features to segmentation. More precisely, ICC>0.7 and ICC>0.85 for G1 and G2, respectively, leads to exactly 204 features, and these two feature sets have Spearman correlation of one; i.e. adding one more ROI nearly decreases ICC values but does not change their order.

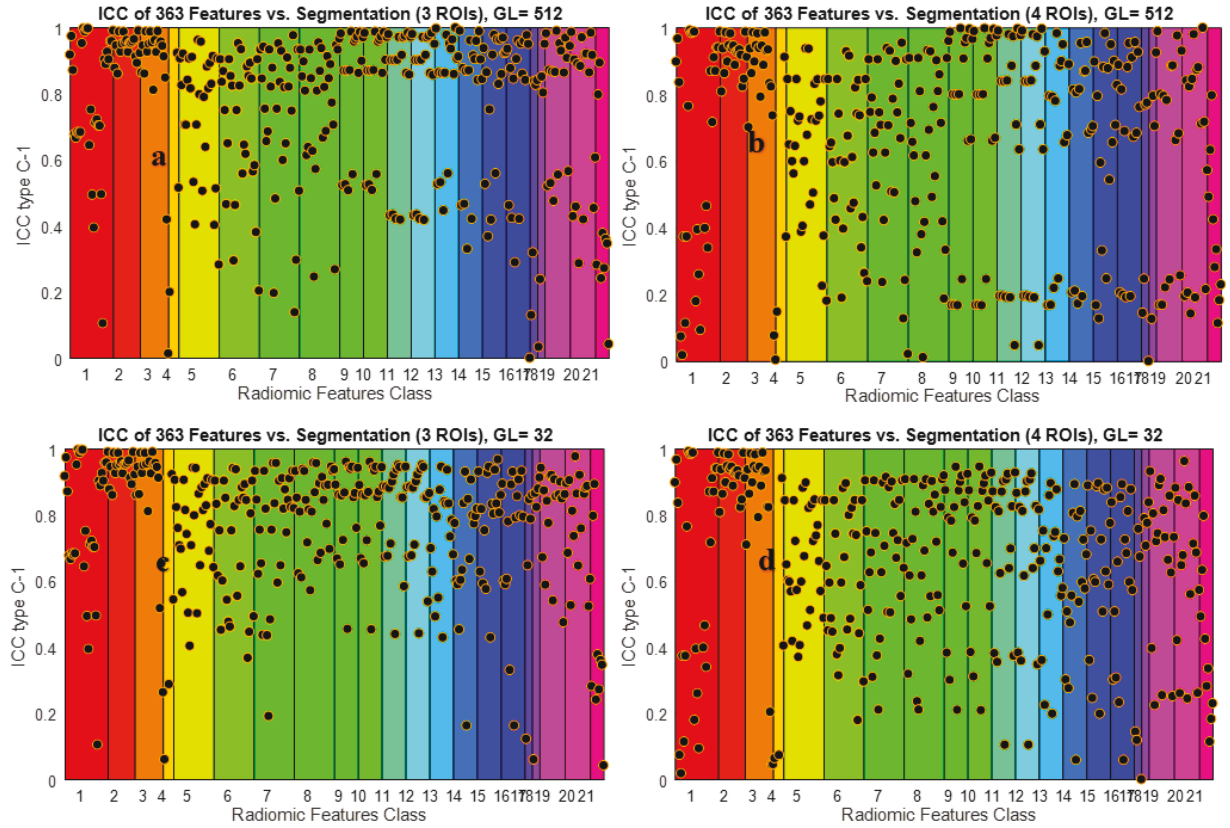

Figure 3-10. ICC type C-1 between all 363 radiomic features and segmentation: a) three segmentations (all except shrunk ROI) with 512 GLs, b) all four ROIs with 512 GLs, c) three segmentations (all except shrunk ROI) with 32 GLs, b) all four ROIs with 32 GLs. Feature classes introduced in Table 1.

A more detailed look at plots in Figure 3-10 shows that the features exhibiting the lowest ICC from higher order feature classes (such as feature groups 9 to 18 from table 1) are those that have higher emphasis on lower GLs, such as low GL run emphasis, short-run low GL emphasis, and long-run low GL emphasis from GLRLM, low GL emphasis, small zone low GL emphasis, and large zone low GL emphasis from GLSZM, and so on. This is rather discouraging since it indicates lower reproducibility of features that primarily focus on information from voxels with lower intensities (RCC cold uptake). On the other hand, some features with the highest ICC are those with an emphasis on high GLs. But despite their high ICC, they should be preferably eliminated too because of their high sensitivity to voxels with high GLs that are may occur due to misregistration/ segmentation. Therefore, a large group of features from the higher order radiomic feature classes with these specifications should be excluded.

Figure 3-11 depicts the Spearman correlation between GL=512 and seven smaller GLs. It shows a relatively consistent Spearman-correlation≥0.5 for gray-levels≥32, suggesting the exclusion of lower GLs. Also, ordering of features was retained for GL larger than 32, but not for 16 and under. At the same time, very high GLs like 256 and 512 should also be avoided due to their sensitivity to the inclusion of high counts and misregistration/segmentation as mentioned above. Therefore, unlike radiomics analysis of the high-uptake images such as in PET, having higher GLs does not necessarily imply better quantification.

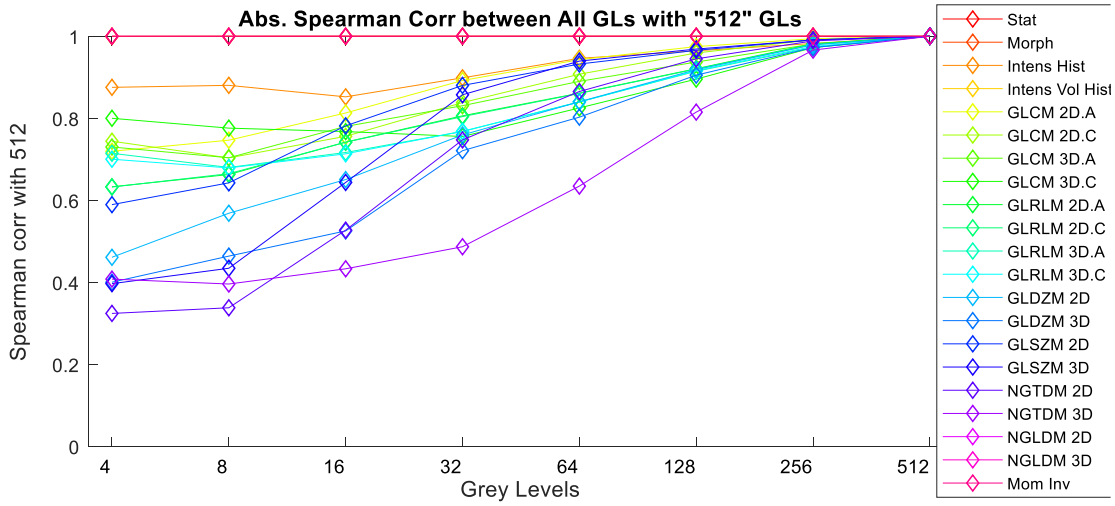

Figure 3-11. Absolute Spearman rank correlation between feature classes calculated with all eight GLs and GL=512. The figure shows consistent corr>0.8 for GL≥32 and all radiomic feature classes except NGTDM 2D

Furthermore, we plotted the Spearman correlation between the top 20 most reproducible features (from the previous steps) and tumor volume in Figure 3-12. We observe a decreasing trend that indicates radiomic features may provide complementary information to tumor volume, which is partially due to the impact of partial volume effect in the reconstructed image.

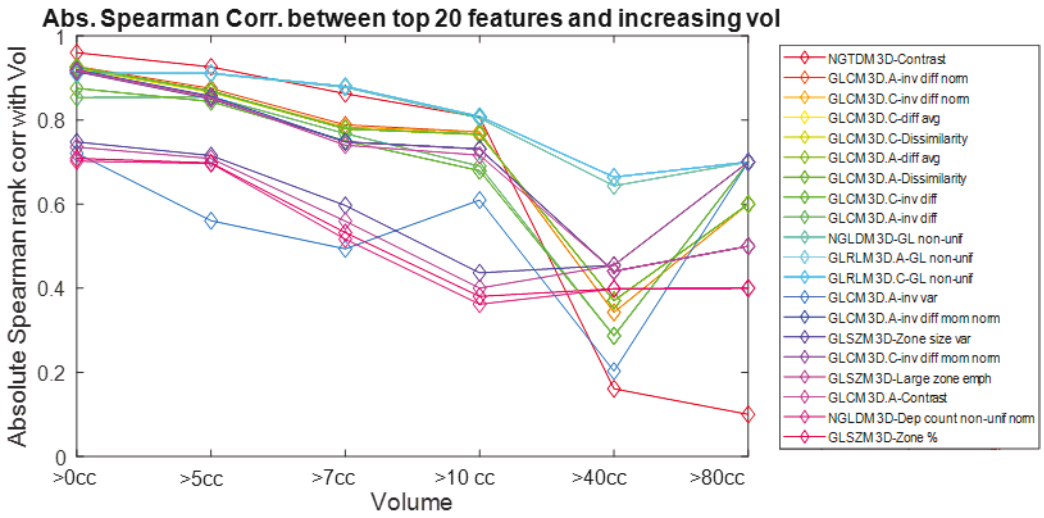

Figure 3-12. Spearman correlation between tumor volume and top 20 most reproducible features with the highest Spearman correlation with volume, GL=512, manual segmentation. Most features exhibit a decreasing trend as the volume increases. Only for volumes>5cc other radiomic features may provide complementary information, that is due to the partial volume effect

3.3.1.4. Discussion

Our aim in this study was to exclude non-reproducible, non-repeatable, and redundant features to avoid data overfitting and to enable proper classification with clinical endpoints. However, our limited dataset size was preventing us to perform further classification or outcome prediction analyses. The current size of our dataset (25 patients) is just enough to perform the kind of analyses we presented in this chapter and is too small to be able to deduce clinically-significant conclusions.

Our results suggested quantization by 32 to 64 GLs for sufficiently capturing heterogeneity information of tumors with cold uptake of ^{99m}Tc-sestamibi. Due to the high uptake present in the normal renal tissue adjacent to tumors ROIs with a cold uptake, most metrics were highly sensitive to segmentation. A subgroup of first-order statistical features should be excluded because voxel intensities in SPECT images are non-normalized. Most morphological, intensity histogram, volume histogram and 2nd-order features including GLCM were reproducible. From the higher-order radiomic features classes, features emphasizing on either low or high GLs should be also excluded. The former group has a very low reproducibility across different segmentations, and the latter group

does not carry much information from a tumor with a cold uptake and is sensitive to misregistration/ segmentation. Instead, features with a rather uniform waiting on all GLs should be adopted. These features are recommended for further investigation of the significant discrimination task amongst RCC tumor subtypes.

3.3.2. Reproducibility of Radiomic features in ^{18}F-DCFPyL PET Imaging of Prostate Cancer

3.3.2.1. Introduction

Prostate cancer (PC) is the second most common cause of cancer-related death in men [453]. PC is often curable; however, many patients experience a residual, recurrent, and metastatic disease that require imaging for diagnosis, lesion detection, therapeutic monitoring, and staging [454]. Prostate-specific membrane antigen (PSMA) is a transmembrane protein that is highly relevant in prostate cancer theranostics because of its marked overexpression in prostate cancer [455]. The past decade has seen significant growth in design, synthesis, and evaluation of radiotracers targeting PSMA for imaging and therapy. An increasing number of studies are demonstrating additional diagnostic value in the primary staging of intermediate- to the high-risk stage and in the late metastatic phase of PC. PSMA PET/CT imaging is capable of localizing lesions for primary PC staging that is not evident on standard-of-care imaging, leading to major impact in early diagnosis and treatment planning [456]. PSMA PET is also shown to demonstrate higher lesion detection rate at lower serum prostate-specific antigen (PSA) levels compared to other PET radiotracers or standard-of-care imaging. This has major implications on biochemical recurrence staging and therapy planning [457, 458].

^{18}F- and ^{68}Ga-labeled inhibitors of PSMA have entered early clinical development for PET imaging of PC since 2012, and have demonstrated great promise for identification of local and distant sites of PC with high sensitivity and high specificity [454, 459]. One of these ^{18}F-labeled PSMA inhibitors, developed at Johns Hopkins University, is 2-(3-(1-carboxy-5-[(6-[^{18}F]fluoro-pyridine-3-carbonyl)-

amino]-pentyl)-ureido)-pentanedioic acid known as [^{18}F]DCFPyL. [^{18}F]DCFPyL or PyL, has shown to be a viable PET agent for PSMA-expressing tissue with high and specific accumulation in primary and metastatic foci [454, 460]. An example of a PyL PET image is shown in Figure 3-13.

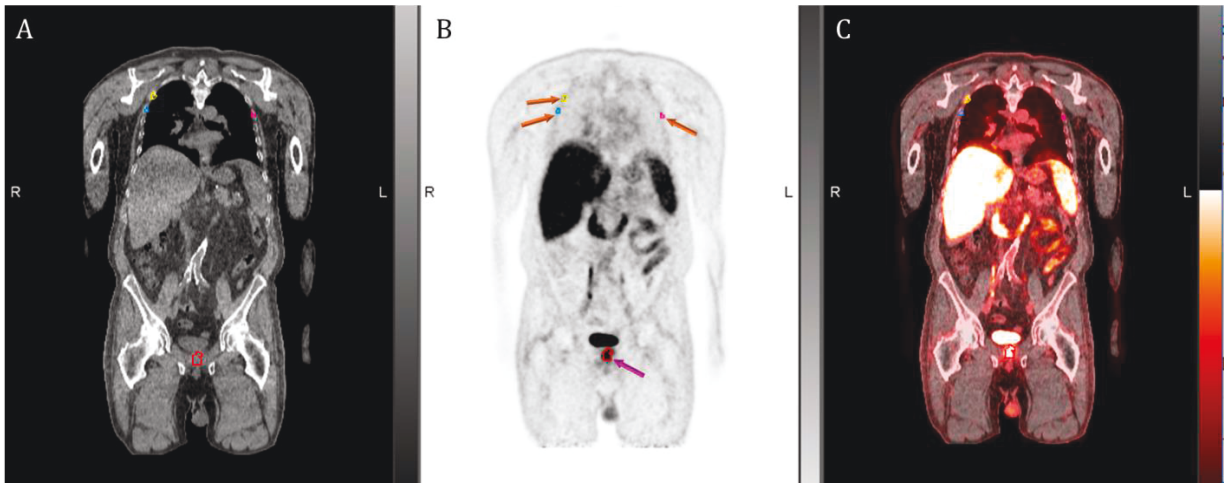

Figure 3-13. An example of a [^{18}F]DCFPyL PET/CT image. "A" shows a coronal CT slice, "B" shows the corresponding PET slice, and "C" shows the fused PET/CT. Primary lesion in prostate gland is shown by a purple arrow in "B", and three metastatic lesions in ribs are shown with orange arrows.

Our motivation in this project is similar to that for the previous study (section 3.3.1.1). The objective is to assess the reproducibility and reliability of quantitative textural imaging features in ^{18}F-DCFPyL PET images, and the relationships between the features.

3.3.2.2. Methods

25 patients within the National Comprehensive Cancer Network with high- or very-high-risk primary PC planned for surgery were imaged with ^{18}F -DCFPyL at 60 min post-injection. To analyze the effect of segmentation in the reproducibility of radiomic features, the prostate tumors were delineated using six different segmentation methods:

- manual delineation performed by a nuclear medicine physician,
- gradient-based semi-automatic segmentation performed by a nuclear medicine physician using MIM® software,

- four threshold-based segmentations ranging from 30% to 60% of SUV_{max}, seed placement by a nuclear medicine physician.

Images were resampled to isotropic cubic voxels of 2 mm for consistency. Two different methods of fixed bin-number discretization (uniform and Max-Lloyd [461]) were considered with seven different GLs (4, 8, 16, 32, 64, 128, and 256) that were used for higher-order textural features extraction. In addition to five conventional quantitative uptake measurements (SUVmax/peak/mean, tumor volume (TV), total lesion activity (TLA)), a total of a selected set of 87 radiomics features were extracted for each ROI per each patient and for each discretization level. These 92 features include: 11 first-order, 9 morphological, 26 GLCM, 12 GLRLM, 13 GLSZM, 5 GLTDM and 10 moment-invariant (MI) features [462] (Table 3-4). Spearman rank correlation analysis was performed for the seven discretization levels and the two methods to determine reliable and practical GLs. We further quantified the intra-class correlation (ICC) to inspect the reproducibility of features across six segmentations and seven GLs of two quantization methods. The statistical relationship between all 92 features was also explored based on the optimum GL and manual segmentation.

3.3.2.3. Results

SUV_{max} derived from primary tumors of 25 patients ranged from 5.6 to 51.8 (15.8±11.5). Tumor volume ranged from 0.5 to 61 cc (7.3cc±11.8cc). Figure 3-14 shows the distribution of SUV_{max} and tumor volume across these 25 patients for each of the 6 segmentation types. The left plot in Figure 3-14 shows an almost identical distribution of the 25 cases across all segmentation, which is due to the very high reproducibility of SUV_{max} with respect to segmentation. The right image shows the proximity of PET-edge segmentation and manual segmentation. It also demonstrates the closes threshold-based segmentation distribution to the manual segmentation is 30% and 40% segmentation.

Table 3-4. List of selected radiomic features categorized based on each feature class used in this study.

Statistical	Histogram	Morphological	GLCM	GLRLM	GLSZM	NGTDM	Moment Invarient
1 SUVmean	7 AUC_CSH	14 MV	25 Energy	51 SRE	64 SZE	77 Coarseness	82 GT2
2 SUVmax	8 Mean	15 TLG	26 Entropy	52 LRE	65 LZE	78 Complexity	83 J1
3 SUVpeak	9 Variance	16 Eccentricity	27 Differenceentropy	53 GLN	66 GLN	79 Contrast	84 Q
4 SUVstd	10 Skewness	17 Solidity	28 Sumentropy	54 RLN	67 ZSN	80 Strength	85 J2
5 SUVvar	11 Kurtosis	18 PercentInactive	29 Variance	55 RP	68 ZP	81 Busyness	86 J3
6 SUVenergy	12 Energy	19 Surface	30 Variance2	56 LGRE	69 LGZE		87 B3
	13 Entropy	20 SVratio	31 Sumvariance	57 HGRE	70 HGZE		88 oJ1
		21 compactness1	32 MaxPossibility	58 SRLGE	71 SZLGE		89 oQ
		22 compactness2	32 MaxPossibility	59 SRHGE	72 SZHGE		90 oJ2
		23 sphericity	33 Contrast	60 LRLGE	73 LZLGE		91 oJ3
		24 SizeROI	34 Dissimilarity	61 LRHGE	74 LZHGE		92 oB3
			35 Homogeneity	62 GLV	75 GLV		
			36 Homogeneity2	63 RLV	76 ZSV		
			37 Correlation				
			38 Correlation2				
			39 Autocorrelation				
			40 Clusterprominence				
			41 Cluster shade				
			42 Cluster tendency				
			43 ICM1				
			44 ICM2				
			45 InvDifMoment				
			46 IDMN				
			47 IDN				
			48 SumAverage				
			49 SumAverage2				
			50 Agreement				

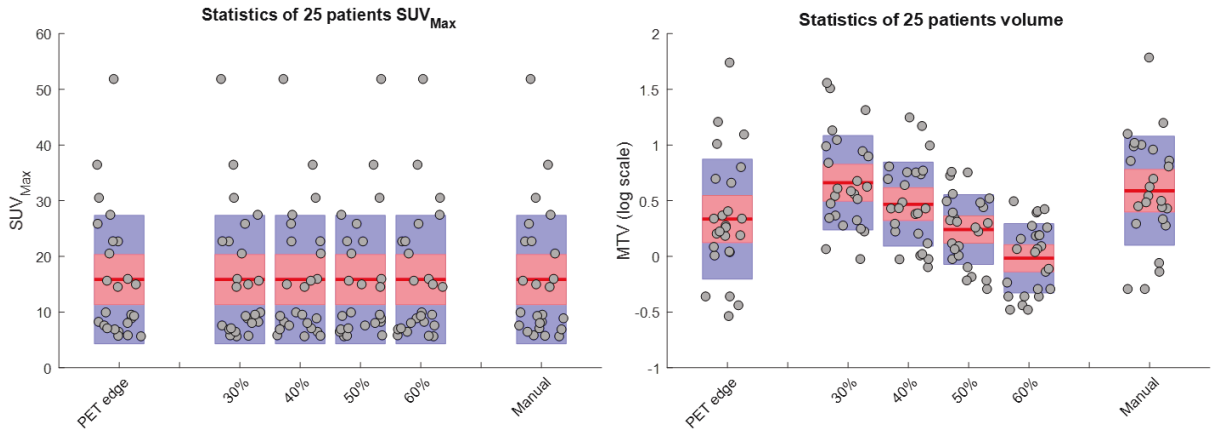

Figure 3-14. Statistical distributions of SUV_{max} and MTV for 25 patients across 6 different segmentation methods, 64 GLs of uniform quantization

Uniform quantization showed consistent Spearman correlations for GLs≥64, while significant details on uptake heterogeneity were lost for GLs≤32. Max-Lloyd quantization had a similar pattern of loss

for GL≤32. We recognized the slight increase in computation time of the Max-Lloyd method, but since it had a similar pattern to the uniform discretization as well as very close Spearman correlation, we suggest to adopt the more commonly-used uniform discretization method as explained in section "Fixed bin-number discretization".

Feature rank correlations against SUV_{max} and tumor volume exceeded 0.7 for 37% and 0.9 for 13% of the feature pairs, indicating limited complementary information for the latter features. Many features were highly correlated with one or more of volume. We selected some of which are highly correlated with some others with p-values≤0.05, but are not correlated with respect to themselves, and arrived at these features: integrated intensity, GT2 (i.e. integrated intensity restricted to voxels with SUV>2), homogeneity and entropy (GLCM) (p-values<1e-6), which we subsequently focused on.

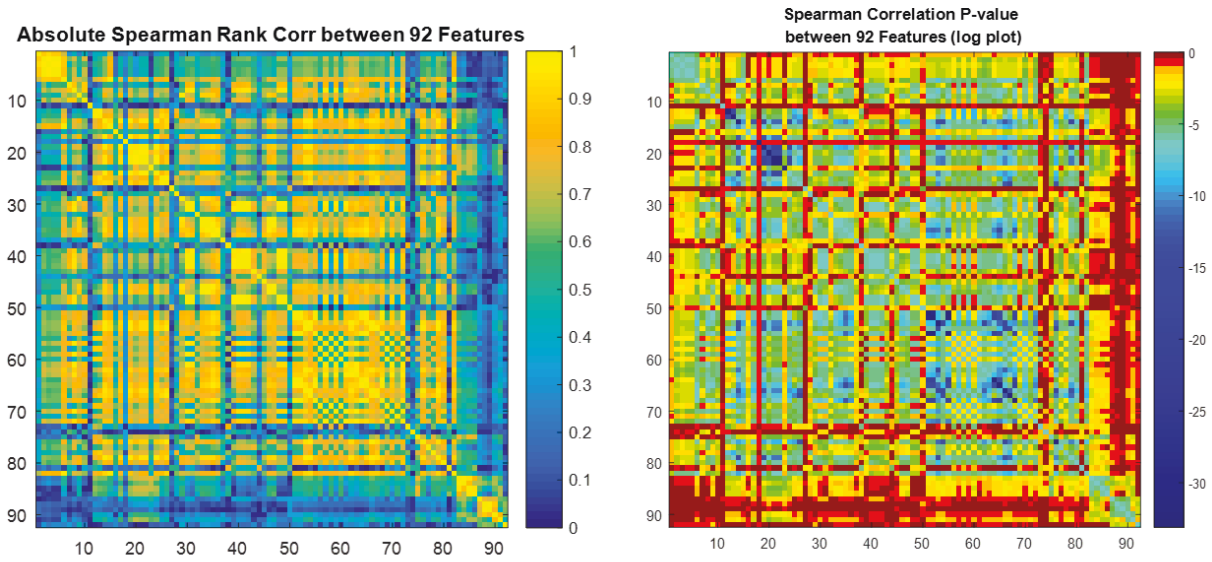

Figure 3-15. Heat map depicting the absolute value of Spearman correlation coefficients between pairs of textural features (left) and their log p-values (right).

Correlations of radiomic features against SUV_{max} and TV tend to decrease substantially with ranges of increasing SUV_{max} and volume. This is because the partial volume effect dominates quantification of smaller tumors, while larger ones tend to exhibit higher anatomical and physiological complexity in the uptake spatial distribution.

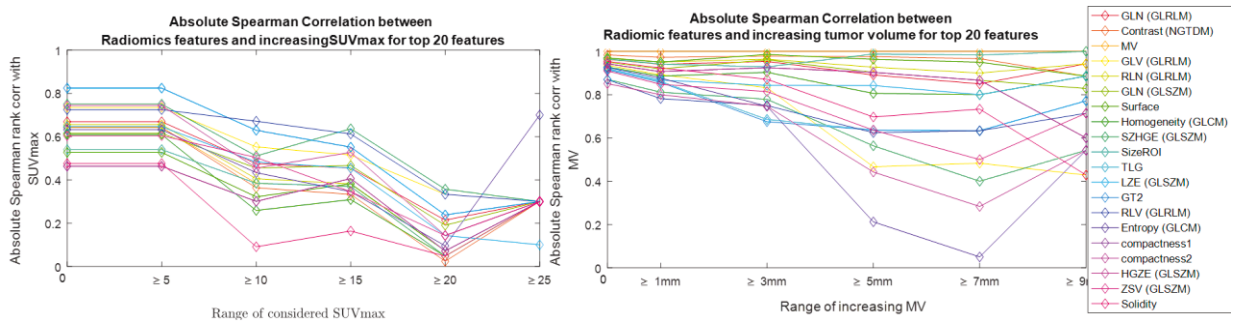

Figure 3-16. Spearman correlation of top 20 most reproducible features with the highest Spearman correlation based on and uniform quantization gray-levels with SUV_{max} (left) and MTV (right), with 64 uniform gray level and manual segmentation. Most features exhibit a decreasing trend as the range shortens.

13 features (6 SUV-based, 3 morphological and 3 GLSZM) had ICC > 0.6 for all 6 segmentation methods. Reproducibility increases substantially considering only manual and gradient-based segmentations, yielding 29 features with ICC > 0.8. Radiomic features exhibiting both high quantization ICC (> 0.85) and high segmentation ICC (> 0.85) include RLN (GLRLM), entropy (GLCM), ZSN and ZSV (GLSZM), J3 (MI) and GT2.

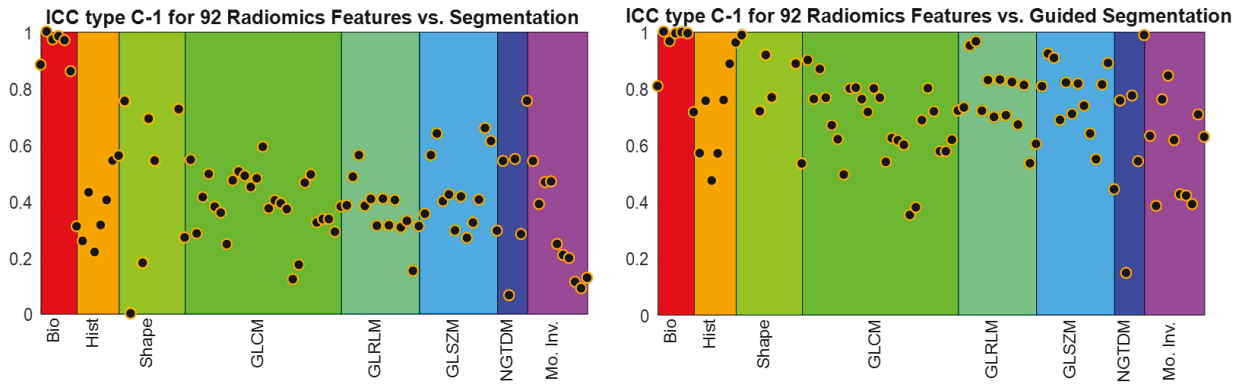

Figure 3-17. ICC between all six segmentations for (left) and between only 2 user-guided segmentations (right) for all 92 features grouped into their feature families.

3.3.2.4. Discussion

Quantization by 64 gray-levels is sufficient to capture heterogeneity information. Threshold contouring significantly impacts reproducibility of features, and user-guided or more advanced segmentations are recommended for more reliable radiomics analysis. Entropy (GLCM), ZSV and ZSN

(GLSZM) exhibit high reproducibility and reliability in this study of ^{18}F-DCFPyL prostate PET. This is an interesting finding since these features are also reported to be reproducible for radiomics studies using FDG PET as well [247]. This finding indicates high reproducibility across radiotracers for the aforementioned features. As a result, these set of features are recommended for investigation of their possible prognostic or predictive value.

3.3.3. Conclusion

In this chapter, we provided more details about the radiomics workflow and explained how the complexity of radiomics analysis impacts reproducibility and reliability. In addition, we presented recommendations on how to configure and then assess radiomics analyses to mitigate these issues. We further explained our efforts in joining the image biomarker standardization initiative (IBSI)—a global group of top universities and cancer centers aimed to standardize the preprocessing and feature calculation of radiomics analyses. We developed a standardized environment for radiomics analysis (SERA) – an entirely in-house toolbox for image preprocessing and feature calculation based on the IBSI guidelines with 100% coverage of the features, released for use by the wider community. Furthermore, we presented two studies on reproducibility analysis of radiomic features: in renal cell carcinoma SPECT and prostate cancer PET imaging. Some radiomic features such as GLCM entropy as well as GLSZM zone-size non-uniformity and zone-size variance were reported as reproducible in both studies. Also, fixed bin-number discretization with 64 to 128 GLs was reported to convey enough details about variabilities inside ROIs that can be captured by different feature classes without discarding important heterogeneity information. Several of the reported features were not only found to be reproducible in our specific studies but also have been reported to be reproducible in studies with other radiotracers, indicating their broader appeal and potentials. Furthermore, features reproducibility was significantly challenged by segmentation methods. We suggest that

more advanced segmentation methods with higher accuracy and less inter- and intra-reader variability may need to be adopted for reproducible radiomics. Overall, the two presented projects are the first studies of their kind in their specific clinical applications and can be used in future efforts to shrink the feature-space and help discover and validate effective imaging biomarkers.

4. Radiomics Analysis of Clinical Myocardial Perfusion Stress SPECT Images to Identify Subclinical Coronary Artery Disease

4.1. Introduction

This chapter aims to enhance the clinical utility of routine clinical myocardial perfusion (MP) SPECT imaging through advanced radiomics analysis. We hypothesize that identification of mild heterogeneities via radiomic analysis can enable identification of subclinical coronary artery disease (CAD) that would carry important diagnostic and prognostic information. In this chapter, we aim to evaluate our exciting and novel hypothesis that MP SPECT radiomic features extracted from clinically normal (non-ischemic) MP SPECT scans correlate with coronary artery calcification (CAC) as extracted from CT imaging. This chapter starts with an introduction to myocardial perfusion imaging using SPECT, as well as coronary artery calcification scoring using CT, as well as clinical motivations for our work; subsequently, we describe our methods, following by results and conclusion.

4.1.1. Myocardial Perfusion stress SPECT test

4.1.1.1. Basics of Myocardial Perfusion Imaging

Myocardial perfusion SPECT (MPS) is established for non-invasive evaluations of patients suspected with coronary artery disease (CAD)[463, 464]. It is probably the most widely-used technique of nuclear cardiology, and its purpose is to assess the adequacy of blood flow to the myocardium [465].

Although MP imaging can be performed with either planar or tomographic techniques [465, 466], nowadays tomographic MP imaging through SPECT scanners has become widely-popular, more accessible and more affordable to patients.

A standard MP SPECT exam consists of two SPECT scans, one under rest and one under stress condition, although recent studies have shown if first the "stress" test is taken, and the image is interpreted as normal, there is no need for the second "rest" test, which is referred to as "stress-only" image [467]. The stress MP test is usually performed following a physical exercise (e.g. on a treadmill), or intravenously administered stress pharmaceuticals such as adenosine or dipyridamole.

MP stress SPECT has an established pathophysiologic basis with radiotracers capturing blood flow. If a patient with CAD is at rest, typically, blood flow through a diseased coronary artery (e.g. narrowed through plaque build-up) is not decreased until coronary stenosis exceeds 90% of the artery. On the other hand, coronary reserve, which refers to the ability to increase coronary blood flow in case of increased metabolic demand, is reduced if coronary stenosis exceeds 50% [468, 469]. As a result, patients suffered from CAD may have a homogenous uptake of myocardial blood flow even in the presence of severely-narrowed coronary artery. But the same degree of narrowing can result in reduced flow reserve when the heart is stressed under exercise, resulting in inhomogeneity of regional MP (Figure 4-1). Such inhomogeneity can be captured using radiotracers that are distributed in the body in proportion to myocardial blood flow [465].

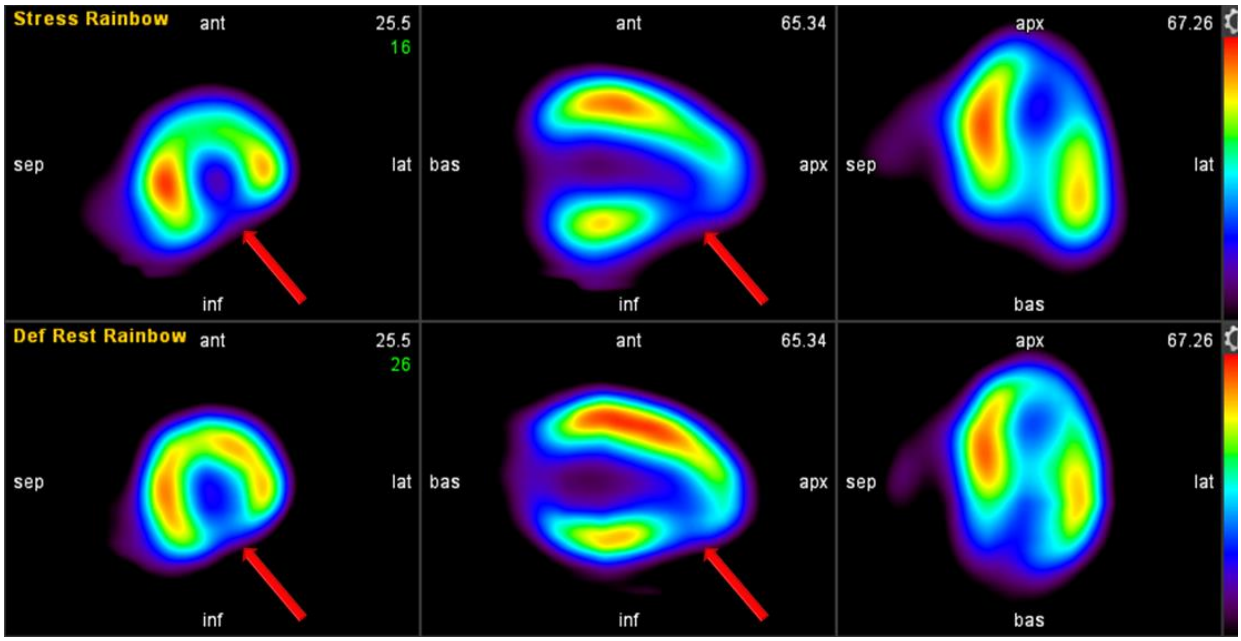

Figure 4-1. MP SPECT image of a patient with myocardial ischemia. The top and the bottom views show images acquired at stress and rest, respectively. The arrow shows decreased blood flow (reversibility) at the inferior wall in the stress image at the place of the arrow compared to the rest—a typical sign of myocardial ischemia.

MP SPECT is typically performed by administrating ^{99m}Tc-sestamibi. Technetium-99m is a gamma-ray emitting radionuclide that rays at 140 keV and is optimally suited for imaging with a gamma camera. It is possible to give patients two doses of ^{99m}Tc-Sestamibi on a single day, while the latter dose must be higher than the first one, e.g. 10mCi for the first scan and 35mCi for the second one. The image acquisition varies between 20 to 45 minutes.

Several visual patterns are common in MP imaging. A normal pattern consists of uniform homogenous uptake of the radiotracer all around the left ventricle. A defect may be identified at areas where the myocardium exhibits relatively lower tracer uptake, such as the one shown with a red arrow in Figure 4-1. A defect is considered reversible if it is present at the stress image but is no longer present, or partially improved, on the rest image. A fixed defect, such as a myocardial infarct, remains unchanged in extent or severity on both rest and stress images.

4.1.1.2. Applications of MP SPECT imaging

MP SPECT imaging has extensive clinical applications. The sensitivity and specificity of stress imaging for detection of CAD using visual assessment have been reported at over 90% with SPECT and 95% with PET (using ^{82}Rb or ^{13}N) [470]. However, ^{13}N requires an on-site cyclotron for production, and ^{82}Rb has a monthly associated cost of generator replacement, plus SPECT cameras have a clear affordability advantage over PET scanners. In addition, MP SPECT has high prognostic value, and whether a patient has a reversible perfusion defect plays a strong role in the assessment of risk [471]. Patients with reversible MP defects have a higher likelihood of CAD events during follow-up compared with patients with fixed defects. Further applications of MP SPECT imaging include preoperative screening before major noncardiac surgical procedures to assess their perioperative risk [472-474], imaging after coronary angioplasty to detect restenosis—a common problem in 20 to 40 percent of the patients undergoing this procedure [475], assessment of MP in acute ischemic syndrome in patients with unstable angina [476], and accurate assessment of myocardial viability [477].

4.1.2. Coronary artery calcification quantitation using coronary artery calcium scoring

Large prospective studies have shown that coronary artery calcification (CAC) scoring is associated with the risk of future cardiovascular events [478-481]. Studies have shown that noninvasive tests for CAD including electrocardiogram (ECG), ultrasound imaging, and even MP SPECT scan, which are used quite often in cardiac patients' assessment and diagnosis, were of limited value to detect this calcification due to their low sensitivity [482]. A minimum of 25% of the patients that experience a non-fatal acute myocardial infarction or sudden death do not have previous symptoms [483], and it is necessary to identify asymptomatic individuals at greater risk of future cardiovascular events to plan for preventive strategies.

Our Motivation: CAC is a highly specific marker of coronary atherosclerosis, and higher CAC scores are associated with increased plaque burden and increased cardiovascular risk [484, 485]. Previous studies demonstrated that a considerable number of stenoses do not result in abnormal perfusion on MP imaging [486, 487], which is why in our work we set the inclusion criteria of "non-ischemic normal" MP stress scans. Furthermore, the CAC score is shown to offer incremental diagnostic information over MP SPECT for identifying patients with significant CAD and *negative* MP imaging results [488]. Therefore, finding a feature or a set of features (a radiomics signature) from MP stress SPECT images that can predict CAC score would be beneficial as it eliminates an additional non-contrast CT for CAC assessment, thus reducing excessive dose to the patient. Unlike MP SPECT, CAC test is not reimbursed by CMS, while it is known to improve risk stratification in asymptomatic individuals [484, 486]; but our study enables CAC assessment from MP SPECT. Moreover, CAC calculation requires sophisticated software and trained radiologists. It is included in the CAD patients' diagnosis package in large institutions such as our Johns Hopkins Hospital but is not readily available in community settings.

CAC scoring is performed using a CT scan based on its axial slices in synchrony with ECG without administrating any contrast agent [489]. Calcification is referred to an area of hyper-attenuation with ≥ 1 mm^2, or ≥ 3 adjacent voxels, of >130 Hounsfield Units (HU) [490]. The main scoring CAC protocol is the Agatston method [490] that is widely-used, especially as a reference for most population databases and risk stratification studies, thus, is the most important method often used in clinical practice. This method uses the weighted-sum of lesions with a density above 130 HU, where the area of calcium is multiplied by a factor related to the maximum plaque attenuation as follows:

- 130 – 199 HU, factor 1
- 200 – 299 HU, factor 2
- 300 – 399 HU, factor 3

- ≥ 400 HU, factor 4

According to the Agatston method, this score is calculated for each of the four three arteries of the heart, namely the left anterior descending (LAD) the right coronary artery (RCA) and the left circumference (LCX), as well as the left main (LM). LM is a small artery that divides into LAD and LCX (Figure 4-2). This calculation, despite being relatively straight-forward, requires special software and the cost associated with its licensing requirements might be another hurdle in the widespread application of CAC scoring in smaller cost-effective radiology centers. An example of CAC present in a CT image shown in Figure 4-3.

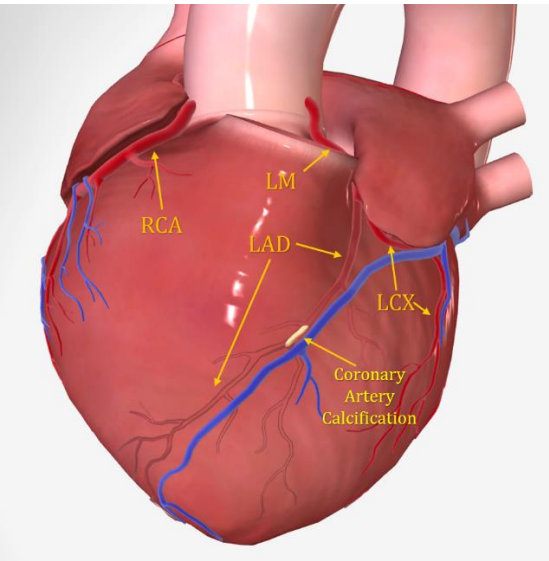

Figure 4-2. Main coronary arteries of the heart: left main (LM) which divides into left anterior descending (LAD) and left circumference (LCX), and right coronary artery (RCA). Coronary artery calcification is present in LAD that results in reduced blood flow to this artery (pale color of the LAD).

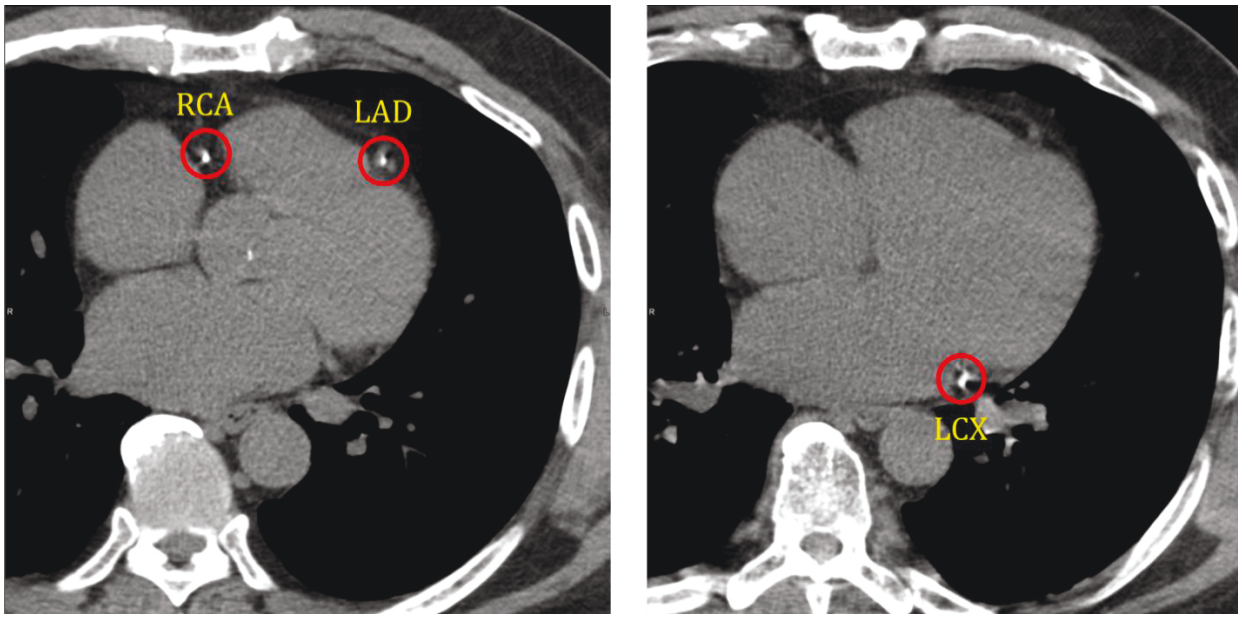

Figure 4-3. Coronary artery calcification in three main arteries of a patient. The left image depicts slices of heart with CAC in their RCA and LAD, and the right image shows CAC in LCX.

4.1.3. An overview of our framework

We elaborated on radiomics analyses in earlier chapters, stating that use and combination of effective metrics that quantify shape and texture/heterogeneity, hold significant promise towards identifying patients with higher risk and to help realize personalized medicine. Here, we develop a pipeline to evaluate various classes of standardized radiomic features. We first inspect these features and narrow them down (i.e. feature selection) towards improved statistical significance and reduced false-discovery rate, which can be helpful to other researchers in nuclear medicine radiomics. Furthermore, by investigating the relationship between radiomic features of clinical MP SPECT images and the CAC scores, we hope to open a new possibility: to use clinical MP SPECT scans for additional assessment of CAC. This has important implications, given that, as mentioned above, CAC assessment is not commonly performed nor reimbursed in a wide community setting, and as such, our proposed framework holds promise for new added usage and value for routine MP SPECT imaging.

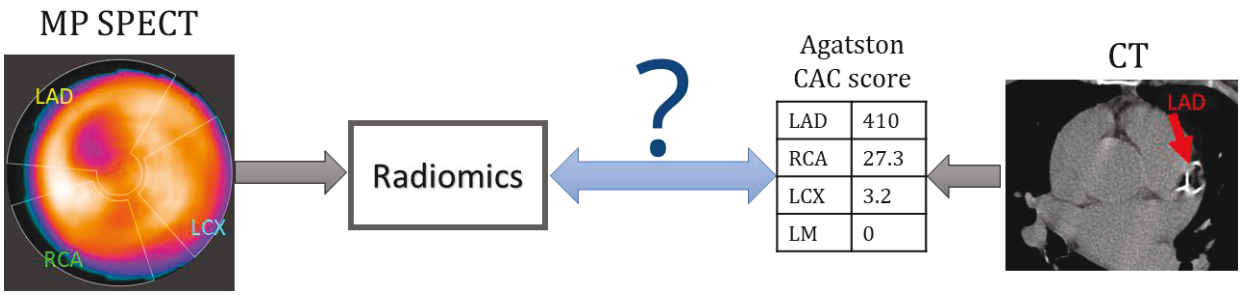

Figure 4-4. Diagram of the problem addressed in this chapter: using radiomics of stress MP SPECT to predict CAC scores from CT.

Radiomics has witnessed significant activity, especially in oncologic MRI, CT, and PET. Yet it has *not* been thoroughly assessed in SPECT and/or cardiac imaging, partially due to their low spatial resolution that may appear less likely to provide valuable texture and heterogeneity information. However, our group has successfully demonstrated the exciting use of radiomics in brain SPECT [291, 292]. At the same time, cardiac SPECT radiomics remains unexplored. Moreover, the prevalence of these scans is significantly higher compared to PET exams, enabling collections of a higher volume of data for such data-oriented MP SPECT research.

An important ingredient to success in the translation of radiomic features to clinical reality is to quantify and ascertain their robustness, which was one of the aims of this study. The result of our reproducibility analysis would be valuable for future MP SPECT radiomics research by others, as well as for the next steps in our own radiomics research towards discovering prognostic cardiac imaging biomarkers, enabling the field of "**imaging-based cardiomics**".

4.2. Radiomics of MPS to Predict Coronary Artery Calcification

We constructed a dataset comprised of n=428 clinical MP stress SPECT images with a normal scan (non-ischemic) with a separate CT for CAC scoring with their detailed CAC reading. We pursue the following three steps, as outlined below. Subsequently, we elaborate on our methodology followed by results.

4.2.1. Three Steps in Our Study

Step 1: Improved quantitative assessment through analysis of standardized radiomic features on MP SPECT images. We start by identification of patients with *normal* MP SPECT test and CAC CT, following by image segmentation.

Step 2: Eliminating non-reproducible and redundant features (feature selection).

Step 3: Use of machine learning to extract CAC information directly from MP SPECT image radiomics, in contrast to the routine use of CT scans.

4.2.2. Methods

4.2.2.1. Patient collection

After obtaining approval from the institutional review board (IRB) at Johns Hopkins University, we searched for patients with stress myocardial perfusion SPECT scans from 2011 till 2015. In this effort which lasted about 4 months, we investigated over 1,800 reports of patients undergone MP stress SPECT, out of which n=428 cases were selected. All of these patients had a CT scan for CAC scoring at the same time as their MP stress SPECT scan in the PACS database. A nuclear medicine physician (NMP) investigated the MP stress SPECT images to be free from i) image artifacts, and ii) overcorrection and iii) spillover from nearby liver or stomach. Our NMP also derived detailed CAC score for each of the 4 arteries of the heart using a clinical software.

The dataset consists of mages collected from scanners with different vendors at the Johns Hopkins Hospital throughout these years, but all were reconstructed with an “attenuation-corrected iterative reconstruction” (AC-IR) algorithm and with a consistent voxel size of 4.8 mm. According to the quality factors of radiomics research, this is an important characteristic of a study to have imaging

acquisition protocols that are "well described and ideally similar across patients", and "methodologic steps taken to incorporate only images of sufficient quality" [436].

We recorded many parameters for each patient, including basic information, clinical factors, scan info, and any possible outcome info, as detailed in Table 4-1.

Table 4-1. Information recorded for every patient during dataset collection.

Basic statistics	Clinical factors	Scan info	Outcome
• Age • Gender • Race • Hight • Weight (at scan time)	• Smoking • Diabetes • Hypertension • Hyperlipidemia • Family history of cardiac disease • Systolic blood pressure (SBP) • Diastolic blood pressure (DBP)	• Scan date • Scan impression (normal, fixed defect, ischemia) • Stress test type (Bruce vs. drug) • Left ventricle ejection fraction (LVEF) • CAC score (LM, LAD, RCA, LCX, Total)	• Death • Cardiac-related progression • Date of death/cardiac related progression • Date of the last follow-up • Cath lab visit

4.2.2.2. Image segmentation

The study involves three different layers of segmentation as applied to MP SPECT images: i) total myocardium, ii) three vascular segments, and iii) 17 polar segments. Feature evaluation and statistical analysis were performed over all three layers. These three segmentation methods are presented in Figure 4-5. The reason we selected two different methods for vascular segmentation is that both methods are widely-used in the clinic. The 3 vascular segment method has a more stringent segment, while the subsets of the 17 polar segments span the whole heart, as can be observed from Figure 4-5.

We used MIM software® and developed a workflow that automatically draws 3D contours over 21 regions of the heart, namely: endocardium, epicardium, 3 vascular segments (as depicted in Figure 4-5.B), and 17 polar segments (as depicted in Figure 4-5.C). The workflow was generating the

myocardium segment using epi- and endocardium segments. This automated segmentation was supervised by a radiologist to assure the contours are correctly placed over corresponding segments. Our workflow then exported the 3D SPECT image and these 22 contours as 3D Matlab® matrices, in which they were subsequently used for analysis.

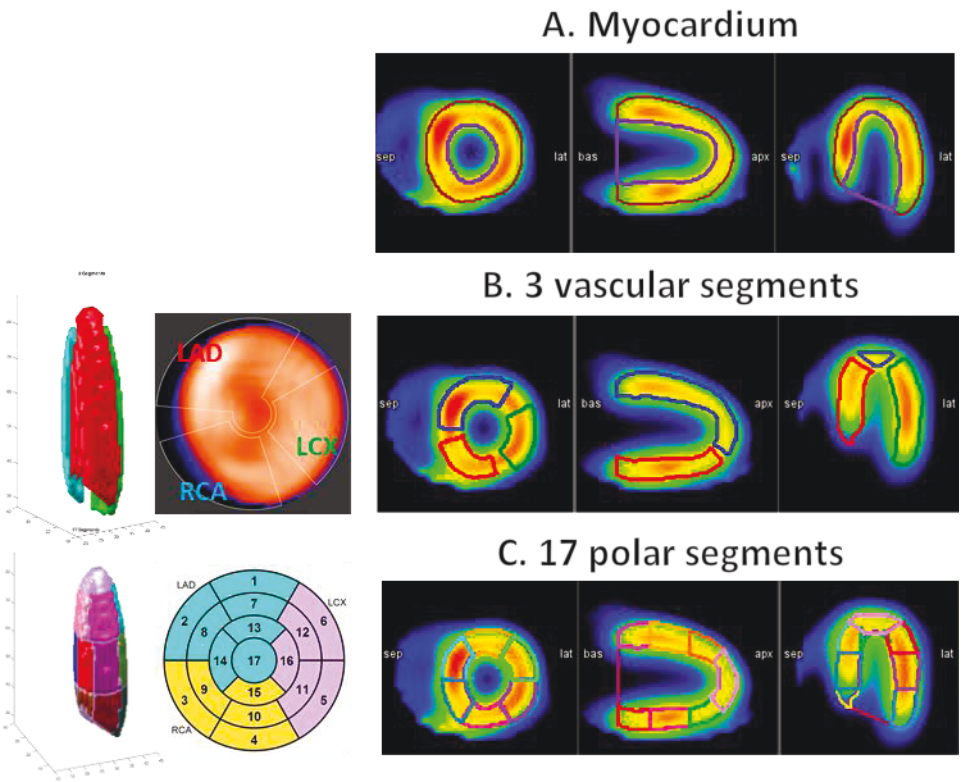

Figure 4-5. Three methods of segmentation used in our study. A) myocardium segment. B) 3 vascular segments of the heart (LAD, RCA and LCX), and C) subsets of 17 polar segments of the heart grouped into LAD, LCX and RCA.

4.2.2.3. Radiomics Framework

We subsequently used our in-house-developed standardized SERA package to derive radiomic features that are standardized and reproducible (section 3.2.3), consistent with guidelines by the Imaging Biomarker Standardization Initiative (IBSI) guideline. IBSI is a global initiative consisting of the world's top universities and cancer centers [429], in which our group is an active participant. More discussions regarding IBSI standardized radiomics and SERA was provided in Chapter 3

(section 3.2.2). SERA calculates 487 standardized radiomic features aiming to standardize the preprocessing and feature evaluation phases and to meet ISBI's standards in order to conduct and pursue reproducible research [450].

Images produced for MP SPECT scan are arbitrary-unit. Therefore, as explained in the last chapter in "Intensity discretization" subsection under section 3.1.3.1, we ought to use the fixed bin number discretization. We considered and investigated a range of GL discretizations, specifically using 4, 8, 16, 32, 64, 128, 256, and 512. All the images in our dataset were reconstructed into 3D images with identical voxel sizes of 4.8×4.8×4.8 mm^3; thus, no resampling and interpolation was needed. We did not perform any GL rounding or re-segmentation. The framework was then ready to calculate 487 features for 8 GLs and 7 segments of the heart.

4.2.2.4. Statistical Analysis

We used statistical analysis to eliminate non-useful features, including features that are identical, non-robust, and redundant. We performed a multistep feature selection to significantly reduce the size of our feature-space of 487×8 features. This process was performed completely independent of outcome (e.g. CAC score, etc.). The selected feature set was subsequently passed on to univariate and multivariate analyses schemes to predict correlate with clinical outcome. We also accounted for false-discovery by employing false-discovery correction methods such as Benjamini-Hochberg [491].

4.2.3. Analyses and Results

4.2.3.1. Analysis of dataset statistics

In this section, we present the data statistics based on variables previously introduced in section 4.2.3.14.2.2.1. We searched for patients with stress myocardial perfusion SPECT scans from 2011 till

2015. In this effort which lasted about 4 months, we investigated over 1,800 reports of patients undergone MP stress SPECT, out of which n=428 cases were selected. All these patients had a CT scan for CAC scoring at the same time as their MP stress SPECT scan in the PACS database. Our NMP also derived detailed CAC score for each of the 4 arteries of the heart using a clinical software.

The dataset consists of mages collected from scanners with different vendors at the Johns Hopkins Hospital throughout these years, but all were reconstructed with an "attenuation-corrected iterative reconstruction" (AC-IR) algorithm and with a consistent voxel size of 4.8 mm. According to the quality factors of radiomics research, this is an important characteristic of a study to have imaging acquisition protocols that are "well described and ideally similar across patients", and "methodologic steps taken to incorporate only images of sufficient quality" [436].

We recorded many parameters for each patient, including basic information, clinical factors, scan info, and any possible outcome info, as detailed in Table 4-1.

Basic statistics

The dataset was comprised of 229 female (49.7%) and 232 male (50.3) subjects. Distributions of patient age, height, weight, and body-mass index based on gender is depicted in Figure 4-6.

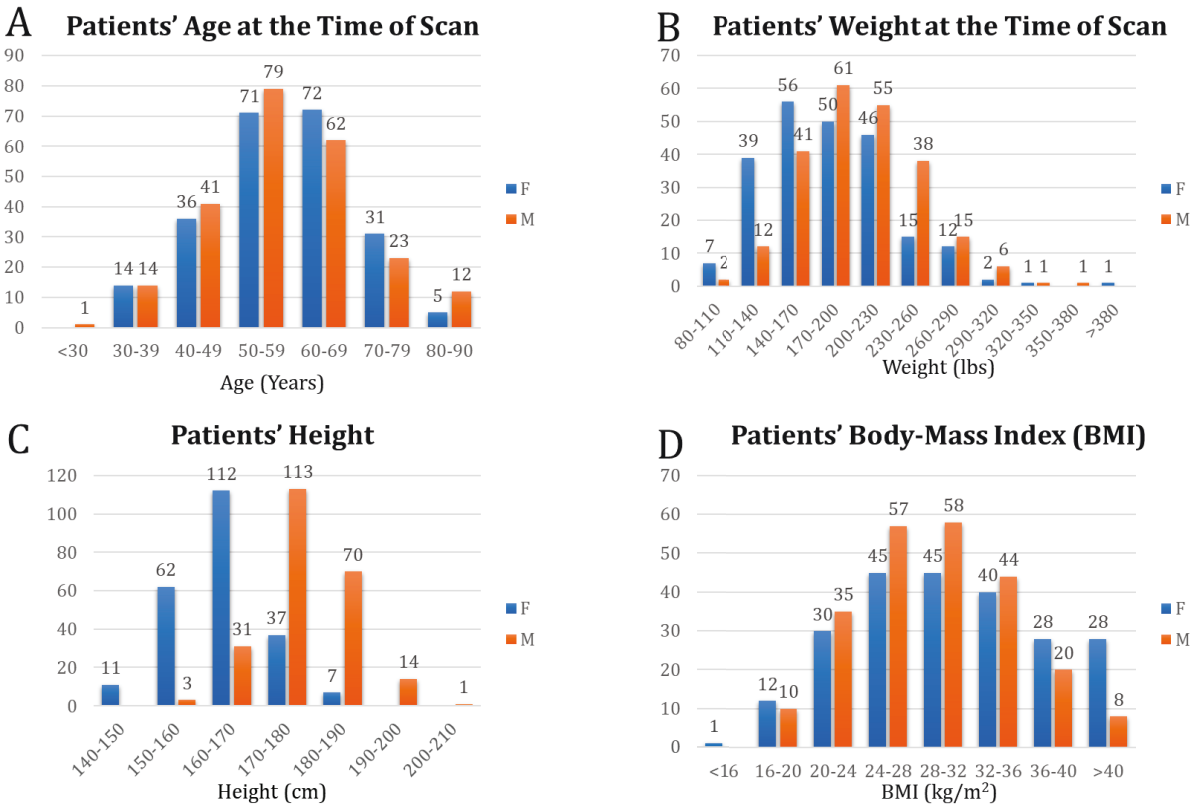

Figure 4-6. Distribution of patients' A) age, B) weight, C) height and D) BMI at the time of scan grouped into male (orange) and female (blue)

We observe a relatively close distribution of age between males and females. The frequency of race is presented in Table 4-2.

Table 4-2. Distribution of the race of the patients

	Frequency	Percent
African American	225	48.8
Asian	10	2.2
Hispanic/Latino	7	1.5
Indian	2	0.4
Middle eastern	19	4.1
Native American	1	0.2
White	197	42.7
Total	**461**	**100**

Clinical factors

Our dataset consists of 428 normal scans. In our dataset, 274 patients received stress compound by injection (264 A2A Adenosine, 7 Dipyridamole, 2 Regadenoson), and 186 patients were stressed on a treadmill based on the Bruce protocol [492]. Figure 4-7 shows the distribution of LVEF for all patients. The ejection fraction compares the amount of blood in the heart to the amount of blood pumped out. It helps to describe how well the heart is pumping blood to the body. The ejection fraction of a normal heart is between 50% to 70%. A higher LVEF may indicate a heart condition such as hypertrophic cardiomyopathy [493, 494].

Table 4-3. Distribution of patients' clinical factors

Factor	Attribute	No. of Cases	% of Total
Smoking			
	Non Smoker	220	47.7
	Current Smoker	137	29.7
	Previous Smoker	104	22.6
Diabetes			
	No	313	67.9
	Yes	148	32.1
Hypertension			
	No	138	29.9
	Yes	323	70.1
Hyperlipidemia			
	No	243	52.7
	Yes	218	47.3
Family History of CAD			
	No	283	61.4
	Yes	178	38.6

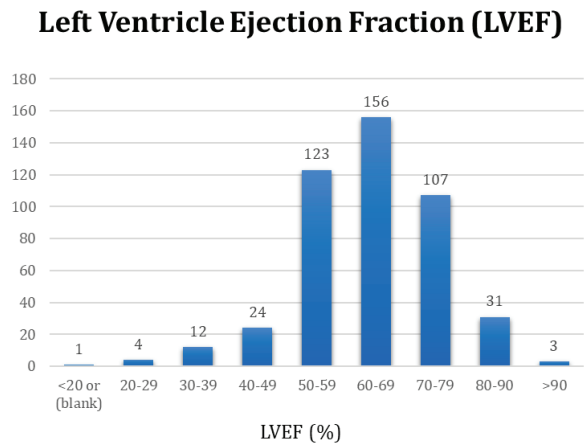

Figure 4-7. Distribution of LVEF in the dataset

A well-known stratification method presented by Berman, *et al.*, divides CAC scores into 5 categories of 0, 0^+ to 10, 10^+ to 100, 100^+ to 400, 400^+ to 1000 and >1000 [486]. The distribution of our patients into these five categories is presented in Figure 4-8.

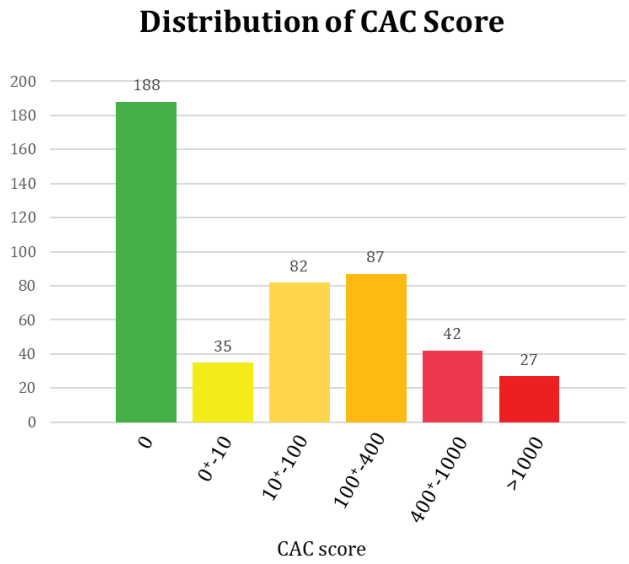

Figure 4-8. Distribution of our patients' CAC score based on widely-used stratification criteria [486].

The plot in Figure 4-8 shows the distribution for all patients. Patients with normal MP stress test were having 178, 32, 77, 80, 38, and 23 cases in each of the five bins of Figure 4-8. This shows 58.4% of the patients with normal MP stress scan had a non-zero CAC score, out of which 33% had a CAC

score ≥100. Studies have shown that MP ischemia is rare in patients with CAC<100, whereas in patients with CAC≥100, the chance of myocardial ischemia increases progressively [486]. Also, one-third of patients with normal MP stress SPECT had a CAC score ≥100. As previously mentioned, CAC scoring is known for its high specificity and very small false-positives, suggesting that assessment of atherosclerosis burden by CAC scoring may be useful in finding CAD when MP stress SPECT fails to report. It also underlies our motivation to develop and study a radiomics-based scheme to extract CAC information directly from MP stress SPECT scans.

Patient outcome

We did not encounter many subsequent death incidents in our dataset. Figure 4-9 shows the distribution of patient death (by any cause) and cardiac-related progression. Out of 38 patients that were referred to Cath Lab following their MP stress SPECT scans, 29 had a normal MP stress SPECT, yet all but 5 patients had a non-zero CAC score, which again demonstrates the added value of CAC scoring in addition to MP stress SPECT.

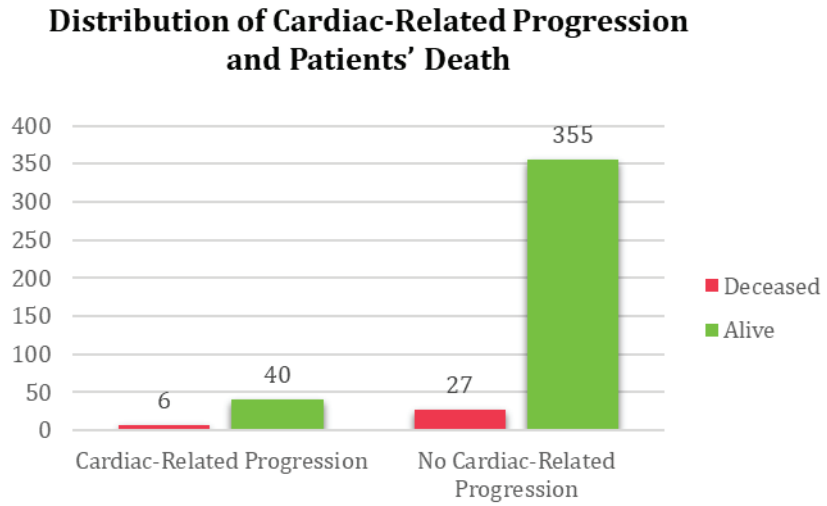

Figure 4-9. Distribution of cardiac-related progression and patients' death for normal patients in our dataset

4.2.3.2. Feature selection

All 487 features in our SERA software from 11 main categories as mentioned in Table 3-1 were initially considered and were calculated for 8 different GLs. In this section, we aim to systematically narrow down this large feature set and arrive at a smaller set of meaningful, robust, non-redundant, and reproducible features for further investigation of their predictive or prognostic value and, at the same time, discourage overfitting. Our feature selection phase can be generally divided into i) pre-feature calculation and ii) post feature calculation, as explained below. Following feature selection, we discuss how to narrow down to an optimum discretization level.

Pre-feature calculation

In the first step, prior to performing any analysis, we eliminate irrelevant feature families based on the nature of our dataset and our knowledge about what each feature captures.

Removing 2D and 2.5D feature families

Our dataset originally consists of images with isotropic voxels. Therefore, there would be no additional information provided to us from 2D or 2.5D feature families. These feature families would have been beneficial when slice thickness (i.e. voxel size in z dimension) was different from the voxel size in x and y dimensions. In that case, resizing and interpolating the images to isotropic voxel sizes may have resulted in modification of the original voxel distribution, causing possible loss or modification of data. In any case, the following feature families were eliminated: 2D and 2.5D GLCM (25 features) and GLRLM (16 features) (both merged and averaged), 2D and 2.5D GLSZM (16 features), GLDZM (16 features), NGTDM (5 features), and (17 features). This removed 272 features, narrowing down our feature space to 215.

Removing useless feature families

MP SPECT images have voxels with arbitrary units (they are not quantitative unlike PET or some SPECT imaging applications). Therefore, any feature that conveys information regarding the exact intensity values of the original ROI is not considered meaningful. As such, intensity-based features (18 features) and local intensity features (2 features) were excluded. Furthermore, the seven segments were created by an automatic segmentation procedure that generates ROIs with similar shapes (all registered to the same reference space). As such, the shape of the segments does not carry any differentiating information, and we are interested mainly in the heterogeneity caused by voxel intensity variations which carry information about the blood flow in different heart segments. As such, morphological features (29 features) were excluded as well. At the end of this step in our analysis, we were left with 166 features out of 487, eliminating the majority via our knowledge of the underlying nature of the features.

Post-feature calculation

Removing feature with identical values

Following this, we searched for features with identical values across all patients for further exclusion. In our dataset, these were 4 features with identical values across all patients: histogram minimum, maximum, and range, and NGLCM dependence count percentage. We now arrive at 162 features.

Removing feature families with more than one variety

In the next step, we calculate the Spearman rank correlation between each feature and all other features to explore the relationship of the features with respect to each other and find redundant and highly correlated features. At this step, we had one subtype of every higher-order feature class (i.e. only 3D, after excluding 2D and 2.5D) except for GLCM and GLRLM, each remaining with two 3D subtypes: 3D merged, and 3D averaged. We investigated the correlation between each variety of

higher-order 3D calculation, i.e. 3D GLCM averaged vs. merged, and also, 3D GLRLM averaged vs. merged. Figure 4-10 shows a heatmap of their correlation. In the diagonal of both heatmaps in Figure 4-10, we observe very high Spearman correlation (between 0.98 to 1) between all the same features within the two feature families, i.e. GLCM-averaged entropy vs. 3D GLCM-merged entropy, etc., indicating the redundancy of features calculated in two varieties (merged vs. averaged). Let us $S_{\{A\}|\{B\}}$ as the Spearman rank correlation between feature families $\{A\}$ and $\{B\}$. We subtracted $S_{\{\text{3D GLCM-averaged}\}|\{\text{All feature families except 3D GLCM-merged}\}}$ from $S_{\{\text{3D GLCM-merged}\}|\{\text{All feature families except 3D GLCM-averaged}\}}$, and did the same for GLRLM, and observed it yields values very close to zero, which further indicates that using one variety vs. the other does not add additional information to our analysis, suggesting exclusion of one variety from both GLCM and GLRLM. Subsequently, to decide which of the two varieties to exclude, we calculated the range of features in both varieties and removed the one with a smaller range, which yield to exclusion of the 3D-merged of both categories and keeping 3D GLCM-averaged and 3D GLRLM-averaged. This further reduced the number of features down to 121. This observation is also consistent with findings in [425], where the authors reported merged features with tighter distribution in a smaller range, and subsequently remove, them from the rest of their study.

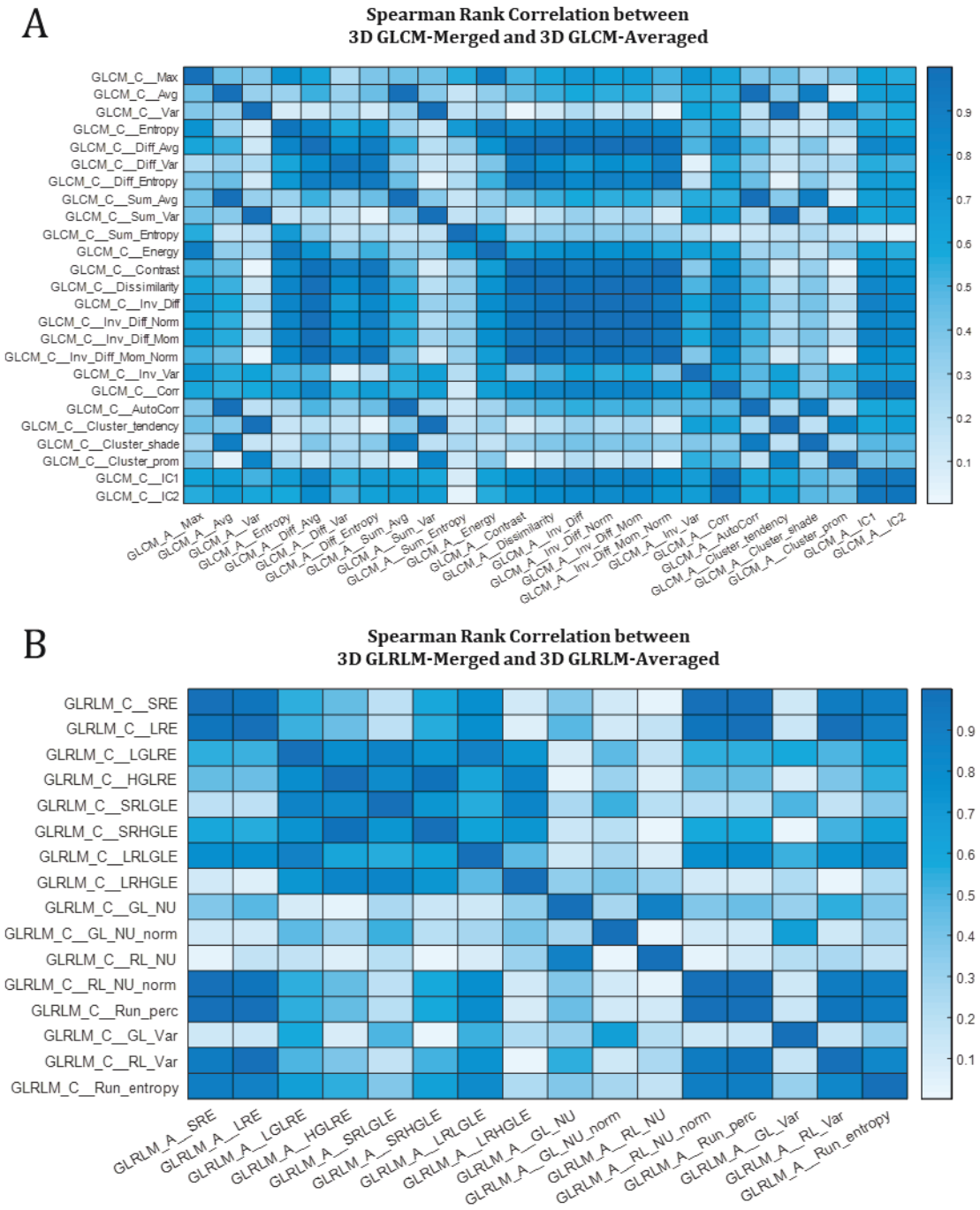

Figure 4-10. Heatmaps of Spearman rank correlation between A) 3D GLCM-averaged vs. 3D GLCM-merged, and B) 3D GLRLM-averaged vs. 3D GLRLM-merged. The diagonal of both plots have values >0.98

Removing redundant features

After using Spearman correlation to reduce the feature space at the feature-family level, we move on to investigate the correlation at the feature level. The next set of features to remove are the feature-pair with Spearman rank correlation coefficient of 1, indicating their redundancy. These features included three pairs: i) "3D GLSZM-zone percentage (ZP)" and "3D GLDZM-ZP", ii) "3D GLSZM-GL non uniformity (NU) normalized" and "3D GLDZM-GL NU normalized", and iii) "3D GLSZM-GL NU" and "3D GLDZM-GL NU", From each pair, we selected the feature with a lower range for exclusion that yielded the removing of the GLDZM features from each pair.

Removing features with a low dynamic range

In the next step, we calculated the percent variance of the features (variance/mean) representing their dynamic range. Subsequently, we removed features with a very low dynamic range less than 10^{-5}, which were five: Histogram-skewness, Histogram-kurtosis, Histogram-min gradient, GLCM-averaged cluster shade, and GLCM-averaged 1st measure of information correlation. Now the dataset has 113 features.

Removing highly-correlated features

In the last step of this phase, using the Spearman correlation of features with respect to each other calculated earlier, we opt to remove highly correlated features as defined by those having a Spearman correlation coefficient $|\rho| \geq 0.95$ as suggested in lthe iterature [495]. These feature-pairs are considered to be highly correlated and likely to provide redundant rather than complementary information. We remove these features through the following recursive operation.

We use the heatmap of feature-pair Spearman correlation to find features with $|\rho| \geq 0.95$. We subsequently record the number of instances a feature fits this criterion. Then, we sort these features based on which feature has more instances of $|\rho| \geq 0.95$ with others in a descending order and call it $\mathcal{F}^{sorted}$. We then start from the first feature in this set. We denote this first feature by $\not{f}^{keep}$, i.e. the

feature to keep, it and save it to $\mathcal{K}$ that denotes the set of features we intend to keep. Subsequently, we mark the highly correlated features with $\mathbb{f}^{keep}$ and save them to an empty set denoted by $\mathcal{R}$, i.e. for removal. We then loop over each feature inside $\mathcal{R}$ and find other highly correlated features with these features and append them to $\mathcal{R}$. Once the procedure is complete, we update $\mathcal{F}^{sorted}$ by removing $\mathbb{f}^{keep}$ and all features inside $\mathcal{R}$. The algorithm then continues recursively with this updated $\mathcal{F}^{sorted}$, letting its first member be $\mathbb{f}^{keep}$ and append it to $\mathcal{K}$, and find features and add them to $\mathcal{R}$ for removal. This process continues until $\mathcal{F}^{sorted}$ becomes empty.

The above algorithm cuts the number of features into a half, removing 57 features from 113, yielding 56 features remained that are not highly correlated with each other and are more likely to provide complementary information.

Selecting the best discretization level (GLs)

The above procedure reduced the feature set from 487×8 to 56×8 features for 8 GLs. Now we focus on discretization levels to systematically remove non-useful GLs. Firstly, we observe that for the three smallest GLs, the number of identical features is higher than the other five GLs. Furthermore, features with smaller dynamic range increase by 22%, 4%, 29%, and 29% compared to GL=64 or 128. Moreover, the two highest GLs have 11% and 22% more feature-pairs with Spearman correlation ≥ 0.9. Therefore, we can safely remove all GLs except 64 and 128.

The procedure in the previous paragraph could have been performed without the analysis of the range and Spearman rank correlation of features. We can safely remove the first three GLs since the intervals that voxel intensities were discretized into are so large that they do not provide enough opportunity to capture the heterogeneity of a region. On the other hand, the two largest GLs produce so many bins to discretize voxels into that many bins will be empty or just have very few representations in the ROI. For instance, the LAD segment consists of averagely 460 voxels. When it

is discretized into 512 GLs, they are actually more bins than voxels, and many bins would be empty or just occur very scarcely. In this case, our higher-order matrices such as GLRLM, GLSZM, GLDZM, etc. in which the number of columns represents different run-lengths, zone sizes, distance zones, etc., would be very long and narrow matrices with very small variability. As a result, these higher GLs should be eliminated, too.

Interestingly, this finding is consistent with our observations and conclusion from the previous chapter (section 3.3.3), as well as some previously published studies on radiomics of PET imaging [425, 443].

Finally, out of the remaining two GLs, 64 and 128, we found very similar behavior from both discretization levels in terms of the range of the features and number of feature-pairs with high Spearman correlation. We decided to choose 64 for the rest of this study, because 1) as mentioned GL=128 does not demonstrate different statistical properties, 2) our results in previous chapter suggested 64 GLs for the other SPECT study – imaging of renal cell carcinoma with ^{99m}Tc, which is the same radiotracer as the one used for MP stress SPECT imaging, and 3) some previous studies have demonstrated that GL=64 provided higher textural feature reproducibility [496] and robustness [443].

Wrapping up feature selection

Through the above procedures, we reduced our feature set of 487×8 to 56. One important note is that these features were excluded in a completely unsupervised manner without any involvement of the clinical outcome (e.g. CAC score, patient survival, etc.). This is an important factor to make our effort statistically sound and more believable.

4.2.3.3. Outcome prediction

In this section, we elaborate on our efforts towards outcome prediction using the narrowed down feature set. We also included our negative findings and unsuccessful attempts, as we believe reporting them helps future researchers, and thus, is of scientific value.

Univariate analysis

We define our outcome as the CAC score of each region of the heart calculated from the CT scan, and we aim to predict this outcome from the radiomic features extracted from the same region of the MP SPECT image, as explained in section 4.2.2.2. We started by investigating whether our selected radiomic features (previous section) directly correlate with the outcome, that is the CAC score. We adopted two approaches to represent the outcome. In the first approach, the actual CAC score with a continuous scale was utilized. In the second approach, we discretized CAC scores of each region of the heart based on the 5-scale clinical stratification criteria explained in section "Clinical factors" and plotted in Figure 4-8. Spearman rank correlations between features of every segment with the CAC score of the same segment were calculated for both CAC approaches (continuous and discrete). We also employed Benjamini-Hochberg false-discovery rate (FDR) correction with q = 0.05 to discourage overfitting. None of the features were able to survive FDR correction and still significantly correlate with outcome under this univariate scheme. Figure 4-11 shows the absolute value of Spearman correlation coefficient values between 56 selected radiomic features and discretized CAC for eight segments, where we can observe the mediocre correlation values. Figure 4-12 shows their corresponding *p*-values (not FDR corrected in this plot). Following Benjamini-Hochberg FDR correction no feature survives. This emphasizes the difficulty of the task at hand, and that it is necessary to adopt a more sophisticated, multivariate algorithm for regression (for continuous CAC outcome) or classification (for discrete CAC outcome).

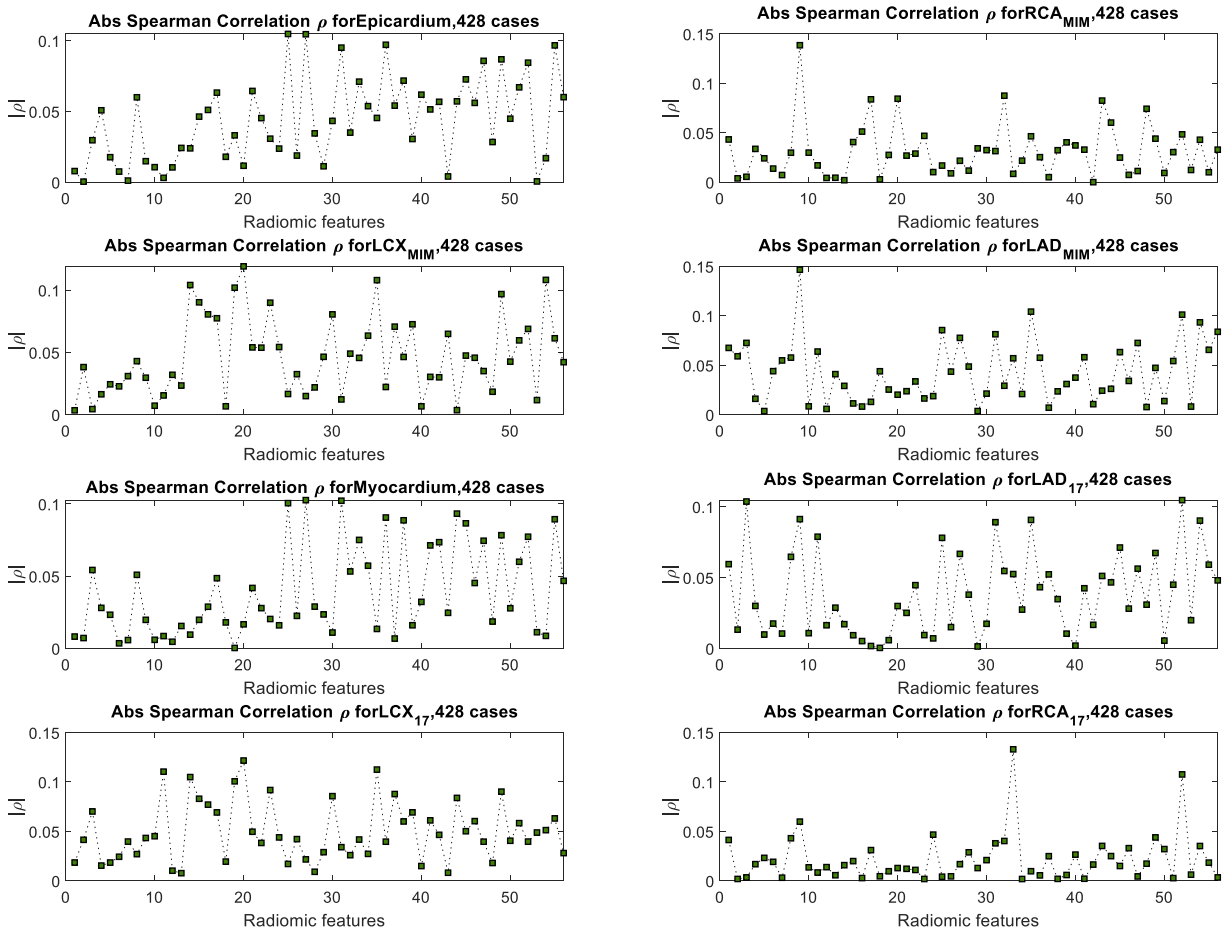

Figure 4-11. Spearman rank correlation between a selected feature of each segment (56 selected features) and the CAC of that segment. The maximum correlation observed in all plots is 0.15, which is mediocre.

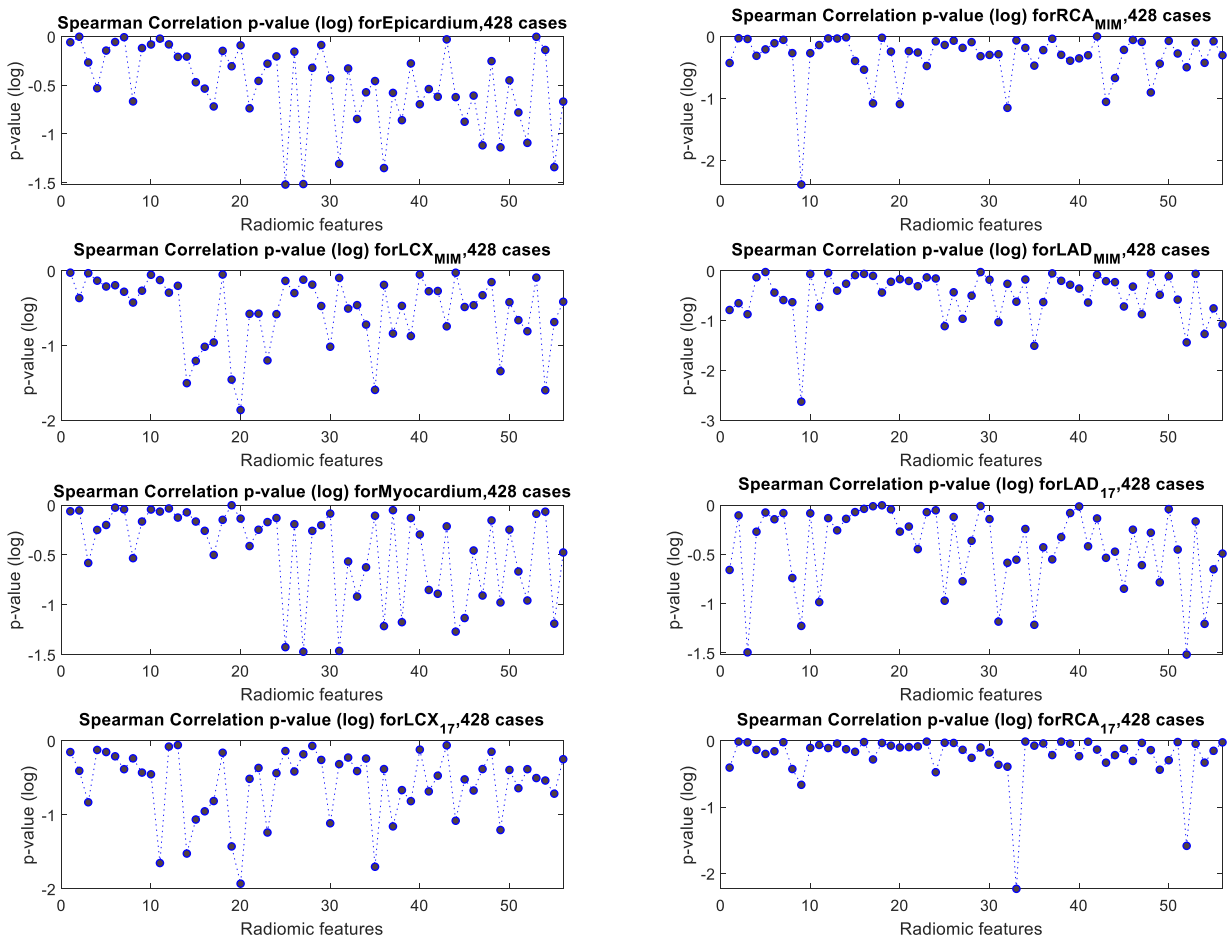

Figure 4-12. Spearman rank correlation p-value between a selected feature of each segment (56 selected features) and the CAC of that segment.

Multivariate analysis

We observed from Figure 4-11 that in general, the correlation values between features and CAC score are relatively low. But despite their low correlation, these selected features had a relatively-higher significance, giving us the hope that while none of these slightly-significant features are highly correlated with the outcome, but a certain multivariate combination of them might actually be predictive and provide significant prediction information. Thus, we pursue a multivariate approach to predicting CAC scores. In this subsection, first we introduce stepwise linear regression, then we describe how we handle feature selection. We then explain how our proposed algorithm manages

data to produce a fair analysis, and finally we run the analysis for three different configurations: i) radiomics features-only, ii) clinical features-only, and iii) radiomic + clinical features and present the results.

Stepwise linear regression

In this stage, we pursue a multivariate analysis approach employing stepwise linear regression. Stepwise regression is a systematic method for adding and removing terms from a linear or generalized linear model based on their statistical significance in explaining the response variable. The method begins with an initial model, which in our case is a linear model, and then compares the explanatory power of incrementally larger and smaller models, which is performed by adding or removing terms by stepwise regression and returning the linear model at the end. The initial fit can be a linear or a constant (intersect) model. After the initial fit, the function examines a set of available features and adds the best one to the model if an F-test for adding the term results in a p-value of P_{enter}, or less. If no terms can be added, it examines the terms currently in the model and removes the worst one if an F-test for removing it has a p-value of P_{remove}, or greater. This process is repeated until no terms can be added or removed. The constant term (intercept) is never removed from the model.

Feature handling

At each step, the method searches for terms to add to or remove from the model based on a criterion, which we selected it to be AIC, a commonly-used estimator of the relative quality of statistical models for a given dataset. AIC estimates the quality of each model relative to other models, providing a mean for model selection. It reduces the chance of overfitting *and* underfitting by providing a balance between goodness of fit and having too many parameters [497].

We can specify the order at which this algorithm starts to add features and later removes them. Instead of an unstructured approach of starting from the arbitrary first feature in the list, we

developed a feature selection method to enter those with higher Spearman correlation to the model first. For this purpose, the Spearman rank-correlation between each individual feature in the *training* set and the outcome (CAC score of the same segment) was calculated. The Spearman correlation coefficients and their corresponding *p*-values were recorded. Then, merely-significant features with a *p*-value smaller than a certain range (e.g. 0.3) were selected and others were discarded. The selected features were then sorted into descending order, based on the value of their Spearman correlation. The input dataset is then rearranged based on this subset of Spearman correlation-sorted features to enter features with the highest correlation to the stepwise algorithm first.

Training/cross-validation/testing setup

The following procedures were performed for each of the cardiac segments separately. First, the given dataset was shuffled and 15% of the data was set aside as the "independent test set". This set was not used until at the very end for independent assessment. Then, the following procedure was performed 20 times: the remaining 85% of the data "training + dev set" was randomly divided into training and cross-validation sets with 75%/25% ratios. The procedure described in the previous subsection has already reduced the number of radiomic features to 56. We use the procedure described in the previous subsection to further reduce the number of features and input more useful features for the regression algorithm first. We subsequently perform stepwise linear regression on the training set. We set P_{enter} as 0.05 and P_{remove} as 0.2.

Once the training is over, we perform cross-validation using the dev set. The aim of cross-validation is to reduce overfitting to the training set. The cross-validation algorithm is configured the same as training, except for the training algorithm the initial fit was a constant (intercept), whereas for cross-validation the initial fit is the output fit from the training dataset. During the above steps, we recorded the model, including the set of features remaining in it, the value of the log-likelihood, *p*-value and the final AIC.

The model fit is typically comprised of several features that survived the stepwise algorithm, and it might be possible that only the intercept term survives. If by coincidence the best model consists of only the intercept term, we skip that and choose the best fit with more than one term.

Following the above procedure, we select the model with the highest AIC of the 20 runs to run on the independent test set blind to the entire operation. To assess its prediction performance, Pearson's correlation was used to assess the relationship between the two distributions (prediction vs. actual), and subsequently recorded the correlation coefficients (ρ) and their corresponding p-values. The above operation was performed for each of the segmented lesions of the heart separately.

But this is not where we come to conclusion yet. We kept the test set aside during the whole analysis to assure a completely independent and blind-to-training assessment; however, our result might still be biased to a specific randomly-selected test set chosen. To even further mitigate such a bias, we took an extra step and run the entire above operation 50 times. That is, randomly shuffling and dividing the dataset into "training + dev" and "test" sets 50 times, then run the stepwise algorithm 20 times over the "training + dev" set. We subsequently perform 50 predictions on 50 independent test sets that give us 50 best regression fits and their p-values, which we subsequently used to derive our conclusion. A flowchart of the algorithm is depicted in Figure 4-13.

Running the multivariate analysis for three configurations

We performed the above entire operation three times: A) with radiomic features only (imaging), B) with clinical features (non-imaging), and C) with both radiomics and clinical features. The ten clinical features employed were i) gender, ii) race, iii) age, iv) smoking, v) diabetes, vi) hypertension, vii) hyperlipidemia, viii) family history of cardiac disease, ix) BMI, and x) LVEF. We also assured that a certain subset of clinical features such as gender, race, diabetes, etc. was treated as "categorical" variables, as opposed to continuous, by the algorithm.

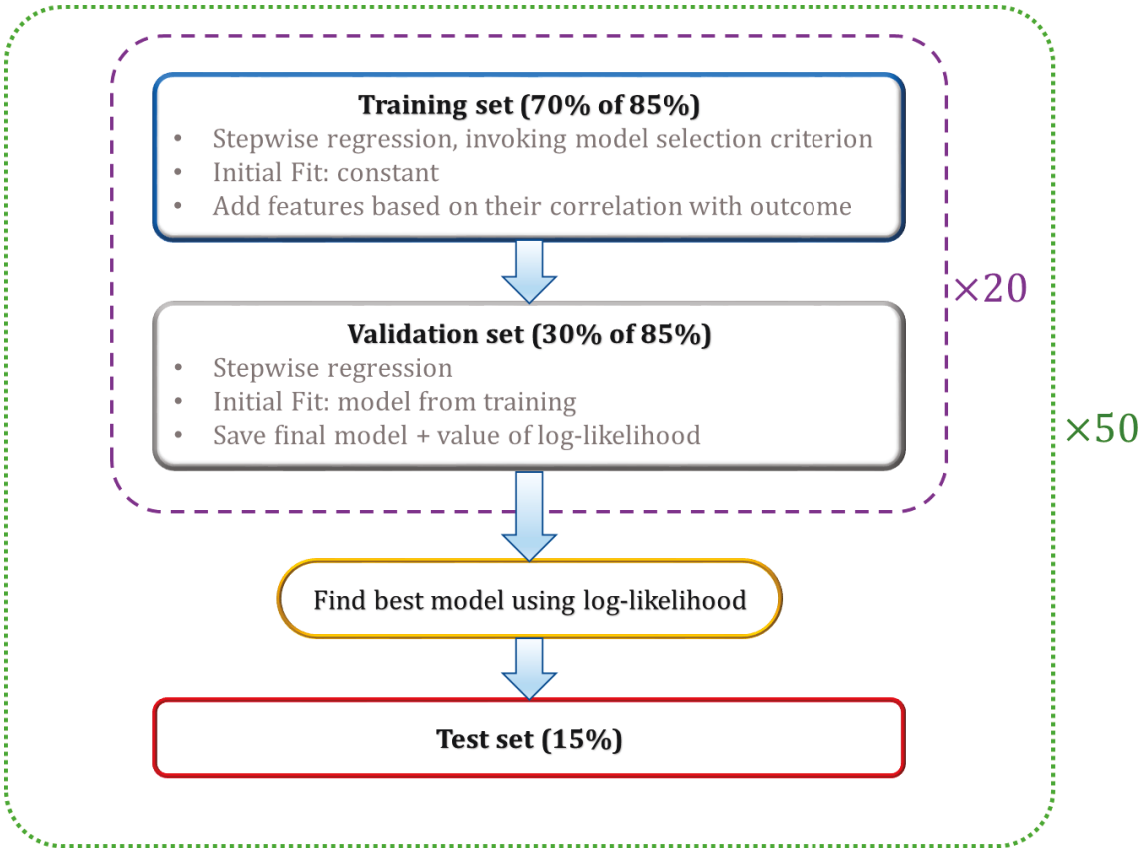

Figure 4-13. A simplistic flowchart of the algorithm.

Results

We used Fisher's method for combining p-values and use the chi-squared distribution test to determine if the stepwise regression method yields a significant fit after running 50 independent trials [498]. Based on this method, under the null hypothesis, the statistics of Fisher's method computed as

$$X = -2\sum_{i=1}^{N} \log(p_i) \qquad 4\text{-}1$$

follows a chi-squared distribution with a degree of freedom of $2 \times N$, where N is the total number of runs (in our case, $N = 50$). Comparing the value of X to the appropriate chi-squared distribution can determine whether the sample is significant. Assuming a significant level of 0.01, the value of the chi-

squared distribution for a degree-of-freedom of 2×50=100 is 135.81. Table 4-4 shows the result of applying Fisher's method to the three configurations, where significant results are shown in bold. We observed that radiomic features were unable to yield a significant model for any of the segmentations, and clinical features were able to result in a significant fit for most of the segments. But the combined clinical + radiomic features results in significant fit across all segments.

Table 4-4. The value of chi-squared distribution for each segment and feature configurations. The value of the chi-squared distribution with degree-of-freedom of 100 is 135.81, and values above this threshold (shown in bold) are considered significant under the null hypothesis.

	RCA_{MIM}	LCX_{MIM}	LAD_{MIM}	Myocardium	LAD_{17}	LCX_{17}	RCA_{17}
Radiomics	95.87	88.02	115.02	111.93	139.25	53.8	53.28
Clinical	84.12	**153.14**	**253.13**	**294.43**	**253.13**	**153.14**	84.12
Combined	**174.53**	**194.73**	**348.97**	**341.39**	**326.97**	**189.2**	**141.6**

shows the distribution of the absolute value of Pearson's correlation coefficient $|\rho|$ for all seven segments. We observe the same pattern across all segments that the combined radiomics + clinical features are more correlated to the CAC scores of that region. Moreover, Figure 4-15 shows the distribution of p-values of the best fit out of the 50 independent runs of the stepwise regression algorithm, each include 20 model fits where the best is selected. This plot shows that adding radiomics to the 10 clinical features will enhance the significance of the regression model and promising a more robust prediction.

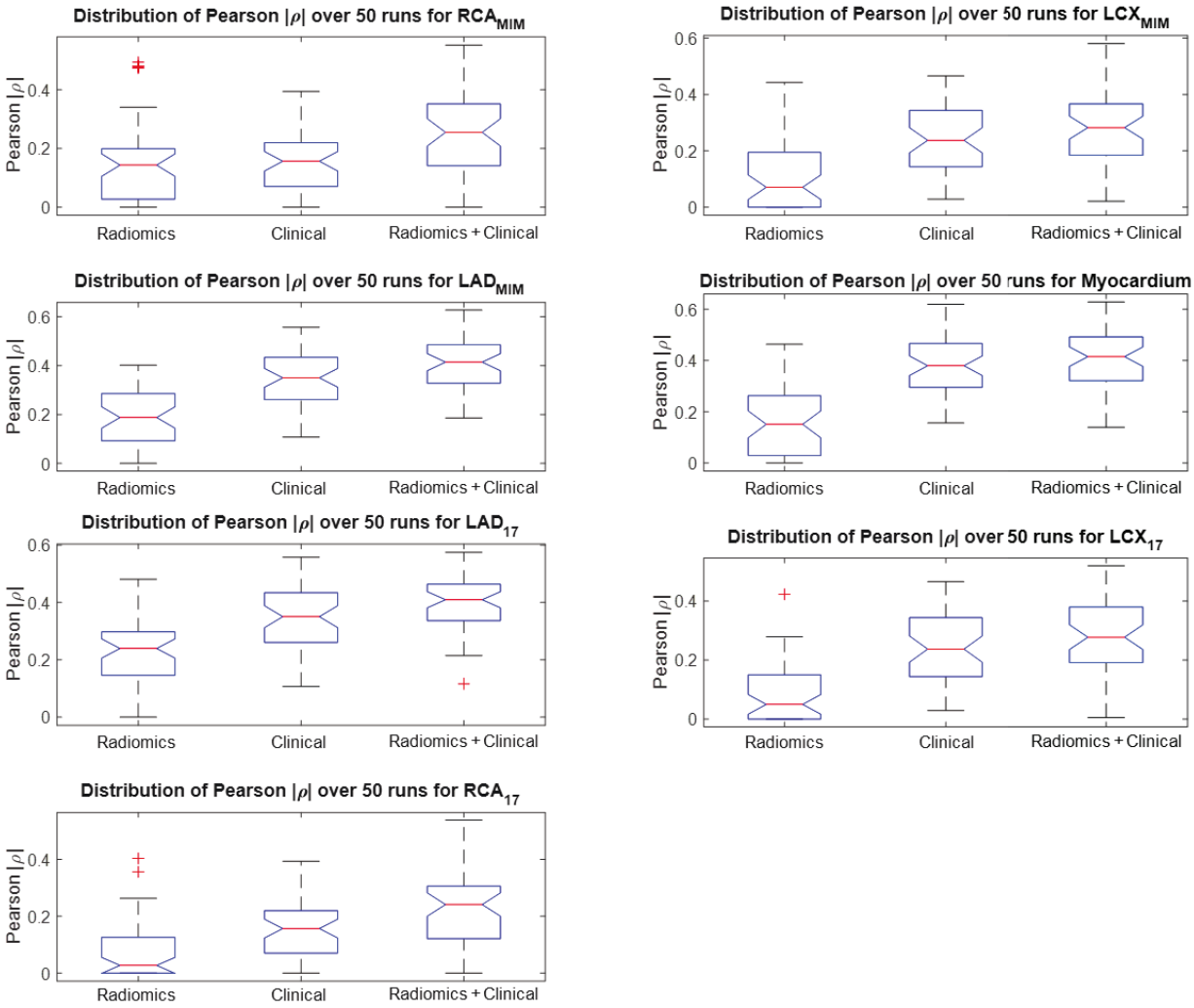

Figure 4-14. Distribution of absolute value of Pearson's ρ of the best fit out of 50 randomized trials of stepwise linear regression for radiomics, clinical and combined features, and for all 7 segmentations (the higher, the better). Adding radiomics to clinical features increases the correlation to the CAC score of the corresponding ROI.

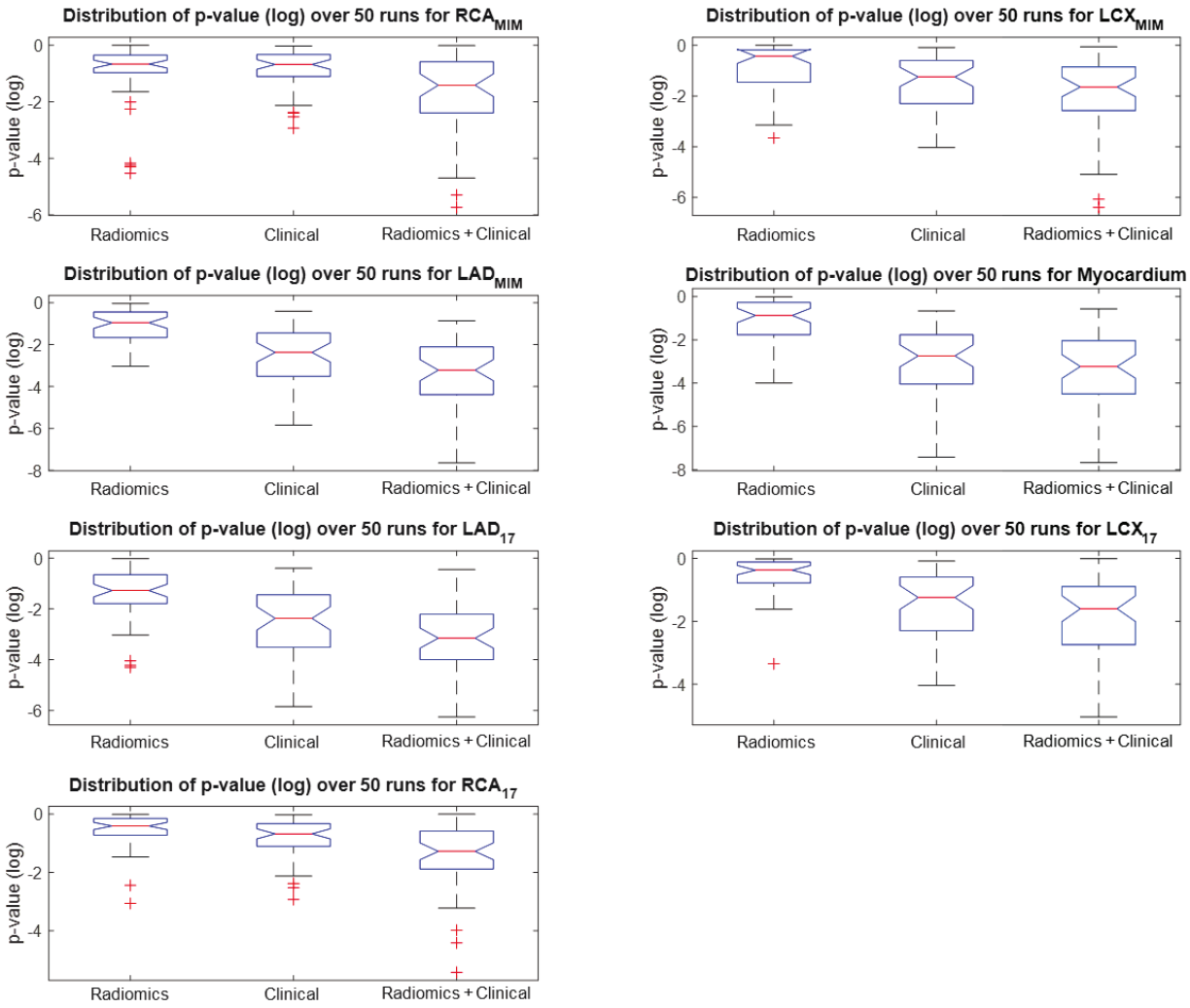

Figure 4-15. Distribution of p-values (log-scale) of the best fit out of 50 randomized trials of stepwise linear regression for radiomics, clinical and combined features, and for all 7 segmentations (the lower, the better). Adding radiomics to clinical features is seen to enhance the regression significance across all segmentations.

In clinical configurations, the most prevalent features in the best fit were somehow consistent across different segmentations and included age, hyperlipidemia, hypertension, and smoking. In the combined configuration, usually, the age variable was the first in the fit, followed by hyperlipidemia, GLSZM-small zone large GL emphasis, and GLDZM-short distance large GL emphasis.

4.2.4. Discussion

The current proposal is the first demonstration of employing radiomics of normal MP stress SPECT to predict CAC score as derived from the CT scan. To our knowledge, no study has been published on radiomics of cardiac SPECT imaging. Moreover, we did not find any study with the same approach as ours that incorporate readily-reconstructed 3D images and preserves the voxels intensities. They focus on using the polar plot for their analyses, which is a 2D projection of the 3D reconstructed image. Recently, few studies have investigated the use of deep learning to predict CAD [499-501]; nonetheless, no studies, to our knowledge exist on predicting CAC scores from SPECT scans, which is, as indicated earlier, a very challenging task.

4.2.4.1. Challenges with the proposed idea

The study of MP stress SPECT radiomics is a challenging task due to several reasons. First, SPECT is a low-resolution imaging modality that results in a substantial loss of heterogeneity information that had the potential to provide extra knowledge about the blood flow and other functionalities of the heart that could have captured by radiomics. Moreover, the lack of quantitation in SPECT imaging further causes a major loss of information, resulting in a mostly-qualitative interpretation of the scan. Of course, the absence of quantitation prevents the utilization of many useful radiomic features. It also impedes performing cross-scan comparisons. Another drawback of non-quantitative SPECT images can be explained by an example of a patient that has calcification in all three main arteries but has a uniform uptake in his SPECT image reported as normal. This can be due to a condition where blood flow is reduced in all three main arteries, resulting in uniformly decreased flow all around the heart. But since blood flow is not quantifiable, this effect cannot be noticed. However, methods to perform quantitative SPECT scan have been published and even recently been commercialized [502, 503]. Quantitative SPECT is shown to carry many clinical implications [504] and promises an increased chance of more accurate and impactful radiomics analysis of the heart.

One other reason that significantly contributes to the challenges in SPECT radiomics is heterogeneity caused by inherent artifacts of SPECT imaging. MP SPECT, specifically, can cause artifacts on the reconstructed image that can appear as reduced uptake in the image, an example of which is shown in Figure 4-16. This effect is called apical thinning and is a well-known phenomenon in MP SPECT. It is often attributed to a reduced myocardial thickness at the apex of the left ventricle. Attenuation correction during the reconstruction appears to exaggerate this effect [505]. Moreover, soft tissue attenuation artifacts also impact MP SPECT images [506]. These artifacts generally appear as fixed defects. Attenuation due to breast tissue usually results in a perfusion defect along the anterior wall of the left ventricle, also affecting the lateral wall, septum, and apex [507]. The effect would be similar to that in Figure 4-16. During our data collection phase, we observed many cases with this effect apparent in their reconstructed image. Undoubtedly, the heterogeneity caused by this effect may be captured by the radiomic features, while it is completely irrelevant to calcifications in arteries.

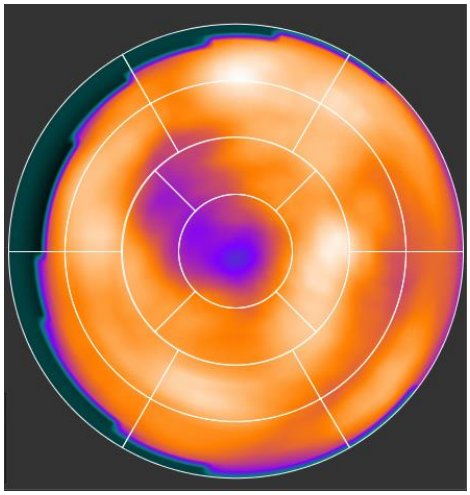

Figure 4-16. A normal MP stress SPECT with apical thinning.

Figure 4-17 shows an example of an MP stress SPECT scan image in a polar plot form, which a 2D projection of the 3D SPECT image into its apex (center circle). This image is interpreted as normal, due to the absence of any reversibility and/or defect. But the CT scan of this patient shows an enormous calcification in the arteries of this patient, having an outstanding CAC score of 2239

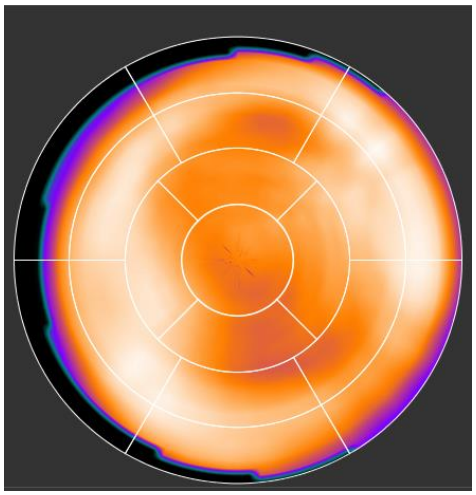

Figure 4-17. A normal MP SPECT with severe calcification. This scan is reported as normal due to relatively uniform uptake with no reversibility and/or fixed defect, but the CAC CT scan shows an extraordinary CAC score of 2239. The promise of our proposed research is to be able to provide assistance in finding such cases with elevated CAC score.

4.2.4.2. Radiomics intuition

We mentioned that radiomic features mostly included in the fit were GLSZM GLSZM-small zone large GL emphasis, and GLDZM-short distance large GL emphasis. Both features emphasize on higher GLs, and higher GLs in a discretized SPECT image depicts higher blood flow. It is interesting to observe and seems intuitive that the radiomic features who capture higher blood flow in each cardiac segment end up being in the fit.

4.2.4.3. Our other efforts

We wish to also point out that we explored more than 10 other regression methods, including different types of regression trees, support vector machine (SVM) regressors, etc., as well as several classification techniques (bagging, SVM, K-nearest neighbor, etc.) to find a significant prediction model, but our investigation did not return any significant results from the aforementioned techniques. Yet we do not exclude the possibility that with further tuning, those algorithms can potentially return significant results.

4.2.4.4. The significance of the study

In the current study, after many feature elimination steps discussed in section 4.2.3.2, and significantly reducing the feature space by a factor of 70, univariate analysis was not able to find any potential correlation with the outcome. On the other hand, our multivariate analysis carefully designed to mitigate the impact of dataset bias on the outcome prediction was able to successfully make a prediction for all segments of the heart. Our statistical analysis in section 4.2.3.1 showed that just around 60% of the patients had a non-zero CAC score and one-third of them had a CAC score≥100 that is shown to progressively increase the chance of myocardial ischemia. As a result, our multivariate analysis has the potential to make a prediction of CAC which is the most prevalent type of atherosclerosis, showing promise for this study.

4.2.5. Conclusion

This chapter investigated the hypothesis that heterogeneity in MP stress SPECT images can possibly convey information regarding calcification in coronary arteries. Many community settings are incapable of providing CAC CT scan for patients, it is not reimbursed by the CMS, and requires sophisticated software. We employed our in-house developed standardized SERA package that can evaluate 487 radiomic features. We segmented MP SPECT images into LAD, RCA, LCX, each with two varieties, as well as the whole myocardium, evaluating features for all 7segments. We also explored 8 different GLs to find the most appropriate setting for our study that yields higher reproducibility, robustness and less redundancy. Our dataset consists of 428 patients with normal (non-ischemic) MP stress SPECT images that were verified to be free from artifact or spillover, in addition to their detailed CAC score acquired from CT, and other clinical parameters. Our focus was on patients with normal stress scan since the possible prediction of coronary artery calcification in those images would have been of clinical significance. Through a multi-step blind-to-outcome unsupervised

feature selection phase, we significantly reduced the feature space 70 folds from 487×8 to 56 features. We also performed the entire operation 50 times to randomly divide our dataset into "training + dev" and "test" sets to mitigate any bias to a specific set of test data. Our univariate analysis using Spearman rank correlation between each feature of the cardiac segment with the corresponding CAC score of that segment was not significant. Our multivariate analysis, however, was able to significantly predict CAC score of all cardiac segments when combining radiomic features with clinical features. Our method has the potential to identify such cases with high coronary artery calcification that can be prompted for more appropriate care, suggesting that radiomics analysis adds diagnostic and prognostic value to standard MPS for wide clinical usage.

5. Summary and Future Work

This section briefly summarizes the main takeaways from the chapters in this dissertation.

5.1. Introduction

Our work aims to advance two frontiers of nuclear medicine imaging, namely image reconstruction, and radiomics, with the goal of enhanced quantitation. We devoted the first chapter to introduce the basics of these two fields. We started by introducing nuclear medicine and tomographic imaging, PET and SPECT, their components, as well as image generation and tomographic reconstruction techniques. We briefly discussed the vast applications of each modality. We subsequently elaborated 12 causes of image degradation. These were required to better illustrate the problem statement of chapter 2 since the promise of PSF modeling is to correct for a subset of those image degradations. The next part introduced medical image quantitation, where we provided a detailed definition of image quantitation, biomarkers, quantitative biomarkers, and provided an example of one of the most commonly-used imaging biomarkers in nuclear medicine. We subsequently discussed radiomics, the large-scale data-oriented study of radiological images for potential discovery of imaging biomarkers that can be used in the clinic and provide additional valuable information to radiologists. The promise of radiomics is to enable personalized medicine, which we discussed in section 1.4. We provided a roadmap of how radiomics can lead to advancements in personalized medicine. We briefly introduced radiomics and mentioned a number of selected applications utilizing

radiomics. We then concluded by an overall motivation for our dissertation and a brief summary of our published work.

5.2. Chapter 2: Adaptive PSF modeling

5.2.1. Summary

This chapter focused on a commonly-used "image reconstruction" technique in PET imaging—PSF modeling. The advantages and challenges of PSF modeling were discussed and a new approach was proposed with the aim to enhance image quantitation. The new approach stepped beyond the past frameworks involving a dichotomy of PSF vs. no-PSF modeling, focusing on a wide-spectrum of PSF models for the potential of enhanced quantitation of standardized uptake values (SUVs). Starting from a simplistic simulation and reconstruction framework, our proposed method was shown to enhance quantitative task performance. Following elaboration of system modeling in PET as well as analytical modeling of image degradation effects, a comprehensive PET simulation and image reconstruction framework were proposed that considered a range of realistically-modeled resolution degradation phenomena that analytically modeled the true PSF. Various measures of noise and bias were defined and subsequently used to demonstrate the efficiency of the proposed scheme. The results of our study were shown to follow that of our preliminary image-based PSF modeling presented earlier in the chapter. Our results demonstrated that for the standard range of iterations employed in the clinic (not excessive), edge enhancement due to overestimation counter-intuitively lowered SUV bias in small tumors, while inter-voxel correlations suppressed image roughness and enhanced uniformity in all tumors, only slightly degrading SUVmean reproducibility in the smallest tumors. Unlike what one may at first imagine that overestimating the PSF would lead to higher overshoots at the edges, this was only observed at higher iterations. In fact, using an overestimated PSF resulted in increased contrast and limited edge overshoot effect at lower iterations, in turn

enabling enhanced SUV quantitation. Another important takeaway of this chapter was that the exact matched system PSF might not be the optimized option for enhanced image reconstruction performance and that slightly overestimated PSF modeling can improve PET image quantitation.

5.2.2. Future work

An interesting real-world implication of this work is in PET scanner calibration. Some commercial PET scanners use Ge-68 point-sources to characterize the PSF as modeled within PSF reconstruction. However, in comparison, the most popular isotope, F-18, has a relatively small positron range. Thus, utilizing a PSF kernel obtained from Ge-68 point-sources with a significant positron range to reconstruct F-18 PET data is approximately equivalent to an overestimated PSF kernel in the reconstruction, because both increase the FWHM of the PSF. It is interesting to note, based on our observations with overestimated PSF kernels, that this may not be a problem in fact, and may effectively lead to improvements in quantitative performance, though further analysis is required to implement the exact model of radiotracers with higher positron range, that is left for future study.

Our aim during the course of our research was to introduce and explore more groundbreaking ideas, as opposed to making slight improvements to existing concepts. Although our comprehensive noise-bias study demonstrated the effectiveness of our proposed model, it can be further investigated against some other important parameters in image reconstruction, including different tumor contrast, adding post-smoothing with different filters, adopting penalized reconstruction, and the incorporation of analytical modeling of positron range.

5.3. Chapter 3: Standardized radiomics

5.3.1. Summary

While radiomics was briefly introduced in Chapter 1, Chapter 3 presented a detailed and systematic view of standardized radiomics workflows. As a typical workflow contains numerous steps, as outlined, including image preprocessing, workflow configuration and feature calculation, the overall complex framework was understood to be highly prone to variability, impacting its robustness and reproducibility. Certain precautions were also discussed to prevent overfitting and to correct for false-discovery rate, which, many past published radiomics efforts suffer from. Our efforts within the IBSI collaboration were illustrated, where it was seen that even for centers using the same images and ROIs, the computed features were not consistent.

We introduced our developed standardized environment for radiomics analysis (SERA) consisting of 487 radiomic features. SERA was used in multiple projects, including studies at the end of this chapter aiming at discovering robust and reproducible radiomic features for two nuclear medicine datasets: renal cell carcinoma ^{99m}Tc-Sestamibi SPECT and ^{18}F-DCFPyL PSMA PET images of prostate cancer. Although both studies were the first demonstration of radiomics for these radiotracers, our results were in agreement with past efforts on more commonly-used tracers such as FDG. Our results also matched the findings in Chapter 4. This chapter further reported on the importance of GL discretization selection, where 64 GLs was seen to convey enough details about image heterogeneity, as well as the preference for more sophisticated segmentation to improve reproducibility of features.

5.3.2. Future work

Immediate future research from the concepts introduced in this chapter is to assess the reported robust and reproducible features in outcome prediction tasks. For our study of radiomics of cold-uptake in RCC SPECT described in section 3.3.1, we tried to use some machine learning techniques to predict benign vs. malignancy of tumors, which were not successful, which we attribute to the

relatively small number of datasets that are more appropriate for reproducibility analysis. Many radiomic studies have been published, yet some of the details of the feature evaluation setting has not yet been carefully studied. For instance, different feature families may be optimized at different discretization levels. The assumption in the efforts presented in this dissertation, as well as other published studies, is to consider one GL for all feature families. However, the effect of discretization levels in different feature families should be investigated. Even further, the effect of discretization in every feature can study too. The concern of the scientific community over radiomics study with relatively-low reproducibility as well as the tendency to report positive results require more efforts on developing and publishing workflows for performing standardized feature selection, and/or standardized feature classification, similar to IBSI that aims at standardized feature evaluation. They are not as many studies with feature selection/classification analysis that are statistically sound. For instance, how to approach the feature classification at the presence of hundreds of features, many of which might not convey any useful information. Or if we are interested to study differentiation of two features that can be performed using the widely-used area under the curve (AUC) analysis of the receiver operative characteristics (ROC) curve, how can it be compared with a third or more variable; i.e. how can we hypothesize ROC analysis of more than two features.

5.4. Chapter 4: Radiomics of Myocardial Perfusion Stress SPECT to predict CAC Score as Captured by CT

5.4.1. Summary

The final chapter contained an end-to-end application of radiomics in nuclear cardiology. This was the first demonstration of utilizing radiomics on cardiac SPECT imaging to derive CT-based CAC information that otherwise is absent from conventional analysis. The importance of MP stress SPECT and CAC scoring from CT in the clinic were described. Given that CAC assessment is not commonly

performed nor reimbursed in a wide community setting, this project has an important implication in the clinic.

SERA-derived radiomic features were utilized in a multi-step feature selection framework, followed by application of machine learning to radiomic feature. Feature selection was completely blind to the outcome, and the selected features were subsequently utilized in machine learning efforts. Results of this chapter demonstrated the possibility that certain information about CAC scoring can be derived from radiomics of MP SPECT, further emphasizing the value of radiomics in extracting visually-unseen information from radiological images. 446035288325

5.4.2. Future work

This was a difficult challenge! And we explained several reasons that contribute to the complexity of this problem in section 4.2.4: Discussion. We started with a problem for which univariate analysis demonstrated very poor performance. We tried a number of regression and classification techniques for outcome prediction in this project. The results demonstrated added value in utilizing extracted radiomic features; significantly predicting CAC score of the LAD segment of the heart using a combination of radiomic features of MP SPECT and clinical features. There is scope for the use of more advanced machine learning techniques for mapping the complex data to the outcome. Our dataset consisted of a medium-sized patient population of 428 patients. Larger datasets can be collected and applied to deep learning frameworks where features are implicitly extracted, which may be better able to handle complex datasets.

Moreover, we note that CAC scoring is only an intermediate step in clinical decision making, aiming to help and improve patient stratification. As such, important future studies can be designed to use the radiomics of SPECT to directly predict patient outcome and to enable clinical decision support, such as recommendations for procedures such as catheterization.

References

REFERENCES

[1] A. Chauffard, “Address in Medicine, ON MEDICAL PROGNOSIS: ITS METHODS, ITS EVOLUTION, ITS LIMITATIONS: Delivered at the Seventeenth International Congress of Medicine,” *Br Med J,* vol. 2, no. 2745, pp. 286-90, Aug 9, 1913.

[2] J. Y. Park, M. S. Choi, Y.-S. Lim, J. W. Park, S. U. Kim, Y. W. Min, G.-Y. Gwak, Y.-H. Paik, J. H. Lee, K. C. Koh, S. W. Paik, and B. C. Yoo, “Clinical features, image findings, and prognosis of inflammatory pseudotumor of the liver: a multicenter experience of 45 cases,” *Gut and liver,* vol. 8, no. 1, pp. 58-63, 2014.

[3] K. Usuda, Y. Saito, M. Sagawa, M. Sato, K. Kanma, S. Takahashi, C. Endo, Y. Chen, A. Sakurada, and S. Fujimura, “Tumor doubling time and prognostic assessment of patients with primary lung cancer,” *Cancer,* vol. 74, no. 8, pp. 2239-2244, 1994.

[4] C. Parmar, R. T. H. Leijenaar, P. Grossmann, E. Rios Velazquez, J. Bussink, D. Rietveld, M. M. Rietbergen, B. Haibe-Kains, P. Lambin, and H. J. W. L. Aerts, “Radiomic feature clusters and Prognostic Signatures specific for Lung and Head & Neck cancer,” *Scientific Reports,* vol. 5, pp. 11044, 06/05/online, 2015.

[5] T. Win, K. A. Miles, S. M. Janes, B. Ganeshan, M. Shastry, R. Endozo, M. Meagher, R. I. Shortman, S. Wan, and I. Kayani, “Tumor heterogeneity and permeability as measured on the CT component

of PET/CT predict survival in patients with non–small cell lung cancer," *Clinical Cancer Research*, 2013.

[6] Y. Ohno, H. Koyama, T. Yoshikawa, K. Matsumoto, N. Aoyama, Y. Onishi, and K. Sugimura, "Diffusion-weighted MRI versus 18F-FDG PET/CT: performance as predictors of tumor treatment response and patient survival in patients with non–small cell lung cancer receiving chemoradiotherapy," *American Journal of Roentgenology,* vol. 198, no. 1, pp. 75-82, 2012.

[7] C. Jahnke, E. Nagel, R. Gebker, T. Kokocinski, S. Kelle, R. Manka, E. Fleck, and I. Paetsch, "Prognostic value of cardiac magnetic resonance stress tests: adenosine stress perfusion and dobutamine stress wall motion imaging," *Circulation,* vol. 115, no. 13, pp. 1769-1776, 2007.

[8] W. G. Hundley, T. M. Morgan, C. M. Neagle, C. A. Hamilton, P. Rerkpattanapipat, and K. M. Link, "Magnetic resonance imaging determination of cardiac prognosis," *Circulation,* vol. 106, no. 18, pp. 2328-2333, 2002.

[9] M. I. Travin, G. V. Heller, L. L. Johnson, D. Katten, A. W. Ahlberg, C. R. Isasi, R. C. Kaplan, C. C. Taub, and D. Demus, "The prognostic value of ECG-gated SPECT imaging in patients undergoing stress Tc-99m sestamibi myocardial perfusion imaging," *Journal of nuclear cardiology,* vol. 11, no. 3, pp. 253-262, 2004.

[10] B. G. Abbott, M. Afshar, A. K. Berger, and J. Frans, "Prognostic significance of ischemic electrocardiographic changes during adenosine infusion in patients with normal myocardial perfusion imaging," *Journal of nuclear cardiology,* vol. 10, no. 1, pp. 9-16, 2003.

[11] G. Pilz, A. Jeske, M. Klos, E. Ali, B. Hoefling, R. Scheck, and P. Bernhardt, "Prognostic value of normal adenosine-stress cardiac magnetic resonance imaging," *The American journal of cardiology,* vol. 101, no. 10, pp. 1408-1412, 2008.

[12] K. Yoshinaga, R. S. B. Beanlands, R. A. deKemp, M. Lortie, J. Morin, M. Aung, R. McKelvie, and R. F. Davies, "Effect of exercise training on myocardial blood flow in patients with stable coronary artery disease," *American Heart Journal,* vol. 151, no. 6, pp. -, Jun, 2006.

[13] Y. Perel, J. Conway, M. Kletzel, J. Goldman, S. Weiss, A. Feyler, and S. L. Cohn, "Clinical impact and prognostic value of metaiodobenzylguanidine imaging in children with metastatic neuroblastoma," *Journal of pediatric hematology/oncology,* vol. 21, no. 1, pp. 13-18, 1999.

[14] S. Abramson, W. Berdon, C. Ruzal-Shapiro, C. Stolar, and J. Garvin, "Cervical neuroblastoma in eleven infants—a tumor with favorable prognosis," *Pediatric radiology,* vol. 23, no. 4, pp. 253-257, 1993.

[15] H. Yamashita, T. Fujikawa, H. Takami, I. Yanai, Y. Okamoto, S. Morinobu, and S. Yamawaki, "Long-term prognosis of patients with major depression and silent cerebral infarction," *Neuropsychobiology,* vol. 62, no. 3, pp. 177-181, 2010.

[16] M. Filippi, and M. A. Rocca, "The role of magnetic resonance imaging in the study of multiple sclerosis: diagnosis, prognosis and understanding disease pathophysiology," *Acta Neurol Belg,* vol. 111, no. 2, pp. 89-98, 2011.

[17] N. R. Galloway, K. A. Tong, S. Ashwal, U. Oyoyo, and A. Obenaus, "Diffusion-weighted imaging improves outcome prediction in pediatric traumatic brain injury," *Journal of neurotrauma,* vol. 25, no. 10, pp. 1153-1162, 2008.

[18] I. F. Ciernik, E. Dizendorf, B. G. Baumert, B. Reiner, C. Burger, J. B. Davis, U. M. Lütolf, H. C. Steinert, and G. K. Von Schulthess, "Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study," *International Journal of Radiation Oncology* Biology* Physics,* vol. 57, no. 3, pp. 853-863, 2003.

[19] Y. E. Erdi, K. Rosenzweig, A. K. Erdi, H. A. Macapinlac, Y.-C. Hu, L. E. Braban, J. L. Humm, O. D. Squire, C.-S. Chui, and S. M. Larson, "Radiotherapy treatment planning for patients with non-small

cell lung cancer using positron emission tomography (PET)," *Radiotherapy and Oncology,* vol. 62, no. 1, pp. 51-60, 2002.

[20] V. Grégoire, K. Haustermans, X. Geets, S. Roels, and M. Lonneux, "PET-based treatment planning in radiotherapy: a new standard?," *Journal of Nuclear Medicine,* vol. 48, no. 1, pp. 68S, 2007.

[21] A. W. Glaudemans, E. F. de Vries, F. Galli, R. A. Dierckx, R. H. Slart, and A. Signore, "The Use of F-FDG-PET/CT for diagnosis and treatment monitoring of inflammatory and infectious diseases," *Clinical and Developmental Immunology,* vol. 2013, 2013.

[22] W. A. Weber, and R. Figlin, "Monitoring cancer treatment with PET/CT: does it make a difference?," *Journal of Nuclear Medicine,* vol. 48, no. 1 suppl, pp. 36S-44S, 2007.

[23] N. E. Avril, and W. A. Weber, "Monitoring response to treatment in patients utilizing PET," *Radiologic Clinics,* vol. 43, no. 1, pp. 189-204, 2005.

[24] W. A. Weber, "Use of PET for monitoring cancer therapy and for predicting outcome," *Journal of Nuclear Medicine,* vol. 46, no. 6, pp. 983, 2005.

[25] S. Ben-Haim, and P. Ell, "18F-FDG PET and PET/CT in the evaluation of cancer treatment response," *Journal of Nuclear Medicine,* vol. 50, no. 1, pp. 88-99, 2009.

[26] M. Hutchings, and S. F. Barrington, "PET/CT for therapy response assessment in lymphoma," *Journal of nuclear medicine*, 2009.

[27] J. Bussink, C. M. van Herpen, J. H. Kaanders, and W. J. Oyen, "PET-CT for response assessment and treatment adaptation in head and neck cancer," *The lancet oncology,* vol. 11, no. 7, pp. 661-669, 2010.

[28] F. G. Dhermain, P. Hau, H. Lanfermann, A. H. Jacobs, and M. J. van den Bent, "Advanced MRI and PET imaging for assessment of treatment response in patients with gliomas," *The Lancet Neurology,* vol. 9, no. 9, pp. 906-920, 2010.

[29] H. C. Thoeny, and B. D. Ross, “Predicting and monitoring cancer treatment response with diffusion-weighted MRI,” *Journal of Magnetic Resonance Imaging,* vol. 32, no. 1, pp. 2-16, 2010.

[30] S. P. Li, and A. R. Padhani, “Tumor response assessments with diffusion and perfusion MRI,” *Journal of Magnetic Resonance Imaging,* vol. 35, no. 4, pp. 745-763, 2012.

[31] K. Doi, “Computer-aided diagnosis in medical imaging: Historical review, current status and future potential,” *Computerized Medical Imaging and Graphics,* vol. 31, no. 4, pp. 198-211, 2007/06/01/, 2007.

[32] M. L. Giger, Z. Huo, M. A. Kupinski, and C. J. Vyborny, “Computer-aided diagnosis in mammography,” *Handbook of medical imaging,* vol. 2, pp. 915-1004, 2000.

[33] M. L. Giger, "Computerized analysis of images in the detection and diagnosis of breast cancer." pp. 411-418.

[34] B. J. Erickson, and B. Bartholmai, “Computer-aided detection and diagnosis at the start of the third millennium,” *Journal of digital imaging,* vol. 15, no. 2, pp. 59-68, 2002.

[35] R. M. Summers, “Road maps for advancement of radiologic computer-aided detection in the 21st century,” *Radiology,* vol. 229, no. 1, pp. 11-13, 2003.

[36] H. Abe, H. MacMahon, J. Shiraishi, Q. Li, R. Engelmann, and K. Doi, "Computer-aided diagnosis in chest radiology." pp. 432-437.

[37] Q. Li, F. Li, K. Suzuki, J. Shiraishi, H. Abe, R. Engelmann, Y. Nie, H. MacMahon, and K. Doi, "Computer-aided diagnosis in thoracic CT." pp. 357-363.

[38] H. Yoshida, and A. H. Dachman, "Computer-aided diagnosis for CT colonography." pp. 419-431.

[39] K. Doi, M. Giger, H. MacMahon, K. Hoffmann, R. Nishikawa, R. Schmidt, K. Chua, S. Katsuragawa, N. Nakamori, and S. Sanada, "Computer-aided diagnosis: development of automated schemes for quantitative analysis of radiographic images." pp. 140-152.

[40] T. W. Freer, and M. J. Ulissey, "Screening Mammography with Computer-aided Detection: Prospective Study of 12,860 Patients in a Community Breast Center," *Radiology,* vol. 220, no. 3, pp. 781-786, 2001.

[41] R. L. Birdwell, P. Bandodkar, and D. M. Ikeda, "Computer-aided detection with screening mammography in a university hospital setting," *Radiology,* vol. 236, no. 2, pp. 451-457, 2005.

[42] T. E. Cupples, J. E. Cunningham, and J. C. Reynolds, "Impact of computer-aided detection in a regional screening mammography program," *American Journal of Roentgenology,* vol. 185, no. 4, pp. 944-950, 2005.

[43] M. J. Morton, D. H. Whaley, K. R. Brandt, and K. K. Amrami, "Screening mammograms: interpretation with computer-aided detection—prospective evaluation," *Radiology,* vol. 239, no. 2, pp. 375-383, 2006.

[44] S. S. Kelkar, and T. M. Reineke, "Theranostics: combining imaging and therapy," *Bioconjugate chemistry,* vol. 22, no. 10, pp. 1879-1903, 2011.

[45] R. Bardhan, S. Lal, A. Joshi, and N. J. Halas, "Theranostic nanoshells: from probe design to imaging and treatment of cancer," *Accounts of chemical research,* vol. 44, no. 10, pp. 936-946, 2011.

[46] E. Y. Yu, and D. A. Mankoff, "Positron emission tomography imaging as a cancer biomarker," *Expert review of molecular diagnostics,* vol. 7, no. 5, pp. 659-672, 2007.

[47] F. Pene, E. Courtine, A. Cariou, and J.-P. Mira, "Toward theragnostics," *Critical care medicine,* vol. 37, no. 1, pp. S50-S58, 2009.

[48] R. P. Baum, and H. R. Kulkarni, "THERANOSTICS: From Molecular Imaging Using Ga-68 Labeled Tracers and PET/CT to Personalized Radionuclide Therapy - The Bad Berka Experience," *Theranostics,* vol. 2, no. 5, pp. 437-447, 2012.

[49] F. Aoun, H. R. Kourie, C. Artigas, and T. Roumeguère, “Next revolution in molecular theranostics: personalized medicine for urologic cancers,” *Future Oncology,* vol. 11, no. 15, pp. 2205-2219, 2015.

[50] J. Zhang, H. R. Kulkarni, A. Singh, and R. P. Baum, “Ten-Year Survival After Peptide Receptor Radionuclide Therapy of a Metastatic Well-differentiated G3 Pancreatic Neuroendocrine Neoplasm,” *Clin Nucl Med,* vol. 43, no. 9, pp. 676-678, Sep, 2018.

[51] K. Y. Choi, G. Liu, S. Lee, and X. Chen, “Theranostic nanoplatforms for simultaneous cancer imaging and therapy: current approaches and future perspectives,” *Nanoscale,* vol. 4, no. 2, pp. 330-342, 2012.

[52] F. Rösch, and R. P. Baum, “Generator-based PET radiopharmaceuticals for molecular imaging of tumours: on the way to THERANOSTICS,” *Dalton transactions,* vol. 40, no. 23, pp. 6104-6111, 2011.

[53] S. R. Cherry, J. A. Sorenson, and M. E. Phelps, *Physics in Nuclear Medicine*: Elsevier Health Sciences, 2012.

[54] J. L. Prince, and J. M. Links, *Medical imaging signals and systems*, 2015.

[55] E. Goshen, L. Beilin, E. Stern, T. Kenig, R. Goldkorn, and S. Ben-Haim, “Feasibility study of a novel general purpose CZT-based digital SPECT camera: initial clinical results,” *EJNMMI physics,* vol. 5, no. 1, pp. 6-6, 2018.

[56] R. Chandra, and A. Rahmim, *Nuclear medicine physics: the basics*: Lippincott Williams & Wilkins, 2017.

[57] P. Bhargava, G. He, A. Samarghandi, and E. S. Delpassand, “Pictorial review of SPECT/CT imaging applications in clinical nuclear medicine,” *American journal of nuclear medicine and molecular imaging,* vol. 2, no. 2, pp. 221-231, 2012.

[58] G. Mariani, L. Bruselli, T. Kuwert, E. E. Kim, A. Flotats, O. Israel, M. Dondi, and N. Watanabe, “A review on the clinical uses of SPECT/CT,” *Eur J Nucl Med Mol Imaging,* vol. 37, no. 10, pp. 1959-85, Oct, 2010.

[59] A. K. Buck, S. Nekolla, S. Ziegler, A. Beer, B. J. Krause, K. Herrmann, K. Scheidhauer, H. J. Wester, E. J. Rummeny, M. Schwaiger, and A. Drzezga, “SPECT/CT,” *J Nucl Med,* vol. 49, no. 8, pp. 1305-19, Aug, 2008.

[60] B. Bybel, R. C. Brunken, F. P. DiFilippo, D. R. Neumann, G. Wu, and M. D. Cerqueira, “SPECT/CT imaging: clinical utility of an emerging technology,” *Radiographics,* vol. 28, no. 4, pp. 1097-113, Jul-Aug, 2008.

[61] C. N. Patel, F. U. Chowdhury, and A. F. Scarsbrook, “Hybrid SPECT/CT: the end of "unclear" medicine,” *Postgrad Med J,* vol. 85, no. 1009, pp. 606-13, Nov, 2009.

[62] G. Mariani, S. A. Gulec, D. Rubello, G. Boni, M. Puccini, M. R. Pelizzo, G. Manca, D. Casara, G. Sotti, and P. Erba, “Preoperative localization and radioguided parathyroid surgery,” *Journal of Nuclear Medicine,* vol. 44, no. 9, pp. 1443-1458, 2003.

[63] H. Palmedo, H. J. Biersack, S. Lastoria, J. Maublant, E. Prats, H. Stegner, P. Bourgeois, R. Hustinx, A. J. Hilson, and A. Bischof-Delaloye, “Scintimammography with technetium-99m methoxyisobutylisonitrile: results of a prospective European multicentre trial,” *European journal of nuclear medicine,* vol. 25, no. 4, pp. 375-385, 1998.

[64] M. Grosso, S. Chiacchio, F. Bianchi, C. Traino, C. Marini, A. Cilotti, G. Manca, D. Volterrani, M. Roncella, and L. Rampin, “Comparison between 99mTc-sestamibi scintimammography and X-ray mammography in the characterization of clusters of microcalcifications: a prospective long-term study,” *Anticancer research,* vol. 29, no. 10, pp. 4251-4257, 2009.

[65] G. Mariani, L. Moresco, G. Viale, G. Villa, M. Bagnasco, G. Canavese, J. Buscombe, H. W. Strauss, and G. Paganelli, "Radioguided sentinel lymph node biopsy in breast cancer surgery," *Journal of Nuclear Medicine,* vol. 42, no. 8, pp. 1198-1215, 2001.

[66] S. Kneifel, P. Bernhardt, H. Uusijärvi, S. Good, L. Plasswilm, C. Buitrago-Téllez, J. Müller-Brand, H. Mäcke, and A. Merlo, "Individual voxelwise dosimetry of targeted 90 Y-labelled substance P radiotherapy for malignant gliomas," *European journal of nuclear medicine and molecular imaging,* vol. 34, no. 9, pp. 1388-1395, 2007.

[67] T. El-Maghraby, H. Moustafa, and E. Pauwels, "Nuclear medicine methods for evaluation of skeletal infection among other diagnostic modalities," *The Quarterly Journal of Nuclear Medicine and Molecular Imaging,* vol. 50, no. 3, pp. 167, 2006.

[68] N. Prandini, E. Lazzeri, B. Rossi, P. Erba, M. G. Parisella, and A. Signore, "Nuclear medicine imaging of bone infections," *Nuclear medicine communications,* vol. 27, no. 8, pp. 633-644, 2006.

[69] D. L. Feingold, F. J. Caliendo, B. T. Chinn, J. R. Notaro, G. C. Oliver, E. P. Salvati, K. B. Wilkins, and T. E. Eisenstat, "Does hemodynamic instability predict positive technetium-labeled red blood cell scintigraphy in patients with acute lower gastrointestinal bleeding? A review of 50 patients," *Diseases of the colon & rectum,* vol. 48, no. 5, pp. 1001-1004, 2005.

[70] N. Yama, E. Ezoe, Y. Kimura, M. Mukaiya, K. Fujimori, Y. Kurimoto, K. Koito, K. Hirata, M. Hareyama, and Y. Asai, "Localization of intestinal bleeding using a fusion of Tc-99m-labeled RBC SPECT and X-ray CT," *Clinical nuclear medicine,* vol. 30, no. 7, pp. 488-489, 2005.

[71] P. Reinartz, J. E. Wildberger, W. Schaefer, B. Nowak, A. H. Mahnken, and U. Buell, "Tomographic imaging in the diagnosis of pulmonary embolism: a comparison between V/Q lung scintigraphy in SPECT technique and multislice spiral CT," *Journal of Nuclear Medicine,* vol. 45, no. 9, pp. 1501-1508, 2004.

[72] T. Sulkin, and C. Cousens, “SPECTCT cerebral perfusion scintigraphy; is the low-dose CT component of diagnostic value?,” *Clinical radiology,* vol. 63, no. 3, pp. 289-298, 2008.

[73] A. Gimelli, G. Rossi, P. Landi, P. Marzullo, G. Iervasi, A. L'Abbate, and D. Rovai, “Stress/rest myocardial perfusion abnormalities by gated SPECT: Still the best predictor of cardiac events in stable ischemic heart disease,” *Journal of Nuclear Medicine,* vol. 50, no. 4, pp. 546, 2009.

[74] A. Rahmim, and H. Zaidi, “PET versus SPECT: strengths, limitations and challenges,” *Nucl Med Commun,* vol. 29, no. 3, pp. 193-207, Mar, 2008.

[75] C. S. Levin, and E. J. Hoffman, “Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution,” *Physics in Medicine & Biology,* vol. 44, no. 3, pp. 781, 1999.

[76] J. Cal-González, J. L. Herraiz, S. España, P. M. G. Corzo, J. J. Vaquero, M. Desco, and J. M. Udias, “Positron range estimations with PeneloPET,” *Physics in Medicine and Biology,* vol. 58, no. 15, pp. 5127-5152, 2013/07/09, 2013.

[77] M. Partridge, A. Spinelli, W. Ryder, and C. Hindorf, “The effect of β+ energy on performance of a small animal PET camera,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* vol. 568, no. 2, pp. 933-936, 2006/12/01/, 2006.

[78] M. Casey, and R. Nutt, “A multicrystal two dimensional BGO detector system for positron emission tomography,” *Ieee transactions on nuclear science,* vol. 33, no. 1, pp. 460-463, 1986.

[79] D. F. C. Hsu, E. Ilan, W. T. Peterson, J. Uribe, M. Lubberink, and C. S. Levin, “Studies of a Next-Generation Silicon-Photomultiplier–Based Time-of-Flight PET/CT System,” *Journal of Nuclear Medicine,* vol. 58, no. 9, pp. 1511-1518, September 1, 2017, 2017.

[80] J. J. Vaquero, and P. Kinahan, “Positron Emission Tomography: Current Challenges and Opportunities for Technological Advances in Clinical and Preclinical Imaging Systems,” *Annu Rev Biomed Eng,* vol. 17, pp. 385-414, 2015.

[81] I. Holl, E. Lorenz, S. Natkaniez, D. Renker, C. Schmelz, and B. Schwartz, “Some studies of avalanche photodiode readout of fast scintillators,” *IEEE transactions on nuclear science,* vol. 42, no. 4, pp. 351-356, 1995.

[82] C. Schmelz, S. Bradbury, I. Holl, E. Lorenz, D. Renker, and S. Ziegler, “Feasibility study of an avalanche photodiode readout for a high resolution PET with nsec time resolution,” *IEEE transactions on nuclear science,* vol. 42, no. 4, pp. 1080-1084, 1995.

[83] V. C. Spanoudaki, and C. S. Levin, “Photo-detectors for time of flight positron emission tomography (ToF-PET),” *Sensors,* vol. 10, no. 11, pp. 10484-10505, 2010.

[84] J. J. van Sluis, J. de Jong, J. Schaar, W. Noordzij, P. van Snick, R. Dierckx, R. Borra, A. Willemsen, and R. Boellaard, “Performance characteristics of the digital Biograph Vision PET/CT system,” *Journal of Nuclear Medicine*, pp. jnumed. 118.215418, 2019.

[85] J. G. Mannheim, A. M. Schmid, J. Schwenck, P. Katiyar, K. Herfert, B. J. Pichler, and J. A. Disselhorst, "PET/mri hybrid systems."

[86] G. Delso, and J. Nuyts, "PET/MRI: Attenuation Correction," *PET/MRI in Oncology*, pp. 53-75: Springer, 2018.

[87] M. Hofmann, B. Pichler, B. Schölkopf, and T. Beyer, “Towards quantitative PET/MRI: a review of MR-based attenuation correction techniques,” *European journal of nuclear medicine and molecular imaging,* vol. 36, no. 1, pp. 93-104, 2009.

[88] A. Martinez-Moller, M. Souvatzoglou, G. Delso, R. A. Bundschuh, C. Chefd'hotel, S. I. Ziegler, N. Navab, M. Schwaiger, and S. G. Nekolla, “Tissue classification as a potential approach for

attenuation correction in whole-body PET/MRI: evaluation with PET/CT data," *Journal of nuclear medicine,* vol. 50, no. 4, pp. 520, 2009.

[89] F. Molina-Duran, D. Dinter, F. Schoenahl, S. O. Schoenberg, and G. Glatting, "Dependence of image quality on acquisition time for the PET/CT Biograph mCT," *Zeitschrift für Medizinische Physik,* vol. 24, no. 1, pp. 73-79, 2014.

[90] V. Y. Panin, A. M. Smith, J. Hu, F. Kehren, and M. E. Casey, "Continuous bed motion on clinical scanner: design, data correction, and reconstruction," *Physics in Medicine & Biology,* vol. 59, no. 20, pp. 6153, 2014.

[91] S. M. Larson, Y. Erdi, T. Akhurst, M. Mazumdar, H. A. Macapinlac, R. D. Finn, C. Casilla, M. Fazzari, N. Srivastava, and H. W. Yeung, "Tumor treatment response based on visual and quantitative changes in global tumor glycolysis using PET-FDG imaging: the visual response score and the change in total lesion glycolysis," *Clinical Positron Imaging,* vol. 2, no. 3, pp. 159-171, 1999.

[92] *Fluorodeoxyglucose (FDG) Positron Emission Tomography (PET) for Solid Tumors* C. f. M. M. Services*,* 2015.

[93] G. K. v. Schulthess, H. C. Steinert, and T. F. Hany, "Integrated PET/CT: Current Applications and Future Directions," *Radiology,* vol. 238, no. 2, pp. 405-422, 2006.

[94] E. M. Rohren, T. G. Turkington, and R. E. Coleman, "Clinical applications of PET in oncology," *Radiology,* vol. 231, no. 2, pp. 305-332, 2004.

[95] P. Conti, and L. Strauss, "The applications of PET in clinical oncology," *J Nucl Med,* vol. 32, no. 4, pp. 623-648, 1991.

[96] H. Jadvar, "PSMA PET in prostate cancer," *Journal of Nuclear Medicine,* vol. 56, no. 8, pp. 1131-1132, 2015.

[97] T. Maurer, M. Eiber, M. Schwaiger, and J. E. Gschwend, "Current use of PSMA–PET in prostate cancer management," *Nature Reviews Urology,* vol. 13, no. 4, pp. 226, 2016.

[98] H. Ahmadzadehfar, E. Eppard, S. Kürpig, R. Fimmers, A. Yordanova, C. D. Schlenkhoff, F. Gärtner, S. Rogenhofer, and M. Essler, "Therapeutic response and side effects of repeated radioligand therapy with 177Lu-PSMA-DKFZ-617 of castrate-resistant metastatic prostate cancer," *Oncotarget,* vol. 7, no. 11, pp. 12477, 2016.

[99] M. Weineisen, M. Schottelius, J. Simecek, R. P. Baum, A. Yildiz, S. Beykan, H. R. Kulkarni, M. Lassmann, I. Klette, and M. Eiber, "68Ga-and 177Lu-labeled PSMA I&T: optimization of a PSMA-targeted theranostic concept and first proof-of-concept human studies," *Journal of Nuclear Medicine,* vol. 56, no. 8, pp. 1169-1176, 2015.

[100] M. Schwaiger, S. Ziegler, and S. G. Nekolla, "PET/CT: challenge for nuclear cardiology," *Journal of Nuclear Medicine,* vol. 46, no. 10, pp. 1664, 2005.

[101] K. L. Gould, "PET perfusion imaging and nuclear cardiology," *Journal of Nuclear Medicine,* vol. 32, no. 4, pp. 579-606, 1991.

[102] V. Dilsizian, S. L. Bacharach, R. S. Beanlands, S. R. Bergmann, D. Delbeke, and R. J. Gropler, "ASNC imaging guidelines for nuclear cardiology procedures: PET myocardial perfusion and metabolism clinical imaging," *J Nucl Cardiol,* vol. 16, no. 4, pp. 651, 2009.

[103] P. Johannsen, J. Jakobsen, P. Bruhn, and A. Gjedde, "Cortical responses to sustained and divided attention in Alzheimer's disease," *Neuroimage,* vol. 10, no. 3, pp. 269-281, 1999.

[104] J. Falconer, J. A. Wada, W. Martin, and D. Li, "PET, CT, and MRI imaging of neuronal migration anomalies in epileptic patients," *Can J Neurol Sci,* vol. 17, no. 1, pp. 35-9, Feb, 1990.

[105] A. B. Young, J. B. Penney, S. Starosta-Rubinstein, D. S. Markel, S. Berent, B. Giordani, R. Ehrenkaufer, D. Jewett, and R. Hichwa, "PET scan investigations of Huntington's disease: cerebral metabolic correlates of neurological features and functional decline," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society,* vol. 20, no. 3, pp. 296-303, 1986.

[106] G. Pagano, F. Niccolini, and M. Politis, “Current status of PET imaging in Huntington’s disease,” *European journal of nuclear medicine and molecular imaging,* vol. 43, no. 6, pp. 1171-1182, 2016.

[107] P. Morrish, J. Rakshi, D. Bailey, G. Sawle, and D. Brooks, “Measuring the rate of progression and estimating the preclinical period of Parkinson’s disease with [18F] dopa PET,” *Journal of Neurology, Neurosurgery & Psychiatry,* vol. 64, no. 3, pp. 314-319, 1998.

[108] A. L. Whone, R. L. Watts, A. J. Stoessl, M. Davis, S. Reske, C. Nahmias, A. E. Lang, O. Rascol, M. J. Ribeiro, and P. Remy, “Slower progression of Parkinson's disease with ropinirole versus levodopa: the REAL-PET study,” *Annals of neurology,* vol. 54, no. 1, pp. 93-101, 2003.

[109] A. Nordberg, “PET imaging of amyloid in Alzheimer's disease,” *The lancet neurology,* vol. 3, no. 9, pp. 519-527, 2004.

[110] L. Mosconi, W. H. Tsui, K. Herholz, A. Pupi, A. Drzezga, G. Lucignani, E. M. Reiman, V. Holthoff, E. Kalbe, and S. Sorbi, “Multicenter standardized 18F-FDG PET diagnosis of mild cognitive impairment, Alzheimer’s disease, and other dementias,” *Journal of nuclear medicine: official publication, Society of Nuclear Medicine,* vol. 49, no. 3, pp. 390, 2008.

[111] G. Chetelat, B. Desgranges, V. De La Sayette, F. Viader, F. Eustache, and J.-C. Baron, “Mild cognitive impairment Can FDG-PET predict who is to rapidly convert to Alzheimer’s disease?,” *Neurology,* vol. 60, no. 8, pp. 1374-1377, 2003.

[112] R. Ossenkoppele, D. R. Schonhaut, M. Schöll, S. N. Lockhart, N. Ayakta, S. L. Baker, J. P. O’Neil, M. Janabi, A. Lazaris, and A. Cantwell, “Tau PET patterns mirror clinical and neuroanatomical variability in Alzheimer’s disease,” *Brain,* vol. 139, no. 5, pp. 1551-1567, 2016.

[113] B. Zeydan, V. J. Lowe, C. G. Schwarz, S. A. Przybelski, N. Tosakulwong, S. M. Zuk, M. L. Senjem, J. L. Gunter, R. O. Roberts, and M. M. Mielke, “Pittsburgh compound-B PET white matter imaging and cognitive function in late multiple sclerosis,” *Multiple Sclerosis Journal,* vol. 24, no. 6, pp. 739-749, 2018.

[114] J. A. Matías-Guiu, C. Oreja-Guevara, M. N. Cabrera-Martín, T. Moreno-Ramos, J. L. Carreras, and J. Matías-Guiu, “Amyloid proteins and their role in multiple sclerosis. considerations in the use of amyloid-PET imaging,” *Frontiers in neurology,* vol. 7, pp. 53, 2016.

[115] W.-D. Heiss, K. Herholz, G. Pawlik, R. Wagner, and K. Wienhard, "Positron emission tomography in neuropsychology," *Methods in Neuropsychology*, pp. 141-149: Elsevier, 1986.

[116] R. Cabeza, and L. Nyberg, “Imaging cognition: An empirical review of PET studies with normal subjects,” *Journal of cognitive neuroscience,* vol. 9, no. 1, pp. 1-26, 1997.

[117] R. L. Wahl, and J. W. Buchanan, *Principles and practice of positron emission tomography*: Lippincott Williams & Wilkins Philadelphia, PA:, 2002.

[118] P. D. Shreve, Y. Anzai, and R. L. Wahl, “Pitfalls in oncologic diagnosis with FDG PET imaging: physiologic and benign variants,” *Radiographics,* vol. 19, no. 1, pp. 61-77, 1999.

[119] K. R. Zasadny, and R. L. Wahl, “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,* vol. 189, no. 3, pp. 847-850, 1993.

[120] S. S. Kety, and C. F. Schmidt, “THE NITROUS OXIDE METHOD FOR THE QUANTITATIVE DETERMINATION OF CEREBRAL BLOOD FLOW IN MAN: THEORY, PROCEDURE AND NORMAL VALUES,” *The Journal of Clinical Investigation,* vol. 27, no. 4, pp. 476-483, 07/01/, 1948.

[121] V. J. Cunningham, S. P. Hume, G. R. Price, R. G. Ahier, J. E. Cremer, and A. K. Jones, “Compartmental analysis of diprenorphine binding to opiate receptors in the rat in vivo and its comparison with equilibrium data in vitro,” *Journal of Cerebral Blood Flow & Metabolism,* vol. 11, no. 1, pp. 1-9, 1991.

[122] A. Lammertsma, C. Bench, S. Hume, S. Osman, K. Gunn, D. Brooks, and R. Frackowiak, “Comparison of methods for analysis of clinical [11C] raclopride studies,” *Journal of Cerebral Blood Flow & Metabolism,* vol. 16, no. 1, pp. 42-52, 1996.

[123] C. S. Patlak, and R. G. Blasberg, "Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Generalizations," *J Cereb Blood Flow Metab,* vol. 5, no. 4, pp. 584-90, Dec, 1985.

[124] A. Gjedde, "High- and Low-Affinity Transport of D-Glucose from Blood to Brain," *Journal of Neurochemistry,* vol. 36, no. 4, pp. 1463-1471, 1981.

[125] C. S. Patlak, R. G. Blasberg, and J. D. Fenstermacher, "Graphical Evaluation of Blood-to-Brain Transfer Constants from Multiple-Time Uptake Data," *Journal of Cerebral Blood Flow & Metabolism,* vol. 3, no. 1, pp. 1-7, 1983.

[126] J. Logan, "A review of graphical methods for tracer studies and strategies to reduce bias," *Nucl Med Biol,* vol. 30, no. 8, pp. 833-44, Nov, 2003.

[127] S. V. Nesterov, E. Deshayes, R. Sciagrà, L. Settimo, J. M. Declerck, X.-B. Pan, K. Yoshinaga, C. Katoh, P. J. Slomka, and G. Germano, "Quantification of myocardial blood flow in absolute terms using 82Rb PET imaging: the RUBY-10 study," *JACC: Cardiovascular Imaging,* vol. 7, no. 11, pp. 1119-1127, 2014.

[128] G. Tomasi, F. Turkheimer, and E. Aboagye, "Importance of quantification for the analysis of PET data in oncology: review of current methods and trends for the future," *Molecular Imaging and Biology,* vol. 14, no. 2, pp. 131-146, 2012.

[129] L. M. Hamberg, G. J. Hunter, N. M. Alpert, N. C. Choi, J. W. Babich, and A. J. Fischman, "The dose uptake ratio as an index of glucose metabolism: useful parameter or oversimplification?," *Journal of Nuclear Medicine,* vol. 35, no. 8, pp. 1308, 1994.

[130] S.-C. Huang, "Anatomy of SUV," *Nuclear medicine and biology,* vol. 27, no. 7, pp. 643-646, 2000.

[131] M. A. Lodge, J. D. Lucas, P. K. Marsden, B. F. Cronin, M. J. O'Doherty, and M. A. Smith, "A PET study of 18FDG uptake in soft tissue masses," *European journal of nuclear medicine,* vol. 26, no. 1, pp. 22-30, 1999.

[132] Y. Sugawara, K. R. Zasadny, H. B. Grossman, I. R. Francis, M. F. Clarke, and R. L. Wahl, “Germ cell tumor: differentiation of viable tumor, mature teratoma, and necrotic tissue with FDG PET and kinetic modeling,” *Radiology,* vol. 211, no. 1, pp. 249-256, 1999.

[133] L. K. Dunnwald, R. K. Doot, J. M. Specht, J. R. Gralow, G. K. Ellis, R. B. Livingston, H. M. Linden, V. Gadi, B. F. Kurland, and E. K. Schubert, “PET Tumor Metabolism in Locally Advanced Breast Cancer Patients Undergoing Neoadjuvant Chemotherapy: Value of Static vs Kinetic Measures of Fluorodeoxyglucose Uptake,” *Clinical Cancer Research*, pp. clincanres. 2649.2010, 2011.

[134] N. A. Karakatsanis, M. A. Lodge, A. K. Tahari, Y. Zhou, R. L. Wahl, and A. Rahmim, “Dynamic whole-body PET parametric imaging: I. Concept, acquisition protocol optimization and clinical application,” *Physics in Medicine & Biology,* vol. 58, no. 20, pp. 7391, 2013.

[135] N. A. Karakatsanis, M. A. Lodge, Y. Zhou, R. L. Wahl, and A. Rahmim, “Dynamic whole-body PET parametric imaging: II. Task-oriented statistical estimation,” *Physics in Medicine & Biology,* vol. 58, no. 20, pp. 7419, 2013.

[136] N. A. Karakatsanis, M. A. Lodge, Y. Zhou, J. Mhlanga, M. A. Chaudhry, A. K. Tahari, W. P. Segars, R. L. Wahl, and A. Rahmim, "Dynamic multi-bed FDG PET imaging: feasibility and optimization." pp. 3863-3870.

[137] D. R. Osborne, and S. Acuff, “Whole-body dynamic imaging with continuous bed motion PET/CT,” *Nuclear medicine communications,* vol. 37, no. 4, pp. 428, 2016.

[138] A. Rahmim, M. A. Lodge, N. A. Karakatsanis, V. Y. Panin, Y. Zhou, A. McMillan, S. Cho, H. Zaidi, M. E. Casey, and R. L. Wahl, “Dynamic whole-body PET imaging: principles, potentials and applications,” *European Journal of Nuclear Medicine and Molecular Imaging,* vol. 46, no. 2, pp. 501-518, February 01, 2019.

[139] V. Y. Panin, F. Kehren, C. Michel, and M. Casey, “Fully 3-D PET reconstruction with system matrix derived from point source measurements.,” *IEEE Trans Med Imaging,* vol. 25, no. 7, pp. 907-921, 2006.

[140] C. D. Ramos, Y. E. Erdi, M. Gonen, E. Riedel, H. W. Yeung, H. A. Macapinlac, R. Chisin, and S. M. Larson, “FDG-PET standardized uptake values in normal anatomical structures using iterative reconstruction segmented attenuation correction and filtered back-projection,” *European journal of nuclear medicine,* vol. 28, no. 2, pp. 155-164, 2001.

[141] P. E. Vaissier, M. C. Goorden, A. B. Taylor, and F. J. Beekman, “Fast Count-Regulated OSEM Reconstruction With Adaptive Resolution Recovery,” *IEEE Trans. Med. Imaging,* vol. 32, no. 12, pp. 2250-2261, 2013.

[142] Y. Jian, *Spatial resolution improvement in positron emission tomography: Physics, statistical models and iterative image reconstruction*: Yale University, 2014.

[143] H. Mohy-ud-Din, “Motion Correction and Pharmacokinetic Analysis in Dynamic Positron Emission Tomography,” 2015.

[144] K. Lange, and R. Carson, “EM reconstruction algorithms for emission and transmission tomography,” *J Comput Assist Tomogr,* vol. 8, no. 2, pp. 306-16, 1984.

[145] L. A. Shepp, and Y. Vardi, “Maximum likelihood reconstruction for emission tomography,” *IEEE transactions on medical imaging,* vol. 1, no. 2, pp. 113-122, 1982.

[146] H. M. Hudson, and R. S. Larkin, “Accelerated image reconstruction using ordered subsets of projection data,” *IEEE transactions on medical imaging,* vol. 13, no. 4, pp. 601-609, 1994.

[147] T. Hsiao, A. Rangarajan, and G. R. Gindi, "Provably convergent OSEM-like reconstruction algorithm for emission tomography." pp. 10-20.

[148] T. Hsiao, A. Rangarajan, and G. Gindi, "A new convergent MAP reconstruction algorithm for emission tomography using ordered subsets and separable surrogates." pp. 409-412.

[149] L. Kaufman, “Maximum likelihood, least squares, and penalized least squares for PET,” *IEEE Transactions on Medical Imaging,* vol. 12, no. 2, pp. 200-214, 1993.

[150] J. Nuyts, and J. A. Fessler, “A penalized-likelihood image reconstruction method for emission tomography, compared to postsmoothed maximum-likelihood with matched spatial resolution,” *IEEE transactions on medical imaging,* vol. 22, no. 9, pp. 1042-1052, 2003.

[151] J. A. Fessler, “Penalized weighted least-squares image reconstruction for positron emission tomography,” *IEEE transactions on medical imaging,* vol. 13, no. 2, pp. 290-300, 1994.

[152] J. A. Fessler, “Mean and variance of implicitly defined biased estimators(such as penalized maximum likelihood): applications to tomography,” *IEEE Transactions on Image Processing,* vol. 5, no. 3, pp. 493-506, 1996.

[153] G. Wang, and J. Qi, “Analysis of penalized likelihood image reconstruction for dynamic PET quantification,” *IEEE transactions on medical imaging,* vol. 28, no. 4, pp. 608-620, 2009.

[154] E. J. Teoh, D. R. McGowan, R. E. Macpherson, K. M. Bradley, and F. V. Gleeson, “Phantom and clinical evaluation of the Bayesian penalized likelihood reconstruction algorithm Q. Clear on an LYSO PET/CT system,” *Journal of nuclear medicine: official publication, Society of Nuclear Medicine,* vol. 56, no. 9, pp. 1447, 2015.

[155] E. U. Mumcuoglu, R. Leahy, S. R. Cherry, and Z. Zhou, “Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images,” *IEEE transactions on Medical Imaging,* vol. 13, no. 4, pp. 687-701, 1994.

[156] J. Qi, and R. M. Leahy, “Resolution and noise properties of MAP reconstruction for fully 3-D PET,” *IEEE transactions on medical imaging,* vol. 19, no. 5, pp. 493-506, 2000.

[157] A. R. De Pierro, and M. E. Yamagishi, “Fast EM-like methods for maximum" a posteriori" estimates in emission tomography,” *IEEE Transactions on Medical Imaging,* vol. 20, no. 4, pp. 280-288, 2001.

[158] J. Nuyts, D. Beque, P. Dupont, and L. Mortelmans, "A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography," *IEEE Transactions on nuclear science,* vol. 49, no. 1, pp. 56-60, 2002.

[159] E. U. Mumcuoglu, R. M. Leahy, S. R. Cherry, and E. Hoffman, "Accurate geometric and physical response modelling for statistical image reconstruction in high resolution PET," *Conference Record IEEE Nuclear Science Symposium,* vol. 3, pp. 1569-1573, 1996.

[160] P. E. Kinahan, D. W. Townsend, T. Beyer, and D. Sashin, "Attenuation correction for a combined 3D PET/CT scanner," *Medical Physics,* vol. 25, no. 10, pp. 2046-2053, 1998.

[161] P. E. Kinahan, B. H. Hasegawa, and T. Beyer, "X-ray-based attenuation correction for positron emission tomography/computed tomography scanners," *Seminars in Nuclear Medicine,* vol. 33, no. 3, pp. 166-179, 2003/07/01/, 2003.

[162] C. Thompson, "The effect of collimation on scatter fraction in multi-slice PET," *IEEE Transactions on Nuclear Science,* vol. 35, no. 1, pp. 598-602, 1988.

[163] T. Spinks, T. Jones, D. Bailey, D. Townsend, S. Grootoonk, P. Bloomfield, M.-C. Gilardi, M. Casey, B. Sipe, and J. Reed, "Physical performance of a positron tomograph for brain imaging with retractable septa," *Physics in Medicine & Biology,* vol. 37, no. 8, pp. 1637, 1992.

[164] D. L. Bailey, and K. P. Willowson, "An evidence-based review of quantitative SPECT imaging and potential clinical applications," *J Nucl Med,* vol. 54, no. 1, pp. 83-9, 2013.

[165] S. Grootoonk, T. Spinks, D. Sashin, N. Spyrou, and T. Jones, "Correction for scatter in 3D brain PET using a dual energy window method," *Physics in Medicine & Biology,* vol. 41, no. 12, pp. 2757, 1996.

[166] S. R. Cherry, S. R. Meikle, and E. J. Hoffman, "Correction and characterization of scattered events in three-dimensional PET using scanners with retractable septa," *Journal of nuclear medicine: official publication, Society of Nuclear Medicine,* vol. 34, no. 4, pp. 671-678, 1993.

[167] J. M. Ollinger, “Model-based scatter correction for fully 3D PET,” *Physics in Medicine & Biology,* vol. 41, no. 1, pp. 153, 1996.

[168] C. C. Watson, “New, faster, image-based scatter correction for 3D PET,” *IEEE Transactions on Nuclear Science,* vol. 47, no. 4, pp. 1587-1594, 2000.

[169] D. Brasse, P. E. Kinahan, C. Lartizien, C. Comtat, M. Casey, and C. Michel, “Correction methods for random coincidences in fully 3D whole-body PET: Impact on data and image quality.,” *J Nucl Med,* vol. 46, no. 5, pp. 859-867, May, 2005.

[170] R. Manavaki, A. J. Reader, C. Keller, J. Missimer, and R. J. Walledge, “Scatter modeling for 3-d PET list-mode EM reconstruction,” in Nuclear Science Symposium Conference Record, 2002 IEEE, 2002, pp. 1863-1868.

[171] J. A. Kimdon, J. Qi, and W. W. Moses, "Effect of random and scatter fractions in variance reduction using time-of-flight information." pp. 2571-2573 Vol.4.

[172] A. Rahmim, J. Tang, M. A. Lodge, S. Lashkari, M. R. Ay, R. Lautamaki, B. M. Tsui, and F. M. Bengel, “Analytic system matrix resolution modeling in PET: an application to Rb-82 cardiac imaging,” *Phys Med Biol,* vol. 53, no. 21, pp. 5947-65, Nov 7, 2008.

[173] M. R. Palmer, Z. Xuping, and J. A. Parker, “Modeling and simulation of positron range effects for high resolution PET imaging,” *Nuclear Science, IEEE Transactions on,* vol. 52, no. 5, pp. 1391-1395, 2005.

[174] B. Bai, A. Ruangma, R. Laforest, Y. C. Tai, and R. M. Leahy, “Positron range modeling for statistical PET image reconstruction,” in Nuclear Science Symposium Conference Record, 2003 IEEE, 2003, pp. 2501-2505.

[175] S. E. Derenzo, “Mathematical Removal of Positron Range Blurring in High Resolution Tomography,” *Nuclear Science, IEEE Transactions on,* vol. 33, no. 1, pp. 565-569, 1986.

[176] A. Del Guerra, *Ionizing radiation detectors for medical imaging*: World Scientific, 2004.

[177] A. Rahmim, J. Qi, and V. Sossi, “Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls,” *Med Phys,* vol. 40, no. 6, pp. 064301, Jun, 2013.

[178] T.-S. Pan, D.-S. Luo, V. Kohli, and M. A. King, “Influence of OSEM, elliptical orbits and background activity on SPECT 3D resolution recovery,” *Physics in Medicine & Biology,* vol. 42, no. 12, pp. 2517, 1997.

[179] K. Wienhard, M. Schmand, M. E. Casey, K. Baker, J. Bao, L. Eriksson, W. F. Jones, C. Knoess, M. Lenox, M. Lercher, P. Luk, C. Michel, J. H. Reed, N. Richerzhagen, J. Treffert, S. Vollmar, J. W. Young, W. D. Heiss, and R. Nutt, “The ECAT HRRT: performance and first clinical application of the new high resolution research tomograph.,” *IEEE Trans Nucl Sci,* vol. 49, no. 1, pp. 104 -110, 2002.

[180] J. M. Anton-Rodriguez, F. A. Kotasidis, G. I. Angelis, P. J. Noonan, C. Kobylecki, P. J. Julyan, and J. C. Matthews, "Comparison of depth of interaction encoding and resolution modelling image reconstruction in High Resolution PET imaging." pp. 1-6.

[181] J. M. Ollinger, “Detector efficiency and Compton scatter in fully 3D PET,” *IEEE transactions on nuclear science,* vol. 42, no. 4, pp. 1168-1173, 1995.

[182] G. Germano, and E. Hoffman, “A study of data loss and mispositioning due to pileup in 2-D detectors in PET,” *IEEE Transactions on Nuclear Science,* vol. 37, no. 2, pp. 671-675, 1990.

[183] R. D. Badawi, and P. Marsden, “Developments in component-based normalization for 3D PET,” *Physics in Medicine & Biology,* vol. 44, no. 2, pp. 571, 1999.

[184] R. D. Badawi, and P. Marsden, “Self-normalization of emission data in 3D PET,” *IEEE Transactions on nuclear science,* vol. 46, no. 3, pp. 709-712, 1999.

[185] J. A. Fessler, "Statistical image reconstruction methods for transmission tomography.," *Handbook of Medical Imaging*, M. S. a. J. M. Fitzpatrick, ed., pp. 1-70, Bellingham: SPIE, 2000.

[186] S. A. Nehmeh, and Y. E. Erdi, "Respiratory motion in positron emission tomography/computed tomography: a review." pp. 167-176.

[187] A. Rahmim, “Advanced motion correction methods in pet,” *Iranian Journal of Nuclear Medicine,* vol. 13, no. 2, pp. 1-17, 2005.

[188] S. Fürst, R. Grimm, I. Hong, M. Souvatzoglou, M. E. Casey, M. Schwaiger, S. G. Nekolla, and S. I. Ziegler, “Motion correction strategies for integrated PET/MR,” *J Nucl Med,* vol. 56, no. 2, pp. 261-9, 2015.

[189] R. Manber, K. Thielemans, B. F. Hutton, A. Barnes, S. Ourselin, S. Arridge, C. O’Meara, S. Wan, and D. Atkinson, “Practical PET respiratory motion correction in clinical PET/MR,” *J Nucl Med,* vol. 56, no. 6, pp. 890-6, 2015.

[190] C. Liu, L. A. Pierce II, A. M. Alessio, and P. E. Kinahan, “The impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging,” *Physics in Medicine & Biology,* vol. 54, no. 24, pp. 7345, 2009.

[191] H. Fayad, F. Lamare, T. Merlin, and D. Visvikis, “Motion correction using anatomical information in PET/CT and PET/MR hybrid imaging,” *The quarterly journal of nuclear medicine and molecular imaging: official publication of the Italian Association of Nuclear Medicine (AIMN)[and] the International Association of Radiopharmacology (IAR),[and] Section of the Society of..* vol. 60, no. 1, pp. 12-24, 2016.

[192] J. R. McClelland, D. J. Hawkes, T. Schaeffter, and A. P. King, “Respiratory motion models: a review,” *Medical image analysis,* vol. 17, no. 1, pp. 19-42, 2013.

[193] A. Rahmim, O. Rousset, and H. Zaidi, “Strategies for motion tracking and correction in PET,” *PET clinics,* vol. 2, no. 2, pp. 251-266, 2007.

[194] A. J. Buckler, L. Bresolin, N. R. Dunnick, D. C. Sullivan, and F. t. Group, “A Collaborative Enterprise for Multi-Stakeholder Participation in the Advancement of Quantitative Imaging,” *Radiology,* vol. 258, no. 3, pp. 906-914, 2011.

[195] A. J. Buckler, D. Paik, M. Ouellette, J. Danagoulian, G. Wernsing, and B. E. Suzek, "A novel knowledge representation framework for the statistical validation of quantitative imaging biomarkers," *J Digit Imaging,* vol. 26, no. 4, pp. 614-29, Aug, 2013.

[196] F. A. Jaffer, and R. Weissleder, "Molecular imaging in the clinical arena," *Jama,* vol. 293, no. 7, pp. 855-62, Feb 16, 2005.

[197] R. G. Abramson, K. R. Burton, J.-P. J. Yu, E. M. Scalzetti, T. E. Yankeelov, A. B. Rosenkrantz, M. Mendiratta-Lala, B. J. Bartholmai, D. Ganeshan, L. Lenchik, and R. M. Subramaniam, "Methods and challenges in quantitative imaging biomarker development," *Academic radiology,* vol. 22, no. 1, pp. 25-32, 2015.

[198] A. J. Buckler, L. Bresolin, N. R. Dunnick, D. C. Sullivan, H. J. Aerts, B. Bendriem, C. Bendtsen, R. Boellaard, J. M. Boone, P. E. Cole, J. J. Conklin, G. S. Dorfman, P. S. Douglas, W. Eidsaunet, C. Elsinger, R. A. Frank, C. Gatsonis, M. L. Giger, S. N. Gupta, D. Gustafson, O. S. Hoekstra, E. F. Jackson, L. Karam, G. J. Kelloff, P. E. Kinahan, G. McLennan, C. G. Miller, P. D. Mozley, K. E. Muller, R. Patt, D. Raunig, M. Rosen, H. Rupani, L. H. Schwartz, B. A. Siegel, A. G. Sorensen, R. L. Wahl, J. C. Waterton, W. Wolf, G. Zahlmann, and B. Zimmerman, "Quantitative imaging test approval and biomarker qualification: interrelated but distinct activities," *Radiology,* vol. 259, no. 3, pp. 875-84, Jun, 2011.

[199] C. R. Meyer, S. G. Armato, C. P. Fenimore, G. McLennan, L. M. Bidaut, D. P. Barboriak, M. A. Gavrielides, E. F. Jackson, M. F. McNitt-Gray, P. E. Kinahan, N. Petrick, and B. Zhao, "Quantitative imaging to assess tumor response to therapy: common themes of measurement, truth data, and error sources," *Transl Oncol,* vol. 2, no. 4, pp. 198-210, Dec, 2009.

[200] B. D. W. Group, "Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework," *Clinical Pharmacology & Therapeutics,* vol. 69, no. 3, pp. 89-95, 2001.

[201] L. G. Kessler, H. X. Barnhart, A. J. Buckler, K. R. Choudhury, M. V. Kondratovich, A. Toledano, A. R. Guimaraes, R. Filice, Z. Zhang, and D. C. Sullivan, "The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions," *Stat Methods Med Res,* vol. 24, no. 1, pp. 9-26, Feb, 2015.

[202] X. Hua, S. Lee, I. Yanovsky, A. D. Leow, Y.-Y. Chou, A. J. Ho, B. Gutman, A. W. Toga, C. R. Jack Jr, and M. A. Bernstein, "Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry: an ADNI study of 515 subjects," *Neuroimage,* vol. 48, no. 4, pp. 668-681, 2009.

[203] R. J. van Klaveren, M. Oudkerk, M. Prokop, E. T. Scholten, K. Nackaerts, R. Vernhout, C. A. van Iersel, K. A. van den Bergh, S. van't Westeinde, and C. van der Aalst, "Management of lung nodules detected by volume CT scanning," *New England Journal of Medicine,* vol. 361, no. 23, pp. 2221-2229, 2009.

[204] R. L. Wahl, H. Jacene, Y. Kasamon, and M. A. Lodge, "From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors," *Journal of nuclear medicine: official publication, Society of Nuclear Medicine,* vol. 50, no. Suppl 1, pp. 122S, 2009.

[205] N. Bhooshan, M. L. Giger, S. A. Jansen, H. Li, L. Lan, and G. M. Newstead, "Cancerous breast lesions on dynamic contrast-enhanced MR images: computerized characterization for image-based prognostic markers," *Radiology,* vol. 254, no. 3, pp. 680-690, 2010.

[206] G. E. Fuleihan, M. A. Testa, J. E. Angell, N. Porrino, and M. S. Leboff, "Reproducibility of DXA absorptiometry: a model for bone loss estimates," *J Bone Miner Res,* vol. 10, no. 7, pp. 1004-14, Jul, 1995.

[207] D. C. Sullivan, "Imaging as a quantitative science," *Radiology,* vol. 248, no. 2, pp. 328-32, Aug, 2008.

[208] J. M. Boone, “Radiological interpretation 2020: toward quantitative image assessment,” *Med Phys,* vol. 34, no. 11, pp. 4173-9, Nov, 2007.

[209] M. K. Choong, and G. Tsafnat, “The implications of biomarker evidence for systematic reviews,” *BMC Med Res Methodol,* vol. 12, pp. 176, Nov 22, 2012.

[210] E. A. Eisenhauer, P. Therasse, J. Bogaerts, L. H. Schwartz, D. Sargent, R. Ford, J. Dancey, S. Arbuck, S. Gwyther, M. Mooney, L. Rubinstein, L. Shankar, L. Dodd, R. Kaplan, D. Lacombe, and J. Verweij, “New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1),” *Eur J Cancer,* vol. 45, no. 2, pp. 228-47, Jan, 2009.

[211] T. Vidaurre, J. Wilkerson, R. Simon, S. E. Bates, and T. Fojo, “Stable disease is not preferentially observed with targeted therapies and as currently defined has limited value in drug development,” *Cancer J,* vol. 15, no. 5, pp. 366-73, Sep-Oct, 2009.

[212] M. Buyse, P. Thirion, R. W. Carlson, T. Burzykowski, G. Molenberghs, and P. Piedbois, “Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. Meta-Analysis Group in Cancer,” *Lancet,* vol. 356, no. 9227, pp. 373-8, Jul 29, 2000.

[213] D. L. Raunig, L. M. McShane, G. Pennello, C. Gatsonis, P. L. Carson, J. T. Voyvodic, R. L. Wahl, B. F. Kurland, A. J. Schwarz, M. Gönen, G. Zahlmann, M. V. Kondratovich, K. O'Donnell, N. Petrick, P. E. Cole, B. Garra, D. C. Sullivan, and Q. T. P. W. Group, “Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment,” *Statistical methods in medical research,* vol. 24, no. 1, pp. 27-67, 2015.

[214] G. J. Kelloff, J. M. Hoffman, B. Johnson, H. I. Scher, B. A. Siegel, E. Y. Cheng, B. D. Cheson, J. O'Shaughnessy, K. Z. Guyton, D. A. Mankoff, L. Shankar, S. M. Larson, C. C. Sigman, R. L. Schilsky, and D. C. Sullivan, “Progress and promise of FDG-PET imaging for cancer patient management and oncologic drug development,” *Clin Cancer Res,* vol. 11, no. 8, pp. 2785-808, Apr 15, 2005.

[215] P. E. Kinahan, and J. W. Fletcher, "Positron emission tomography-computed tomography standardized uptake values in clinical practice and assessing response to therapy," *Seminars in ultrasound, CT, and MR,* vol. 31, no. 6, pp. 496-505, 2010.

[216] R. Boellaard, "Standards for PET image acquisition and quantitative data analysis," *J Nucl Med,* vol. 50 Suppl 1, pp. 11S-20S, 2009.

[217] C. K. Kim, N. C. Gupta, B. Chandramouli, and A. Alavi, "Standardized uptake values of FDG: body surface area correction is preferable to body weight correction," *J Nucl Med,* vol. 35, no. 1, pp. 164-167, 1994.

[218] Y. Sugawara, K. R. Zasadny, A. W. Neuhoff, and R. L. Wahl, "Reevaluation of the standardized uptake value for FDG: variations with body weight and methods for correction," *Radiology,* vol. 213, no. 2, pp. 521-5, Nov, 1999.

[219] A. K. Tahari, D. Chien, J. R. Azadi, and R. L. Wahl, "Optimum lean body formulation for correction of standardized uptake value in PET imaging," *Journal of nuclear medicine : official publication, Society of Nuclear Medicine,* vol. 55, no. 9, pp. 1481-1484, 2014.

[220] I. A. E. AGENCY, *Standard Operating Procedures for PET/CT: A Practical Approach for Use in Adult Oncology*, Vienna: INTERNATIONAL ATOMIC ENERGY AGENCY, 2013.

[221] F. R. Vogenberg, C. Isaacson Barash, and M. Pursel, "Personalized medicine: part 1: evolution and development into theranostics," *P & T : a peer-reviewed journal for formulary management,* vol. 35, no. 10, pp. 560-576, 2010.

[222] K. Jain, "Personalized medicine," *Current opinion in molecular therapeutics,* vol. 4, no. 6, pp. 548-558, 2002.

[223] G. S. Ginsburg, and J. J. McCarthy, "Personalized medicine: revolutionizing drug discovery and patient care," *Trends in Biotechnology,* vol. 19, no. 12, pp. 491-496, 2001/12/01/, 2001.

[224] X. Zhang, “Precision Medicine,” *Personalized Medicine, Omics and Big Data: Concepts and Relationships. J Pharmacogenomics Pharmacoproteomics,* vol. 6, pp. e144, 2015.

[225] A. Pandey, and M. Mann, “Proteomics to study genes and genomes,” *Nature,* vol. 405, no. 6788, pp. 837, 2000.

[226] O. Fiehn, "Metabolomics—the link between genotypes and phenotypes," *Functional genomics*, pp. 155-171: Springer, 2002.

[227] K. W. Turteltaub, M. Ascher, R. Langlois, I. Fodor, J. Kercher, K. M. Laughlin, D. Nelson, W. Colston, and F. P. Milanovich, *Pathomics: Final Report,* UCRL-TR-226770; TRN: US201207%%47 United States 10.2172/1036867 TRN: US201207%%47 LLNL English, ; Lawrence Livermore National Lab. (LLNL), Livermore, CA (United States), 2006.

[228] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R. G. P. M. van Stiphout, P. Granton, C. M. L. Zegers, R. Gillies, R. Boellard, A. Dekker, and H. J. W. L. Aerts, “Radiomics: Extracting more information from medical images using advanced feature analysis,” *European Journal of Cancer,* vol. 48, no. 4, pp. 441-446, 2012/03/01/, 2012.

[229] J. Saltz, J. Almeida, Y. Gao, A. Sharma, E. Bremer, T. DiPrima, M. Saltz, J. Kalpathy-Cramer, and T. Kurc, “Towards Generation, Management, and Exploration of Combined Radiomics and Pathomics Datasets for Cancer Research,” *AMIA Summits on Translational Science Proceedings,* vol. 2017, pp. 85-94, 07/26, 2017.

[230] A. M. Rutman, and M. D. Kuo, “Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging,” *European journal of radiology,* vol. 70, no. 2, pp. 232-241, 2009.

[231] M. A. Mazurowski, “Radiogenomics: what it is and why it is important,” *Journal of the American College of Radiology,* vol. 12, no. 8, pp. 862-866, 2015.

[232] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, “Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA: A Cancer Journal for Clinicians,* vol. 68, no. 6, pp. 394-424, 2018.

[233] G. H. D. Exchange, "Global Burdon of Disease Study 2017," 2017.

[234] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, “Global cancer statistics, 2012,” *CA: a cancer journal for clinicians,* vol. 65, no. 2, pp. 87-108, 2015.

[235] C. Parmar, “Machine learning applications for Radiomics: towards robust non-invasive predictors in clinical oncology,” Maastricht University, 2017.

[236] R. L. Cohen, and J. Settleman, “From cancer genomics to precision oncology—tissue’s still an issue,” *Cell,* vol. 157, no. 7, pp. 1509-1514, 2014.

[237] M. Gerlinger, A. J. Rowan, S. Horswell, J. Larkin, D. Endesfelder, E. Gronroos, P. Martinez, N. Matthews, A. Stewart, and P. Tarpey, “Intratumor heterogeneity and branched evolution revealed by multiregion sequencing,” *New England journal of medicine,* vol. 366, no. 10, pp. 883-892, 2012.

[238] T. A. Yap, M. Gerlinger, P. A. Futreal, L. Pusztai, and C. Swanton, “Intratumor heterogeneity: seeing the wood for the trees,” *Science translational medicine,* vol. 4, no. 127, pp. 127ps10-127ps10, 2012.

[239] M. Hatt, C. Cheze-le Rest, A. van Baardwijk, P. Lambin, O. Pradier, and D. Visvikis, “Impact of Tumor Size and Tracer Uptake Heterogeneity in 18F-FDG PET and CT Non–Small Cell Lung Cancer Tumor Delineation,” *Journal of Nuclear Medicine,* vol. 52, no. 11, pp. 1690-1697, November 1, 2011, 2011.

[240] R. T. H. M. Larue, G. Defraene, D. D. Ruysscher, P. Lambin, and W. v. Elmpt, “Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures,” *The British Journal of Radiology,* vol. 90, no. 1070, pp. 20160665, 2017.

[241] V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, K. Forster, H. J. W. L. Aerts, A. Dekker, D. Fenstermacher, D. B. Goldgof, L. O. Hall, P. Lambin, Y. Balagurunathan, R. A. Gatenby, and R. J. Gillies, "Radiomics: the process and the challenges," *Magnetic Resonance Imaging,* vol. 30, no. 9, pp. 1234-1248, 2012/11/01/, 2012.

[242] A. T. Haase, K. Henry, M. Zupancic, G. Sedgewick, R. A. Faust, H. Melroe, W. Cavert, K. Gebhard, K. Staskus, and Z.-Q. Zhang, "Quantitative image analysis of HIV-1 infection in lymphoid tissue," *Science,* vol. 274, no. 5289, pp. 985-989, 1996.

[243] P. Lambin, R. G. Van Stiphout, M. H. Starmans, E. Rios-Velazquez, G. Nalbantov, H. J. Aerts, E. Roelofs, W. Van Elmpt, P. C. Boutros, and P. Granone, "Predicting outcomes in radiation oncology—multifactorial decision support systems," *Nature reviews Clinical oncology,* vol. 10, no. 1, pp. 27-40, 2013.

[244] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: Images Are More than Pictures, They Are Data," *Radiology,* vol. 278, no. 2, pp. 563-77, Feb, 2016.

[245] S. Yip, S. F. , and H. Aerts, J. W. L. , "Applications and limitations of radiomics," *Physics in Medicine and Biology,* vol. 61, no. 13, pp. R150, 2016.

[246] H. L. Aerts, "The potential of radiomic-based phenotyping in precision medicine: A review," *JAMA Oncology,* vol. 2, no. 12, pp. 1636-1642, 2016.

[247] M. Hatt, F. Tixier, L. Pierce, P. E. Kinahan, C. C. Le Rest, and D. Visvikis, "Characterization of PET/CT images using texture analysis: the past, the present… any future?," *European Journal of Nuclear Medicine and Molecular Imaging*, pp. 1-15, 2016.

[248] R. Thawani, M. McLane, N. Beig, S. Ghose, P. Prasanna, V. Velcheti, and A. Madabhushi, "Radiomics and radiogenomics in lung cancer: A review for the clinician," *Lung Cancer,* vol. 115, pp. 34-41, 2018/01/01/, 2018.

[249] L. Oakden-Rayner, G. Carneiro, T. Bessen, J. C. Nascimento, A. P. Bradley, and L. J. Palmer, "Precision radiology: predicting longevity using feature engineering and deep learning methods in a radiomics framework," *Scientific reports,* vol. 7, no. 1, pp. 1648, 2017.

[250] M. L. Giger, "Machine learning in medical imaging," *Journal of the American College of Radiology,* vol. 15, no. 3, pp. 512-520, 2018.

[251] J. Lao, Y. Chen, Z.-C. Li, Q. Li, J. Zhang, J. Liu, and G. Zhai, "A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme," *Scientific reports,* vol. 7, no. 1, pp. 10353, 2017.

[252] W. Wu, C. Parmar, P. Grossmann, J. Quackenbush, P. Lambin, J. Bussink, R. Mak, and H. J. Aerts, "Exploratory study to identify radiomics classifiers for lung cancer histology," *Frontiers in oncology,* vol. 6, pp. 71, 2016.

[253] H. J. Yoon, I. Sohn, J. H. Cho, H. Y. Lee, J.-H. Kim, Y.-L. Choi, H. Kim, G. Lee, K. S. Lee, and J. Kim, "Decoding tumor phenotypes for ALK, ROS1, and RET fusions in lung adenocarcinoma using a radiomics approach," *Medicine,* vol. 94, no. 41, 2015.

[254] Y. Liu, J. Kim, Y. Balagurunathan, Q. Li, A. L. Garcia, O. Stringfield, Z. Ye, and R. J. Gillies, "Radiomic features are associated with EGFR mutation status in lung adenocarcinomas," *Clinical lung cancer,* vol. 17, no. 5, pp. 441-448. e6, 2016.

[255] Y. Zhu, H. Li, W. Guo, K. Drukker, L. Lan, M. L. Giger, and Y. Ji, "Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma," *Scientific reports,* vol. 5, pp. 17787, 2015.

[256] T. P. Coroller, V. Agrawal, V. Narayan, Y. Hou, P. Grossmann, S. W. Lee, R. H. Mak, and H. J. Aerts, "Radiomic phenotype features predict pathological response in non-small cell lung cancer," *Radiotherapy and Oncology,* vol. 119, no. 3, pp. 480-486, 2016.

[257] K. Nie, L. Shi, Q. Chen, X. Hu, S. Jabbour, N. Yue, T. Niu, and X. Sun, "Rectal cancer: assessment of neoadjuvant chemo-radiation outcome based on radiomics of multi-parametric MRI," *Clinical cancer research*, pp. clincanres. 2997.2015, 2016.

[258] T. P. Coroller, V. Agrawal, E. Huynh, V. Narayan, S. W. Lee, R. H. Mak, and H. J. Aerts, "Radiomic-based pathological response prediction from primary tumors and lymph nodes in NSCLC," *Journal of Thoracic Oncology,* vol. 12, no. 3, pp. 467-476, 2017.

[259] A. Ahmed, P. Gibbs, M. Pickles, and L. Turnbull, "Texture analysis in assessment and prediction of chemotherapy response in breast cancer," *Journal of Magnetic Resonance Imaging,* vol. 38, no. 1, pp. 89-101, 2013.

[260] J. Liu, Y. Mao, Z. Li, D. Zhang, Z. Zhang, S. Hao, and B. Li, "Use of texture analysis based on contrast-enhanced MRI to predict treatment response to chemoradiotherapy in nasopharyngeal carcinoma," *Journal of Magnetic Resonance Imaging,* vol. 44, no. 2, pp. 445-455, 2016.

[261] H. Li, Y. Zhu, E. S. Burnside, K. Drukker, K. A. Hoadley, C. Fan, S. D. Conzen, G. J. Whitman, E. J. Sutton, and J. M. Net, "MR imaging radiomics signatures for predicting the risk of breast cancer recurrence as given by research versions of MammaPrint, Oncotype DX, and PAM50 gene assays," *Radiology,* vol. 281, no. 2, pp. 382-391, 2016.

[262] K. Gnep, A. Fargeas, R. E. Gutiérrez-Carvajal, F. Commandeur, R. Mathieu, J. D. Ospina, Y. Rolland, T. Rohou, S. Vincendeau, and M. Hatt, "Haralick textural features on T2-weighted MRI are associated with biochemical recurrence following radiotherapy for peripheral zone prostate cancer," *Journal of Magnetic Resonance Imaging,* vol. 45, no. 1, pp. 103-117, 2017.

[263] Y.-q. Huang, C.-h. Liang, L. He, J. Tian, C.-s. Liang, X. Chen, Z.-l. Ma, and Z.-y. Liu, "Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer," *Journal of Clinical Oncology,* vol. 34, no. 18, pp. 2157-2164, 2016.

[264] M. Vallières, C. R. Freeman, S. R. Skamene, and I. El Naqa, "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities," *Physics in Medicine & Biology,* vol. 60, no. 14, pp. 5471, 2015.

[265] T. P. Coroller, P. Grossmann, Y. Hou, E. R. Velazquez, R. T. Leijenaar, G. Hermann, P. Lambin, B. Haibe-Kains, R. H. Mak, and H. J. Aerts, "CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma," *Radiotherapy and Oncology,* vol. 114, no. 3, pp. 345-350, 2015.

[266] E. Huynh, T. P. Coroller, V. Narayan, V. Agrawal, Y. Hou, J. Romano, I. Franco, R. H. Mak, and H. J. Aerts, "CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer," *Radiotherapy and Oncology,* vol. 120, no. 2, pp. 258-266, 2016.

[267] Y. Huang, Z. Liu, L. He, X. Chen, D. Pan, Z. Ma, C. Liang, J. Tian, and C. Liang, "Radiomics signature: A potential biomarker for the prediction of disease-free survival in early-stage (i or ii) non—small cell lung cancer," *Radiology,* vol. 281, no. 3, pp. 947-957, 2016.

[268] C. Parmar, P. Grossmann, D. Rietveld, M. M. Rietbergen, P. Lambin, and H. J. Aerts, "Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer," *Frontiers in oncology,* vol. 5, pp. 272, 2015.

[269] E. Segal, C. B. Sirlin, C. Ooi, A. S. Adler, J. Gollub, X. Chen, B. K. Chan, G. R. Matcuk, C. T. Barry, and H. Y. Chang, "Decoding global gene expression programs in liver cancer by noninvasive imaging," *Nature biotechnology,* vol. 25, no. 6, pp. 675, 2007.

[270] M. Nishino, S. E. Dahlberg, L. E. Fulton, S. R. Digumarthy, H. Hatabu, B. E. Johnson, and L. V. Sequist, "Volumetric tumor response and progression in EGFR-mutant NSCLC patients treated with erlotinib or gefitinib," *Academic radiology,* vol. 23, no. 3, pp. 329-336, 2016.

[271] J.-S. Hsu, M.-S. Huang, C.-Y. Chen, G.-C. Liu, T.-C. Liu, I.-W. Chong, S.-H. Chou, and C.-J. Yang, "Correlation between EGFR mutation status and computed tomography features in patients with

advanced pulmonary adenocarcinoma," *Journal of thoracic imaging,* vol. 29, no. 6, pp. 357-363, 2014.

[272] A. Wibmer, H. Hricak, T. Gondo, K. Matsumoto, H. Veeraraghavan, D. Fehr, J. Zheng, D. Goldman, C. Moskowitz, and S. W. Fine, "Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores," *European radiology,* vol. 25, no. 10, pp. 2840-2850, 2015.

[273] H. Itakura, A. S. Achrol, L. A. Mitchell, J. J. Loya, T. Liu, E. M. Westbroek, A. H. Feroze, S. Rodriguez, S. Echegaray, and T. D. Azad, "Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities," *Science translational medicine,* vol. 7, no. 303, pp. 303ra138-303ra138, 2015.

[274] M. Nicolasjilwan, Y. Hu, C. Yan, D. Meerzaman, C. A. Holder, D. Gutman, R. Jain, R. Colen, D. L. Rubin, and P. O. Zinn, "Addition of MR imaging features and genetic biomarkers strengthens glioblastoma survival prediction in TCGA patients," *Journal of Neuroradiology,* vol. 42, no. 4, pp. 212-221, 2015.

[275] P. Grossmann, D. A. Gutman, W. D. Dunn, C. A. Holder, and H. J. Aerts, "Imaging-genomics reveals driving pathways of MRI derived volumetric tumor phenotype features in Glioblastoma," *BMC cancer,* vol. 16, no. 1, pp. 611, 2016.

[276] S. Chicklore, V. Goh, M. Siddique, A. Roy, P. K. Marsden, and G. J. Cook, "Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis," *European journal of nuclear medicine and molecular imaging,* vol. 40, no. 1, pp. 133-140, 2013.

[277] D.-H. Kim, J.-h. Jung, S. H. Son, C.-Y. Kim, S. Y. Jeong, S.-W. Lee, J. Lee, and B.-C. Ahn, "Quantification of intratumoral metabolic macroheterogeneity on 18F-FDG PET/CT and its prognostic significance in pathologic N0 squamous cell lung carcinoma," *Clinical nuclear medicine,* vol. 41, no. 2, pp. e70-e75, 2016.

[278] F. Khalvati, A. Wong, and M. A. Haider, “Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models,” *BMC medical imaging,* vol. 15, no. 1, pp. 27, 2015.

[279] Y. Cui, K. K. Tha, S. Terasaka, S. Yamaguchi, J. Wang, K. Kudo, L. Xing, H. Shirato, and R. Li, “Prognostic imaging biomarkers in glioblastoma: development and independent validation on the basis of multiregion and quantitative analysis of MR images,” *Radiology,* vol. 278, no. 2, pp. 546-553, 2015.

[280] B. Ganeshan, S. Abaleke, R. C. Young, C. R. Chatwin, and K. A. Miles, “Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage,” *Cancer imaging,* vol. 10, no. 1, pp. 137, 2010.

[281] S. Basu, L. O. Hall, D. B. Goldgof, Y. Gu, V. Kumar, J. Choi, R. J. Gillies, and R. A. Gatenby, "Developing a classifier model for lung tumors in CT-scan images." pp. 1306-1312.

[282] R. Jain, L. M. Poisson, D. Gutman, L. Scarpace, S. N. Hwang, C. A. Holder, M. Wintermark, A. Rao, R. R. Colen, and J. Kirby, “Outcome prediction in patients with glioblastoma by using imaging, clinical, and genomic biomarkers: focus on the nonenhancing component of the tumor,” *Radiology,* vol. 272, no. 2, pp. 484-493, 2014.

[283] F. Feng, P. Wang, K. Zhao, B. Zhou, H. Yao, Q. Meng, L. Wang, Z. Zhang, Y. Ding, and L. Wang, “Radiomic features of hippocampal subregions in Alzheimer’s disease and amnestic mild cognitive impairment,” *Frontiers in aging neuroscience,* vol. 10, 2018.

[284] M. S. de Oliveira, M. L. Balthazar, A. D'Abreu, C. L. Yasuda, B. P. Damasceno, F. Cendes, and G. Castellano, “MR imaging texture analysis of the corpus callosum and thalamus in amnestic mild cognitive impairment and mild Alzheimer disease,” *AJNR Am J Neuroradiol,* vol. 32, no. 1, pp. 60-6, Jan, 2011.

[285] K. R. Anandh, C. M. Sujatha, and S. Ramakrishnan, “Segmentation and Analysis of Corpus Callosum in Alzheimer MR Images using Total Variation Based Diffusion Filter and Level Set Method,” *Biomed Sci Instrum,* vol. 51, pp. 355-61, 2015.

[286] A. Chincarini, F. Sensi, L. Rei, G. Gemme, S. Squarcia, R. Longo, F. Brun, S. Tangaro, R. Bellotti, N. Amoroso, M. Bocchetta, A. Redolfi, P. Bosco, M. Boccardi, G. B. Frisoni, and F. Nobili, “Integrating longitudinal information in hippocampal volume measurements for the early detection of Alzheimer's disease,” *Neuroimage,* vol. 125, pp. 834-847, Jan 15, 2016.

[287] A. Chaddad, C. Desrosiers, and T. Niazi, “Deep radiomic analysis of MRI related to Alzheimer’s Disease,” *IEEE Access,* vol. 6, pp. 58213-58221, 2018.

[288] J. Zhang, L. Tong, L. Wang, and N. Li, “Texture analysis of multiple sclerosis: a comparative study,” *Magnetic resonance imaging,* vol. 26, no. 8, pp. 1160-1166, 2008.

[289] X. Ma, L. Zhang, D. Huang, J. Lyu, M. Fang, J. Hu, Y. Zang, D. Zhang, H. Shao, and L. Ma, “Quantitative radiomic biomarkers for discrimination between neuromyelitis optica spectrum disorder and multiple sclerosis,” *Journal of Magnetic Resonance Imaging*, 2018.

[290] A. Chaddad, C. Desrosiers, L. Hassan, and C. Tanougast, “Hippocampus and amygdala radiomic biomarkers for the study of autism spectrum disorder,” *BMC neuroscience,* vol. 18, no. 1, pp. 52, 2017.

[291] A. Rahmim, P. Huang, N. Shenkov, S. Fotouhi, E. Davoodi-Bojd, L. Lu, Z. Mari, H. Soltanian-Zadeh, and V. Sossi, “Improved prediction of outcome in Parkinson's disease using radiomics analysis of longitudinal DAT SPECT images,” *NeuroImage: Clinical,* vol. 16, pp. 539-544, 2017.

[292] A. Rahmim, Y. Salimpour, S. Jain, S. A. L. Blinder, I. S. Klyuzhin, G. S. Smith, Z. Mari, and V. Sossi, “Application of texture analysis to DAT SPECT imaging: Relationship to clinical assessments,” *NeuroImage: Clinical,* vol. 12, no. Supplement C, pp. e1-e9, 2016/02/01/, 2016.

[293] S. Ashrafinia, P. Dalaie, R. Yan, P. Ghazi, C. Marcus, M. Taghipour, P. Huang, M. Pomper, T. Schindler, and A. Rahmim, "Radiomics Analysis of Clinical Myocardial Perfusion SPECT to Predict Coronary Artery Calcification," *Journal of Nuclear Medicine,* vol. 59, no. supplement 1, pp. 512, May 1, 2018, 2018.

[294] S. Ashrafinia, P. Dalaie, R. Yan, P. Huang, M. Pomper, T. Schindler, and A. Rahmim, "Application of Texture and Radiomics Analysis to Clinical Myocardial Perfusion SPECT Imaging," *Journal of Nuclear Medicine,* vol. 59, no. supplement 1, pp. 94, May 1, 2018, 2018.

[295] I. El Naqa, and R. K. Ten Haken, "Can radiomics personalise immunotherapy?," *The Lancet Oncology,* vol. 19, no. 9, pp. 1138-1139, 2018.

[296] G. Palma, S. Monti, V. D'Avino, M. Conson, R. Liuzzi, M. C. Pressello, V. Donato, J. O. Deasy, M. Quarantelli, and R. Pacelli, "A voxel-based approach to explore local dose differences associated with radiation-induced lung damage," *International Journal of Radiation Oncology* Biology* Physics,* vol. 96, no. 1, pp. 127-133, 2016.

[297] C. Veiga, D. Landau, J. R. McClelland, J. A. Ledermann, D. Hawkes, S. M. Janes, and A. Devaraj, "Long term radiological features of radiation-induced lung damage," *Radiotherapy and Oncology,* vol. 126, no. 2, pp. 300-306, 2018.

[298] P. Lohmann, G. Stoffels, G. Ceccon, M. Rapp, M. Sabel, C. P. Filss, M. A. Kamp, C. Stegmayr, B. Neumaier, and N. J. Shah, "Radiation injury vs. recurrent brain metastasis: combining textural feature radiomics analysis and standard parameters may increase 18F-FET PET accuracy without dynamic scans," *European radiology,* vol. 27, no. 7, pp. 2916-2927, 2017.

[299] G. J. R. Cook, M. Siddique, B. P. Taylor, C. Yip, S. Chicklore, and V. Goh, "Radiomics in PET: principles and applications," *Clinical and Translational Imaging,* vol. 2, no. 3, pp. 269-276, 2014.

[300] N. L. S. T. R. Team, "The national lung screening trial: overview and study design," *Radiology,* vol. 258, no. 1, pp. 243-253, 2011.

[301] I. Shiri, H. Maleki, G. Hajianfar, H. Abdollahi, S. Ashrafinia, M. Ghelich Oghli, O. M., and A. Rahmim, "PET/CT radiomic sequencer for prediction of EGFR and KRAS mutation status in NSCLC patients."

[302] I. Shiri, H. Maleki, H. Abdollahi, G. Hajianfar, S. Ashrafinia, G. O. Mostafa, M. Oveisi, and A. Rahmim, "Next generation radiogenomics sequencing for prediction of EGFR and KRAS mutation status in NSCLC patients using hybrid imaging and machine learning approaches," *Clinical Cancer Research*, 2018.

[303] S. Ashrafinia, H. Mohy-ud-Din, N. A. Karakatsanis, A. K. Jha, M. E. Casey, D. J. Kadrmas, and A. Rahmim, "Generalized PSF modeling for optimized quantitation in PET imaging," *Physics in medicine and biology,* vol. 62, no. 12, pp. 5149, 2017.

[304] S. Ashrafinia, N. Karakatsanis, H. Mohy-ud-Din, and A. Rahmim, "Towards continualized task-based resolution modeling in PET imaging," vol. 9033, pp. 903327-903327-6, 2014.

[305] S. Ashrafinia, H. Mohy-ud-Din, N. KARAKATSANIS, M. Casey, M. Lodge, and A. Rahmim, "PSF Overestimation Improves PET Image SUV Quantification," *Journal of Nuclear Medicine,* vol. 57 (suppl. 2), no. supplement 2, pp. 1965, May 1, 2016, 2016.

[306] S. Ashrafinia, E. Mena, H. Mohy-ud-Din, A. Jha, R. Subramaniam, and A. Rahmim, "Adaptive PSF Modeling for Enhanced Heterogeneity Quantification in Oncologic PET Imaging," *Journal of Nuclear Medicine,* vol. 57, no. supplement 2, pp. 479, May 1, 2016, 2016.

[307] S. Ashrafinia, H. Mohy-ud-Din, N. Karakatsanis, D. Kadrmas, and A. Rahmim, "Enhanced quantitative PET imaging utilizing adaptive partial resolution modeling," *J NUCL MED MEETING ABSTRACTS,* vol. 55, no. 1_MeetingAbstracts, pp. 371-, May 1, 2014, 2014.

[308] A. K. Jha, E. Mena, B. Caffo, S. Ashrafinia, A. Rahmim, E. Frey, and R. M. Subramaniam, "Practical no-gold-standard evaluation framework for quantitative imaging methods: application to lesion segmentation in positron emission tomography," *Journal of Medical Imaging,* vol. 4, no. 1, pp. 011011-011011, 2017.

[309] A. Rahmim, C. R. Schmidtlein, A. Jackson, S. Sheikhbahaei, C. Marcus, S. Ashrafinia, M. Soltani, and R. M. Subramaniam, “A novel metric for quantification of homogeneous and heterogeneous tumors in PET for enhanced clinical outcome prediction,” *Phys. Med. Biol.,* vol. 61, pp. 227-242, 2016.

[310] A. Rahmim, K. Bak-Fredslund, S. Ashrafinia, C. Schmidtlein, R. Subramaniam, A. Morsing, S. Keiding, J. Horsager, and O. Munk, “Quantification of Colorectal Liver Metastases using FDG PET Volumetric and Heterogeneity Features for Improved Prediction of Clinical Outcome,” *Journal of Nuclear Medicine,* vol. 59, no. supplement 1, pp. 350-350, 2018.

[311] A. Zwanenburg, “EP-1677: multicentre initiative for standardisation of image biomarkers,” *Radiotherapy and Oncology,* vol. 123, pp. S914-S915, 2017.

[312] A. Rahmim, S. Ashrafinia, S. Rowe, C. Schmidtlein, M. Vendelbo, T. El-Galaly, L. Gormsen, and O. Munk, “Quantification of Lymphoma using FDG PET Heterogeneity Features for Improved Prediction of Clinical Outcome,” *Journal of Nuclear Medicine,* vol. 59, no. supplement 1, pp. 594-594, 2018.

[313] S. Ashrafinia, K. Jones, M. Gorin, S. Rowe, M. S. Javadi, M. pomper, M. allaf, and A. Rahmim, "Reproducibility of cold uptake radiomics in 99mTc-Sestamibi SPECT imaging of renal cell carcinoma."

[314] S. Ashrafinia, P. Ghazi, C. V. Marcus, M. Taghipour, R. Yan, I. Valenta, M. Pomper, T. H. Schindler, and A. Rahmim, "Robustness and Reproducibility of Radiomic Features in 99mTc-Sestamibi SPECT imaging of Myocardial Perfusion."

[315] I. Shiri, P. Ghafarian, S. Ashrafinia, A. Bitarafan-Rajabi, M. Ay, and A. Rahmim, "Optimal quantitative SUV metrics over wide range of lesion sizes in advanced image reconstruction (TOF and PSF) for PET." pp. S427-S428.

[316] I. Shiri, A. Rahmim, S. Ashrafinia, P. Geramifar, and A. Bitarafan-Rajabi, "Impact of matrix size on metabolic tumor volume (MTV) and total lesion glycolysis (TLG) in PSF-based PET image." pp. S439-S439.

[317] S. Ashrafinia, M. DiGianvittorio, S. Rowe, M. Gorin, L. Lu, M. Lodge, M. Pomper, M. Allaf, and A. Rahmim, "Reproducibility and Reliability of Radiomic Features in 18F-DCFPyL PET/CT Imaging of Prostate Cancer," *Journal of Nuclear Medicine,* vol. 58, no. supplement 1, pp. 503-503, 2017.

[318] Y. Gao, M. Bilgel, S. Ashrafinia, L. Lu, O. Rousset, S. Resnick, D. Wong, and A. Rahmim, "Evaluation of non-local methods with and without anatomy information for improved quantitative amyloid PET imaging," *Journal of Nuclear Medicine,* vol. 59, no. supplement 1, pp. 1779-1779, 2018.

[319] K. Leung, W. Marashdeh, R. Wray, S. Ashrafinia, A. Rahmim, M. Pomper, and A. Jha, "A deep-learning-based fully automated segmentation approach to delineate tumors in FDG-PET images of patients with lung cancer," *Journal of Nuclear Medicine,* vol. 59, no. supplement 1, pp. 323-323, 2018.

[320] R. Yan, S. Ashrafinia, S. Park, J. Lee, L. C. Chu, C. T. Lin, A. Hussien, N. Malguria, J. Steingrimsson, and A. Rahmim, "The Use of Low-Dose CT Intra-and Extra-Nodular Image Texture Features to Improve Small Lung Nodule Diagnosis in Lung Cancer Screening." pp. 1-4.

[321] M. Soltani, M. Sefidgar, H. Bazmara, S. Sheikhbahaei, C. Marcus, S. Ashrafinia, R. Subramaniam, and A. Rahmim, "WE-AB-204-07: Spatiotemporal Distribution of the FDG PET Tracer in Solid Tumors: Contributions of Diffusion and Convection Mechanisms," *Medical Physics,* vol. 42, no. 6Part37, pp. 3660-3660, 2015.

[322] A. Rahmim, C. R. Schmidtlein, A. Jackson, C. Marcus, S. Ashrafinia, M. Soltani, and R. M. Subramaniam, "Novel quantitative generalized total effective entrapment (gTEE) metric for enhanced clinical outcome prediction: application to PET/CT imaging of pancreatic cancer," *J. Nucl. Med.,* vol. 56 (suppl. 3), pp. 647, 2015.

[323] Y. Zhou, M. Lodge, J. Crandall, N. Karakatsanis, M. Casey, S. Ashrafinia, X. Chen, R. Subramaniam, R. Wahl, and A. Rahmim, “Evaluation of whole-body parametric PET/CT image generation from high resolution high sensitivity mCT scanner,” *Journal of Nuclear Medicine,* vol. 56, no. supplement 3, pp. 372-372, 2015.

[324] A. Rahmim, C. Schmidtlein, A. Jackson, C. Marcus, S. Ashrafinia, M. Soltani, and R. Subramaniam, "A generalized total effective entrapment metric (gTEE) to quantify burden of homogeneous and heterogeneous tumors in PET imaging for enhanced clinical outcome prediction." pp. 1-9.

[325] A. Zwanenburg, M. Abdalah, A. Apte, S. Ashrafinia, J. Beukinga, M. Bogowicz, C. Dinh, M. Götz, M. Hatt, and R. Leijenaar, “PO-0981: Results from the Image Biomarker Standardisation Initiative,” *Radiotherapy and Oncology,* vol. 127, pp. S543-S544, 2018.

[326] S. Ashrafinia, M. Salehi Sadaghiani, P. Dalaie, R. Tulbah, Y. Yin, K. Leung, R. VanDenBerg, J. P. Leal, M. Gorin, M. Pomper, A. Rahmim, and S. Rowe, "Characterization of Segmented 18F-DCFPyL PET/CT Lesions in the Context of PSMA-RADS Structured Reporting."

[327] J. Kim, S. Seo, S. Ashrafinia, A. Rahmim, and V. Sossi, "Training of deep convolutional neural nets to extract radiomic signatures of tumors."

[328] P. Dalaie, S. Ashrafinia, M. Salehi Sadaghiani, M. Pomper, L. B. Solnes, A. Rahmim, and S. Rowe, "Predicting Cardiac Function Worsening using Myocardial Perfusion SPECT and Coronary Artery Calcium Scoring."

[329] K. Leung, S. Ashrafinia, M. Salehi Sadaghiani, P. Dalaie, R. Tulbah, Y. Yin, K. Leung, R. VanDenBerg, J. P. Leal, M. Gorin, Y. Du, M. Pomper, S. Rowe, and A. Rahmim, "A fully automated deep-learning based method for lesion segmentation in 18 F-DCFPyL PSMA PET images of patients with prostate cancer."

[330] I. Shiri, S. Ashrafinia, G. Hajianfar, E. Jenabi, M. Oveisi, and A. Rahmim, "Radiogenomics analysis of PET/CT images in lung cancer patients: Conventional radiomics versus deep learning."

[331] I. Shiri, S. Ashrafinia, G. Hajianfar, E. Jenabi, M. Oveisi, and A. Rahmim, "Application of ensemble machine learning to PET/CT radiomic features for improved prediction of recurrence in non-small cell lung cancer."

[332] K. Leung, W. Marashdeh, R. Wray, S. Ashrafinia, M. Pomper, A. Rahmim, and A. jha, "A Modular Deep-Learning Based Fully Automated Framework for Tumor Segmentation in PET Images," *Radiology*, 2019, Under review.

[333] A. Zwanenburg, M. Abdalah, H. J. Aerts, V. Andrearczyk, A. Apte, S. Ashrafinia, M. Vallières, and E. al., "Standardized quantitative radiomics for high-throughput image-based phenotyping," *Medical Image Analysis*, 2019 Under review.

[334] I. Shiri, H. maleki, H. Abdollahi, G. Hajianfar, S. Ashrafinia, G. O. Mostafa, M. Oveisi, and A. Rahmim, "Next generation radiogenomics sequencing for prediction of EGFR and KRAS mutation status in NSCLC patients using hybrid imaging and machine learning approaches," *Clinical Cancer Research*, 2019, Under review.

[335] M. Soret, S. L. Bacharach, and I. Buvat, "Partial-volume effect in PET tumor imaging," *Journal of Nuclear Medicine,* vol. 48, no. 6, pp. 932-945, Jun, 2007.

[336] *Performance Measurement of Positron Emission Tomographs* N. E. M. Association*,* 2012.

[337] H. Zaidi, and M. L. Montandon, "The new challenges of brain PET imaging technology," *Current Medical Imaging Reviews,* vol. 2, pp. 3-13, 2006.

[338] E. J. Hoffman, S.-C. Huang, and M. E. Phelps, "Quantitation in positron emission computed tomography: 1. Effect of object size," *Journal of computer assisted tomography,* vol. 3, no. 3, pp. 299-308, 1979.

[339] J. C. Mazziotta, M. E. Phelps, D. Plummer, and D. E. Kuhl, "Quantitation in positron emission computed tomography. V. Physical-anatomical effects," *Journal of Computer Assisted Tomography,* vol. 5, no. 5, pp. 734- 43, 1981.

[340] J. B. Chawluk, A. Alavi, R. Dann, H. I. Hurtig, S. Bais, M. J. Kushner, R. A. Zimmerman, and M. Reivich, "Positron emission tomography in aging and dementia: effect of cerebral atrophy," *Journal of Nuclear Medicine,* vol. 28, no. 4, pp. 431-437, 1987.

[341] N. K. Tanna, M. I. Kohn, D. N. Horwich, P. R. Jolles, R. A. Zimmerman, W. M. Alves, and A. Alavi, "Analysis of brain and cerebrospinal fluid volumes with MR imaging: impact on PET data correction for atrophy. Part II. Aging and Alzheimer dementia," *Radiology,* vol. 178, no. 1, pp. 123-130, 1991.

[342] A. J. Da Silva, H. R. Tang, K. H. Wong, M. C. Wu, M. W. Dae, and B. H. Hasegawa, "Absolute quantification of regional myocardial uptake of 99mTc-sestamibi with SPECT: experimental validation in a porcine model," *Journal of Nuclear Medicine,* vol. 42, no. 5, pp. 772-779, 2001.

[343] G. G. Bural, D. A. Torigian, W. Chamroonrat, K. Alkhawaldeh, M. Houseni, G. El-Haddad, and A. Alavi, "Quantitative assessment of the atherosclerotic burden of the aorta by combined FDG-PET and CT image analysis: a new concept," *Nuclear medicine and biology,* vol. 33, no. 8, pp. 1037-1043, 2006.

[344] M. Hickeson, M. Yun, A. Matthies, H. Zhuang, L.-E. Adam, L. Lacorte, and A. Alavi, "Use of a corrected standardized uptake value based on the lesion size on CT permits accurate characterization of lung nodules on FDG-PET," *European journal of nuclear medicine and molecular imaging,* vol. 29, no. 12, pp. 1639-1647, 2002.

[345] O. G. Rousset, A. Rahmim, A. Alavi, and H. Zaidi, "Partial volume correction strategies in PET," *PET Clinics,* vol. 2, pp. *235-249*, 2007.

[346] R. M. Kessler, J. J. Ellis, and M. Eden, "Analysis of emission tomographic scan data: limitations imposed by resolution and background," *Journal of computer assisted tomography,* vol. 8, no. 3, pp. 514-522, 1984.

[347] O. G. Rousset, Y. Ma, and A. C. Evans, “Correction for partial volume effects in PET: principle and validation,” *J Nucl Med,* vol. 39, no. 5, pp. 904-11, May, 1998.

[348] A. Le Pogam, M. Hatt, P. Descourt, N. Boussion, C. Tsoumpas, F. E. Turkheimer, C. Prunier-Aesch, J. L. Baulieu, D. Guilloteau, and D. Visvikis, “Evaluation of a 3D local multiresolution algorithm for the correction of partial volume effects in positron emission tomography,” *Medical Physics,* vol. 38, no. 9, pp. 4920-4933, Sep, 2011.

[349] M. Shidahara, C. Tsoumpas, A. Hammers, N. Boussion, D. Visvikis, T. Suhara, I. Kanno, and F. E. Turkheimer, “Functional and structural synergy for resolution recovery and partial volume correction in brain PET,” *Neuroimage,* vol. 44, no. 2, pp. 340-348, Jan 15, 2009.

[350] N. Boussion, M. Hatt, F. Lamare, Y. Bizais, A. Turzo, C. Cheze-Le Rest, and D. Visvikis, “A multiresolution image based approach for correction of partial volume effects in emission tomography.,” *Phys Med Biol,* vol. 51, pp. 1857-1876, 2006.

[351] B. K. Teo, Y. Seo, S. L. Bacharach, J. A. Carrasquillo, S. K. Libutti, H. Shukla, B. H. Hasegawa, R. A. Hawkins, and B. L. Franc, “Partial-volume correction in PET: Validation of an iterative postreconstruction method with phantom and patient data,” *Journal of Nuclear Medicine,* vol. 48, no. 5, pp. 802-810, 2007.

[352] A. S. Kirov, J. Z. Piao, and C. R. Schmidtlein, “Partial volume effect correction in PET using regularized iterative deconvolution with variance control based on local topology,” *Physics in Medicine and Biology,* vol. 53, no. 10, pp. 2577-2591, May 21, 2008.

[353] N. Boussion, C. C. Le Rest, M. Hatt, and D. Visvikis, “Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging,” *European journal of nuclear medicine and molecular imaging,* vol. 36, no. 7, pp. 1064-1075, Jul, 2009.

[354] A. Iriarte, R. Marabini, S. Matej, C. O. S. Sorzano, and R. M. Lewitt, “System models for PET statistical iterative reconstruction: A review,” *Computerized Medical Imaging and Graphics,* vol. 48, pp. 30-48, 3//, 2016.

[355] P. J. Slomka, T. Pan, and G. Germano, “Recent Advances and Future Progress in PET Instrumentation,” *Seminars in Nuclear Medicine,* vol. 46, no. 1, pp. 5-19, 1//, 2016.

[356] B. W. Jakoby, Y. Bercier, M. Conti, M. E. Casey, B. Bendriem, and D. W. Townsend, “Physical and clinical performance of the mCT time-of-flight PET/CT scanner,” *Physics in Medicine and Biology,* vol. 56, no. 8, pp. 2375-2389, 2011.

[357] E. Rapisarda, V. Bettinardi, K. Thielemans, and M. C. Gilardi, “Image-based point spread function implementation in a fully 3D OSEM reconstruction algorithm for PET,” *Physics in Medicine and Biology,* vol. 55, no. 14, pp. 4131-4151, Jul 21, 2010.

[358] M. Miller, J. Zhang, K. Binzel, J. Griesmer, T. Laurence, M. Narayanan, D. Natarajamani, S. Wang, and M. Knopp, “Characterization of the Vereos Digital Photon Counting PET System,” *Journal of Nuclear Medicine,* vol. 56, no. supplement 3, pp. 434, May 1, 2015, 2015.

[359] J. Qi, J. Zhou, H. YE, and W. Wang, "Image-based point-spread-function modelling in time-of-flight positron-emission-tomography iterative list-mode reconstruction," Google Patents, 2015.

[360] C. Cloquet, F. C. Sureau, M. Defrise, G. Van Simaeys, N. Trotta, and S. Goldman, “Non-Gaussian space-variant resolution modelling for list-mode reconstruction,” *Physics in Medicine and Biology,* vol. 55, no. 17, pp. 5045-5066, Sep 7, 2010.

[361] F. A. Kotasidis, J. C. Matthews, G. I. Angelis, P. J. Noonan, A. Jackson, P. Price, W. R. Lionheart, and A. J. Reader, “Single scan parameterization of space-variant point spread functions in image space via a printed array: the impact for two PET/CT scanners,” *Physics in Medicine and Biology,* vol. 56, no. 10, pp. 2917-2942, May 21, 2011.

[362] F. C. Sureau, A. J. Reader, C. Comtat, C. Leroy, M.-J. Ribeiro, I. Buvat, and R. Trebossen, "Impact of image-space resolution modeling for studies with the High-Resolution Research Tomograph.," *J Nucl Med,* vol. 49, no. 6, pp. 1000-1008, June 1, 2008, 2008.

[363] A. J. Reader, P. J. Julyan, H. Williams, D. L. Hastings, and J. Zweit, "EM algorithm system modeling by image-space techniques for PET reconstruction.," *IEEE Trans Nuc Sci,* vol. 50, no. 5, pp. 1392-1397, 2003.

[364] A. Rahmim, M. Lenox, C. Michel, A. J. Reader, and V. Sossi, "Space-variant and anisotropic resolution modeling in list-mode EM reconstruction," *IEEE Nucl. Sci. Symp. Conf. Record* vol. 5, pp. 3074-3077, 2003.

[365] P. Antich, R. Parkey, S. Seliounine, N. Slavine, E. Tsyganov, and A. Zinchenko, "Application of expectation maximization algorithms for image resolution improvement in. a small animal PET system," *IEEE Transactions on Nuclear Science,* vol. 52, no. 3, pp. 684-690, Jun, 2005.

[366] F. A. Kotasidis, G. I. Angelis, J. Anton-Rodriguez, P. Markiewicz, W. R. Lionheart, A. J. Reader, and J. C. Matthews, "Image-Based Spatially Variant and Count Rate Dependent Point Spread Function on the HRRT," *IEEE Transactions on Nuclear Science,* vol. 61, no. 3, pp. 1192-1202, 2014.

[367] V. Y. Panin, F. Kehren, H. Rothfuss, D. Hu, C. Michel, and M. E. Casey, "PET reconstruction with system matrix derived from point source measurements.," *IEEE Trans Nuc Sci,* vol. 53, no. 1, pp. 152-159, 2006.

[368] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar, "High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner.," *Phys Med Biol,* vol. 43, no. 4, pp. 1001-1013, Apr, 1998.

[369] J. Qi, R. M. Leahy, H. Chinghan, T. H. Farquhar, and S. R. Cherry, "Fully 3D Bayesian image reconstruction for the ECAT EXACT HR+," *Nuclear Science, IEEE Transactions on,* vol. 45, no. 3, pp. 1096-1103, 1998.

[370] A. M. Alessio, P. E. Kinahan, and T. K. Lewellen, “Modeling and incorporation of system response functions in 3-D whole body PET.,” *IEEE Trans Med Imaging,* vol. 25, no. 7, pp. 828-837, 2006.

[371] R. Lecomte, D. Schmitt, and G. Lamoureux, “Geometry Study of a High Resolution PET Detection System Using Small Detectors,” *Nuclear Science, IEEE Transactions on,* vol. 31, no. 1, pp. 556-561, 1984.

[372] D. Schmitt, B. Karuta, C. Carrier, and R. Lecomte, “Fast point spread function computation from aperture functions in high-resolution positron emission tomography,” *Medical Imaging, IEEE Transactions on,* vol. 7, no. 1, pp. 2-12, 1988.

[373] V. V. Selivanov, Y. Picard, J. Cadorette, S. Rodrigue, and R. Lecomte, “Detector response models for statistical iterative image reconstruction in high resolution PET.,” *IEEE Trans Nucl Sci,* vol. 47, no. 3, pp. 1168-1175, 2000.

[374] D. Strul, R. B. Slates, M. Dahlbom, S. R. Cherry, and P. K. Marsden, “An improved analytical detector response function model for multilayer small-diameter PET scanners,” *Physics in Medicine and Biology,* vol. 48, no. 8, pp. 979-994, Apr 21, 2003.

[375] B. Bai, R. Laforest, A. M. Smith, and R. M. Leahy, “Evaluation of MAP image reconstruction with positron range modeling for 3D PET,” in Nuclear Science Symposium Conference Record, 2005 IEEE, 2005, pp. 2686-2689.

[376] A. Ruangma, B. Bai, J. S. Lewis, X. Sun, M. J. Welch, R. Leahy, and R. Laforest, “Three-dimensional maximum a posteriori (MAP) imaging with radiopharmaceuticals labeled with three Cu radionuclides,” *Nucl Med Biol,* vol. 33, no. 2, pp. 217-26, Feb, 2006.

[377] A. Rahmim, J. Tang, M. A. Lodge, S. Lashkari, M. R. Ay, and F. M. Bengel, “Resolution modeled PET image reconstruction incorporating space-variance of positron range: Rubidium-82 cardiac PET imaging,” in IEEE Nucl. Sci. Symp. Conf. Record, 2008, pp. 3643-3650.

[378] A. Alessio, and L. MacDonald, "Spatially variant positron range modeling derived from CT for PET image reconstruction." pp. 3637-3640.

[379] T. Yamaya, N. Hagiwara, T. Obi, M. Yamaguchi, N. Ohyama, K. Kitamura, T. Hasegawa, H. Haneishi, E. Yoshida, N. Inadama, and H. Murayama, "Transaxial system models for jPET-D4 image reconstruction.," *Phys Med Biol,* vol. 22, no. 5339-5355, 2005.

[380] H. Takahashi, T. Yamaya, T. Kobayashi, K. Kitamura, T. Hasegawa, H. Murayama, and M. Suga, "System modeling of small bore DOI-PET scanners for fast and accurate 3D image reconstruction." pp. 3478-3481.

[381] R. E. Carson, W. C. Barker, J.-S. Liow, and C. A. Johnson, "Design of a motion-compensation OSEM list-mode algorithm for resolution-recovery reconstruction for the HRRT," in IEEE Nuclear Science Symposium Conference Record, Portland, OR, 2003, pp. 3281-3285.

[382] A. M. Alessio, C. W. Stearns, T. Shan, S. G. Ross, S. Kohlmyer, A. Ganin, and P. E. Kinahan, "Application and Evaluation of a Measured Spatially Variant System Model for PET Image Reconstruction," *IEEE Trans Med Imag,* vol. 29, no. 3, pp. 938-949, 2010.

[383] J. Vercher-Conejero, A. Sher, J. Kolthammer, A. Kohan, C. Rubbert, P. Wojtylak, S. Partovi, M. Gaeta, and J. O'Donnell, "Contribution of the modeling of point spread function and time-of-flight technology on the image quality of a PET/CT system," *J Nuc Med,* vol. 54, no. 2_MeetingAbstracts, pp. 101-, May 1, 2013, 2013.

[384] S. Tong, A. M. Alessio, and P. E. Kinahan, "Noise and signal properties in PSF-based fully 3D PET image reconstruction: an experimental evaluation," *Physics in Medicine and Biology,* vol. 55, no. 5, pp. 1453-1473, Mar 7, 2010.

[385] A. Rahmim, and J. Tang, "Impact of resolution modeling on detectability: Analysis in terms of noise-equivalent quanta," *J Nuc Med,* vol. 54, no. Suppl. 2, pp. 2119, May 1, 2013, 2013.

[386] N. Karakatsanis, A. Mehranian, M. Lodge, M. Casey, A. Rahmim, and H. Zaidi, "Clinical evaluation of direct 4D whole-body PET parametric imaging with time-of-flight and resolution modeling capabilities," *IEEE Nucl. Sci. Symp. Conf. Record*, 2015.

[387] A. Alessio, A. Rahmim, and C. G. Orton, "Resolution modeling enhances PET imaging (Point/Counterpoint)," *Med. Phys.,* vol. 40, pp. 120601, 2013.

[388] A. Rahmim, and J. Tang, "Noise propagation in resolution modeled PET imaging and its impact on detectability," *Physics in Medicine and Biology,* vol. 58, no. 19, pp. 6945, 2013.

[389] R. Boellaard, N. C. Krak, O. S. Hoekstra, and A. A. Lammertsma, "Effects of Noise, Image Resolution, and ROI Definition on the Accuracy of Standard Uptake Values: A Simulation Study," *J Nucl Med,* vol. 45, no. 9, pp. 1519-1527, September 1, 2004, 2004.

[390] E. D. Bernardi, M. Mazzoli, F. Zito, and G. Baselli, "Resolution Recovery in PET During AWOSEM Reconstruction: A Performance Evaluation Study," *IEEE Transactions on Nuclear Science,* vol. 54, no. 5, pp. 1626-1638, 2007.

[391] D. J. Kadrmas, M. E. Casey, N. F. Black, J. J. Hamill, V. Y. Panin, and M. Conti, "Experimental Comparison of Lesion Detectability for Four Fully-3D PET Reconstruction Schemes," *IEEE Transactions on Medical Imaging,* vol. 28, no. 4, pp. 523-534, Apr, 2009.

[392] E. J. Soares, C. L. Byrne, and S. J. Glick, "Noise characterization of block-iterative reconstruction algorithms: I. Theory.," *IEEE Trans Med Imaging,* vol. 19, no. 4, pp. 261-270, Apr, 2000.

[393] E. J. Soares, S. J. Glick, and J. W. Hoppin, "Noise characterization of block-iterative reconstruction algorithms: II. Monte Carlo simulations.," *IEEE Trans Med Imaging,* vol. 24, no. 1, pp. 112-121, 2005.

[394] H. H. Barrett, D. W. Wilson, and B. M. W. Tsui, "Noise properties of the EM algorithm. I. Theory.," *Phys Med Biol,* vol. 39, no. 5, pp. 833-846, 1994.

[395] D. W. Wilson, B. M. W. Tsui, and H. H. Barrett, "Noise properties of the EM algorithm. II. Monte Carlo simulations," *Phys Med Biol,* vol. 39, no. 5, pp. 847-871, 1994.

[396] S. Tong, A. M. Alessio, K. Thielemans, C. Stearns, S. Ross, and P. E. Kinahan, "Properties and Mitigation of Edge Artifacts in PSF-Based PET Reconstruction," *IEEE Transactions on Nuclear Science,* vol. 58, no. 5, pp. 2264-2275, Oct, 2011.

[397] A. Rahmim, J.-C. Cheng, and V. Sossi, "Improved noise propagation in statistical image reconstruction with resolution modeling," in Nuclear Science Symposium Conference Record, 2005 IEEE, 2005, pp. 2576-2578.

[398] N. J. Hoetjes, F. H. P. van Velden, O. S. Hoekstra, C. J. Hoekstra, N. C. Krak, A. A. Lammertsma, and R. Boellaard, "Partial volume correction strategies for quantitative FDG PET in oncology," *European journal of nuclear medicine and molecular imaging,* vol. 37, no. 9, pp. 1679-1687, Aug, 2010.

[399] S. A. Blinder, K. Dinelle, and V. Sossi, "Scanning rats on the high resolution research tomograph (HRRT): A comparison study with a dedicated micro-PET," *Med Phys,* vol. 39, no. 8, pp. 5073-83, Aug, 2012.

[400] D. L. Snyder, M. I. Miller, L. J. Thomas, and D. G. Politte, "Noise and edge artifacts in maximum-likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imaging,* vol. MI-6, pp. 313-319, 1987.

[401] D. G. Politte, and D. L. Snyder, "The Use of Constraints to Eliminate Artifacts in Maximum-Likelihood Image Estimation for Emission Tomography," *IEEE Transactions on Nuclear Science,* vol. 35, no. 1, pp. 608-610, Feb, 1988.

[402] N. A. Karakatsanis, M. A. Lodge, A. K. Tahari, Y. Zhou, R. L. Wahl, and A. Rahmim, "Dynamic whole body PET parametric imaging: I. Concept, acquisition protocol optimization and clinical application," *Phys. Med. Bio.,* vol. , no. , pp. , 2013.

[403] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," *Medical Physics,* vol. 37, no. 9, pp. 4902-4915, 2010.

[404] D. Feng, S. C. Huang, and X. Wang, "Models for computer simulation studies of input functions for tracer kinetic modeling with positron emission tomography," *Int J Biomed Comput,* vol. 32, no. 2, pp. 95-110, Mar, 1993.

[405] A. Rahmim, J. Tang, M. Lodge, S. Lashkari, M. Ay, R. Lautamaki, B. Tsui, and F. Bengel, "Improved clinical and quantitative dynamic Rb-82 cardiac imaging with resolution modeling," *J Nuc Med,* vol. 49, no. MeetingAbstracts_1, pp. 62P-, May 1, 2008, 2008.

[406] A. Rahmim, J. Tang, M. R. Ay, and F. Bengel, "4D Respiratory Motion-Corrected Rb-82 Myocardial Perfusion PET Imaging," *IEEE Nucl. Sci. Symp. Conf. Record* pp. 3312-3316, 2010.

[407] B. J. Kemp, C. Kim, J. J. Williams, A. Ganin, and V. J. Lowe, "NEMA NU 2-2001 performance measurements of an LYSO-based PET/CT system in 2D and 3D acquisition modes," *J Nucl Med,* vol. 47, no. 12, pp. 1960-7, Dec, 2006.

[408] T. Torizuka, S. Nobezawa, S. Momiki, N. Kasamatsu, T. Kanno, E. Yoshikawa, M. Futatsubashi, H. Okada, and Y. Ouchi, "Short dynamic FDG-PET imaging protocol for patients with lung cancer," *Eur J Nucl Med,* vol. 27, no. 10, pp. 1538-42, Oct, 2000.

[409] S. Okazumi, K. Isono, K. Enomoto, T. Kikuchi, M. Ozaki, H. Yamamoto, H. Hayashi, T. Asano, and M. Ryu, "Evaluation of Liver Tumors Using fluorine- 18-Fluorodeoxyglucose PET: Characterization of Tumor and Assessment of Effect of Treatment," *Journal of Nuclear Medicine,* vol. 33, no. 3, pp. 333-339, March 1, 1992, 1992.

[410] A. Dimitrakopoulou-Strauss, L. G. Strauss, T. Heichel, H. Wu, C. Burger, L. Bernd, and V. Ewerbeck, "The Role of Quantitative 18F-FDG PET Studies for the Differentiation of Malignant and Benign Bone Lesions," *Journal of Nuclear Medicine,* vol. 43, no. 4, pp. 510-518, April 1, 2002, 2002.

[411] S. Tong, A. M. Alessio, and P. E. Kinahan, “Noise and signal properties in PSF-based fully 3D PET image reconstruction: an experimental evaluation,” *Phys Med Biol,* vol. 55, no. 5, pp. 1453-73, Mar 7, 2010.

[412] X. Niu, E. Asma, H. Ye, W. Wang, and D. Gagnon, "Count-level Dependent Image Domain PSF Kernel Width Selection for Fully 3D PET Image Reconstruction."

[413] L. Le Meunier, P. J. Slomka, D. Dey, A. Ramesh, L. E. J. Thomson, S. W. Hayes, J. D. Friedman, V. Cheng, G. Germano, and D. S. Berman, “Enhanced definition PET for cardiac imaging,” *Journal of Nuclear Cardiology,* vol. 17, no. 3, pp. 414-426, Jun, 2010.

[414] J. F. Eary, F. O'Sullivan, J. O'Sullivan, and E. U. Conrad, “Spatial Heterogeneity in Sarcoma (18)F-FDG Uptake as a Predictor of Patient Outcome,” *Journal of Nuclear Medicine,* vol. 49, no. 12, pp. 1973-1979, Dec, 2008.

[415] I. El Naqa, P. W. Grigsby, A. Apte, E. Kidd, E. Donnelly, D. Khullar, S. Chaudhari, D. Yang, M. Schmitt, R. Laforest, W. L. Thorstad, and J. O. Deasy, “Exploring feature-based approaches in PET images for predicting cancer treatment outcomes,” *Pattern Recognition,* vol. 42, no. 6, pp. 1162-1171, Jun, 2009.

[416] F. H. P. van Velden, P. Cheebsumon, M. Yaqub, E. F. Smit, O. S. Hoekstra, A. A. Lammertsma, and R. Boellaard, “Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies,” *European Journal of Nuclear Medicine and Molecular Imaging,* vol. 38, no. 9, pp. 1636-1647, Sep, 2011.

[417] F. Tixier, C. C. Le Rest, M. Hatt, N. Albarghach, O. Pradier, J. P. Metges, L. Corcos, and D. Visvikis, “Intratumor Heterogeneity Characterized by Textural Features on Baseline (18)F-FDG PET Images Predicts Response to Concomitant Radiochemotherapy in Esophageal Cancer,” *Journal of Nuclear Medicine,* vol. 52, no. 3, pp. 369-378, Mar 1, 2011.

[418] M. C. Asselin, J. P. B. O'Connor, R. Boellaard, N. A. Thacker, and A. Jackson, "Quantifying heterogeneity in human tumours using MRI and PET," *European Journal of Cancer,* vol. 48, no. 4, pp. 447-455, Mar, 2012.

[419] D. Vriens, J. A. Disselhorst, W. J. G. Oyen, L. F. de Geus-Oei, and E. P. Visser, "Quantitative Assessment of Heterogeneity in Tumor Metabolism Using FDG-PET," *International Journal of Radiation Oncology Biology Physics,* vol. 82, no. 5, pp. E725-E731, Apr 1, 2012.

[420] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R. G. P. M. van Stiphout, P. Granton, C. M. L. Zegers, R. Gillies, R. Boellard, A. Dekker, H. J. W. L. Aerts, and Q.-C. Consortium, "Radiomics: Extracting more information from medical images using advanced feature analysis," *European Journal of Cancer,* vol. 48, no. 4, pp. 441-446, Mar, 2012.

[421] V. Kumar, Y. H. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, K. Forster, H. J. W. L. Aerts, A. Dekker, D. Fenstermacher, D. B. Goldgof, L. O. Hall, P. Lambin, Y. Balagurunathan, R. A. Gatenby, and R. J. Gillies, "Radiomics: the process and the challenges," *Magnetic Resonance Imaging,* vol. 30, no. 9, pp. 1234-1248, Nov, 2012.

[422] S. Chicklore, V. Goh, M. Siddique, A. Roy, P. K. Marsden, and G. J. R. Cook, "Quantifying tumour heterogeneity in F-18-FDG PET/CT imaging by texture analysis," *European Journal of Nuclear Medicine and Molecular Imaging,* vol. 40, no. 1, pp. 133-140, Jan, 2013.

[423] F. Tixier, M. Hatt, C. Valla, V. Fleury, C. Lamour, S. Ezzouhri, P. Ingrand, R. Perdrisot, D. Visvikis, and C. C. Le Rest, "Visual Versus Quantitative Assessment of Intratumor F-18-FDG PET Uptake Heterogeneity: Prognostic Value in Non-Small Cell Lung Cancer," *Journal of Nuclear Medicine,* vol. 55, no. 8, pp. 1235-1241, Aug, 2014.

[424] H. J. W. L. Aerts, E. R. Velazquez, R. T. H. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M. M. Rietbergen, C. R. Leemans, A. Dekker, J. Quackenbush, R. J. Gillies, and P. Lambin, "Decoding tumour phenotype by noninvasive

imaging using a quantitative radiomics approach," *Nature Communications,* vol. 5, pp. 4006, Aug, 2014.

[425] M. Hatt, M. Majdoub, M. Vallieres, F. Tixier, C. C. Le Rest, D. Groheux, E. Hindie, A. Martineau, O. Pradier, R. Hustinx, R. Perdrisot, R. Guillevin, I. El Naqa, and D. Visvikis, "F-18-FDG PET Uptake Characterization Through Texture Analysis: Investigating the Complementary Nature of Heterogeneity and Functional Tumor Volume in a Multi-Cancer Site Patient Cohort," *Journal of Nuclear Medicine,* vol. 56, no. 1, pp. 38-44, Jan, 2015.

[426] L. Lu, W. Lv, J. Jiang, J. Ma, Q. Feng, A. Rahmim, and W. Chen, "Robustness of radiomic features in [11C]Choline and [18F]FDG PET/CT imaging of nasopharyngeal carcinoma: impact of segmentation and discretization " *Molec. Imag. Biol.,* vol. 18, pp. 935-945, , 2016.

[427] F. van Velden, G. Kramer, V. Frings, I. Nissen, E. Mulder, O. Hoekstra, E. Smit, and R. Boellaard, "Impact of reconstruction algorithms on the repeatability of various PET/CT radiomics features in lung cancer patient," *Journal of Nuclear Medicine,* vol. 56, no. 3, May 1, 2015.

[428] M. Soufi, A. Kamali-Asl, P. Geramifar, and A. Rahmim, " A novel framework for automated segmentation and labeling of homogeneous versus heterogeneous lung tumors in [18F]FDG PET imaging," *Molec. Imag. Biol.,* vol. In Press (doi:10.1007/s11307-016-1015-0), 2016.

[429] A. L. Zwanenburg, Stefan; Vallières, Martin; Löck, Steffen;, "Image biomarker standardisation initiative," *arXiv:1612.07003*, 12/2016, 2016.

[430] B. A. Altazi, G. G. Zhang, D. C. Fernandez, M. E. Montejo, D. Hunt, J. Werner, M. C. Biagioli, and E. G. Moros, "Reproducibility of F18-FDG PET radiomic features for different cervical tumor segmentation methods, gray-level discretization, and reconstruction algorithms," *Journal of applied clinical medical physics,* vol. 18, no. 6, pp. 32-48, 2017.

[431] C. Bailly, C. Bodet-Milin, S. Couespel, H. Necib, F. Kraeber-Bodéré, C. Ansquer, and T. Carlier, "Revisiting the robustness of PET-based textural features in the context of multi-centric trials," *PLoS One,* vol. 11, no. 7, pp. e0159984, 2016.

[432] M. Shafiq-ul-Hassan, G. G. Zhang, K. Latifi, G. Ullah, D. C. Hunt, Y. Balagurunathan, M. A. Abdalah, M. B. Schabath, D. G. Goldgof, and D. Mackin, "Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels," *Medical physics,* vol. 44, no. 3, pp. 1050-1062, 2017.

[433] K. Varnas, C. Halldin, and H. Hall, "Autoradiographic distribution of serotonin transporters and receptor subtypes in human brain," *Hum Brain Mapp,* vol. 22, no. 3, pp. 246-60, Jul, 2004.

[434] R. M. Haralick, and K. Shanmugam, "Textural features for image classification," *IEEE Transactions on systems, man, and cybernetics*, no. 6, pp. 610-621, 1973.

[435] M. M. Galloway, "Texture analysis using grey level run lengths," *NASA STI/Recon Technical Report N,* vol. 75, 1974.

[436] M. Vallières, A. Zwanenburg, B. Badic, C. Cheze Le Rest, D. Visvikis, and M. Hatt, "Responsible Radiomics Research for Faster Clinical Translation," *Journal of Nuclear Medicine,* vol. 59, no. 2, pp. 189-193, February 1, 2018, 2018.

[437] F. J. Brooks, "On some misconceptions about tumor heterogeneity quantification," *European journal of nuclear medicine and molecular imaging,* vol. 40, no. 8, pp. 1292-1294, 2013.

[438] F. Orlhac, M. Soussan, J.-A. Maisonobe, C. A. Garcia, B. Vanderlinden, and I. Buvat, "Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis," *Journal of Nuclear Medicine,* vol. 55, no. 3, pp. 414-422, 2014.

[439] N.-M. Cheng, Y.-H. D. Fang, and T.-C. Yen, "The promise and limits of PET texture analysis," *Annals of nuclear medicine,* vol. 27, no. 9, pp. 867-869, 2013.

[440] A. Zwanenburg, “EP-1677: Multicentre initiative for standardisation of image biomarkers,” *Radiotherapy and Oncology,* vol. 123, pp. S914-S915.

[441] F. Tixier, M. Hatt, C. C. Le Rest, A. Le Pogam, L. Corcos, and D. Visvikis, “Reproducibility of Tumor Uptake Heterogeneity Characterization Through Textural Feature Analysis in 18F-FDG PET,” *Journal of Nuclear Medicine,* vol. 53, no. 5, pp. 693-700, May 1, 2012, 2012.

[442] R. T. H. Leijenar, S. Carvalho, E. R. V. Zquez, W. J. C. Van Elmpt, C. Parmar, O. S. Hoekstra, C. J. Hoekstra, R. Boelard, A. L. A. J. Dekker, R. J. Gillies, H. J. W. L. Aerts, and P. Lambin, “Stability of FDG-PET Radiomics features: An integrated analysis of test-retest and inter-observer variability,” *Acta oncologica (Stockholm, Sweden),* vol. 52, no. 7, pp. 1391-1397, 09/09, 2013.

[443] M. Hatt, F. Tixier, C. Cheze Le Rest, O. Pradier, and D. Visvikis, “Robustness of intratumour 18F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma,” *European Journal of Nuclear Medicine and Molecular Imaging,* vol. 40, no. 11, pp. 1662-1671, 2013.

[444] C. Parmar, “Machine learning applications for Radiomics: Towards robust non-invasive predictors in clinical oncology ”, Datawyse / Universitaire Pers Maastricht, Universitaire Pers Maastricht, Netherlands, 2017.

[445] C. Parmar, P. Grossmann, J. Bussink, P. Lambin, and H. J. W. L. Aerts, “Machine Learning methods for Quantitative Radiomic Biomarkers,” *Scientific Reports,* vol. 5, pp. 13087, 08/17, 2015.

[446] M. A. Gorin, S. P. Rowe, A. S. Baras, L. B. Solnes, M. W. Ball, P. M. Pierorazio, C. P. Pavlovich, J. I. Epstein, M. S. Javadi, and M. E. Allaf, “Prospective Evaluation of 99mTc-sestamibi SPECT/CT for the Diagnosis of Renal Oncocytomas and Hybrid Oncocytic/Chromophobe Tumors,” *European Urology,* vol. 69, no. 3, pp. 413-416, 3//, 2016.

[447] P. Lambin, R. T. H. Leijenaar, T. M. Deist, J. Peerlings, E. E. C. de Jong, J. van Timmeren, S. Sanduleanu, R. T. H. M. Larue, A. J. G. Even, A. Jochems, Y. van Wijk, H. Woodruff, J. van Soest, T.

Lustberg, E. Roelofs, W. van Elmpt, A. Dekker, F. M. Mottaghy, J. E. Wildberger, and S. Walsh, "Radiomics: the bridge between medical imaging and personalized medicine," *Nat Rev Clin Oncol,* vol. advance online publication, 10/04/online, 2017.

[448] M. Zhou, J. Scott, B. Chaudhury, L. Hall, D. Goldgof, K. W. Yeom, M. Iv, Y. Ou, J. Kalpathy-Cramer, S. Napel, R. Gillies, O. Gevaert, and R. Gatenby, "Radiomics in Brain Tumor: Image Assessment, Quantitative Feature Descriptors, and Machine-Learning Approaches," *American Journal of Neuroradiology*, 2017.

[449] J. W. Lee, and S. M. Lee, "Radiomics in Oncological PET/CT: Clinical Applications," *Nuclear Medicine and Molecular Imaging*, 2017/10/20, 2017.

[450] F. S. Collins, and L. A. Tabak, "NIH plans to enhance reproducibility," *Nature,* vol. 505, no. 7485, pp. 612-613, 2014.

[451] F. H. P. van Velden, G. M. Kramer, V. Frings, I. A. Nissen, E. R. Mulder, A. J. de Langen, O. S. Hoekstra, E. F. Smit, and R. Boellaard, "Repeatability of Radiomic Features in Non-Small-Cell Lung Cancer [18F]FDG-PET/CT Studies: Impact of Reconstruction and Delineation," *Molecular Imaging and Biology,* vol. 18, no. 5, pp. 788-795, 2016//, 2016.

[452] L. Lu, W. Lv, J. Jiang, J. Ma, Q. Feng, A. Rahmim, and W. Chen, "Robustness of Radiomic Features in [11C]Choline and [18F]FDG PET/CT Imaging of Nasopharyngeal Carcinoma: Impact of Segmentation and Discretization," *Molecular Imaging and Biology*, pp. 1-11, 2016.

[453] R. Siegel, D. Naishadham, and A. Jemal, "Cancer statistics, 2012," *CA: a cancer journal for clinicians,* vol. 62, no. 1, pp. 10-29, 2012.

[454] Z. Szabo, E. Mena, S. P. Rowe, D. Plyku, R. Nidal, M. A. Eisenberger, E. S. Antonarakis, H. Fan, R. F. Dannals, Y. Chen, R. C. Mease, M. Vranesic, A. Bhatnagar, G. Sgouros, S. Y. Cho, and M. G. Pomper, "Initial Evaluation of [18F]DCFPyL for Prostate-Specific Membrane Antigen (PSMA)-Targeted PET

Imaging of Prostate Cancer," *Molecular Imaging and Biology,* vol. 17, no. 4, pp. 565-574, August 01, 2015.

[455] H. Jadvar, "Radiotheranostics in Prostate Cancer: Introduction and Overview," *Journal of Nuclear Medicine,* vol. 57, no. Supplement 3, pp. 1S-2S, October 1, 2016, 2016.

[456] M. O. Demirkol, Ö. Acar, B. Uçar, S. R. Ramazanoğlu, Y. Sağlıcan, and T. Esen, "Prostate-specific membrane antigen-based imaging in prostate cancer: Impact on clinical decision making process," *The Prostate,* vol. 75, no. 7, pp. 748-757, 2015.

[457] M. Eiber, K. Herrmann, J. Calais, B. Hadaschik, F. L. Giesel, M. Hartenbach, T. Hope, R. Reiter, T. Maurer, and W. A. Weber, "Prostate cancer molecular imaging standardized evaluation (PROMISE): proposed miTNM classification for the interpretation of PSMA-ligand PET/CT," *Journal of Nuclear Medicine,* vol. 59, no. 3, pp. 469-478, 2018.

[458] S. P. Rowe, K. J. Pienta, M. G. Pomper, and M. A. Gorin, "Proposal for a structured reporting system for prostate-specific membrane antigen–targeted PET imaging: PSMA-RADS version 1.0," *Journal of Nuclear Medicine,* vol. 59, no. 3, pp. 479-485, 2018.

[459] A. Afshar-Oromieh, C. Zechmann, M. Eder, M. Eisenhut, and U. Haberkorn, "[68Ga] Gallium labelled PSMA ligand as new PET-tracer for the diagnosis of prostate cancer: Normal uptake of healthy tissue," *Journal of Nuclear Medicine,* vol. 53, no. supplement 1, pp. 1414-1414, 2012.

[460] Y. Chen, M. Pullambhatla, C. A. Foss, Y. Byun, S. Nimmagadda, S. Senthamizhchelvan, G. Sgouros, R. C. Mease, and M. G. Pomper, "2-(3-{1-Carboxy-5-[(6-[18F] fluoro-pyridine-3-carbonyl)-amino]-pentyl}-ureido)-pentanedioic acid,[18F] DCFPyL, a PSMA-based PET imaging agent for prostate cancer," *Clinical Cancer Research,* vol. 17, no. 24, pp. 7645-7653, 2011.

[461] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory,* vol. 28, no. 2, pp. 129-137, 1982.

[462] J. Flusser, and T. Suk, "Pattern recognition by affine moment invariants," *Pattern recognition,* vol. 26, no. 1, pp. 167-174, 1993.

[463] Y. Arad, K. J. Goodman, M. Roth, D. Newstein, and A. D. Guerci, "Coronary Calcification, Coronary Disease Risk Factors, C-Reactive Protein, and Atherosclerotic Cardiovascular Disease Events: The St. Francis Heart Study," *Journal of the American College of Cardiology,* vol. 46, no. 1, pp. 158-165, 2005/07/05/, 2005.

[464] P. Slomka, Y. Xu, D. Berman, and G. Germano, "Quantitative analysis of perfusion studies: strengths and pitfalls," *Journal of Nuclear Cardiology,* vol. 19, no. 2, pp. 338-346, 2012.

[465] B. L. Zaret, and F. J. Wackers, "Nuclear cardiology," *New England Journal of Medicine,* vol. 329, no. 11, pp. 775-783, 1993.

[466] J. Maddahi, K. van Train, F. Prigent, E. V. Garcia, J. Friedman, E. Ostrzega, and D. Berman, "Quantitative single photon emission computed thallium-201 tomography for detection and localization of coronary artery disease: optimization and prospective validation of a new technique," *Journal of the American College of Cardiology,* vol. 14, no. 7, pp. 1689-1699, 1989.

[467] S. M. Chang, F. Nabi, J. Xu, U. Raza, and J. J. Mahmarian, "Normal stress-only versus standard stress/rest myocardial perfusion imaging: similar patient mortality with reduced radiation exposure," *Journal of the American College of Cardiology,* vol. 55, no. 3, pp. 221-230, 2010.

[468] K. L. Gould, R. J. Westcott, P. C. Albro, and G. W. Hamilton, "Noninvasive assessment of coronary stenoses by myocardial imaging during pharmacologic coronary vasodilatation: II. Clinical methodology and feasibility," *The American journal of cardiology,* vol. 41, no. 2, pp. 279-287, 1978.

[469] F. Klocke, "Perfusion imaging: Interpretation of regional differences during flow augmentation and detection of hibernating myocardium," *Louis: CV Mosby*, pp. 149-54, 1993.

[470] D. P. Zipes, P. Libby, R. O. Bonow, D. L. Mann, and G. F. Tomaselli, *Braunwald's Heart Disease E-Book: A Textbook of Cardiovascular Medicine*: Elsevier Health Sciences, 2018.

[471] S. Kaul, C. A. Boucher, J. B. Newell, D. A. Chesler, J. M. Greenberg, R. D. Okada, H. W. Strauss, R. E. Dinsmore, and G. M. Pohost, "Determination of the quantitative thallium imaging variables that optimize detection of coronary artery disease," *Journal of the American College of Cardiology,* vol. 7, no. 3, pp. 527-537, 1986.

[472] C. A. Boucher, D. C. Brewster, R. C. Darling, R. D. Okada, H. W. Strauss, and G. M. Pohost, "Determination of cardiac risk by dipyridamole-thallium imaging before peripheral vascular surgery," *New England Journal of Medicine,* vol. 312, no. 7, pp. 389-394, 1985.

[473] O. Parodi, C. Marcassa, R. Casucci, G. Sambuceti, E. Verna, M. Galli, E. Inglese, P. Marzullo, S. Pirelli, and G. Bisi, "Accuracy and safety of technetium-99m hexakis 2-methoxy-2-isobutyl isonitrile (sestamibi) myocardial scintigraphy with high dose dipyridamole test in patients with effort angina pectoris: a multicenter study," *Journal of the American College of Cardiology,* vol. 18, no. 6, pp. 1439-1444, 1991.

[474] S. Nishimura, J. J. Mahmarian, T. M. Boyce, and M. S. Verani, "Equivalence between adenosine and exercise thallium-201 myocardial tomography: A multicentre, prospective, crossover trial," *Journal of the American College of Cardiology,* vol. 20, no. 2, pp. 265-275, 1992.

[475] H. S. Hecht, R. E. Shaw, T. R. Bruce, C. Ryan, S. H. Stertzer, and R. K. Myler, "Usefulness of tomographic thallium-201 imaging for detection of restenosis after percutaneous transluminal coronary angioplasty," *The American journal of cardiology,* vol. 66, no. 19, pp. 1314-1318, 1990.

[476] J. Grégoire, and P. Théroux, "Detection and assessment of unstable angina using myocardial perfusion imaging: comparison between technetium-99m sestamibi SPECT and 12-lead electrocardiogram," *The American journal of cardiology,* vol. 66, no. 13, pp. E42-E46, 1990.

[477] D. S. Kayden, S. Sigal, R. Soufer, J. Mattera, B. L. Zaret, and J. T. Frans, "Thallium-201 for assessment of myocardial viability: quantitative comparison of 24-hour redistribution imaging with imaging after reinjection at rest," *Journal of the American College of Cardiology,* vol. 18, no. 6, pp. 1480-1486, 1991.

[478] P. Raggi, M. C. Gongora, A. Gopal, T. Q. Callister, M. Budoff, and L. J. Shaw, "Coronary artery calcium to predict all-cause mortality in elderly men and women," *Journal of the American College of Cardiology,* vol. 52, no. 1, pp. 17-23, 2008.

[479] M. J. Budoff, L. J. Shaw, S. T. Liu, S. R. Weinstein, P. H. Tseng, F. R. Flores, T. Q. Callister, P. Raggi, D. S. Berman, and T. P. Mosler, "Long-term prognosis associated with coronary calcification: observations from a registry of 25,253 patients," *Journal of the American College of Cardiology,* vol. 49, no. 18, pp. 1860-1870, 2007.

[480] P. Greenland, L. LaBree, S. P. Azen, T. M. Doherty, and R. C. Detrano, "Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals," *Jama,* vol. 291, no. 2, pp. 210-215, 2004.

[481] R. Detrano, A. D. Guerci, J. J. Carr, D. E. Bild, G. Burke, A. R. Folsom, K. Liu, S. Shea, M. Szklo, and D. A. Bluemke, "Coronary calcium as a predictor of coronary events in four racial or ethnic groups," *New England Journal of Medicine,* vol. 358, no. 13, pp. 1336-1345, 2008.

[482] A. Schmidt, T. Stefenelli, E. Schuster, and G. Mayer, "Informational contribution of noninvasive screening tests for coronary artery disease in patients on chronic renal replacement therapy," *Am J Kidney Dis,* vol. 37, no. 1, pp. 56-63, Jan, 2001.

[483] P. Greenland, S. C. Smith Jr, and S. M. Grundy, "Improving coronary heart disease risk assessment in asymptomatic people: role of traditional risk factors and noninvasive cardiovascular tests," *Circulation,* vol. 104, no. 15, pp. 1863-1867, 2001.

[484] Y. Arad, K. J. Goodman, M. Roth, D. Newstein, and A. D. Guerci, "Coronary calcification, coronary disease risk factors, C-reactive protein, and atherosclerotic cardiovascular disease events: the St. Francis Heart Study," *J Am Coll Cardiol,* vol. 46, 2005.

[485] R. Haberl, A. Becker, A. Leber, A. Knez, C. Becker, C. Lang, R. Bruning, M. Reiser, and G. Steinbeck, "Correlation of coronary calcification and angiographically documented stenoses in patients with suspected coronary artery disease: results of 1,764 patients," *J Am Coll Cardiol,* vol. 37, 2001.

[486] D. S. Berman, N. D. Wong, H. Gransar, R. Miranda-Peats, J. Dahlbeck, S. W. Hayes, J. D. Friedman, X. Kang, D. Polk, R. Hachamovitch, L. Shaw, and A. Rozanski, "Relationship between stress-induced myocardial ischemia and atherosclerosis measured by coronary calcium tomography," *Journal of the American College of Cardiology,* vol. 44, no. 4, pp. 923-930, 2004/08/18/, 2004.

[487] Z.-X. He, T. D. Hedrick, C. M. Pratt, M. S. Verani, V. Aquino, R. Roberts, and J. J. Mahmarian, "Severity of Coronary Artery Calcification by Electron Beam Computed Tomography Predicts Silent Myocardial Ischemia," *Circulation,* vol. 101, no. 3, pp. 244-251, 2000.

[488] T. Schepis, O. Gaemperli, P. Koepfli, M. Namdar, I. Valenta, H. Scheffel, S. Leschka, L. Husmann, F. R. Eberli, T. F. Luscher, H. Alkadhi, and P. A. Kaufmann, "Added Value of Coronary Artery Calcium Score as an Adjunct to Gated SPECT for the Evaluation of Coronary Artery Disease in an Intermediate-Risk Population," *Journal of Nuclear Medicine,* vol. 48, no. 9, pp. 1424-1430, September 1, 2007, 2007.

[489] K. Nasir, and M. Clouse, "Role of nonenhanced multidetector CT coronary artery calcium testing in asymptomatic and symptomatic individuals," *Radiology,* vol. 264, no. 3, pp. 637-49, Sep, 2012.

[490] A. S. Agatston, W. R. Janowitz, F. J. Hildner, N. R. Zusmer, M. Viamonte, and R. Detrano, "Quantification of coronary artery calcium using ultrafast computed tomography," *Journal of the American College of Cardiology,* vol. 15, no. 4, pp. 827-832, 1990.

[491] Y. Benjamini, and Y. Hochberg, “Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing,” *Journal of the Royal Statistical Society Series B-Methodological,* vol. 57, no. 1, pp. 289-300, 1995.

[492] L. A. Kaminsky, and M. H. Whaley, “Evaluation of a new standardized ramp protocol: the BSU/Bruce Ramp protocol,” *Journal of Cardiopulmonary Rehabilitation and Prevention,* vol. 18, no. 6, pp. 438-444, 1998.

[493] M. Pfisterer, A. Battler, and B. Zaret, “Range of normal values for left and right ventricular ejection fraction at rest and during exercise assessed by radionuclide angiocardiography,” *European heart journal,* vol. 6, no. 8, pp. 647-655, 1985.

[494] D. H. MacIver, and M. Townsend, “A novel mechanism of heart failure with normal ejection fraction,” *Heart,* vol. 94, no. 4, pp. 446-449, 2008.

[495] S. S. Yip, J. Kim, T. P. Coroller, C. Parmar, E. R. Velazquez, E. Huynh, R. H. Mak, and H. J. Aerts, “Associations Between Somatic Mutations and Metabolic Imaging Phenotypes in Non–Small Cell Lung Cancer,” *Journal of Nuclear Medicine,* vol. 58, no. 4, pp. 569-576, 2017.

[496] F. Tixier, M. Hatt, C. C. Le Rest, A. Le Pogam, L. Corcos, and D. Visvikis, “Reproducibility of Tumor Uptake Heterogeneity Characterization Through Textural Feature Analysis in F-18-FDG PET,” *Journal of Nuclear Medicine,* vol. 53, no. 5, pp. 693-700, May, 2012.

[497] D. R. Anderson, *Model based inference in the life sciences: a primer on evidence*: Springer Science & Business Media, 2007.

[498] R. A. Fisher, "Statistical methods for research workers," *Breakthroughs in statistics*, pp. 66-70: Springer, 1992.

[499] J. Betancur, F. Commandeur, M. Motlagh, T. Sharir, A. J. Einstein, S. Bokhari, M. B. Fish, T. D. Ruddy, P. Kaufmann, A. J. Sinusas, E. J. Miller, T. M. Bateman, S. Dorbala, M. Di Carli, G. Germano, Y. Otaki, B. K. Tamarappoo, D. Dey, D. S. Berman, and P. J. Slomka, “Deep Learning for Prediction

of Obstructive Disease From Fast Myocardial Perfusion SPECT," *A Multicenter Study,* vol. 11, no. 11, pp. 1654-1663, 2018.

[500] J. Betancur, L.-H. Hu, F. Commandeur, T. Sharir, A. Einstein, S. Bokhari, F. Mathews, T. Ruddy, P. Kaufmann, A. Sinusas, E. Miller, T. Bateman, S. Dorbala, M. DiCarli, G. Germano, Y. Otaki, B. Tamarappoo, D. Dey, D. Berman, and P. Slomka, "Automatic Deep Learning Analysis of Upright-Supine High-Speed SPECT Myocardial Perfusion Imaging for Prediction of Obstructive Coronary Artery Disease: A Multicenter Study," *Journal of Nuclear Medicine,* vol. 59, no. supplement 1, pp. 507, May 1, 2018, 2018.

[501] J. A. Betancur, L.-H. Hu, F. Commandeur, T. Sharir, A. J. Einstein, M. B. Fish, T. D. Ruddy, P. Kaufmann, A. J. Sinusas, E. J. Miller, T. M. Bateman, S. Dorbala, M. Di Carli, G. Germano, Y. Otaki, J. X. Liang, B. K. Tamarappoo, D. Dey, D. S. Berman, and P. J. Slomka, "Deep Learning Analysis of Upright-Supine High-Efficiency SPECT Myocardial Perfusion Imaging for Prediction of Obstructive Coronary Artery Disease: A Multicenter Study," *Journal of Nuclear Medicine*, September 27, 2018, 2018.

[502] B. Tsui, E. C. Frey, X. Zhao, D. Lalush, R. Johnston, and W. McCartney, "The importance and implementation of accurate 3D compensation methods for quantitative SPECT," *Physics in medicine & biology,* vol. 39, no. 3, pp. 509, 1994.

[503] K. Willowson, D. L. Bailey, and C. Baldock, "Quantitative SPECT reconstruction using CT-derived corrections," *Physics in Medicine & Biology,* vol. 53, no. 12, pp. 3099, 2008.

[504] D. L. Bailey, and K. P. Willowson, "An evidence-based review of quantitative SPECT imaging and potential clinical applications," *Journal of nuclear medicine,* vol. 54, no. 1, pp. 83-89, 2013.

[505] N. J. Purser, I. S. Armstrong, H. A. Williams, C. M. Tonge, and R. S. Lawson, "Apical thinning: real or artefact?," *Nucl Med Commun,* vol. 29, no. 4, pp. 382-9, Apr, 2008.

[506] E. G. DePuey, and A. Rozanski, “Using gated technetium-99m-sestamibi SPECT to characterize fixed myocardial defects as infarct or artifact,” *Journal of Nuclear Medicine,* vol. 36, no. 6, pp. 952-955, 1995.

[507] S. Burrell, and A. MacDonald, “Artifacts and pitfalls in myocardial perfusion imaging,” *Journal of nuclear medicine technology,* vol. 34, no. 4, pp. 193-211, 2006.

Curriculum Vitae

Saeed Ashrafinia was born in Tehran, Islamic Republic of Iran. He received his B.Sc. from Sharif University of Technology, Tehran, Iran, in electrical engineering specialized in control engineering. He then moved to Canada and received his M.Sc. from Simon Fraser University in electrical engineering, with his research focusing on wireless communication and numerical optimization. In 2013, he joined Tomographic Image Reconstruction and Analysis Lab at the division of nuclear medicine, department of radiology at Johns Hopkins University and started his research in PET image reconstruction. He then started his Ph.D. in electrical engineering at Johns Hopkins University and continued his work in the same field. In the second half of his Ph.D. program, he diverted his research towards image processing and radiomics in nuclear medicine imaging, incorporating machine learning and data science approach for enhanced image quantitation and personalized medicine. Through extensive collaboration efforts with both technical and clinical faculty, Saeed has over 40 peer-reviewed publications, 16 of them as the first author (the complete publication record is available at section 1.5 of this dissertation). Saeed has also developed two software packages available to the scientific community, including a Matlab-based PET simulation and reconstruction, as well as a standardized environment for radiomics analysis (SERA). He is a recipient of Bradley-Alavi fellowship from society of nuclear medicine and molecular imaging (SNMMI) in 2018 for his work on cardiac SPECT radiomics, which later was also recognized for SNMMI Young Investigator Award. He also received multiple awards and grants from IEEE and American association of physicists in medicine (AAPM). Saeed is currently continuing his work on data-oriented approaches to nuclear medicine imaging, via utilization of radiomics and deep learning for the aim of enhanced quantitation, improved diagnosis, prognosis, treatment response, and patient survival.