

3-30-2016

# Increasing 18F-FDG PET/CT Capabilities in Radiotherapy for Lung and Esophageal Cancer via Image Feature Analysis

Jasmine Alexandria Oliver  
*University of South Florida*

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [Bioimaging and Biomedical Optics Commons](#), and the [Nuclear Commons](#)

## Scholar Commons Citation

Oliver, Jasmine Alexandria, "Increasing 18F-FDG PET/CT Capabilities in Radiotherapy for Lung and Esophageal Cancer via Image Feature Analysis" (2016). *Graduate Theses and Dissertations*.  
<http://scholarcommons.usf.edu/etd/6123>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact [scholarcommons@usf.edu](mailto:scholarcommons@usf.edu).

Increasing 18F-FDG PET/CT Capabilities in Radiotherapy for Lung and Esophageal  
Cancer via Image Feature Analysis

by

Jasmine A. Oliver

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy in Applied Physics  
with a concentration in Medical Physics  
Department of Physics  
College of Arts and Sciences  
University of South Florida

Co-Major Professor: Geoffrey G. Zhang, Ph.D.  
Co-Major Professor: Garrett Matthews, Ph.D.  
Eduardo G. Moros, Ph.D.  
Sarah Hoffe, M.D.

Date of Approval:  
March 11, 2015

Keywords: Radiomics, Texture Analysis, Imaging, Fiducials

Copyright © 2016, Jasmine A. Oliver

## **Dedication**

I would like to dedicate this work to my parents, Leonard and Belinda Oliver and to my younger sister Jacquelyn Oliver. Thank you for your love, support, and encouragement. You have been my motivators, cheerleaders, and confidants throughout this journey. You taught me to “plan my work and work my plan”. You challenged me to be the best that I could be. You prayed with and for me. You gave all of yourself and laid the foundation for my success. For these reasons, I love you so much. This work is also dedicated to my grandparents, Minnie Bell, the late Herman Bell, the late Ethel Williams, and the late Lewis Williams, Sr. who passed the torch. Their legacy lives on in all of their descendants.

## Acknowledgments

First, I would like to acknowledge my anchor and my rock; my Lord and Savior Jesus Christ. I would like to acknowledge my advisor Dr. Geoffrey Zhang for contributing many hours toward this research and toward my development. I'd especially like to recognize Dr. Zhang's computational expertise and kindness. I truly appreciate your contributions as my advisor. I would also like to acknowledge my mentor and program director, Dr. Eduardo Moros, who has assisted me throughout this journey in more ways than I can count, some of which include his efforts towards the CAMPEP accreditation of the medical physics program at USF and the provision of research funding for two semesters. I would like to acknowledge my co-advisor, Dr. Garrett Matthews for the time he contributed to this work and to my development. Dr. Mikalai Budzevich is acknowledged for his mentorship, guidance, advice, and many hours of effort with this research. Your contributions were invaluable. Thank you for your help, computational expertise, and the laughs and coffee during rough days.

I would like to acknowledge Dr. Kujtim Latifi for his expertise, teaching, research opportunities, and IT skills. I would like to recognize and thank all of my committee members, especially Dr. Sarah Hoffe who took time out of her busy schedule in the clinic and the medical school to help me complete this research as well as write a grant proposal. I'd also like to acknowledge Dr. Hoffe for allowing me to work with her esophageal dataset. Special recognition goes to Dr. Robert Gillies for the time and effort that was spent toward preparing a NIH grant proposal as well as many opportunities to attend and present at Radiomics Retreats and network with other investigators in image feature analysis. Dr. Dylan Hunt is acknowledged for his expertise in understanding the imaging process and image noise as well as the "Stump the Chump" sessions. Dr. Puja Venkat is recognized for her collaboration, insight and assistance with the esophageal patient dataset and contouring. The late Mrs. Kuykendall is acknowledged

as a team member for time and expertise devoted to this work. She was the lead PET technologist during the start of this study. Dr. Jongphil Kim is acknowledged for his contributions in the statistical analysis in Chapter 4 of this work.

I want to thank my parents, who have been my rock, my motivators, and my confidants. I love you and I am so thankful that God blessed me with you. To my sister Jacquelyn, I love you more than you know. To my Allen Temple AME Church family, thank you for your support and your words of encouragement. Thank you to Mrs. Brenda Jackson-Lee for ensuring that I was always taken care of. Special thanks to Mr. Bernard Batson who has contributed significantly to my professional development and helped fund my journey to the Ph.D. He has been a true mentor and played a pivotal role in this Ph.D. process. I'd like to acknowledge my sponsors during the journey to the Ph.D. including the National Science Foundation (NSF) Florida Georgia Alliance for Minority Participation (FGLSAMP) Bridge to the Doctorate Fellowship (HRD# 1108950), Mr. Charles Jackson and the Florida Education Fund McKnight Doctoral Fellowship, a USF Foundation (through Moffitt Cancer Center) sponsored Graduate Research Assistantship and the University of South Florida Department of Physics sponsored Teaching Assistantship. I'd also like to acknowledge MARS, Incorporated for funding my trip to the 64<sup>th</sup> Annual Lindau Meeting of the Nobel Laureates in Lindau, Germany and Dr. Garrett Matthews for his recommendation for this illustrious trip. Lastly, I would like to acknowledge the Bridge to the Doctorate (BD) cohort for their inspiration, friendship, and encouragement.

## Table of Contents

List of Tables .....	iv
List of Figures .....	vi
Abstract.....	viii
Chapter One: Introduction .....	1
1.1 Motivation .....	1
1.2 Imaging Modalities.....	2
1.2.1 Positron Emission Tomography.....	2
1.2.2 Computed Tomography .....	6
1.3 Motion .....	7
1.4 Image Noise .....	8
1.5 Image Feature Analysis .....	9
1.6 ROI Delineation & Tumor Segmentation .....	11
1.7 Objectives.....	11
1.8 Dissertation Structure .....	12
Chapter Two: Background .....	13
2.1 Image Feature Analysis in PET/CT .....	13
2.2 Introduction of the Problem.....	13
2.3 Texture Analysis, Image Feature Analysis, or Radiomics? .....	16
2.4 Clinical Application, Uses, and Relevance of Image Feature Analysis .....	17
2.5 Image Feature Analysis in PET .....	17
2.5.1 Using Heterogeneity as an Indicator in 18F-FDG PET .....	18
2.5.2 Tumor Delineation and Volumes .....	19
2.5.3 Image Feature Studies Assessing Features as Predictive or Prognostic Indicators.....	21
2.5.4 Test-Retest .....	21
2.5.5 Limitations and New Approaches .....	22
2.6 Image Feature Analysis in CT.....	24
2.7 The Future of Image Feature Analysis.....	26
Chapter Three: Materials & Methods.....	28
3.1 ROI Delineation and Tumor Segmentation .....	28
3.2 Feature Extraction .....	30
3.3 Features .....	32
3.3.1 Shape Descriptors and First Order Features .....	32
3.3.2 GLCM Features .....	34
3.3.3 Run Length Matrix Features.....	38
3.3.4 Grey Level Size Zone Matrix Features .....	39
Chapter Four: Variability of Image Features Computed from Conventional and Respiratory-Gated PET/CT Images of Lung Cancer .....	43
4.1 Introduction.....	43
4.2 Materials and Methods .....	44

4.2.1 Preliminaries .....	44
4.2.2 ROI Delineation and Tumor Segmentation .....	45
4.2.3 Feature Extraction .....	46
4.2.4 RG (4D) PET Phase Analysis .....	48
4.2.5 Long Axis Calculation and Rotation Analysis.....	49
4.3 Results .....	49
4.3.1 3D and RG PET/CT Image Feature Analysis .....	49
4.3.2 RG (4D) PET Phase Analysis .....	54
4.3.3 Overall Feature Results.....	54
4.3.4 Long Axis Tumor Length, Rotation, and Center of Mass Motion (CMM) .....	55
4.4 Discussion .....	56
4.4.1 Limitations.....	60
4.5 Conclusions.....	60
Chapter Five: Sensitivity of Image Features to Noise in Conventional and Respiratory Gated PET/CT Images of Lung Cancer: Uncorrelated Noise Effects .....	
5.1 Introduction.....	62
5.2 Materials and Methods .....	63
5.2.1 Phantom Study .....	63
5.2.1.1 Noise Application.....	63
5.2.1.2 Signal-to-Noise Ratio (SNR) and Noise Power Spectrum (NPS).....	64
5.2.2 Patient Study.....	67
5.2.2.1 Feature Extraction .....	68
5.2.2.2 Statistical Analysis on Patient Data.....	68
5.3 Results .....	70
5.3.1 Phantom Study .....	70
5.3.2 Patient Study.....	70
5.4 Discussion .....	76
5.5 Conclusions.....	81
Chapter Six: Fiducials vs. 18F-FDG PET/CT for Esophageal Cancer GTV Delineation for Radiotherapy Treatment Planning .....	
6.1 Introduction.....	82
6.2 Materials and Methods .....	84
6.2.1 Patient Population .....	84
6.2.2 Measurement of MTV.....	84
6.2.3 Measurement of Fiducial Distance .....	85
6.3 Results .....	86
6.3.1 Large Discordances .....	88
6.4 Discussion .....	90
6.5 Conclusion.....	93
Chapter Seven: Conclusions and Future Work.....	
7.1 Research Outcomes.....	94
7.2 Future Work.....	95
List of References .....	97
Appendix A: List of Publications .....	116

Appendix B: Copyright Permissions .....	121
Appendix C: IRB MCC 16784 Determination Letter.....	122
Appendix D: IRB MCC 16567 Determination Letter.....	125
About the Author .....	End Page



## List of Tables

Table 3.1	Extracted Image Features including Shape features, Intensity features, GLCM features, RLM features, and GLSZM features .....	31
Table 3.2	Shape Features .....	32
Table 3.3	Intensity and First-Order Features .....	33
Table 3.4	Additional Intensity and First-Order Features .....	34
Table 3.5	GLCM Features .....	36
Table 3.6	Additional GLCM Features.....	37
Table 3.7	RLM Features.....	40
Table 3.8	Gray-Level Size-Zone Matrix Features .....	41
Table 4.1	Features Presenting Average Differences Between 3D and RG PET Image Features.....	51
Table 4.2	Features Presenting Average Differences Between 3D and RG CT Image Features .....	52
Table 4.3	Percent Differences ( $\%Diff_i^{3D/RG}$ ) between Image Features of 3D and RG, PET and CT Images and Conglomerate Image Features of RG PET phases for all cases ( $\%Diff_{ij}^{Mean}$ ).....	52
Table 4.4	Image Features with Common Average Differences in 3D/RG PET and CT .....	53
Table 4.5	Long Axis Lengths of Lung Tumors on 3D PET Images and RG PET Images at Exhale and Inhale. ....	56
Table 5.1	CT and PET Feature Classifications .....	69
Table 5.2	CCC Strength of Agreement Scale .....	69
Table 5.3	R1 Features ( $\%Diff < 10\%$ ) for CT and PET.....	74
Table 5.4	NR1 Features ( $\%Diff > 100\%$ ) for PET and CT.....	75
Table 6.1	Patient Characteristics .....	84
Table 6.2	Fiducial vs MTV Border .....	89

Table 6.3 FID Pre PET Fiducial vs MTV Border.....	90
Table 6.4 FID Post PET Fiducial vs MTV Border.....	90

## List of Figures

Figure 1.1 Molecular structures of glucose and <sup>18</sup> F-fluorodeoxyglucose .....	3
Figure 1.2 The PET Mechanism .....	4
Figure 1.3 The GE Discover STE PET/CT Scanner at Moffitt Cancer Center .....	5
Figure 1.4 Respiratory-gated imaging .....	7
Figure 2.1 PET coronal image slice, tumor with contour and its grey-level co-occurrence matrix .....	15
Figure 3.1 CT segmentation of one patient viewed in 2D (the ROI extends in 3D).....	30
Figure 3.2 (a) Sample figure; (b) Sample GLCM matrix in the 0° direction based on a 5x5 pixel image with 5 grey-levels. ....	35
Figure 3.3 (a) Sample Image; (b) RLM Matrix calculated based on sample image with 5 grey levels in the 0° direction .....	39
Figure 3.4 (a) Sample Image; (b) Sample GLSZM Matrix based on sample image with 5 grey levels. ....	41
Figure 4.1 Method of tumor rotation calculation. ....	50
Figure 4.2 Average differences between 3D and Respiratory-Gated Image Features.....	51
Figure 4.3 Concordance correlation coefficients for each feature with mean and standard deviation for each feature subtype for (A) 3D/4D CT and (B) 3D/4D PET .....	53
Figure 4.4 Feature dependency on respiration phase for selected features. ....	55
Figure 5.1 Method of SNR Calculation for (a) CT and (b) PET using Ge-68 solid epoxy phantom. ....	66
Figure 5.2 Method to measure CT noise power spectrum using Ge-68 solid epoxy phantom .....	66
Figure 5.3 One coronal slice of a PET/CT image with and without noise. ....	67
Figure 5.4 Signal to noise ratios for: (A) 3D and 4D CT (one phase) of phantom with motion table and static (3D only). (B) 3D and 4D PET phantom with motion table and static (3D only). ....	72

Figure 5.5 NPS of Ge-68 solid epoxy phantom for GE Discovery STE PET/CT scanner with increasing image noise.....	73
Figure 5.6 Average percent differences between noise and original images across feature sub-types for low, medium, and high noise in (A) 3D CT, (B) 4D CT, (C) 3D PET, and (D) 4D PET .....	75
Figure 5.7 Median CCCs for all CT features .....	76
Figure 5.8 Median CCCs for all PET features .....	77
Figure 6.1 Method of determining MT threshold for each esophageal tumor.....	85
Figure 6.2 Method identifying the fiducial and marking the centroid at the inferior border of the MTV .....	87
Figure 6.3 Inferior discordance and superior discordance histograms. ....	88
Figure 6.4 Case with largest discord (6.87 cm) between superior fiducial and superior MTV border.....	89
Figure 6.5 Example of discordance between fiducials and MTV at the inferior and superior location .....	89

## Abstract

Positron Emission Tomography (PET) is an imaging modality that has become increasingly beneficial in Radiotherapy by improving treatment planning (1). PET reveals tumor volumes that are not well visualized on computed tomography CT or MRI, recognizes metastatic disease, and assesses radiotherapy treatment (1). It also reveals areas of the tumor that are more radiosensitive allowing for dose painting - a non-homogenous dose treatment across the tumor (1). However, PET is not without limitations. The quantitative unit of PET images, the Standardized Uptake Value (SUV), is affected by many factors such as reconstruction algorithm, patient weight, and tracer uptake time (2). In fact, PET is so sensitive that a patient imaged twice in a single day on the same machine and same protocol will produce different SUV values. The objective of this research was to increase the capabilities of PET by exploring other quantitative PET/CT measures for Radiotherapy treatment applications.

The technique of quantitative image feature analysis, nowadays known as radiomics, was applied to PET and CT images. Image features were then extracted from PET/CT images and how the features differed between conventional and respiratory-gated PET/CT images in lung cancer was analyzed. The influence of noise on image features was analyzed by applying uncorrelated, Gaussian noise to PET/CT images and measuring how significantly noise affected features. Quantitative PET/CT measures outside of image feature analysis were also investigated. The correlation of esophageal metabolic tumor volumes (tumor volume demonstrating high metabolic uptake) and endoscopically implanted fiducial markers was studied.

It was found that certain image features differed greatly between conventional and respiratory-gated PET/CT. The differences were mainly due to the effect of respiratory motion including affine motion, rotational motion and tumor deformation. Also, certain feature groups

were more affected by noise than others. For instance, contour-dependent shape features exhibited the least change with noise. Comparatively, GLSZM features exhibited the greatest change with added noise.

Discordance was discovered between the inferior and superior tumor fiducial markers and metabolic tumor volume (MTV). This demonstrated a need for both fiducial markers and MTV to provide a comprehensive view of a tumor.

These studies called attention to the differences in features caused by factors such as motion, acquisition parameters, and noise, etc. Investigators should be aware of these effects. PET/CT radiomic features are indeed highly affected by noise and motion. For accurate clinical use, these effects must be account by investigators and future clinical users. Further investigation is warranted towards the standardization of PET/CT radiomic feature acquisition and clinical application.

## Chapter One: Introduction<sup>1</sup>

### 1.1 Motivation

Lung cancer has the highest estimated death for all cancers in 2015 – causing more deaths than colorectal, breast, and prostate cancers combined. At 28% for males and 26% for females, lung cancer deaths lead prostate cancer by 18% and breast cancer by 11% (3). Non-small cell lung cancer (NSCLC), the most common type of lung cancer, accounts for 83% of all lung cancers and 70% of cancers present with advanced or metastatic disease at diagnosis (4). Typically, early stage NSCLC is treated surgically (stage 1-2)(4). Locally advanced cancers are typically treated with chemotherapy and concurrent radiation therapy (4). Chemotherapy is generally prescribed for metastatic disease and palliation. Targeted drugs to improve response rates (4).

Esophageal cancer accounts for only 4% of estimated deaths for 2015 but the five year survival rate for esophageal cancer patients is only 18% (3). There are approximately 17,000 estimated new esophageal cancer cases this year alone and almost 16,000 estimated deaths (3). Most esophageal patients present with locally advanced or metastatic disease. To improve outcomes, neoadjuvant therapy (chemo-radiotherapy followed by surgery) has been recommended for patients with locally advanced disease prior to surgery. At present, data on individual sensitivity to chemotherapy and radiation therapy is lacking, thus patients are advised to undergo standard of care chemoradiation based on their clinical, rather than molecular, factors (5). Clinicians rely on radiographic indicators to assess response but, in the absence of progression at restaging, patients proceed to esophageal resection (surgical removal of the tumor); the quality of life implications and medical cost are profound if patients have a

---

<sup>1</sup> Portions of this chapter have been previously published in *Translational Oncology*, 2015, 8(6): 524-534, and have been reproduced with permission from Elsevier. The author of this dissertation is the first author of the previously published work. Additional portions of this chapter have been submitted for publication.

pathologic complete response and yet have undergone removal of their esophagus. Conversely, if patients are found at the time of surgery to have had no response, their outcomes are no better than if they went directly to surgery upfront.

For all cancers, diagnostic procedures (such as biopsy) are costly and invasive, and can be accompanied by severe side effects (6). Image feature analysis is a non-invasive, cost-effective technique that extracts tumor identifiers that could lead to more individualized treatment planning by providing predictive and prognostic biomarkers for cancer patients (7). It is routine in lung and esophageal cancer for patients to undergo diagnostic PET/CT imaging prior to initial treatment. In esophageal cancers at our institution (H. Lee Moffitt Cancer Center & Research Institute) it is also routine for patients to receive endoscopically placed fiducial markers prior to radiotherapy treatment. These fiducial markers delineate the visible tumor burden. PET delineates metabolic tumor burden and metabolic metastasis.

New methodologies and techniques to extract quantifiable information from current technologies (such as image feature analysis) may increase personalized treatment planning and reduce toxicity by guiding therapy and monitoring disease progression and treatment response in lung and esophageal cancers.

## **1.2 Imaging Modalities**

### **1.2.1 Positron Emission Tomography**

Positron Emission Tomography (PET) is a molecular imaging modality that provides information about the functional biological processes of the body. In radiation oncology, PET is used in: 1) diagnosis and staging; 2) target definition; and 3) treatment response assessment (8). In a study by the National Oncological PET Registry, FDG PET was shown to affect patient management in 36.5% of cases in a study of 22,975 (8).

In cancer imaging, PET can be used to provide information such as tumor heterogeneity, proliferation, hypoxia, vascular perfusion and angiogenesis, and apoptosis and cellular



signaling, depending on the positron emitter used (8). The most common radioactive tracer in PET imaging for radiotherapy treatment,  $^{18}\text{F}$ -fluorodeoxyglucose (FDG), is used to highlight areas with high metabolic activity. Tumors are typically characterized by increased FDG uptake due to increased glycolysis and glucose transport compared to normal tissues (8). Fluorine-18 is a radioactive isotope that decays via positron emission (97%) and electron capture (3%). This isotope is substituted for a hydroxyl (OH) group on a glucose molecule at the 2' position to form  $^{18}\text{F}$ -FDG (Figure 1.1).

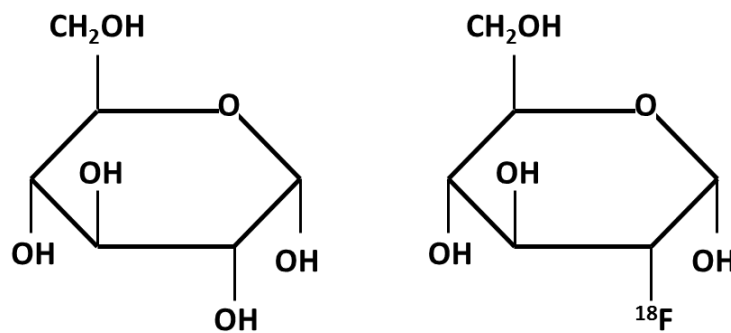


Figure 1.1 Molecular structures of glucose and  $^{18}\text{F}$ -fluorodeoxyglucose.

The radioactive tracer accumulates in areas of increased glycolysis and glucose transport (8) – “hot” areas usually include typically tumors, inflammation, brain, and bladder. After  $^{18}\text{F}$  decays via positron emission, the positrons lose their kinetic energy while travelling a distance of up to 2.4 mm (Maximal positron range in water. The maximal positron energy is 634 keV (9). Once the kinetic energy is spent, the positron annihilates with an electron producing two 511 keV gamma rays that travel nearly opposite from each other. Ideally, the rays would travel at a perfect  $180^\circ$  angle from each other (Figure 1.2). However, these gamma rays are not always collinear due to scatter. The minimum possible resolution of clinical PET scanners is 1.83 mm full width at half maximum (FWHM) due to the non-collinearity of the gamma rays, the distance the positron travels before annihilation (10), detector size, and other factors(11). Pre-clinical PET scanners can have a minimum resolution of 0.67 mm FWHM. At Moffitt Cancer

Center the minimum PET resolution is 4 mm. The PET/CT scanner at Moffitt Cancer Center is shown in Figure 1.3 (see page 5).

A ring of detectors surrounding the patient detects the gamma rays and creates a line in space (line of response) connecting the two interactions. PET detectors are composed of scintillation crystals (2-3 cm depth) coupled with photomultiplier tubes (9). Scintillation crystals (commonly bismuth germinate -  $\text{Bi}_4\text{Ge}_3\text{O}_{12}$  (BGO) or cerium-doped leutium –  $\text{Lu}_2\text{SiO}_5$  (LSO)) when struck by incoming gamma rays (511 keV) absorb the energy and scintillate.

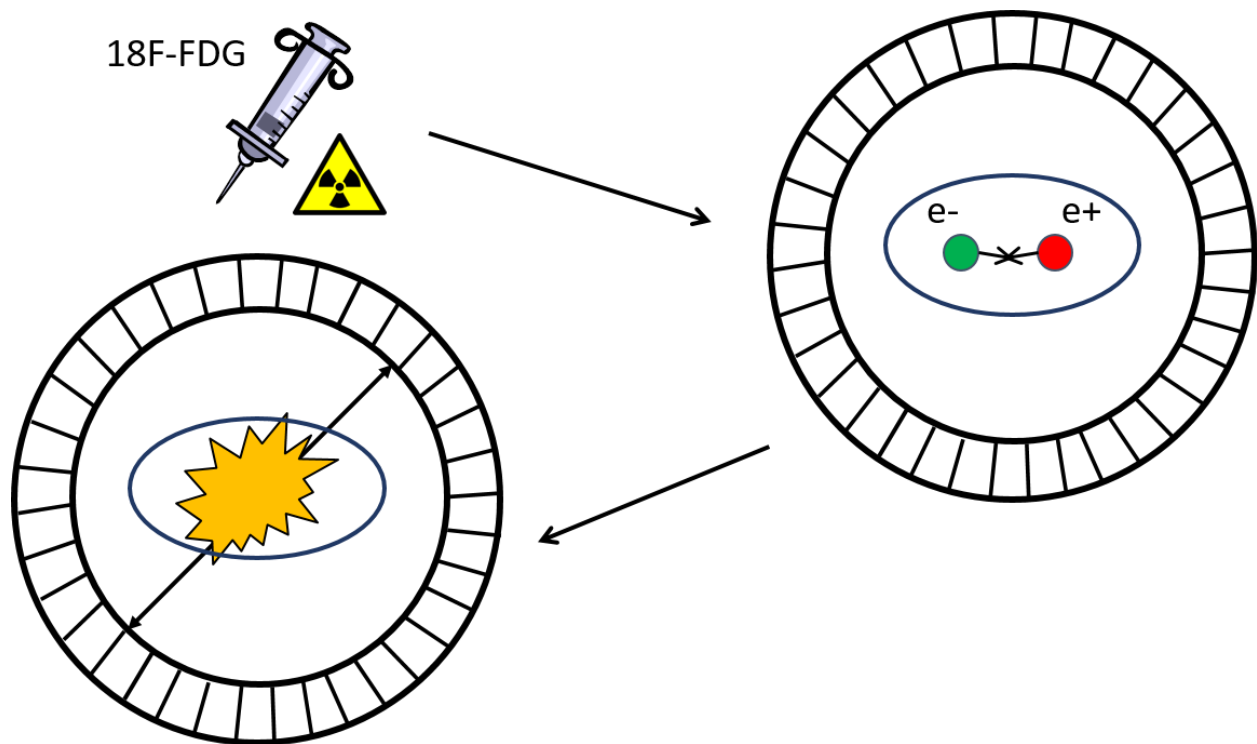


Figure 1.2 The PET Mechanism. The patient is injected with  $^{18}\text{F}$ -FDG. As positrons are emitted, each positron annihilates with an electron creating two 511 keV Gamma rays which are detected by the PET gantry.

Standardized Uptake Value (SUV) is the standard PET measure in image analysis (equation 1.1). However, it is clear that SUV is dependent on many technical as well as physiologic factors (12). It is proportionately dependent upon FDG uptake, which in turn is affected by dose calibration, clock (decay) synchronization, patient weight and blood sugar level, documentation of unused tracer remains, and other set-up specifics (12). Moreover, SUV

indirectly depends on the method of obtaining raw data, radionuclide uptake time, hardware platform, and the applied reconstruction algorithm, lesion size, motion, and user region-of-interest ROI selection (9, 10, 12).

$$SUV = \frac{\text{activity concentration in a voxel}}{\text{administered activity} / \text{patient mass}} \quad (1.1)$$

Studies have shown SUV's predictive ability of therapy response and survival (13, 14), although particularly in NSCLC discrepancy remains regarding whether maximum or minimum SUV is a predictor based on treatment modality (15). The prognostic ability of SUV parameters has also been shown, but conflict emerges when defining the best cut-off value (13). Nonetheless, SUV can be unreliable. A clinical study done on the test-retest reproducibility of SUV demonstrated greater than expected SUV variability within a single institution (12) and 10%-25% SUV variability was detected in a multicenter consortium prior to biological effects or protocol influences (16). Based on these results and the SUV's reliance on various non-standardized factors, it is clear that there is a need for additional indicators that are more robust than SUV or complementary to SUV-based findings. Quantifiable and robust image features may be candidates for such indicators.



Figure 1.3 The GE Discovery STE PET/CT Scanner at Moffitt Cancer Center.

To avoid errors related to radiation attenuation and Compton scattering in the patient tissue, the patient receives a Computed Tomography (CT) scan prior to the PET scan that acts as an attenuation map for PET reconstruction (17).

### 1.2.2 Computed Tomography

CT scanners, though paired with PET scanners are also stand-alone diagnostic imaging devices and the most commonly used imaging modality for radiation therapy treatment planning. The first CT scanner was invented in 1972 by Godfrey Hounsfield of EMI Laboratories, England and Allan Cormack Of Tufts University, Massachusetts. CT scanners were not installed for clinical use until 1974-1976 (9). These scanners are considered transmission devices (PET scanners are emission devices). To produce a CT image, x-rays are transmitted through the body at multiple angles (projections) by rotating an x-ray tube around the body (18). These images are produced on the order of milliseconds and eliminate the need for many invasive diagnostic surgeries (9). Because CT images work by absorbing x-rays in tissue they provide the electron density and are used for attenuation correction in PET images. The electron density is converted into CT numbers. When normalized with values of -1000 for air, 0 for water, and more than 1000 for bone, CT numbers are known as Hounsfield Units (equation 1.2) (19).

$$HU(x, y, z) = 1000 \frac{(\mu(x,y,z) - \mu_w)}{\mu_w} \quad (1.2)$$

where  $\mu(x, y, z)$  is average linear attenuation coefficient for a voxel of tissue at location  $(x,y,z)$  and  $\mu_w$  is the linear attenuation coefficient for water (9). The relationship between CT numbers and electron density is nonlinear because of the different atomic numbers of tissues. This affects whether the Compton or Photoelectric effect is the dominant interaction for beam attenuation at x-ray energies commonly used in CT. For instance, the relationship between lung and soft tissue is linear, but the relationship between soft tissue and bone is nonlinear. Because CT images provide electron densities of tissues, they are used in the dose calculation process for radiation treatment planning. They provide tissue heterogeneity corrections and are used to

delineate the target volume (19). CT is used for tumor localization, radiation dose calculation in radiotherapy planning, and treatment assessment.

### 1.3 Motion

Due to the relatively long scan time of PET images (2-3 minutes per bed position with 1-16 bed positions), motion degrades PET images. It affects both the quantification (SUV) and tumor volume. Because of respiratory motion, there can be an averaging effect in the thoracic region for PET images (20), resulting in an altered image of the tumor. Conventional or three-dimensional (3D) imaging is the most common imaging protocol. During 3D imaging, photon emission loci are convoluted with anatomical motion. Four-dimensional (4D) imaging accounts for tumor motion. In this work, we study both 3D and respiratory-gated 4D images (Figure 1.4).

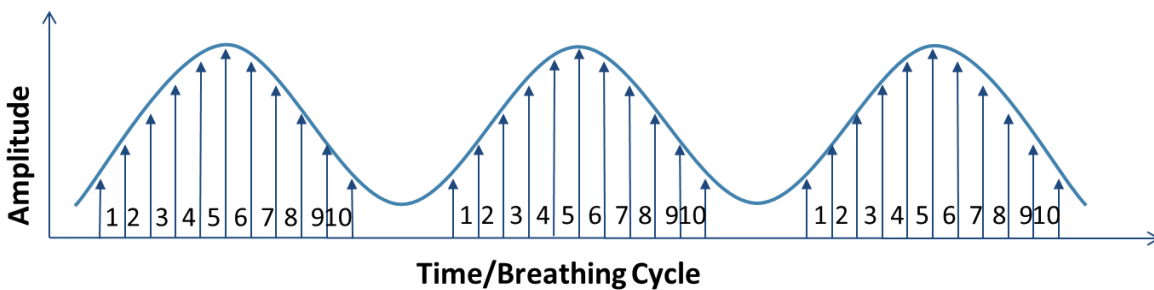


Figure 1.4 Respiratory-gated imaging. Data is binned into phases according to the respiratory cycle. Ten bins are used in our institution.

Figure 1.4 demonstrates the mechanism of respiratory gated imaging. One method to accomplish this uses an infrared reflective marker block placed on a patient's abdomen (an internal motion surrogate). An infrared revolutions-per-minute (RPM) optical monitor tracks the location of the marker and associates this with the position of the abdomen at a specific point of the breathing cycle. Detected events (detected photons) are binned according to the position of the abdomen at that point on the breathing cycle. This results in one image for each bin (10 images at our institution). These 10 images can be played in cine model thereby demonstrating tumor motion and deformation over the breathing cycle.

## 1.4 Image Noise

According to Hasegawa, “Noise is often defined as the uncertainty in a signal due to random fluctuations in that signal.”(21) Image noise is caused by a variety of modality specific factors. PET and CT images possess varying levels of image noise due to the different mechanisms of detection and image reconstruction. One of the most considerable sources of image noise (common to both PET and CT), is random variations in photon counting caused by the statistical nature of x-rays, which is related to the number of photons detected (correlated to signal and image texture). This is commonly referred to as quantum noise. Electronic noise, or “dark noise”, is another common source of image noise. This source of noise is inherent to the detector and independent of the number of photons detected. It is caused by the electronic components that make up the detector.

CT image noise is mainly random, statistical noise caused by the finite number of x-rays in projection measurements (22) and the Poisson nature of x-rays (23). In CT, common sources of noise include body attenuation, detector inefficiency, electronic noise, roundoff errors, artifacts, and structural noise (density variations in the object). On the other hand, PET image noise is caused by the random nature of radioactive decay (23, 24). PET images are affected by partial volume effects, tumor motion, source to background ratio, patient weight, protocol (three-dimensional or respiratory-gated), and signal loss (e.g., induced by respiratory motion) (25). Scattering coincidences, random coincidences, and the random corrections degrade the PET signal-to-noise ratio (SNR). Other contributors of image noise in PET images include: detectors, electronics and recorder systems, reconstruction algorithms, convolution kernels, modes of attenuation correction, and radioactive decay correction (26). Electronic noise is a common source of noise in both PET and CT modalities. This particular source of noise is considered spatial frequency independent (27). It is typically considered uncorrelated noise.

Image noise is an unavoidable component of medical imaging. Smoothing filters can be used to reduce noise, but they cannot eliminate noise completely, and they may also reduce the signal of interest as well.

## **1.5 Image Feature Analysis**

Texture is a global pattern resulting from repetition of local sub patterns (28). Features associated with image texture describe the relationships between the gray-scale intensity of pixels (or voxels in 3D) on a local or global image scale. These features have been used for classification and segmentation purposes, identifying regions of interest in an image and estimating heterogeneity (29).

Image features are quite vast in number and can be subdivided into shape features, first, second, and higher order features. First order features provide information about gray-scale intensities and are derived from intensity distributions and histograms. Whereas second order texture features are derived from gray-tone spatial dependency matrices which are constructed from the intensity value of an image as described below the spatial relationships between voxel intensities (29, 30). In this paper we refer to all of these as image features.

In medical imaging, CT image texture analysis has been studied extensively, dating back to the early 1980s (31). More recent studies in CT image analysis have uncovered feature correspondence with lung tumor aggressiveness and tumor heterogeneity, demonstrated potential as a marker for survival in NSCLC and revealed relationships between features, tumor stage, and metabolism (32-34). The reproducibility and robustness of specific identifiers in NSCLC CT images have also been studied (35).

The application of image feature analysis to PET images has been explored more recently. Prior studies in PET/CT image texture analysis have demonstrated its potential as a predictor of tumor and normal tissue response to therapy; a quantifier of tumor heterogeneity and radiosensitivity; and an indicator for adaptive therapy schemes (13, 36, 37). Conclusive and

beneficial results regarding tumor response to treatment using multi-modality imaging (38) have been drawn, and PET texture analysis is currently being explored for application to predictive models for treatment outcome, partnered with genomics and proteomics patterns (36). Additionally, various features have been investigated using test-retest and inter-observer stability in FDG-PET (39). PET image analysis has also been shown to predict response to radiochemotherapy in esophageal cancer (40) and to quantify tumor heterogeneity as a response predictor (41). Partnered with a multimodality modeling system, PET texture analysis could lead to more individualized treatment planning in lung radiotherapy (38).

Moreover, certain FDG-PET-based texture features have demonstrated association with non-response to chemoradiotherapy for NSCLC tumors (7). These and many other studies are part of a more general systematic approach, namely radiomics, which is an emerging framework relating image features to molecular medicine where large amounts of quantitative features (400+) are extracted for diagnostic, prognostic, and predictive information (36, 42-44). In other words, PET/CT image feature analysis is an emerging and promising quantitative imaging field.

Although radiomics show much promise, there have been several studies showing image features dependency on various factors in the production of images. For example, in a recent study, 45 of 50 texture features showed 10 - 200% variability across acquisition protocols and reconstruction algorithms (45). Therefore several investigators have pointed to the need for standardization in texture analysis (46-48). The usefulness of radiomics depends on the reliability of feature values, so it is important to characterize feature behavior under many potential clinical conditions. Our goal in this study was to explore image feature value variability between respiratory-gated (RG) and conventional (3D) PET images acquired on the same patient during a single PET scan session as well as to explore the dependence of image features on image noise.



## 1.6 ROI Delineation & Tumor Segmentation

There is a wide variety of methods for ROI delineation and tumor segmentation in PET imaging (42, 49). A study by Foster *et al.* reviews the current state-of-the-art PET segmentation methods (49). Common segmentation algorithms in PET imaging include: manual segmentation, threshold based segmentation, gradient or boundary based segmentation, region-based segmentation, stochastic and learning-based segmentation algorithms, and joint segmentation methods (49). There is a wide range of opinions about what method is the golden standard, especially in application to image feature analysis (42) and further research in image segmentation and agreed-upon metrics is merited (49). A report by the American Association of Physicists in Medicine (AAPM) TG211 is being composed which analyzes the pros and cons of the different methods of auto-segmentation in PET.

In this work, we used the fixed threshold method for PET. This method delineates ROIs by a threshold intensity value or percentage of an intensity value. In radiation oncology, commonly used PET threshold values of 40-43% of the  $SUV_{max}$  and 2.5 SUV (49). Although standard thresholding segmentation methods are the easiest to apply, they are not always the most reliable because of the factors that affect SUV. Factors that will degrade tumor delineation are patient motion, inter-observer variability, partial volume effect, and noise. These factors affect any segmentation method because they affect the PET image. However, these factors can highly affect structures delineated via threshold. Regardless, we deemed the thresholding method suitable for our study.

## 1.7 Objectives

The objective of this study was to increase the capabilities and uses of PET/CT in radiotherapy in lung and esophageal cancers using quantitative techniques such as image feature analysis. The aims of this work were to: 1) evaluate the effect that motion and respiratory-gated imaging has on image features; 2) evaluate the effect that uncorrelated noise

has on image features; 3) evaluate the discrepancies between metabolic tumor volume (MTV) and fiducials for esophageal tumors.

## **1.8 Dissertation Structure**

Chapter 2 is a review of image feature analysis in lung and esophageal cancers. Chapter 3 introduces the materials and methods used throughout the study. The method of segmentation and extraction of image features are described in detail. Chapter 4 explains the procedure, results and conclusions of the variability of image features computed from conventional and respiratory-gated PET/CT images of lung cancer. Chapter 5 details uncorrelated noise effects on image features extracted from conventional and respiratory-gated PET/CT images of lung cancer. Chapter 6 details the correlation between metabolic tumor volume and fiducials for locally advanced esophageal cancer. Chapter 7 is a summary of the work detailed in the dissertation and discussion of possible future work.

## Chapter Two: Background<sup>2</sup>

### 2.1 Image Feature Analysis in PET/CT

Clinical imaging by Positron Emission Tomography (PET) and Computed Tomography (CT) is evolving into a quantitative discipline where a large number of metrics are computed in the intensity and gray-level matrix domains; this discipline has been termed Radiomics (42, 50). The purpose of this dissertation is to review previous studies in image feature analysis and radiomics applied to Radiotherapy and explore the clinical applications and impact that image feature analysis can have on patient treatment individualization by elucidating the optimal treatment method prior to treatment via predictive or prognostic indicators. Radiomics of CT & PET images have shown promise as a diagnostic, prognostic, and predictive tool in cancer treatment (42, 51-54). It is also being combined with other “omics” (e.g., genomics, transcriptomics, proteomics, metabolomics) into decision support systems (55).

### 2.2 Introduction of the Problem

Lung cancer is the leading cause of death in cancer patients with estimated death rates of 27% for males and 26% for females in 2016 (56). This rate is higher than both prostate and breast cancer rates. According to the American Cancer Society, lung and esophageal cancers have a 5-year survival rate of 17% and 18%, respectively. These are the lowest reported cancer survival rates (liver and lung are tied for the lowest rate of 17%) (56). It is routine practice for lung and esophageal cancer patients to receive diagnostic PET/CT images prior to treatment (57). Image feature analysis offers a non-invasive, cost effective method to extract additional quantitative data from medical images.

---

<sup>2</sup> Portions of this chapter have been submitted for publication.

At 18%, esophageal cancer has one of the lowest five-year survival rates of all cancers and is comparable to that of lung cancer. Recent randomized data supports neoadjuvant chemoradiation followed by resection for locally advanced disease (58). There are, however, subsets of patients who are cured with definitive chemoradiation without surgery who thus successfully undergo organ preservation (59). Clinical validation of sensitivity analyses to chemotherapy and radiation therapy is ongoing to try to personalize an individual tumor's response (5). Standard of care treatment at the present time does not reliably define how to predict a patient's likelihood of a complete response to chemoradiation, thus the recommendation to proceed to surgical resection since even a negative PET/CT (Positron Emission Tomography/Computerized Tomography) at 6 weeks post treatment can still harbor residual microscopic disease (60). Accordingly, it has been hypothesized (54) that there are quantitative imaging methods to determine what treatment regimen will be most suitable for a particular tumor. The future benefit of such a strategy would be to avoid esophagectomy in those patients validated to have a clinical complete pathologic response and avoid futile chemoradiation neoadjuvantly in those patients who do not have sensitive tumors.

A study done by Burton et al. (1998) showed that 44% of all malignant tumors diagnosed at autopsy were undiagnosed or misdiagnosed. Of these, the top mis- and undiagnosed tumors were from the respiratory tract (33%) and the gastrointestinal tract (23%) (61). The overall aim of image feature analysis is to increase the quantitative information about a tumor towards more personalized treatment. Many investigators try to accomplish this aim by correlating features with outcomes to be used as predictive or prognostic indicators for more individualized patient treatment. If predictive or prognostic indicators about a patient's tumor are available prior to treatment, then the optimal treatment regimen can be selected prior to treatment. This saves a patient time, money, gives a better chance of survival, and may spare the patient unnecessary treatment. For instance, in locally advanced esophageal cancer it is standard-of-care for

patients to undergo neoadjuvant chemoradiation therapy followed by surgery. If the oncologist knows prior to treatment that a tumor will not respond to chemoradiation they can recommend surgery upfront. This saves a patient the pain of undergoing chemoradiation and allowing the tumor to grow stronger over that period of time while giving the patient their best chance of survival. Investigators are a long way away from this goal, but a lot of research has been done in this area in recent years (62).

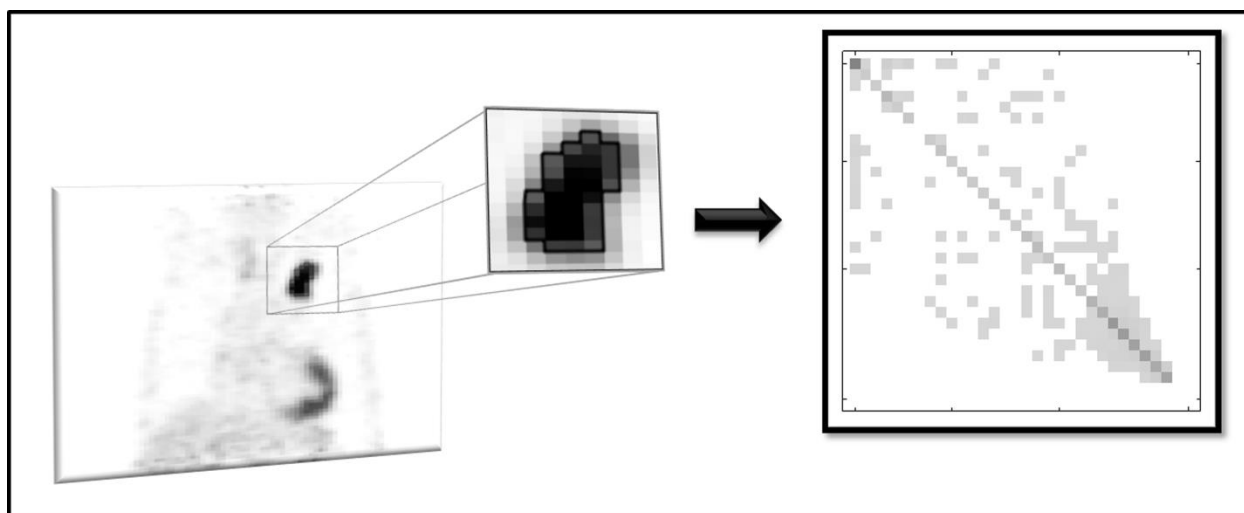


Figure 2.1. PET coronal image slice, tumor ROI with contour and its grey-level co-occurrence matrix (GLCM). This is a 2D example for illustration purposes. The GLCM matrix was calculated with Wolfram Mathematica 9.0 for this illustration (Wolfram Research, Inc., Mathematica, Version 9.0, Champaign, IL (2012)).

The technique of image feature analysis extracts indicators (features) that provide information about the heterogeneity, shape, and intensity of a segmented tumor region that is not traditionally extracted from medical images. These features may also be extracted from matrices generated from PET/CT images and are known as texture features. Images are composed of pixels and texture features analyze pixel pairs and pixel areas such as in figure 2.1 (29, 63-65). An example of a matrix is the grey-level co-occurrence matrix (GLCM) where the number of times 2 pixels with intensity  $i$  and  $j$  occur adjacent to each other in a particular direction (29). Other matrices and techniques include: Laws features, Minkowski functionals,

wavelets, Lapacian transforms and fractal dimensions. These techniques are usually used in the discipline of radiomics (66).

### **2.3 Texture Analysis, Image Feature Analysis, or Radiomics?**

Although commonly used interchangeably, there is a distinction between texture analysis, image feature analysis and radiomics. Texture analysis refers to the extraction of features from an image that provides information on the relationship and patterns between two or more pixels. In application to tumor images, for example, texture analysis reveals information about tumor heterogeneity. Texture as defined by Zucker et. al (1980), is a global pattern resulting from repetition of local subpatterns (28). These subpatterns are directly related to the intensity values stored in the pixels of the image. Texture analysis was originally created in 1973 by Haralick *et al.* with the invention of the Grey-Level Co-occurrence Matrix (GLCM) (29). Image texture describes relationships between pixels (or voxels) on a local or global image scale and has been used for classification and segmentation purposes (29, 63), identifying regions of interest in an image and quantifying heterogeneity. Every image is composed of a different orientation of pixel data and therefore has a different "texture".

Image feature analysis is a broad concept that includes texture analysis, but also includes shape descriptors and features derived from intensity histograms. Radiomics is an organized discipline of image feature analysis that extracts large numbers statistically ( $p < 0.05$ ) of mineable quantitative image features from CT, PET, and magnetic resonance imaging (MRI) images for decision support. The goal of radiomics is to "convert images into mineable data" – thus shifting clinical image analysis from a qualitative to a quantitative field (36, 50). Gillies *et al.* provided a very detailed description of the current state of radiomics, the challenges and future of radiomics. Throughout the remainder of this review, features will be spoken of collectively as image features unless they were specifically named texture features (textural features) or radiomic features in their respective studies.

## 2.4 Clinical Application, Uses, and Relevance of Image Feature Analysis

Image feature analysis of positron emission tomography (PET) and computed tomography (CT) images has the potential to provide identifiers that may improve treatment individualization in oncology. It has been studied for application in many aspects of oncology including chemotherapy, pathology, & radiotherapy. Many studies have demonstrated that image feature analysis, also known as texture analysis, has predictive or prognostic indicators (67). This work aims to review various studies which have tested the predictability, repeatability, and limitations of image feature analysis for lung and esophageal cancer in Radiotherapy applications. Radiomics as predictive or prognostic biomarkers (66) will also be reviewed.

## 2.5 Image Feature Analysis in PET

Currently, PET scanners in combination with CT is used for diagnosing patients, determining cancer stage, and identifying areas of high metabolic activity including the primary tumor location (68). The Standardized Uptake Value (SUV), which is used for simple semi-quantitative analysis, depends on a variety of factors, such as the selected ROI, size of the tumor, sugar level of the patient, etc (69). Radiation oncologists rely heavily on the SUV value when analyzing PET for staging. However, studies have shown that SUV values differ across reconstruction algorithms (45) and from day to day (2).

PET in radiotherapy has improved NSCLC diagnosis of “definite lesions” by 41% (70) and is also used for dose-painting and tissue sparing. Investigators have applied image feature analysis in PET to detect heterogeneity indicators in relationship with predictive or prognostic biomarkers in a number of cancers including non-small cell lung cancer (NSCLC), esophagus, and cervix. This review focuses mainly on NSCLC and esophageal cancer.

Before the first study on image feature analysis in PET imaging by El Naqa *et al.* in 2009 (14), clinical image feature analysis was generally applied to CT, MRI and ultrasound images (71). Since 2009 there has been a burst of interest and investigation in image feature analysis

applied to PET images. Many studies have identified features that can be used as predictive or prognostic identifiers for treatment assessment.

### **2.5.1 Using Heterogeneity as an Indicator in 18F-FDG PET**

Many investigators are interested in the application of image feature analysis to 18F-FDG PET imaging because it provides a blueprint for the heterogeneity of the metabolic activity of a tumor. Because of the intrinsic nature of PET, texture pattern within PET tumor images represents metabolic heterogeneity and biological activity. Since tumors possess characteristic phenotypes and genotypes, it is believed that tumor heterogeneity provides information about the radioresistance and aggressiveness of a tumor. Ganeshan *et al.* (72, 73) demonstrated that tumor heterogeneity can identify adverse biological features in CT images. However, Cook *et al.* (7) resolved that in 18F-FDG PET, these biological correlations are still, to a large extent, unknown.

El Naqa *et al.*'s study (14), the first published study on the application of image feature analysis to PET images, was performed on cervix and head and neck cancers. Although this study was preliminary, it was the first to suggest that image feature analysis of PET images could provide better tools and power for use in clinical prognosis. This study also introduced a new group of features, the intensity-volume histogram metrics. The purpose of the study was to extract prognostic factors that would be more reliable and informative than SUV statistics towards the prediction of treatment outcomes. A later study from the same group, by Vaidya *et al.*, was performed with the purpose of predicting local failure in radiotherapy for NSCLC patients (38). A total of 32 variables for 27 patients were studied. They found that multimodality feature modeling provided better performance than single modality features. This finding supports the current shift towards radiomics studies. These studies will be discussed later in this report.



In a study of pre- and post-treatment 18F-FDG PET on patients diagnosed with NSCLC, small cell lung cancer (SCLC), diffuse large B-cell lymphoma, metastatic papillary renal cell carcinoma, cerebellar hemangioblastoma, neurofibroma, lymphomatoid granulomatosis, lung neoplasm, and other cancer types, textural features in combination with SUV measurements improved the prediction accuracy of morphological changes significantly (74). In this study, investigators created a framework where tumor regions on PET were automatically segmented. They found that entropy, max probability, contrast, and SRE were the most valuable features (74).

A study by Cook *et al.* investigated the predictive and prognostic abilities of 4 features from the neighborhood gray-tone difference matrices (NGTDM) for NSCLC patients undergoing concomitant radiochemotherapy (7). Their features were able to differentiate responders from nonresponders (defined by CT Response Evaluation Criteria in Solid Tumors (RECIST) (75)). The textural features performed better than the standard quantitative indicators: metabolic tumor volume (MTV), total lesion glycolysis (TLG), Standardized Uptake Value (SUV) max, SUV mean or peak SUV. Overall, Cook *et al.* found that the abnormal texture in 18F-FDG PET for NSCLC defined by the features busyness, contrast, coarseness, and complexity from the NGTDM was associated with nonresponse to chemoradiotherapy by RECIST guidelines and associated with poorer prognosis (7).

Another study on 18F-FDG PET image feature analysis assessed the relationship between textural features and  $SUV_{max}$  with histological tumor grade, tumor location and stage (AJCC). Significant correlations were found between energy and entropy (features derived from the grey-level co-occurrence matrix, GLCM) and SUV and tumor stage (76).

## 2.5.2 Tumor Delineation and Volumes

There have also been discrepancies about which tumor delineation method to use for feature analysis. Currently used segmentation methods include threshold methods, model-

based methods, and methods that account for background are currently in use. Hatt *et al.* studied the robustness of image features across tumor delineation methods and the partial volume effect (67). They also assessed the predictive value of features with different delineation methods. They found that the features intensity variability and size zone variability, regional descriptors from the grey-level size-zone matrix (GLSZM) were the most sensitive to tumor delineation method and partial volume effect (PVE). Those features that were the most robust were entropy and homogeneity as well as other local parameters. Zone percentage (ZP) and high intensity emphasis (HIE) were also robust. There was only a small change in their absolute or predictive values when classifying patient response to therapy (67).

Another study by Hatt *et al.* on multi-cancer sites recommended that the minimum volumes for use in image feature analysis were 10 cc (77). They also investigated complementary prognostic values between volume and heterogeneity as well as investigating the influence of quantization preprocessing and the method of texture feature calculation on features. They found that most features that have been shown to quantify tumor heterogeneity are correlated with tumor volume and that prognostic values increased as tumor volumes increased.

Brooks and Grigsby showed that tumor volume size affects metabolic heterogeneity studies. They proposed that studies of intratumoral uptake heterogeneity metrics should be applied to tumors above 45 cc to provide unbiased results (78).

Orlhac *et al.* investigated the relationship between texture indices, SUV, metabolic volume, and TLG. They found that the minimum resampling of 32 gray levels should be used for texture analysis and that many texture features are highly correlated with MTV (79).

### 2.5.3 Image Feature Studies Assessing Features as Predictive or Prognostic Indicators

Many studies have demonstrated the potential for image features as predictive or prognostic indicators. Chicklore *et al.* produced an excellent review detailing the studies through 2013 that had been published towards image feature analysis as an indicator for tumor heterogeneity (13). Tan *et al.* assessed FDG PET features for predicting pathologic response to neoadjuvant chemoradiotherapy in esophageal cancer (80). They found that these features were useful predictors of pathological response to chemoradiotherapy. These features were found to have the same or higher AUC values than  $SUV_{max}$  and  $SUV_{peak}$  (traditional measures). Those features were inertia, correlation, and cluster prominence,  $SUV_{mean}$  decline, and skewness pre-CRT (80).

Tixier *et al.* (40) also studied whether textural features predicted response to chemoradiotherapy. However, in this case, response was measured using pre and post treatment CT scans according to the Response Evaluation Criteria in Solid Tumors (RECIST) rather than the American Joint Committee on Cancer *Cancer Staging Manual* (6th edition) pathology as in (80). In this study, features were extracted on a global scale as well as local and regional levels. The results from this study demonstrated that features were able to differentiate non-responders from responders. Local homogeneity and entropy were the best identifiers from the co-occurrence matrix. The best features in predicting response to treatment were intensity variability and size zone variability, regional descriptors. Not only were these features able to differentiate non-responders from responders, they were able to distinguish partial responders from complete responders.

### 2.5.4 Test-Retest

Although studies have demonstrated that certain features can be used as indicators, it is important to know if these features are reproducible. There have been a number of test-retest

studies on image features derived from PET images. These studies are very important to distinguish those features that can be used clinically with confidence. Leijenaar *et al.* tested the stability of PET radiomics features in NSCLC. The majority of features, 71%, demonstrated high test-retest stability. Ninety-one percent of features had a high inter-observer stability, as indicated by the intra-class correlation coefficient. Typically, features with high test-retest stability also demonstrated high inter-observer stability (39).

A following article by Tixier *et al.* tested the reproducibility of heterogeneity characterization via image feature analysis in FDG PET images (41). They tested features extracted from double baseline FDG PET scans and found that some features had better reproducibility than traditional factors such as mean and maximum SUV. Interestingly, regional parameters that characterized larger homogenous areas of high intensity may predict response to therapy because of high uptake is usually associated with aggressiveness. They suggest that investigators carefully select their features to provide both a complete and reproducible characterization of the spatial heterogeneity of tumor FDG uptake (41).

Willaime *et al.* (37) journeyed away from the traditional PET radiotracer to 18F-fluorothymidine PET (FLT PET), a cell proliferation radiotracer. They assessed if features extracted from breast cancer PET images could be used to characterize intra-tumor heterogeneity of in vivo cell proliferation. They also assessed if these features could be used to predict response to chemotherapy in breast cancer patients. Eight texture features exhibited variability below 30% and seven of those had reasonably small inter-patient variability. Among these eight features were entropy (first order), entropy (GLCM), contrast, dissimilarity, homogeneity.

### **2.5.5 Limitations and New Approaches**

Despite the optimistic results of many previously described studies, there are many sobering studies which investigate the limitations of image feature analysis in PET. Galavis *et*

*al.* demonstrated that image features are affected by modes of acquisition and reconstruction parameters (45). After testing 50 features, 45 of 50 demonstrated variability ranging from 10% - 200 % between acquisition modes and reconstruction parameters. In a more recent study, Yan *et al.* found that image features have different sensitivities to reconstruction settings (81). They also found that grid size, iteration number and FWHM had an impact on image features with grid size have the largest impact. An interesting investigation by Chalkidou *et al.* demonstrated that PET/CT image feature studies contain a probability of type-I error of 34-99% with an average of 76%. This means that on average, 76% of studies found a false predictive correlation between features and outcome (82). This is supported by a perspective from Buvat *et al.* which brings up other areas that hinder PET image feature analysis such as spatial resolution and noise (71). Apparently these factors bring their own textural patterns and signal correlation to PET images (71). Another important factor is anatomical motion during imaging. Oliver *et al.* investigated the variability of image features between respiratory-gated and conventional protocols (83) finding that 26.2 % of PET features had percent differences less than 5% between protocols. Leijenaar *et al.* demonstrated that image features are affected by SUV discretization (47) and Nyflot *et al.* stressed that additional standardization is needed for PET textural features (48). These and other investigators call for the standardization of image features.

Cheng *et al.* alerted investigators that although certain PET features have demonstrated prognostic significance for certain solid tumors (NSCLC, esophageal squamous cell, cervical cancer, and oropharyngeal squamous cell), standardization is needed to take PET feature analysis from the research lab to oncological practice (46). Cheng also resolved that well-designed multicenter studies in which PET texture features from different institutions could help with standardization by removing the inconsistencies of PET features (46). Many of which are caused by differing reconstruction parameters and acquisition modes (45, 46). They also resolved that one PET textural feature alone will not have accurate information regarding clinical

outcomes. Instead, groups of multimodal biomarkers should be used since they will have better prognostic accuracy. This can be achieved through Radiomics (46).

## 2.6 Image Feature Analysis in CT

CT is the most commonly used diagnostic imaging modality in radiotherapy. It is used for tumor localization and its unique imaging method provides the electronic density data essential for treatment planning. CT data are also used for PET attenuation in PET/CT dual modalities. Image feature analysis in CT images dates back to the 1980s. In recent years, studies have assessed features derived from CT images towards revealing prognostic phenotypes, relationships between features, tumor metabolism and stage, and correlations between tumor hypoxia, angiogenesis, and aggressiveness (32, 34, 73, 84).

Aerts *et al.* used radiomic features derived from CT images to reveal prognostic phenotypes in lung cancer and head and neck cancer. The study consisted of 7 independent datasets - a total of 1019 patients. It was discovered that CT image features have strong prognostic information associated with gene-expression patterns. They also showed that radiomic features were complimentary to TNM staging for outcome prediction (84).

A study by Ganeshan *et al.* demonstrated evidence for a relationship between NSCLC texture features derived from CT images and tumor metabolism and stage (32). A following study from the same group tested if tumor heterogeneity, described by texture features from CT images in NSCLC, were survival markers (33). They found these features to have the *potential* to provide independent predictions of survival in NSCLC. Ganeshan *et al.* also found that heterogeneity in oesophageal cancers defined by texture analysis of CT images were associated with increased tumor metabolism and advanced tumor stage (both defined via 18F-FDG PET) (72). In a later study, Ganeshan *et al.* also found that tumor heterogeneity described using CT image texture in NSCLC had significant correlations with tumor hypoxia markers and

angiogenesis (73). These studies indicate that features may provide prognostic and functional information towards personalized treatment in esophageal cancer and NSCLC.

Hunter *et al.* tested the reproducibility of image features across 3 CT machines from 2 institutions in NSCLC (35). They found that features derived from average 4D CT machines were the best candidates for correlation in the clinic and had the best reproducibility compared to end-exhale 4D CT and breath-hold 3D CT (35). This study was modeled after a prior study by Kumar *et al.* (85) which demonstrated that the variance for SUV in FDG-PET/CT inside a single institution was much greater than expected. Relative differences were  $\pm 49\%$  for  $SUV_{max}$  and  $\pm 44\%$  for  $SUV_{mean}$ .

A study by Ravanelli *et al.* was the first to test whether texture analysis can be used as independent predictors of chemotherapy response in NSCLC (86). They found that there was potential for such features as predictive indicators however, a larger study size was needed to validate the preliminary results. They also created an index,  $GL*U$ , composed of two features, mean grey level and uniformity (32, 33, 72, 86). Texture analysis in CT has also been tested for its application to radiation-induced lung damage. Ruyscher *et al.* demonstrated that changes in Hounsfield Units (HU), the quantitative measure of CT images could identify the entire range of radiosensitivity on a quantitative scale (87).

The reproducibility of CT image features is a very relevant topic and one that has been studied by Balagurunathan *et al.* This group tested the reproducibility of CT image features in NSCLC (88). Patients received 2 CT scans within a 15 minute period. According to the concordance correlation coefficient (CCC), certain features were identified with reproducibility and the ability to predict a prognostic score on the samples (88). Grove *et al.* (89) developed 2 new robust quantitative imaging features, convexity and entropy ratio. These features were able to score tumors and identify imaging phenotypes exhibiting survival differences in lung cancer demonstrating the existing connection between patient survival and features. In future, these

features may be used diagnostically for lung adenocarcinomas. Image feature analysis of CT images is not a new topic but the discipline of radiomics is reinvigorating research in this area.

## 2.7 The Future of Image Feature Analysis

In recent years, networks have been established that encourage the standardization of quantitative imaging and the sharing of results between researchers, clinical and industry professionals (90). Two major initiatives, the Quantitative Imaging Network (QIN) and the Quantitative Imaging Biomarkers Alliance™ (QIBA™), were developed to promote research in quantitative imaging for clinical trials.

The Quantitative Imaging Network is a framework established to further the development of quantitative imaging methods and possible biomarkers to measure tumor response in clinical trials (91). Currently there is an image analysis working group that “provides guidance, coordination, consensus building, and awareness regarding the development of algorithms and methods for quantitative analysis.” This working group is composed of dynamic contrast-enhanced (DCE)-MRI and PET/CT subgroups.

The Quantitative Biomarkers Alliance (QIBA), sponsored by the Radiological Society of North America (RSNA) and the National Institute for Biomedical Imaging and Bioengineering was developed in 2007 by RSNA to promote quantitative image assessment in radiology through the convergence of researchers, healthcare professionals, and industry. Since that time they have established a protocol for FDG-PET imaging (90) which provides details from the beginning: subject handling and image data acquisition, to the end: image interpretation and reporting, and quality control.

Amongst investigators in image feature analysis, radiomics is considered “the next frontier in clinical decision making” (66). Gillies *et al.* stressed the importance of benchmarks for radiomics studies: extraction, analysis, and presentation of features (66). Radiomics is becoming the new face of image feature analysis in medical imaging. The annual Radiomics



Workshop organized by Dr. Gillies from the Moffitt Cancer Center, which began in 2009, has evolved into a meeting of over 100 investigators from 5 countries and 27 institutions which encourages the sharing of current radiomic findings and collaboration in the field.

Image feature analysis applied to PET/CT images in lung and esophageal cancer have proven to have prognostic and predictive indicators that may be used for treatment assessment. However, studies testing the limitations of these features are underway which will lead to better standardization and reliability of quantitative image features for decision support.

## Chapter Three: Materials & Methods<sup>3</sup>

### 3.1 ROI Delineation and Tumor Segmentation

Throughout this work, image sets were imported and viewed with Mirada Medical Software (Mirada RTx, Mirada Medical, Oxford, UK), a proprietary software providing tools for tumor analysis and visualization of Digital Imaging and Communications in Medicine (DICOM) image data for 3D and RG PET/CT images. The image viewing software provided tumor visualization, easy access to X, Y, or Z slices of 3D PET data, and region-of-interest (ROI) delineation.

Segmentation method varied between studies. In Chapter 3, lung tumors are segmented based on a 40% maximum threshold setting determined by Mirada RTx. PET image tumors were contoured at 40% maximum intensity inside a defined volume of interest. Tumors on CT were contoured with CT region segmentation. This method of segmentation was deemed acceptable for application in this chapter because we were comparing contours from the same image with differing noise levels.

In Chapters 2 & 4, a background-adapted threshold method of segmentation defined by Dholakia *et al.*, which accounted for background uptake, was applied to eliminate subjective errors and inter-observer variability (92). This method involved placing a 3-cm spherical contour inside the liver and extracting the mean SUV and standard deviation to calculate a threshold value for the lung tumor (equation 3.1):

$$MTV_{threshold} = SUV_{mean} + 2 * SUV_{SD} \quad (3.1)$$

---

<sup>3</sup> Portions of this chapter have been previously published in *Translational Oncology*, 2015, 8(6): 524-534, and have been reproduced with permission from Elsevier. The author of this dissertation is the first author of the previously published work. All figures are used with copyright permission. (See Appendix B).

where  $SUV_{mean}$  is the mean SUV value of the 3-cm sphere and  $SUV_{SD}$  is the standard deviation of the 3-cm sphere's SUV values. This segmentation method will be referred to as the Dholakia method. We are aware of the many segmentation methods in the literature and that no one method is generally regarded optimal for general medical applications (49). With lung tumors, major structures and surrounding tissues were minimal. Consequently there was little uptake outside the tumor volume. If, however, a tumor was close to the diaphragm or pleura, nearby metabolic structures were also segmented. Due to the adaptive segmentation method 6 PET lung lesions were rendered too tiny and were not evaluated. In esophageal tumors, the Dholakia method gave rise to areas of the heart and/or stomach that were included in the contour. In these cases, a physician-resident reviewed and manually edited these MTVs.

In CT images, tumors were contoured with CT threshold, a proprietary algorithm using Mirada RTx (see Figure 3.1). The three-dimensional contours were drawn separately on the 3D CT image and on one phase (phase 1 or phase 10) of the corresponding RG CT image for each patient. In our clinic, CT contours are used for treatment planning purposes whereas PET ensures the entire metabolic tumor volume is included in the gross tumor volume (GTV). CT contours were only collected for lung tumors.

Images were exported from Mirada to a local drive for analysis in DICOM file format. Structures (contours) were exported from Mirada in DICOM RT file format, an extension of the DICOM file format. An open source program, RegGui v0.42., in MATLAB (MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States) was used to produce a MetalImage file (.mhd and .raw) using CT images. A MetalImage is a raw file format, a text-based tagged file format for medical images that stores 3D data in one file. The .mhd file includes the header information and .raw file includes the data.

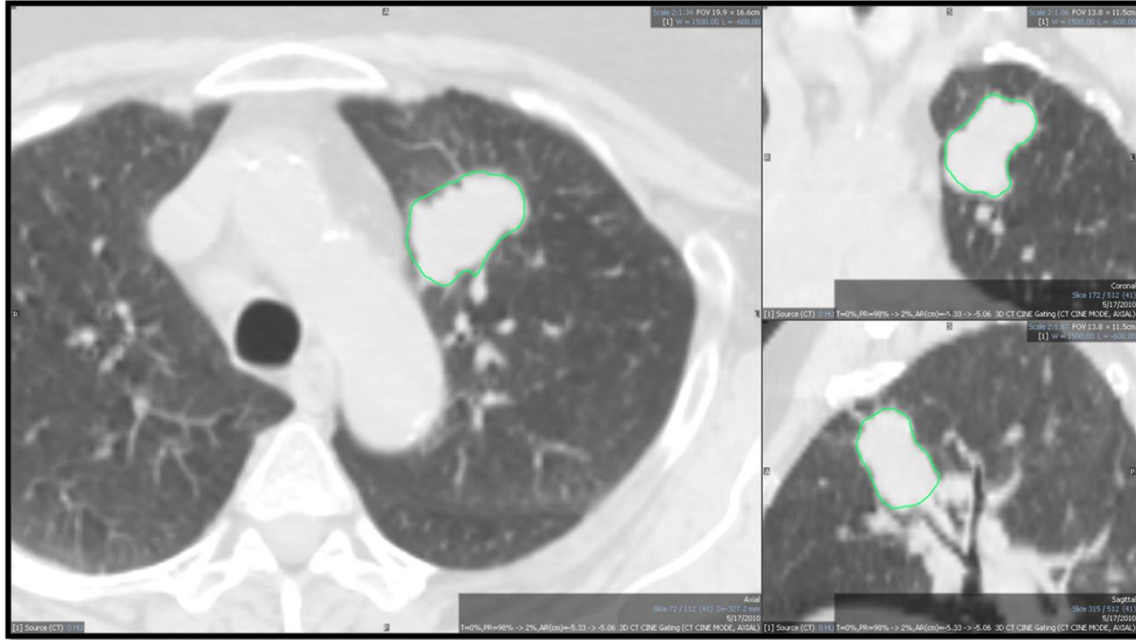


Figure 3.1: CT segmentation of one patient viewed in 2D (the ROI extends in 3D). This CT image is viewed in the window preset for the lung in Mirada DBx (RTx, Mirada Medical, Oxford, UK).

### 3.2 Feature Extraction

An internally-developed application imported the ROI data files. We imported *.mhd* file format for CT and DICOM file format for PET images, and extracted image intensity statistics, shape descriptors, co-occurrence matrices, run length matrices, and other second order features from each ROI for a total of 81 image features (Tables 3.1-3.6). Some groups have shown the instability of certain features from 3D images (40, 41, 45). Others have used large numbers of features (>200, (36)). Nevertheless, we deemed 81 features sufficient to assess the variability between 3D and RG feature values. A total list of image features is listed in Table 3.1.

In PET, the image intensity represented the number of registered counts per voxel, whereas for CT, intensity represented the Hounsfield Units in each voxel. All intensity levels were used and normalization was applied only to the co-occurrence and run length matrices (in the form of binning; 128 bins). Opposite to standard practice, intensity values for PET images were not converted to SUV. Instead, stored image intensity values were analyzed directly.

Table 3.1 Extracted Image Features including Shape Features, Intensity Features, GLCM features, RLM features, and GLSZM features.

<b>Shape (11)</b>	<b>Intensity (22)</b>	<b>GLCM<sup>a</sup> (26)</b>	<b>RLM<sup>b</sup> (11)</b>	<b>GLSZM<sup>c</sup> (11)</b>
Volume (voxels)	Minimum Intensity	Energy	Short Run Emphasis (SRE)	Short Area Emphasis (SAE)
Volume (cc)	Maximum Intensity	Contrast	Long Run Emphasis (LRE)	Large Area Emphasis(LAE)
Surface Area	Mean Intensity	Local Homogeneity	Low Gray-Level Run Emphasis (LGRE)	Low-Intensity Emphasis (LIE)
Surface Area/Volume	Peak Intensity	Entropy	High Gray-Level Run Emphasis (HGRE)	High-Intensity Emphasis (HIE)
Sphericity	Standard Deviation	Correlation	Short Run Low Gray-Level Emphasis (SRLGE)	Low-Intensity Small-Area Emphasis (LISAE)
Compactness	Skewness	Autocorrelation	Short Run High Gray-Level Emphasis (SRHGE)	High-Intensity Small-Area Emphasis (HISAE)
Spherical Disproportion	Coefficient of Variation	Correlation	Long Run Low Gray-Level Emphasis (LRLGE)	Low-Intensity Low-Area Emphasis (LILAE)
Long	Kurtosis	Dissimilarity	Long Run High Gray-Level Emphasis (LRHGE)	High-Intensity Low-Area Emphasis(HILAE)
Short	TGV	Cluster Tendency	Gray-Level Non-uniformity (GLNU)	Intensity Variability (IV)
Eccentricity	RMS	Cluster Shade	Run Length Non-uniformity (RLNU)	Size-Zone Variability (SZV)
Convexity	I30	Cluster Prominence	Run Percentage (RPC)	Zone Percentage (ZP)
	I10-I90	Mean		
	V40	Variance		
	V70	Max Probability		
	V80	Inverse Variance		
	V10-V90	Inverse Variance P		
	Energy	Inverse Difference Moment		
	Entropy	Inverse Difference		
	Contrast	Sum Average		
	Local Homogeneity	Sum Variance		
	Histogram Entropy	Sum Entropy		
	Uniformity	Difference Average		
		Difference Variance		
		Difference Entropy		
		Information Measure of Correlation 1		
		Information Measure of Correlation 2		

<sup>a</sup>Gray-Level Co-occurrence Matrix (GLCM). <sup>b</sup>Run Length Matrix (RLM). <sup>c</sup>Gray-Level Size-Zone Matrix (GLSZM).

### 3.3 Features

Image features range from shape descriptors to texture descriptors. As previously described, first order features provide information about gray-scale intensities. These features are derived from intensity distributions and histograms. Second order texture features are derived from gray-tone spatial dependency matrices which are constructed from the intensity value of an image. They describe the spatial relationships between voxel intensities (29, 30). Second order features provide information about the heterogeneity of a ROI.

Table 3.2 Shape Features

<i>Feature</i>	<i>Description</i>	<i>Mathematical Description</i>
Convexity	A measure of the spiculation of the ROI (ratio of true ROI volume to convex ROI volume)	--
Volume	A measure of the size of the tumor in three dimensions.	--
Surface Area	A measure of the area of the surface of the tumor volume.	--
Surface Area/Volume	A measure of the surface area divided by the volume.	--
Long Axis	A measure of the longest tumor diameter.	--
Short Axis	A measure of the shortest tumor diameter.	--
Sphericity	A measure of the spherical shape (roundness) of the ROI	$(\sqrt[3]{\pi} * \sqrt[3]{6V^2})/A$
Compactness	A ratio between the length of the tumor boundary and the area (93).	$V/(\sqrt{\pi} * \sqrt[3]{A^2})$
Spherical Disproportion	A measure of the spherical shape of the tumor.	$A/4\sqrt{\pi} * \sqrt[3]{\left(\frac{3V}{4\pi}\right)^2}$
Eccentricity	A measure of the non-circularity of the tumor. An eccentricity of 0 is a circle and 1 is a line (14).	$1 - \sqrt{\frac{b^2}{a^2}}$

Where  $a$  is the semi-major axis and  $b$  is the semi-minor axis.  $V$  is tumor volume and  $A$  is the tumor area.

#### 3.3.1 Shape Descriptors and First Order Features

Shape descriptors and first order features provide contour dependent information about a ROI. Shape descriptors are calculated directly from the segmented ROI. First order features are extracted from image intensity statistics and are calculated from volume intensity histograms. A full list and description of the

shape descriptors and first order features that were used throughout this work are listed in Tables 3.2 - 3.4. One particular shape descriptor, convexity, is calculated by taking a ratio of the real convex volume (a volume created by connecting the furthest tumor edges) to the actual tumor volume. Six first-order features were based on intensity volume histograms – following the method originally introduced by El Naqa et al. based on dose-volume histograms (14).

Table 3.3 Intensity and First-Order Features

<b>Feature</b>	<b>Description</b>	<b>Mathematical Description</b>
MinI	Minimum intensity value in the volume-of-interest (VOI).	$Min(I)$
MaxI	Maximum intensity value in the VOI.	$Max(I)$
MeanI	Mean intensity value in the VOI. (The first standardized moment.) (94)	$\frac{1}{N} \sum_{v=1}^N I(v)$
SD	Variation from the average intensity in VOI. (SD <sup>2</sup> is the second standardized moment.) (94)	$\sqrt{\frac{1}{N} \sum_{i=1}^N (I(v) - I_{mean})^2}$
Skewness	Measures symmetry of the intensity distribution. (Third standardized moment.) (94)	$\frac{1}{N} \sum_{v=1}^N (I(v) - I_{mean})^3 / SD^3$
Kurtosis	Measures shape of the peak of the intensity distribution. (Fourth standardized moment.) (94)	$\frac{1}{N} \sum_{v=1}^N (I(v) - I_{mean})^4 / SD^4$
Coefficient of Variation	A normalized measure of the dispersion of the VOI. The ratio of the standard deviation to the mean value of the VOI intensities.	$\frac{\sigma}{\mu}$
TGV	Represents the total summed intensity of the VOI.	$\sum_{v=1}^N I(v)$
RMS	The square root of the sum of the squares of the voxel intensities.	$\sqrt{\sum_{v=1}^N I(v)^2}$
Energy	A measure of the homogeneity of the intensity histogram.	$\sum_i^N \sum_j^N \sum_k^M [P(i, j, k)]^2$
Contrast	A measure of the intensity variation of the intensity histogram.	$\sum_i^L \sum_j^N \sum_k^M P(i, j, k)$

Where  $I(v)$  is the intensity of a voxel,  $N$  is the number of voxels in a volume-of-interest (VOI), and  $p(i)$  is the probability of the occurrence of the grey-level  $i$  (37). Where  $w(i)$  is the width of the  $i$ th bin of the histogram, and TGV is total summed intensity.

Table 3.3 (Continued)

<b>Feature</b>	<b>Description</b>	<b>Mathematical Description</b>
Local Homogeneity	Opposite of contrast.	$\sum_i^L \sum_j^N \sum_k^M \frac{P(i, j, k)}{(1 + i^2 + j^2 + k^2 - jk - jk - ik - ik)}$
Histogram Entropy	A measure of entropy of the intensity histogram of a VOI (95).	$-\sum_{i=1}^N p(i) \log\left(\frac{p(i)}{w(i)}\right)$
Uniformity	Histogram based assessment of the number of voxels at each intensity value.	Sum of squares of number of voxels for each intensity value.
Entropy	A measure of the information content (37).	$-\sum_{i=1}^{N_g} p(i) \log_2 p(i)$

Where  $I(v)$  is the intensity of a voxel,  $N$  is the number of voxels in a volume-of-interest (VOI), and  $p(i)$  is the probability of the occurrence of the grey-level  $i$  (37). Where  $w(i)$  is the width of the  $i$ th bin of the histogram, and TGV is total summed intensity.

Table 3.4 Additional Intensity and First-Order Features

<b>Feature</b>	<b>Description</b>
I30	Intensity ranging from the lowest 30% to highest intensity volume.
I10-I90	Intensity ranging from lowest to 10% highest intensity volume minus intensity ranging from lowest to 90% highest intensity volume.
V40	Percentage volume with at least 40% intensity
V70	Percentage volume with at least 70% intensity
V80	Percentage volume with at least 80% intensity
V10-V90	Percentage volume with at least 10% intensity minus percentage volume with at least 90% intensity
I30	Intensity ranging from the lowest 30% to highest intensity volume.
PeakI	Peak intensity in the VOI.

### 3.3.2 GLCM Features

GLCM features, originally described by Haralick *et al*, provide texture information about the spatial relationships of voxels and “summarize the relative frequency distribution” of an image (29). These features describe relationships between pixel pairs and describe how often a pixel of gray level  $i$  occurs in a defined spatial relationship to a pixel gray level  $j$ . A matrix



containing these relationships is formed in a discrete number of grey level bins (usually 32, 64, 128, or 256). These gray levels are listed on both the row (i) and column (j) of the matrix (see Figure 3.2). An algorithm counts the number of instances that two pixel intensities are adjacent to each other in 13 directions. A matrix is created with this data (see Figure 3.2b). Features are shown in Tables 3.5 and 3.6. Figure 3.2 (see page 37) demonstrates a sample image and sample GLCM in the 0° direction (horizontal). This image has 5 intensity levels and each intensity is given its own bin. In the zero direction, the matrix is formed by noting the number of time pixels of gray level i and j are adjacent. For instance, gray levels 3 and 2 occur in adjacent pixels twice. The GLCM is a symmetrical matrix and is normalized by dividing each matrix element by the number of possible pixel pairs,  $(N_x - 1) * N_y$ , where  $N_x$  represents the height and  $N_y$  represents the width of the image. Thus the GLCM matrix element for (3,2) and (2,3) is listed as 2/20.

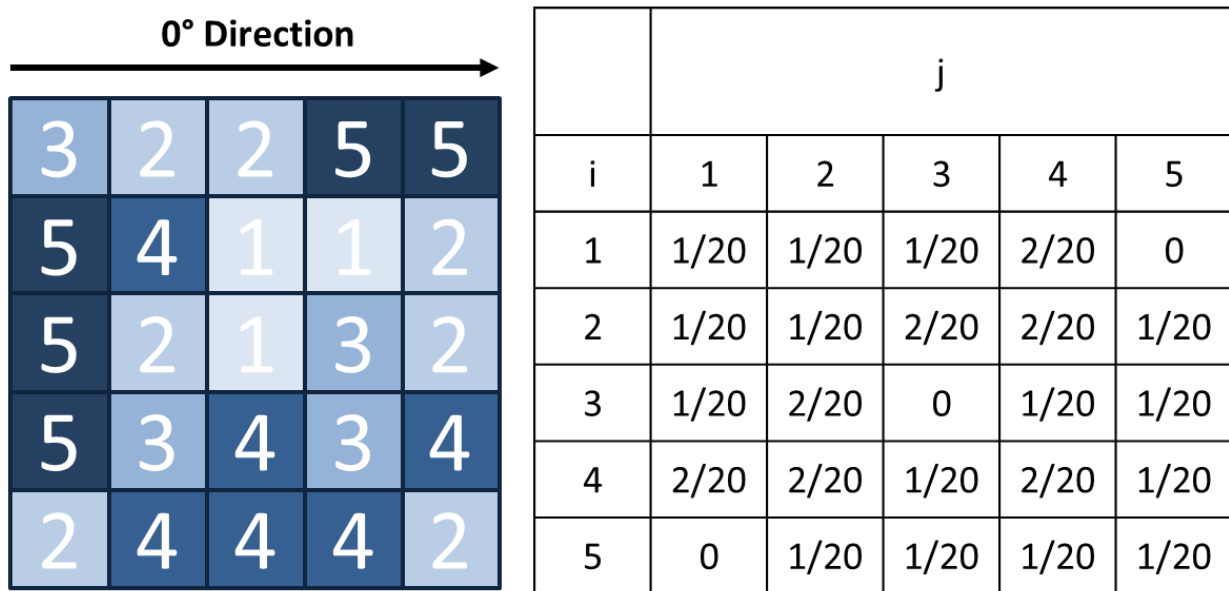


Figure 3.2 (a) Sample figure; (b) Sample GLCM matrix in the 0° direction (horizontal) based on a 5x5 pixel image with 5 grey-levels.

Table 3.5 GLCM Features.

Feature	Description	Mathematical Description
Energy	Also defined as Angular Second Moment. This feature describes the homogeneity of an image. 0 represents complete heterogeneity. 1 represents complete homogeneity (96).	$\sum_{i=1}^N \sum_{j=1}^N p(i,j)^2$
Entropy	Measures the pair contributions and information content.	$\sum_{i=1}^N \sum_{j=1}^N p(i,j) \log p(i,j)$
Contrast	Measures local intensity variation. This favors matrix values that are away from the diagonal. A value of 0 demonstrates no contrast.	$\sum_{n=0}^{N-1} n^2 \left\{ \sum_{i=1}^N \sum_{j=1}^N p(i,j) \right\},  i-j  = n$
Homogeneity	Measures local intensity variation. This favors matrix values that are close to the diagonal. A value of 0 demonstrates dissimilarity.	$\sum_{i=1}^N \sum_{j=1}^N \frac{p(i,j)}{1 +  i-j }$
Local Homogeneity	Measures the relation of GLCM intensities to the diagonal GLCM matrix. A value of 1 represents total homogeneity. A value of 0 represents non-homogeneity (96). Also defined as Inverse Difference Moment.	$\sum_{i=1}^N \sum_{j=1}^N \frac{p(i,j)}{1 + (i-j)^2}$
Correlation	Measures correlation between co-occurrence matrix values. Measures gray level linear dependence between pixels.	$\frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
Cluster Shade	Measures of the skewness of the matrix. High values = non-symmetry	$\sum_i \sum_j \left( (i - \mu_i)^3 + (j - \mu_j)^3 \right) p(i,j)$
Cluster Prominence	Measures of the skewness of the matrix. High values = non-symmetry. A peak in the matrix near mean value = small variation in gray-scales.	$\sum_i \sum_j \left( (i - \mu_i)^4 + (j - \mu_j)^4 \right) p(i,j)$
Cluster Tendency	Measures groups of pixels with similar intensity values (97).	$\sum_i \sum_j (i + j - 2\mu)^k p(i,j)$
Mean (Co-occurrence Matrix)	The mean value of the co-occurrence matrix values.	$\sum_i \sum_j p(i,j)$
Variance	Places high weight on values differing from the average matrix value.	$\sum_i \sum_j (i - \mu)^2 p(i,j) + \sum_i \sum_j (j - \mu)^2 p(i,j)$

Where  $p(i,j)$  is the  $(i,j)$ th entry in a normalized grey-level co-occurrence matrix and  $p(i,j)$  is the  $(i,j)$ th entry in a non-normalized grey-level co-occurrence matrix.  $N$  is the number of grey-levels,  $p_x$  is the  $i$ th entry obtained by summing the rows of  $p(i,j)$ ,  $p_y$  is the  $j$ th entry obtained by summing the columns of  $p(i,j)$ .  $H_X$  and  $H_Y$  are entropies of  $p_x$  and  $p_y$ . GLCM features were originally developed by Haralick et al. (29, 98).

Table 3.5 (Continued)

Feature	Description	Mathematical Description
Dissimilarity	Contrast weighted linearly with the difference of grey-level values (distance from diagonal where 0=similarity) (37)	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g}  i-j p(i,j)$
Max Probability	Detects the most frequent pattern (93). Expected to be high if there is a high occurrence of the most predominant pixel pair (97).	$\text{Max}_{i,j}^{N_g, N_g} p(i,j)$

Where  $p(i,j)$  is the  $(i,j)$ th entry in a normalized grey-level co-occurrence matrix and  $p(i,j)$  is the  $(i,j)$ th entry in a non-normalized grey-level co-occurrence matrix.  $N$  is the number of grey-levels,  $p_x$  is the  $i$ th entry obtained by summing the rows of  $p(i,j)$ ,  $p_y$  is the  $j$ th entry obtained by summing the columns of  $p(i,j)$ .  $H_X$  and  $H_Y$  are entropies of  $p_x$  and  $p_y$ . GLCM features were originally developed by Haralick et al. (29, 98).

Table 3.6. Additional GLCM Features.

Feature	Mathematical Description
Sum Average	$\sum_{i=2}^{2N_g} ip_{x+y}(i)$
Sum Variance	$\sum_{i=2}^{2N_g} (i + \sum_i \sum_j p(i,j) \log(p(i,j)))^2 \log\{p_{x+y}(i)\}$
Sum Entropy	$- \sum_i \sum_j p_{x+y}(i) \log\{p_{x+y}(i)\}$
Difference Average	$\sum_{i=2}^{2N_g} ip_{x-y}(i)$
Difference Variance	$\sum_{i=2}^{2N_g} (i + \sum_i \sum_j p(i,j) \log(p(i,j)))^2 \log\{p_{x-y}(i)\}$
Difference Entropy	$- \sum_i \sum_j p_{x-y}(i) \log\{p_{x-y}(i)\}$
Inverse Variance	$\sum_i \sum_j \frac{P(i,j)}{(i-j)^2}$
Inverse Variance P	$\sum_i \sum_j \frac{p(i,j)}{(i-j)^2}$

Where  $p(i,j)$  is the  $(i,j)$ th entry in a normalized grey-level co-occurrence matrix and  $p(i,j)$  is the  $(i,j)$ th entry in a non-normalized grey-level co-occurrence matrix.  $N$  is the number of grey-levels,  $p_x$  is the  $i$ th entry obtained by summing the rows of  $p(i,j)$ ,  $p_y$  is the  $j$ th entry obtained by summing the columns of  $p(i,j)$ .  $H_X$  and  $H_Y$  are entropies of  $p_x$  and  $p_y$ . GLCM features were originally developed by Haralick et al. (29, 98).

Table 3.6 (Continued)

Feature	Mathematical Description
Inverse Difference	$\sum_{i=1}^N \sum_{j=1}^N \frac{p_{ij}}{1 + (i - j)}$
Information Measure of Correlation 1	$HXY = - \sum_i^{N_g} \sum_j^{N_g} p(i, j) \log\{p(i, j)\}$ $HXY1 = - \sum_i^{N_g} \sum_j^{N_g} p(i, j) \log\{p_x(i)p_y(j)\},$ $HXY2 = - \sum_i^{N_g} \sum_j^{N_g} p_x(i)p_y(j) \log\{p_x(i)p_y(j)\},,$ $\frac{HXY - HXY1}{\max\{hx, hy\}}$
Information Measure of Correlation 2	$\sqrt[2]{1 - e^{[-2.0(HXY2 - HXY)']}}$
Autocorrelation	$\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i * j * p(i, j)$

Where  $p(i, j)$  is the  $(i, j)$ th entry in a normalized grey-level co-occurrence matrix and  $p(i, j)$  is the  $(i, j)$ th entry in a non-normalized grey-level co-occurrence matrix.  $N$  is the number of grey-levels,  $p_x$  is the  $i$ th entry obtained by summing the rows of  $p(i, j)$ ,  $p_y$  is the  $j$ th entry obtained by summing the columns of  $p(i, j)$ .  $HX$  and  $HY$  are entropies of  $p_x$  and  $p_y$ . GLCM features were originally developed by Haralick et al. (29, 98).

### 3.3.3 Run Length Matrix Features

The Run Length Matrix (RLM) was originally introduced by Galloway (64). The RLM is a matrix consisting of dimensions  $L \times R$  where  $L$  is the number of grey levels and  $R$  is the number of possible runs. A run is a set of adjacent voxels that possess the same grey level in a particular direction. A sample RLM is demonstrated in Figure 3.3. The RLM consists of 5 traditional features proposed by Galloway (64), 2 features proposed by Chu *et al.* (99), and 4 features proposed by Dasarathy and Holder (100). These features are introduced in Table 3.7. Figure 3.3 demonstrates a sample image and sample RLM in the  $0^\circ$  direction (horizontal). This image has 5 intensity levels. The gray level 3 occurs in a run 1 pixel long 4 times. Thus, the GLCM matrix element  $(R, L)$  for  $(3, 1)$  is 4.

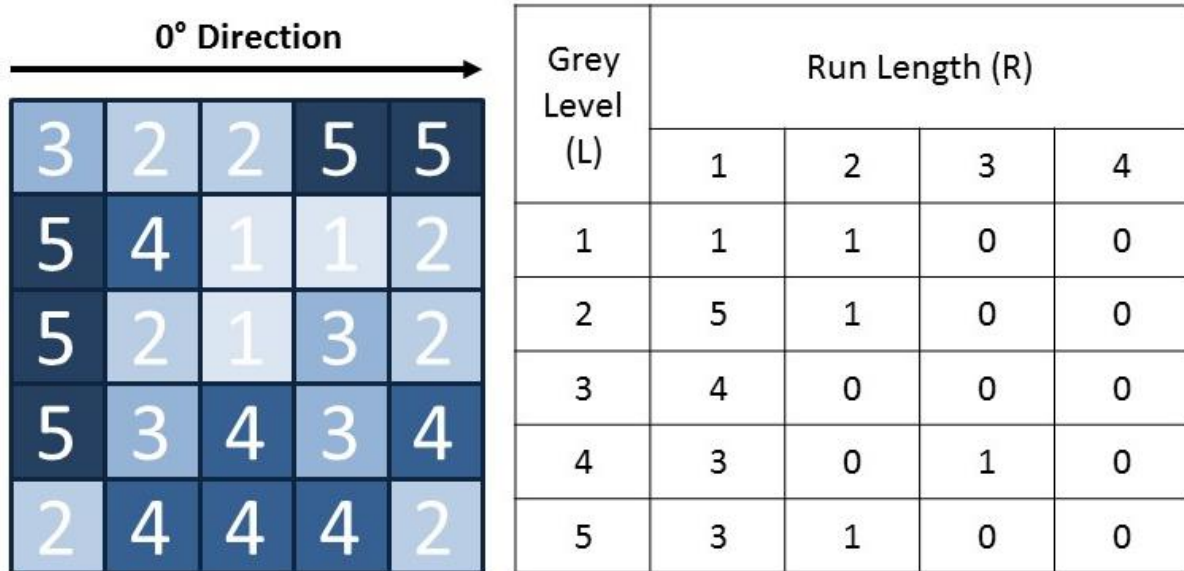


Figure 3.3 (a) Sample Image; (b) RLM Matrix calculated based on sample image with 5 grey levels in the 0° direction.

### 3.3.4 Grey Level Size Zone Matrix Features

The Grey Level Size Zone Matrix (GLSZM) was developed by Thibault (63, 101) according to the run length matrix principle. This principle states that the value of matrix  $z(i,j)$  equals the number of zones of size  $i$  and gray level  $j$  (101). The GLSZM is different from the run length matrix in that it examines “zones” instead of “pairs”. It records the size of groups of adjacent pixels containing the same intensity levels (see Figure 3.4). GLSZM features are analogs of RLM features and are described in Table 3.8. Figure 3.4 demonstrates a sample 2D image and sample GLSZM. Because the GLSZM measures zones it does not need to be measured in any particular direction. Each direction will produce the same GLSZM. The sample image has 5 intensity levels. GLSZM matrix element (2,1) has a value of 3 because the gray level 2 has a size of 1 in 3 instances.

Table 3.7 RLM Features

<b>Feature</b>	<b>Description</b>	<b>Mathematical Description</b>
SRE <sup>a</sup>	Measures short run distribution; (short run emphasis).	$\frac{1}{n} \sum_{i=1}^M \sum_{j=1}^N \frac{R(i,j)}{j^2}$
LRE <sup>a</sup>	Measures long run distribution; (long run emphasis).	$\frac{1}{n} \sum_{i=1}^M \sum_{j=1}^N R(i,j)j^2$
LGRE <sup>b</sup>	Measures low grey-level distribution; (low grey-level run emphasis).	$\frac{1}{n} \sum_{i=1}^M \sum_{j=1}^N \frac{R(i,j)}{i^2}$
HGRE <sup>b</sup>	Measures high grey-level distribution; (high grey-level run emphasis).	$\frac{1}{n} \sum_{i=1}^M \sum_{j=1}^N R(i,j)i^2$
SRLGE <sup>c</sup>	Measures short runs and low grey-level distribution; (short run low grey-level emphasis).	$\frac{1}{n} \sum_{i=1}^M \sum_{j=1}^N \frac{R(i,j)}{i^2 j^2}$
LRLGE <sup>c</sup>	Measures long runs and low grey-level distribution; (long run low grey-level emphasis).	$\frac{1}{n} \sum_{i=1}^M \sum_{j=1}^N \frac{R(i,j)j^2}{i^2}$
SRHGE <sup>c</sup>	Measures short runs and high grey-level distribution; (short run high grey-level emphasis).	$\frac{1}{n} \sum_{i=1}^M \sum_{j=1}^N \frac{R(i,j)i^2}{j^2}$
LRHGE <sup>c</sup>	Measures long runs and high grey-level distribution; (long run high grey-level emphasis).	$\frac{1}{n} \sum_{i=1}^M \sum_{j=1}^N R(i,j)i^2 j^2$
RLNU <sup>a</sup>	Measures the non-uniformity of the run lengths; (run length non-uniformity).	$\frac{1}{n} \sum_{i=1}^N \left( \sum_{j=1}^M R(i,j) \right)^2$
GLNU <sup>a</sup>	Measures the non-uniformity of the grey-levels; (grey-level non-uniformity).	$\frac{1}{n} \sum_{j=1}^M \left( \sum_{i=1}^N R(i,j) \right)^2$
RPC <sup>a</sup>	Ratio of total number of runs to total number of pixels in the image. Measures homogeneity and run distribution; (run percentage).	$\frac{n}{n_p}$

<sup>a</sup>Where  $R(i,j)$  is an element of the Run Length Matrix,  $n$  is the total number of runs,  $n_p$  is the number of pixels in the image,  $N$  is the longest run and  $M$  is the number of grey levels. <sup>a</sup>RLM features developed by Galloway et al. (64). <sup>b</sup>RLM features developed by Chu et al. (99). <sup>c</sup>RLM features developed by Dasarathy and Holder (100).

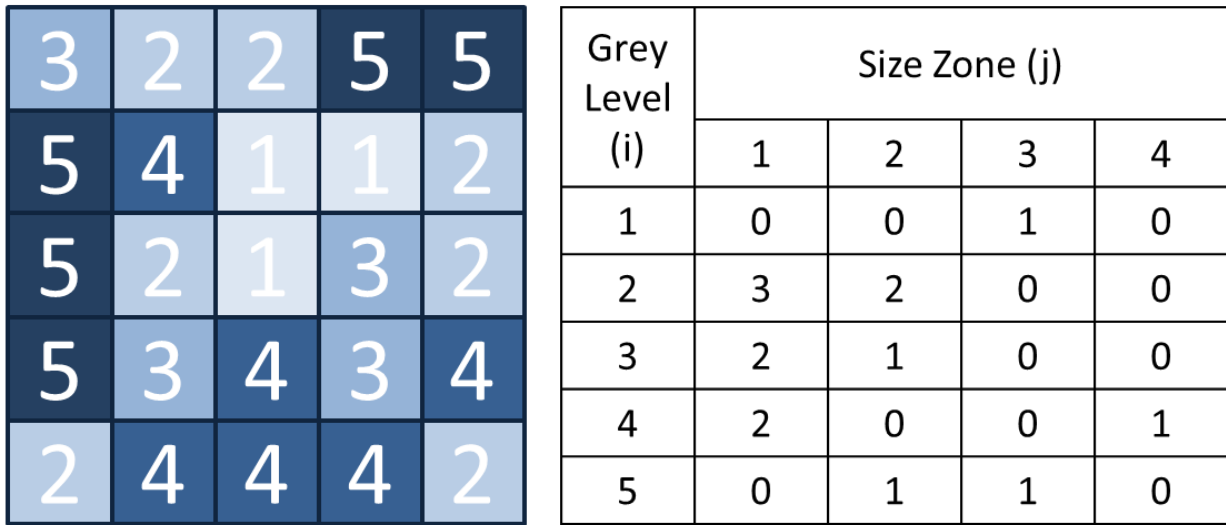


Figure 3.4 (a) Sample Image; (b) Sample GLSZM Matrix based on sample image with 5 grey levels.

Table 3.8 Gray-Level Size-Zone Matrix Features.

Feature	Description	Mathematical Description
Small-area emphasis (SAE)	Emphasizes small areas	$\frac{1}{\Omega} \sum_{i=1}^m \sum_{j=1}^n \frac{z(i,j)}{j^2}$
Large-area emphasis	Emphasizes large areas	$\frac{1}{\Omega} \sum_{i=1}^m \sum_{j=1}^n j^2, z(i,j)$
Low-intensity emphasis	Emphasizes low intensity areas.	$\frac{1}{\Omega} \sum_{i=1}^m \sum_{j=1}^n \frac{z(i,j)}{i^2}$
High-intensity emphasis	Emphasizes high intensity areas	$\frac{1}{\Omega} \sum_{i=1}^m \sum_{j=1}^n i^2, z(i,j)$
Low-intensity small-area emphasis	Emphasizes small areas with low intensity	$\frac{1}{\Omega} \sum_{i=1}^m \sum_{j=1}^n \frac{z(i,j)}{i^2, j^2}$
High-intensity small-area emphasis	Emphasizes small areas with high intensity	$\frac{1}{\Omega} \sum_{i=1}^m \sum_{j=1}^n i^2, j^2, z(i,j)$
Low-intensity large-area emphasis	Emphasizes large areas with low intensity	$\frac{1}{\Omega} \sum_{i=1}^m \sum_{j=1}^n \frac{j^2, z(i,j)}{i^2}$

Where  $z(i,j)$  is the  $(i,j)$ th entry in the  $z$  matrix,  $\Omega$  is the number of homogeneous areas,  $n$  is the number of distinct intensity values within the tumor and  $m$  is the size of the largest homogeneous area in the  $z$  matrix (101).

Table 3.8 (Continued)

<b>Feature</b>	<b>Description</b>	<b>Mathematical Description</b>
High-intensity large-area emphasis	Emphasizes large areas with high intensity	$\frac{1}{\Omega} \sum_{i=1}^m \left[ \sum_{j=1}^n z(i,j) \right]^2$
Intensity variability	Variance of intensity levels	$\frac{1}{\Omega} \sum_{i=1}^m \sum_{j=1}^n \frac{i^2, z(i,j)}{j^2}$
Zone Percentage	A ratio between the number of zones and the total number of voxels (total possible zones)	$\Omega / \sum_{i=1}^m \sum_{j=1}^n j^2, z(i,j)$
Size-zone variability	Variance of the size-zone	$\frac{1}{\Omega} \sum_{i=1}^m \left[ \sum_{j=1}^n \frac{z(i,j)}{i^2} \right]^2$

Where  $z(i,j)$  is the  $(i,j)$ th entry in the  $z$  matrix,  $\Omega$  is the number of homogeneous areas,  $n$  is the number of distinct intensity values within the tumor and  $m$  is the size of the largest homogeneous area in the  $z$  matrix (101).



## Chapter Four: Variability of Image Features Computed From Conventional and Respiratory-Gated PET/CT Images of Lung Cancer<sup>4</sup>

### 4.1 Introduction

Positron Emission Tomography (PET) is a beneficial technology in the process of cancer diagnosis and staging (102, 103), monitoring tumor response to treatment (104), detecting necrosis and tumor heterogeneity, identifying the primary tumor location (68), and delineating tumors from atelectasis (105), particularly in lung cancers. In fact, studies have shown that the use of PET/CT improves confidence in diagnosis, increases the number of “definite” lesions by 41% in patients with Non-Small Cell Lung Cancer (NSCLC) (70) and improves delineation accuracy for gross tumor volumes (GTV) in radiation therapy (68, 105).

Conventional (3D) PET images are influenced by motion because of their relatively long acquisition times. The acquired coincidence counts measured and used to form the images are spatiotemporally averaged over multiple breathing cycles (106), consequently, for a point inside a mobile tumor the signal is convoluted along its trajectory of motion. Respiratory-gated PET/CT aims to account for respiratory motion and thereby respiration-induced image blurring. One way to discern the effect of motion on feature values is by comparing image feature values between conventional and respiratory-gated acquisition protocols, although other important factors stemming from the differences in the imaging protocols such as image noise are also at play. Thus far, only one study accounting for motion in PET images has been reported by Yip et al. (107), which was limited to only 5 features. This report represents the first study that evaluates how 3D and RG acquisitions affect a large number of image features currently being used and

---

<sup>4</sup> Portions of this chapter have been previously published in *Translational Oncology*, 2015, 8(6): 524-534, and have been reproduced with permission from Elsevier (see Appendix B). The author of this dissertation is the first author of the previously published work. All figures are used with copyright permission.

tested in several medical applications. Robust features that emerge from this study may be suitable candidates for future quantitative imaging applications involving mobile tumors.

## 4.2 Materials and Methods

### 4.2.1 Preliminaries

Twenty-three lung cancer patients were retrospectively selected for a study of image feature variation between 3D & RG (RG is alternatively known as 4D) PET/CT images, and feature value variation among the 10 phases of a respiratory-gated scan. The main selection criterion was for these patients to have both 3D and RG PET/CT scans performed during the same imaging session. The 3D images were acquired in free breathing conditions for patients with regular breathing as required for Radiotherapy planning. There were thirteen female and ten male patients ranging from age 47 to 83. All lung cancer patients were diagnosed with Non-Small Cell Lung Cancer (NSCLC). For each patient,  $^{18}\text{F}$ -FDG was the administered radiotracer and the respiratory-gated PET scan was acquired during the same session as the 3D PET scan in the same position per protocol for radiation treatment planning. A radiologist approved 4D PET/CT protocol was applied for image acquisition (108). The 3D PET/CT protocol was adapted for our institution from the Netherlands Protocol (109). In routine clinical practice following these protocols, the average scan start times after the tracer administration were  $118 \pm 17.3$  minutes (standard deviation) for 3D PET and  $117 \pm 36.0$  minutes for RG PET with average administered activity of  $11.9 \pm 2.0$  mCi. A study on SUV variance in clinical FDG-PET/CT found that  $\text{SUV}_{\text{max}}$  and  $\text{SUV}_{\text{mean}}$  were independent of variations in the uptake period (85).

PET/CT data were obtained using a GE Discovery STE PET/CT Scanner (for 21 cases) and a GE Discovery 600 PET/CT Scanner (for 2 cases). The 3D CT was a standard step and shoot CT (not helical) and the respiratory-gated PET counts were binned into 10 phases with 3D CT attenuation correction applied to 4D PET data (standard protocol at our institution). The standard reconstruction protocol was the ordered subset expectation maximization (OSEM) algorithm with 20 or 28 subsets and 2 iterations. Full width at half maximum (FWHM) and field

of view (FOV) varied between patients: full width at half maximum of 4.29 mm, 7 mm or 10 mm, and field of view of 50, 60 or 70 cm were standard PET settings. Standard of practice procedures at our institution were followed and this study was approved with waived informed consent by the University of South Florida Institutional Review Board.

#### 4.2.2 ROI Delineation and Tumor Segmentation

Following patient selection, the tumor region was identified using an advanced image viewing software (Mirada RTx, Mirada Medical, Oxford, UK), allowing identification of the primary tumor location and exportation of Digital Imaging and Communications in Medicine (DICOM) image data for 3D and RG PET/CT images. The image viewing software provided tumor visualization, easy access to X, Y, or Z slices of 3D PET data, and region-of-interest (ROI) delineation. A background-adapted thresholding method of segmentation defined by Dholakia *et al.*, which accounted for background uptake, was applied to eliminate subjective errors and inter-observer variability (92). This method involved placing a 3-cm spherical contour inside the liver and extracting the mean SUV and standard deviation to calculate a threshold value for the lung tumor (Equation 4.1):

$$MTV_{threshold} = SUV_{mean} + 2 * SUVSD \quad (4.1)$$

where SUV<sub>mean</sub> is the mean SUV value of the 3-cm sphere and SUVSD is the standard deviation of the 3-cm sphere's SUV values. We are aware of the many segmentation methods in the literature and that no one method is generally regarded optimal for general medical applications (49). Since we were working with lung tumors, major structures and surrounding tissues were minimal. Consequently there was little uptake outside the tumor volume. If, however, a tumor was close to the diaphragm or pleura, nearby metabolic structures were also segmented. Due to the adaptive segmentation method 6 PET lesions were rendered too tiny and were not evaluated.

In CT images, tumors were contoured with CT threshold, a proprietary algorithm using Mirada RTx (RTx, Mirada Medical, Oxford, UK) (see Figure 2.1). The three-dimensional

contours were drawn separately on the 3D CT image and on one phase (phase 1 or phase 10) of the corresponding RG CT image for each patient. In our clinic, CT contours are used for treatment planning purposes whereas PET ensures the entire metabolic tumor volume is included in the gross tumor volume (GTV).

#### **4.2.3 Feature Extraction**

An internally-developed application imported the ROI data file and extracted image intensity statistics, shape descriptors, co-occurrence matrices, run length matrices, and other second order features from each ROI for a total of 56 image features. Although some authors have shown the instability of certain features from 3D images (40, 41, 45), we decided to include them here to analyze their stability for RG images. Moreover, some groups have used a large number of features (>200; (36)). Nevertheless, we deemed 56 features sufficient to assess the variability between 3D and RG feature values.

Shape descriptors were calculated directly from the segmented ROI. First order features (extracted from image intensity statistics) were calculated from volume intensity histograms. Second order grey-level co-occurrence matrix (GLCM) features, originally described by Haralick *et al.* (29, 98) were implemented with feature descriptions provided by Liang (96). The Haralick definition of second order statistics (based on grey-level matrix metric), nearest neighbor spatial dependence matrices, provided texture information from the spatial relationship of image voxels (29).

The GLCM feature calculations were implemented as follows: the intensities of image voxels were binned into 256 gray-scale levels for PET (128 gray-scale levels for CT) with equal intervals. The resulting 2D co-occurrence matrix was 256 x 256 (128 x 128) with unit (1) pixel distance. Co-occurrence matrices were calculated in 13 directions across a 3D image and the resultant matrix was the average of the matrices in the 13 directions. Given a set of cubical voxels, the 13 directions were: 3 axial directions, 2 diagonal directions per axial plane x 3 axial planes, and 4 diagonal directions cross cube (110). These 13 directions were chosen so that the

resulting matrix would represent the entire tumor texture without bias. The elements of the matrix were integers. Next, a probability matrix was calculated by dividing each element by the total sum of the matrix so that the sum of the probability matrix was 1. The features were then calculated using the probability matrix.

Galloway's original run length features were also implemented (64). Feature definitions were acquired from Galloway, Chu *et al.*, and Dasarathy and Holder (64, 99, 100). The run length matrix (RLM) had dimensions of L x R, where L was the number of grey-scale levels (256 for PET; 128 for CT) and R were the possible runs (determined case-by-case). The elements of the matrix were integers which represented runs. A run was defined as a set of pixels that possessed the same gray level in a specified direction (111). The RLM was calculated in 13 directions across an image (similar to the co-occurrence matrix) (112). The feature values were the summed values of all 13 directions normalized by the total runs in each direction. No probability matrix was involved for the run-length features.

In PET, the image intensity was the number of registered counts per voxel. For CT, intensity represented the Hounsfield Units in each voxel. All intensity levels were used. Normalization was applied only to the co-occurrence and run length matrices (in the form of binning; 128 bins). Additionally, intensity values for PET images were not converted to SUV. Instead, stored image intensity values were analyzed directly. For each patient, image features were extracted from the 3D PET ROI, 3D CT ROI, all phases (bins) of the respiratory-gated PET ROI, and one phase (bin) of the respiratory-gated CT ROI.

Following feature extraction, 3D and RG PET/CT image feature differences were calculated using Equation 4.2,

$$\%Diff_i^{3D/RG} = \frac{|RGTV_i - 3DTV_i|}{\text{Max}(|RGTV_i|, |3DTV_i|)} \quad (4.2)$$

where  $RGTV_i$  is the respiratory-gated image feature value for feature  $i$  and  $3DTV_i$  is the 3D image feature value for feature  $i$ . This method was chosen because it accounts for features

that changed sign between 3D and RG cases. The maximum possible percent difference using Equation 2 is 200% and differences greater than 100% were deemed large. The percent difference across cases was then averaged for each image feature and a paired, two-tailed t-test was applied to 3D and RG feature data to compare the two datasets. We assumed normal distributions and that the t-test was applicable. The concordance correlation coefficient (CCC) was calculated for all features across 3D and RG feature data to determine the correlation between the two datasets. The scale used for determining strength-of-agreement was as follows: high strength-of-agreement – CCC > 0.99, substantial strength-of-agreement – CCC: 0.95-0.99, moderate strength-of-agreement – CCC: 0.90-0.95, poor strength-of-agreement – CCC < 0.90 (113).

#### 4.2.4 RG (4D) PET Phase Analysis

The previously described procedure of image feature extraction was applied to all respiratory-gated PET bins. Mean percent difference was used to compare features between phases,

$$\%Diff_{ij}^{Mean} = \left| \frac{(TV_{ij} - \mu_{TV,j})}{\mu_{TV,j}} \right| \quad (4.3)$$

where  $i$  represents the bin,  $j$  represents the specific image feature,  $TV_{ij}$  represents the value for bin  $i$  and feature  $j$ , and  $\mu_{TV,j}$  represents the mean value for image feature  $j$ . Image feature values were also normalized by average value across all bins,

$$NormValue_{ij} = \frac{TV_{ij}}{\mu_{TV,j}} \quad (4.4)$$

The subscript definitions for Equation 4.3 also apply to Equation 4.4. Additionally, a paired, two-tailed t-test was applied to RG inhale (phase 1) feature data and RG exhale (phase 5) feature data to compare the two datasets. The CCC was calculated for phase 1 and phase 5 of the feature data to determine correlation between the two datasets.

#### 4.2.5 Long Axis Calculation and Rotation Analysis

The long axis length (through the center of mass) was calculated with an internally-developed program for each bin of the respiratory-gated cycle (PET only). The tumor's center of mass location was calculated for the inhale and exhale phases of the respiratory-gated PET image sets (Equation 4.5).

$$CM(x, y, z) = \left[ \frac{\sum_i I_i x_i}{\sum_i I_i}, \frac{\sum_i I_i y_i}{\sum_i I_i}, \frac{\sum_i I_i z_i}{\sum_i I_i} \right] \quad (4.5)$$

where  $CM(x, y, z)$  is the center of mass for a tumor in a PET phase and  $I_i$  is the number of counts per voxel  $i$ . The center of mass motion (CMM) was calculated as the displacement between the center of mass for inhale phase and center of mass for exhale phase. The difference in long axis length and CMM were used to assess changes in internal tumor morphology. Tumor angle was defined as the angle between the long axis of the tumor and the XY plane (see Figure 4.1). A Pearson's correlation test was applied to identify correlation in tumor angle and long axis length between inhale and exhale images.

### 4.3 Results

#### 4.3.1 3D and RG PET/CT Image Feature Analysis

Features from both PET and CT images demonstrated dependency on whether the acquisition was 3D, which is conventional (also called *static*), or respiratory-gated, (RG or 4D), where the coincidence counts are binned in multiple phases/bins composing the respiratory cycle. Large differences in some features were found between 3D PET/CT and one of the phases/bins of the corresponding respiratory-gated data set. The percent differences between 3D and respiratory-gated modalities were usually larger for CT than for PET. For PET, 10 of 56 features had a percentage difference (between 3D PET and RG PET for each patient) of less than 5% for more than half of the cases. In comparison, 11 of 56 CT features had a percentage difference (between 3D CT and RG CT for each patient) of less than 5% for more than half the cases. The percent differences between 3D PET and RG PET varied from 0% to 193%. The

outlier of 193% was kurtosis. For 4 of 17 cases, kurtosis demonstrated the greatest percent difference between 3D and RG PET. Image feature average differences between 3D PET and RG PET are shown in Table 4.1.

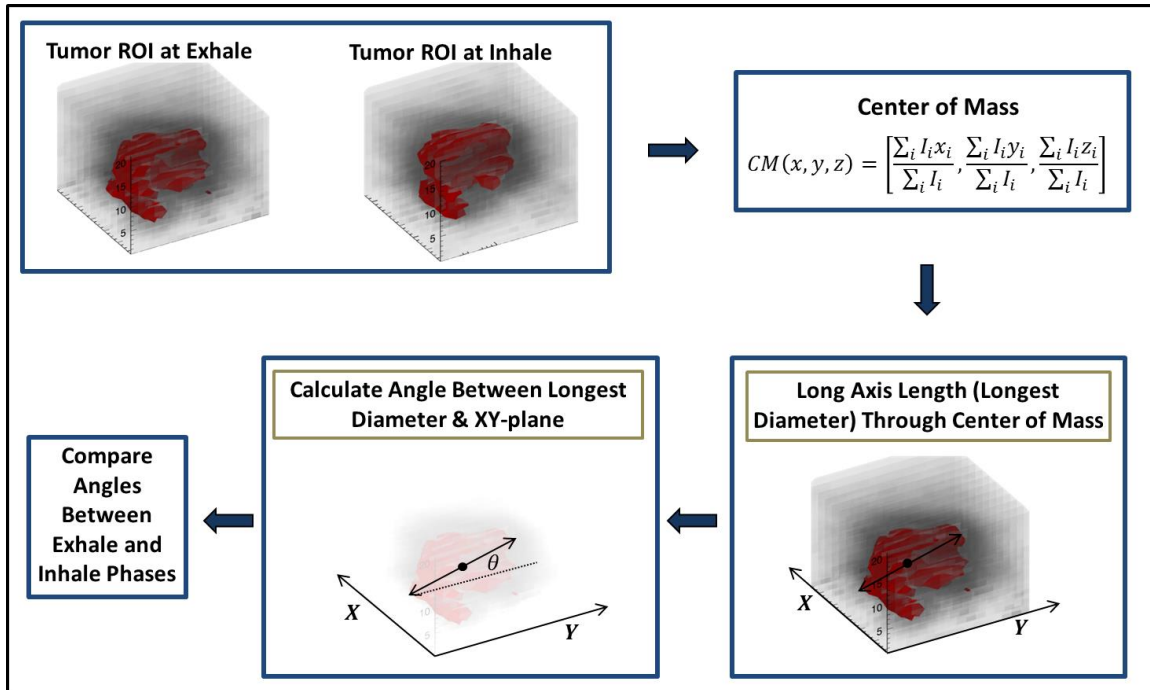


Figure 4.1 Method of tumor rotation calculation. First, the tumor volume is delineated at exhale (phase 1) and inhale (phase 5) on RG PET images. Second, the center of mass of each volume is calculated. The long axis length (longest diameter) through the center of mass of the tumor is calculated. Then, the angle between the long axis length and the XY-plane is calculated. This angle is compared between the exhale (phase 1) and inhale (phase 2) to determine the pseudo-tumor rotation.

Percent differences between 3D CT and RG CT varied from 0% to 176%; kurtosis again being the outlier. Figure 4.2 shows selected feature percent differences and Table 4.2 shows image feature average differences between 3D and RG CT. Forty-six percent of the CT features between 3D CT and RG CT presented average percent differences larger than 20%. In some cases average percent differences were larger than 50%. Table 4.3 displays the number and percent of total features with specific percent differences for CT, PET, and PET RG phases.



### 3D/4D PET/CT Feature Differences

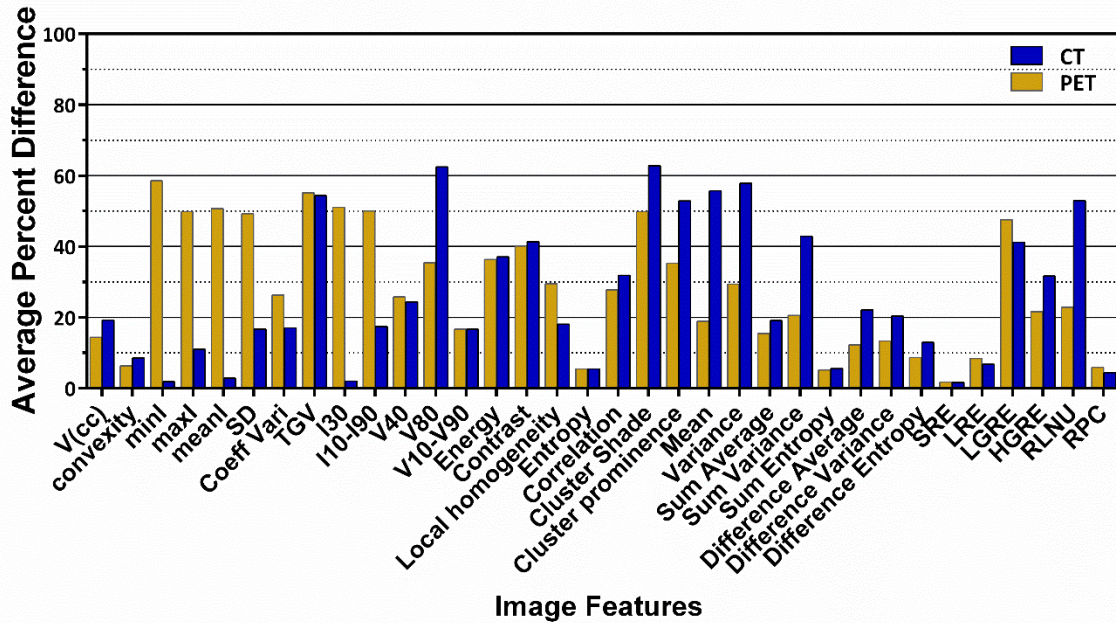


Figure 4.2 Average differences between 3D and Respiratory-Gated Image Features. %Diff<sub>i</sub><sup>3D/RG</sup> between selected image features from 3D PET/CT and RG PET/CT.

Table 4.1 Features Presenting Average Differences Between 3D and RG PET Image Features.

<2% Difference	<5% Difference	<10% Difference	<15% Difference	<20% Difference	>50% Difference
SRE	Sphericity	Surface Area/Volume	Volume	V10-V90	Minimum Intensity
	Spherical Disproportion	Compactness	Surface Area	Contrast (1st order)	Mean Intensity
	Entropy (1st order)	Convexity	Long Axis	Co-occurrence Mean	Kurtosis
	Information Measure of Correlation 2	Entropy (2nd order)	Short Axis	Sum Average	TGV
	RPC	Sum Entropy	Local Homogeneity (1st order)	Information Measure of Correlation 1	RMS
		Difference Entropy	Difference Average		I30
			Difference Variance		I10-I90
					LRLGE

Table 4.2 Features Presenting Average Differences Between 3D and RG CT Image Features.

<b>&lt;2% Difference</b>	<b>&lt;5% Difference</b>	<b>&lt;10% Difference</b>	<b>&lt;15% Difference</b>	<b>&lt;20% Difference</b>	<b>&gt;50% Difference</b>
Minimum Intensity	Mean Intensity	Convexity	Surface Area/Volume	Volume	Kurtosis
SRE	RMS	LRE	Sphericity	SD	TGV
	I30		Compactness	Coefficient of Variation	V70
	RPC		Spherical Disproportion	I10-I90	V80
			Difference Entropy	Local Homogeneity (2nd Order)	Energy (1 <sup>st</sup> Order)
				Sum Average	Cluster Shade
					Cluster Prominence
					Co-occurrence Mean
					Co-occurrence Variance
					GLNU
					RLNU

Overall, 249 of 952 (26.2%) of all PET features (56 features per patient) had a percent difference of less than 5% between 3D and RG protocols whereas, 342 of 1288 (26.6%) of all CT features (56 features per patient) had a percent difference of less than 5% between 3D and RG scans. Table 4.4 shows features that had percent differences between 3D and RG protocols for all cases for both PET and CT modalities.

Table 4.3 Percent Differences ( $\%Diff_i^{3D/RG}$ ) between Image Features of 3D and RG, PET and CT Images and Conglomerate Image Features of RG PET phases for all cases ( $\%Diff_{ij}^{Mean}$ ).

<i>Percent Difference</i>	<i>CT</i>		<i>PET</i>		<i>PET RG Phases</i>	
	<i>No. of Features<sup>a</sup> (1288 total)</i>	<i>% Total Features<sup>a</sup></i>	<i>No. of Features<sup>a</sup> (952 total)</i>	<i>% Total Features<sup>a</sup></i>	<i>No. of Features<sup>a</sup> (9464 total)</i>	<i>% of Total Features<sup>a</sup></i>
<5%	342	26.6%	249	26.2%	5051	53.4%
<10%	498	38.7%	405	42.5%	7258	76.7%
<15%	617	47.9%	515	54.1%	8043	85.0%
<20%	697	54.1%	585	61.4%	8410	88.9%
>20%	591	45.9%	367	38.6%	998	10.5%

<sup>a</sup>Total number of features refers to 56 image features per tumor.

According to the CCC strength-to-agreement scale by McBride et al., PET and CT feature subtypes demonstrated poor correlation between 3D and RG images (113) (see Figure 4.3). This was demonstrated by CCC strength-to-agreement values less than 0.90 for each feature subtype (shape, first order, GLCM, and RLM). However, there were specific features that demonstrated substantial strength-of-agreement. These were from the shape and first-order features in PET and shape features only in CT.

**Table 4.4 Image Features with Common Average Differences in 3D/RG PET and CT.**

<i>Percent Difference</i>	<i>Common Features</i>
<2%	SRE
<5%	--
<10%	Convexity, 1 <sup>st</sup> and 2 <sup>nd</sup> order Entropy, Sum Entropy, LRE, RPC
<15%	Surface Area/Volume, Sphericity, Compactness, Spherical Disproportion, Difference Entropy, Information Measure of Correlation 2
<20%	Volume, Long Axis Length, V10-V90, Sum Average
>50%	Kurtosis, TGV

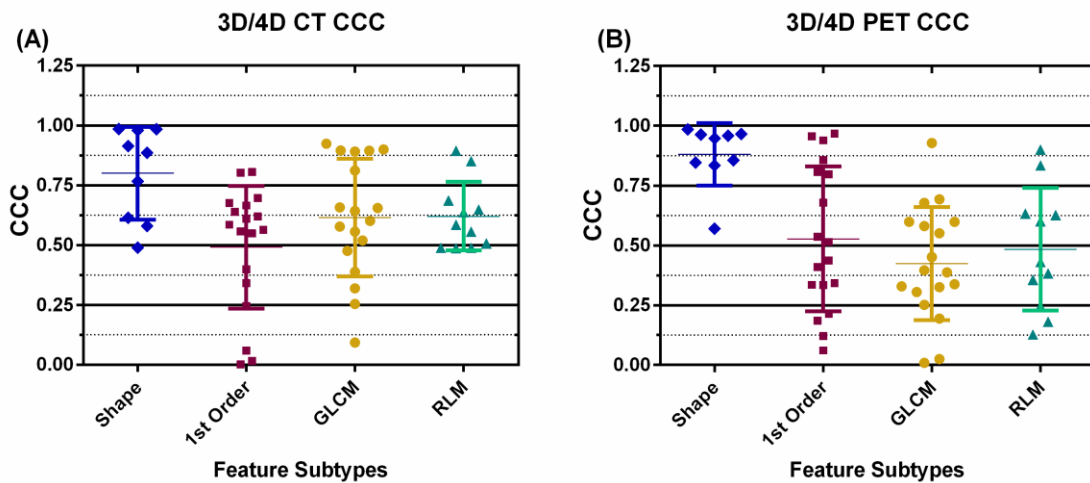


Figure 4.3 Concordance correlation coefficients for each feature with mean and standard deviation for each feature subtype for (A) 3D/4D CT and (B) 3D/4D PET.

The paired, two-tailed, t-test for 3D PET and RG PET features revealed 17 PET features with p values <0.05 (indicating that these datasets are different). The t-test for 3D CT and RG

CT features revealed 12 CT features with p values <0.05. Features with p values <0.05 for both PET and CT were entropy (1<sup>st</sup> order), compactness, and information measure of correlation 1 and 2.

#### **4.3.2 RG (4D) PET Phase Analysis**

Results indicated a weak dependency (relative to the differences between 3D and RG presented above) of all PET features on respiration phase in RG scans of 10 phases (see Figure 4.4). The most robust features (less than 5% difference among RG phases) belonged to select features from all categories (shape descriptors, first and second order features). Sphericity, spherical disproportion, information measure of correlation 2, SRE, and LRE were within 10% difference of the average value for all cases across all phases. Normalized image features across 10 phases for RG PET demonstrated that for all patients, 77% (7258:9464) of image features (56 features per phase per patient) varied less than 10% from the average values and 10.5% (998:9464) demonstrated more than 20% difference from average values (Table 3). Features with the largest difference (>50%) were kurtosis, LGRE, SRLGE, and LRLGE. The paired, two-tailed, t-test for RG PET inhale and RG PET exhale feature data revealed one PET feature, namely short axis length, with p value <0.05. The CCC revealed that the shape features had the highest CCC strength-to-agreement between image datasets from phases 1 and phase 5 (mean CCC strength-to-agreement 0.95; moderate). First order features and GLCM had mean strength-to-agreement values of 0.93 (moderate) and RLM features exhibited mean CCC strength-to-agreement of 0.86 (poor).

#### **4.3.3 Overall Feature Results**

Comparisons of results among respiratory-phases and 3D-to-RG PET features, we concluded that the features with least variability overall for PET images were sphericity, spherical disproportion, 1<sup>st</sup> order entropy, information measure of correlation 2, and SRE. Features demonstrating the greatest variability were kurtosis and LRLGE. For CT images, features with the least variability were minimum intensity, mean intensity, RMS, SRE, and RPC,

while features with the greatest variability were kurtosis, V70, V80, energy (1<sup>st</sup> order), cluster shade, cluster prominence, co-occurrence mean, co-occurrence variance, GLNU, and RLNU.

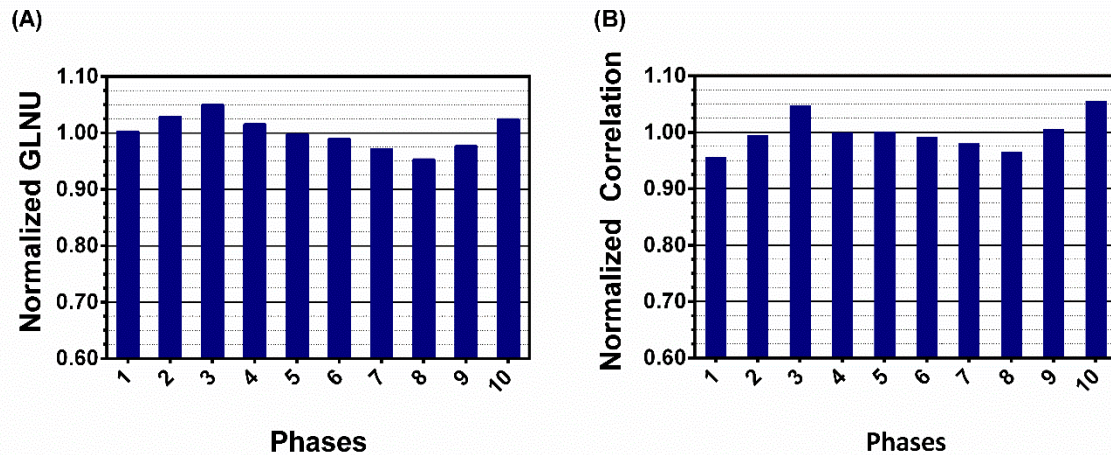


Figure 4.4 Feature dependency on respiration phase for selected features. (A) Normalized GLNU across 10 phases of RG PET image sets. (B) Normalized Correlation across 10 phases of a RG PET image set.

#### 4.3.4 Long Axis Tumor Length, Rotation, and Center of Mass Motion (CMM)

The long axis tumor length and rotation results demonstrated that tumors exhibited deformation over respiratory-gated phases. A Pearson's correlation test demonstrated that there was a weak correlation between the tumor angle with respect to the XY plane at inhale and the same angle in the corresponding 3D image ( $R=0.350$ ), and a weak correlation between the tumor angle at the exhale phase and the corresponding 3D image ( $R=0.319$ ). There was a weak correlation between 3D image tumor volume and the 3D image tumor angle ( $R=-0.399$ ) and long axis length was not correlated to the breathing cycle. Table 5 shows that the long axis length of the tumor was inconsistent across inhalation phase (phase 1), 3D scan and exhalation phase (phase 5). The long axis lengths of the tumor for 3D, phase 1, and phase 5 were highly correlated (3D and phase 1:  $R=0.936$ , 3D and phase 5:  $R=0.954$ , phase 1 and phase 5:  $R=0.986$ ), but long axis lengths between phase 1 and phase 5 varied indicating a possible change in tumor shape during the respiratory cycle. The largest difference was case 11 with long axis lengths of 124.5 mm and 139.9 mm for phase 1 and phase 5, respectively, while the

long axis length for 3D was 147.4 mm. There was a weak to moderate correlation between tumor angle at the inhale phase and the exhale phase ( $R=0.438$ ) indicating tumor rotation during the respiratory cycle. Moreover, the long axis angle changed from positive to negative indicating tumor rotation. There was also a weak to moderate negative correlation between average percent difference in 3D and RG images (in PET) and center of mass motion ( $R=-0.445$ ).

Table 4.5 Long Axis Lengths of Lung Tumors on 3D PET Images and RG PET Images at Exhale and Inhale.

Case	Length (mm)			Angle (relative to XY axis)			Volume (cc)			CMM (mm)
	3D	Exhale	Inhale	3D	Exhale	Inhale	3D	Exhale	Inhale	
1	31.58	46.29	35.65	18.10	-25.08	-39.95	12.57	12.79	12.36	3.70
2	67.19	69.78	66.97	-25.98	-48.58	-22.99	45.77	40.45	40.02	4.21
3	62.73	76.31	71.78	38.72	0.00	-43.10	82.56	81.77	77.24	6.22
4	24.67	25.85	24.92	7.62	-14.65	-7.54	4.69	4.44	4.14	1.87
5	41.00	41.98	40.38	-13.84	22.92	34.53	23.47	24.94	21.81	13.30
6	49.21	44.18	43.71	-36.73	31.20	-8.61	30.41	30.18	29.75	2.99
7	125.95	133.22	126.72	-65.32	-68.87	-61.33	140.78	119.62	113.54	3.35
8	55.03	46.76	46.66	-40.82	-24.81	-57.25	24.84	20.84	19.90	1.29
9	28.11	21.67	21.96	13.45	26.91	-17.33	6.45	4.38	4.31	0.28
10	29.05	24.78	21.96	19.74	-15.30	17.33	6.85	4.67	4.74	1.71
11	147.42	124.53	139.92	-36.79	6.03	-37.42	571.04	419.89	427.80	0.16
12	63.53	68.11	63.27	27.60	22.59	14.98	64.35	58.92	53.38	4.08
13	46.96	48.25	55.16	16.17	-7.79	-62.77	26.21	24.21	28.17	2.05
14	54.12	54.02	54.02	17.58	14.01	14.01	33.25	33.34	32.91	0.49
15	10.94	47.98	47.98	0.00	-15.82	-15.82	35.40	27.38	30.46	0.29
16	24.59	20.40	19.86	23.51	-18.70	-19.23	6.75	3.81	2.80	4.60
17	41.55	47.61	39.99	18.35	-33.33	19.09	26.01	22.70	22.99	2.84

#### 4.4 Discussion

RG PET scans can provide a “snapshot” of the tumor within a phase along the breathing cycle, thereby greatly reducing the effects of motion on a tumor’s shape, volume, and image feature values. In contrast, 3D PET scans convolute the absorbed activity distribution over the motion/deformation pattern a tumor and its surroundings experience during multiple respiration cycles (25). Consequently, a 3D (*static*) PET may fail to provide accurate position, volume and

absorbed activity distribution for a mobile tumor. This is especially true in the thoracic region and regions with substantial internal motion. This agrees with Adams' finding that respiratory motion affects the SUV with changes up to 30% and that any moving lesion would be inaccurately measured due to the effects of blurring (69). In addition, patient motion/breathing is known to cause image artifacts due to a mismatch in registration between CT attenuation correction and emission scans (109). Internal motion, as the results support, notably affected the image feature values of PET and CT images. The percent differences between 3D and RG CT were generally greater than those in PET. CT images have higher spatial resolution than PET images, and therefore more voxels for texture formation and thus a greater sensitivity to motion. In addition, 3D CT may also be affected by motion depending on the acquisition protocol (114).

In addition to the affine tumor motion caused by respiration, we identified deformation of tumors (characterized by varying tumor axis lengths and angles with respect to the XY-plane between 3D PET, RG PET at inhale, and RG PET at exhale). Conceivably, rotations and deformations also affect image feature values. Our results demonstrated a weak correlation between the long axis angles of RG images at inhalation and exhalation. There was also an inconsistency of long axis length between 3D images, RG images at inhale and RG images at exhale; thus indicating that tumor shape and rotation varied between phases. The degree to which rotations and/or deformations affect image features, and in particular texture values, requires further investigation. Interestingly, there was no correlation between CMM, tumor volume, or long axis length with 3D/RG feature value differences based on Pearson's correlation tests. There was, however, a weak to moderate correlation between CMM and average percent difference. Nonetheless, it is clear from our data that the feature value differences between RG phases are smaller than the differences between 3D images and RG images at a given phase. In other words, the rotational motion and/or deformation of the tumors in our patient cohort had a smaller effect on image feature values than the averaging effects of

the static acquisition. Yip et al. also investigated variability of texture features between 3D and RG imaging (107). In contrast to our study, they tested only five image features (contrast, busyness, coarseness, maximal correlation coefficient, and long run low gray). They found that differences between 3D and RG PET were significant (107) after having accounted for noise differences due to different acquisition times. This agrees with our findings that certain features (e.g., kurtosis and LRLGE) demonstrated large variability between 3D and RG protocols. There were, however, certain features in our study (e.g., SRE, 1<sup>st</sup> order entropy, and RPC) that did not demonstrate large variability between protocols.

Figure 4.2 and Tables 4.1-4.3 show differences between feature values from 3D and RG protocols. The features with the smallest change across PET for all RG bins and for 3D PET were sphericity, spherical disproportion, entropy (1<sup>st</sup> and 2<sup>nd</sup> order), sum entropy, information measure of correlation 2, SRE, LRE, and RPC. Interestingly, a study by Galavis *et al.* on the variability of PET texture features caused by different acquisition modes and reconstruction parameters demonstrated that 1<sup>st</sup> order entropy exhibited small variation ( $\leq 5\%$ ) while 2<sup>nd</sup> order entropy, and sum entropy exhibited intermediate variability (10%-25%) (45). Our results were comparable, showing that 1<sup>st</sup> order entropy exhibited variation smaller than 5% and that 2<sup>nd</sup> order entropy and sum entropy exhibited less than 10% difference between 3D and RG PET protocols. Sum entropy, 2<sup>nd</sup> order entropy, and the information measure of correlation 2 are based on entropy calculations which measure randomness in a pattern. A portion of the randomness can be attributed to the noise intrinsic to the scanner while the remaining can be attributed to statistical differences in counts (quantum noise). Hence 3D images are less noisy than RG images since percentage image noise is given by  $(1/\sqrt{N}) \cdot 100$ , where N is the count density (counts/cm<sup>2</sup>). Thus, 3D/RG feature differences are a combination of both tumor motion and count statistics. This suggests that it would be informative to normalize for count density. Unfortunately, this study was retrospective and list-mode data were not accessible for normalization. Nevertheless, the number of counts and therefore the noise among RG images



from the 10 phases can be assumed similar. Therefore, the differences in feature values from phase to phase may be attributed to the effect of motion and/or deformation.

The features LRE, SRE, and RPC, which demonstrated small change across PET for all RG bins and 3D PET are features from the run length matrix. LRE measures the long run emphasis distribution. Correspondingly, SRE measures the short run emphasis distribution. Run percentage is the ratio of the number of runs to the number of pixels in an image (Table 4.5). We conclude that the cumulative number and length of short runs and cumulative number and length of long runs does not vary significantly between 3D and RG images, and that the total number of runs does not vary significantly between 3D and RG images. These conclusions may depend on the algorithms used to calculate these features. For example, in this paper we averaged runs from 13 directions; other definitions are possible.

Feature differences between 3D and RG in PET and CT images that showed large differences (>50%) were typically features from intensity volume histograms such as kurtosis and TGV. Thus, the intensity histogram distributions between 3D and RG features were quite different in terms of symmetry about their means and the degrees of “peakness” of their distributions. Cluster shade and cluster prominence exhibited large differences in CT. These features measure the skewness of the GLCM (93). According to Ion, a high cluster shade value reveals an asymmetric image (93).

Overall, it is clear that image feature values are different between 3D and RG images. As discussed above this is due to both the smearing effects of tumor motion, both affine and non-affine, and noise intrinsic to image acquisition—the former apparently having larger effect (107). This is also supported by the relative variation in feature values from different phases of the RG scans even though the tumor VOIs varied from phase to phase due to motion and deformation. Thus, the motion convoluted into the 3D images seems to have a greater effect on feature values than noise, given that the RG images are intrinsically noisier due to lower counts (acquisition times). This study suggests that it would be important to account for motion in

quantitative image feature analysis, regardless of modality (PET or CT), as attempted by other investigators (38). Alternatively, if the definition of any one feature includes details of the acquisition protocol, then 3D and RG features may be treated as “different” sets of features. Further studies are needed to elucidate the potential usefulness of this alternative definition.

#### **4.4.1 Limitations**

Though our results clearly demonstrated that image feature values were different between 3D and RG protocols, there were limitations to the study. First of all, we were unable to normalize for count density between 3D and RG protocols. Another limitation was the nonconformity of the uptake time with the protocol. This was mainly due to clinical logistics. Also, partial volume effects were not taken into account. Since 3D and respiratory-gated data on same patient were acquired on the same scanner, and hence partial volume effects were similar in both sets of images except for the effect of motion, we did not take these affects into account. In addition, binning artifacts and breathing irregularities were assumed negligible since only patients with regular breathing patterns are candidates for RG PET for radiotherapy in our institution (115). Another limitation was that 4D PET received 3D CT attenuation correction. This is currently standard procedure at our institution. Lastly, our patient size was limited, but comparable to other published studies (37, 38, 107). We plan to address these limitations in future studies.

#### **4.5 Conclusions**

This study investigated the variation of image features between 3D and respiratory-gated PET/CT images of lung tumors. To our knowledge, this is the first study that evaluates how 3D and RG acquisitions affect a large number of image features currently being used and tested in several medical applications. The data showed that image feature analysis using a static acquisition (3D) versus a respiratory-gated acquisition (to account for motion of the ROI) revealed notably different feature values. The results support that these differences are mainly due to the effect that respiratory motion has on image features. We have also concluded that

rotational motion and deformation of the tumor also affect the features of an image. However, the effect of rotational motion and deformation from phase to phase appear to be smaller than the averaging/smearing effects of static acquisition. In sum, this study calls attention to the differences in 3D and RG image feature values for mobile tumors. The predictive and/or prognostic power of RG versus 3D image feature values will be explored in future studies.

## Chapter Five: Sensitivity of Image Features to Noise in Conventional and Respiratory-Gated PET/CT Images of Lung Cancer: Uncorrelated Noise Effects<sup>5</sup>

### 5.1 Introduction

Clinical imaging by Positron Emission Tomography (PET) and Computed Tomography (CT) is evolving into a quantitative discipline where a large number of metrics are computed in the intensity and gray-level matrix domains; this discipline has been termed Radiomics (36, 42). Radiomics of CT & PET images have shown promise as a diagnostic, prognostic, and predictive tool in cancer treatment (13, 33, 36, 40, 116). It is also being combined with other “omics” (e.g., genomics, transcriptomics, proteomics, metabolomics) into decision support systems (84). However, features are sensitive to various acquisition conditions (scanner type, image reconstruction algorithm, etc.) (13, 33, 36, 40, 45, 116). One major confounding factor introduced by these conditions is the presence of various random contributions to the signal, commonly referred to as noise. However, few authors have examined the impact that quantum or electronic noise can have on Radiomic features. In this paper, we examine the influence of electronic noise, a signal independent contributor to image noise, on Radiomics.

An image feature is a quantity that provides quantitative information about an image. It can be derived directly from the image (1<sup>st</sup> order), or from heterogeneity matrices that are derived from the image (2<sup>nd</sup> order). Image features or “metrics” that describe image texture and heterogeneity analyze relationships between voxel pairs or groups of voxels. When noise is introduced into an image the fundamental relationships between voxels are altered. As a result, the image metrics are also altered and the texture or heterogeneity of the object may be misrepresented. If not accounted for, this noise can have significant implications on the clinical utility of image features. Although there are protocols for the standardization of PET/CT

---

<sup>5</sup> Portions of this chapter have been submitted for publication.

imaging, image noise varies between scanners, manufacturers, and institutions (109, 117, 118). Thus, the impact of noise on image features may adversely affect multi-institutional studies involving radiomics.

Although the focus of this study is on image noise, motion is also a factor that affects image quality. It can affect SUV by up to 30% and causes image artifacts because of registration mismatches in the attenuation correction (CT) and emission scans (69, 109). In PET, respiratory-gated (RG or 4D) images tend to have higher levels of noise because of the smaller number of counts (due to shorter acquisition times per bed position), but the quality of RG images are impacted less by motion. Both conventional (3D) and respiratory-gated (4D) images are included in this study.

Since the goal of Radiomics is the clinical application of image features, it is important to carefully characterize image features and to understand how they might be influenced by various clinical situations with varying levels of noise. The goal of this study is to evaluate the effects of noise on image features.

## **5.2 Materials and Methods**

### **5.2.1 Phantom Study**

A standard ACR accreditation phantom with a Germanium (Ge)-68 cylindrical insert (Benchmark by RadQual LLC, Weare, NH, SN: BMCY06813067103), was placed on a motion table with 2.4 cm motion amplitude and a 4 second period to simulate lung tumor motion due to the respiratory cycle. The phantom was imaged with three protocols: 1) 3D PET/CT with motion, 2) 4D PET/CT with RG motion, and 3) 3D PET/CT without motion (static).

#### **5.2.1.1 Noise Application**

To assess the implications of electronic noise on image features, uncorrelated Gaussian noise with varying standard deviations was added to PET and CT patient and phantom images. A custom program was used to apply noise with varying standard deviation to phantom images using the following Gaussian function ( $\rho_g$ , Equation 5.1).

$$\left[ p_g(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \right] \quad (5.1)$$

where  $\mu$  is mean noise added,  $\sigma$  is standard deviation, and  $z$  is gray level. CT noise images were created with standard deviations of 10, 20, 50, 80, and 120 Hounsfield Units (HU). These will be referred to as GN10, GN20, GN50, GN80, and GN120 henceforth. PET noise images were created with standard deviations of 2.5%, 4.0% and 6.0% of the maximum intensity (not SUV). These will be referred to as low noise, medium noise, and high noise. PET images had varying standard deviations due to the variation in maximum intensities (not an issue in phantoms but very significant in patient images). Although the GN120 noise level may represent greater levels of electronic noise than expected in a scanner, we believe this was necessary to clearly distinguish noise-affected features. The low noise levels of GN10 and GN20 were included to demonstrate how small noise levels affect image feature analysis. The higher noise levels were included to show the gross effects of noise on feature analysis. As demonstrated by Latifi et al., low-dose 4D CT settings sometimes involve high levels of noise (119).

The specific activity (SA) of the ACR phantom was calculated using the activity on the date of source production (108274.4 Bq/cm<sup>3</sup> on 03/20/2013), the date of measurement (02/18/2014), and the volume of the source (58.1 cm<sup>3</sup>). The measured specific activity was calculated using Mirada DBx (Mirada RTx, Mirada Medical, Oxford, UK). Mean counts (105527 Bq/cm<sup>3</sup>) and standard deviation (2927.7 Bq/cm<sup>3</sup>) were obtained from the cylindrical ROI (24.4 cm<sup>3</sup>) inside the volume-of-interest (VOI). The reciprocal of the coefficient of variation, or SNR, which was 2.7% from the noise contribution in the phantom data, was calculated.

### **5.2.1.2 Signal-to-Noise Ratio (SNR) and Noise Power Spectrum (NPS)**

To understand the noise inherent in the images and to quantify the noise added to the images, the signal-to-noise ratio and noise power spectrum of the scanner was calculated. The noise power spectrum (NPS), representative of the noise texture of an image (used primarily in CT modalities) was calculated using the Ge-68 phantom. The signal-to-noise ratio (SNR),

representative of the amplitude of noise in an image was also calculated on the phantom for both PET and CT modalities to verify and quantify the addition of noise to the images.

The SNR of a GE Discovery STE PET/CT Scanner was measured with the Ge-68 phantom with activities of 0.62 mCi and 0.79 mCi. The phantom was scanned with 70 cm field-of-view (FOV), 120 kV, 210 mA, 28 subsets, 2 iterations, and FWHM of 7 mm for 3D PET/CT and 60 cm FOV, 120 kV, 200 mA, 28 subsets, 2 iterations, and FWHM of 7mm for 4D PET/CT. To calculate the CT SNR, medical imaging software (Mirada RTx, Mirada Medical, Oxford, UK) was used to draw five 4 cm spheres onto the phantom image (figure 5.1a). For PET SNR, two 4 cm spheres were drawn on the phantom image. One sphere was inside the Ge-68 source and the other was in a non-radioactive region inside the phantom (figure 5.1b). The SNR was calculated using Equations 5.2-5.4:

$$signal \rightarrow [ s = \sum_{i=1}^n \overline{HU}_i ] \quad (5.2)$$

$$noise \rightarrow \left[ \bar{\sigma} = \sqrt{\sum_{i=1}^n \sigma_i^2} \right] \quad (5.3)$$

$$SNR \rightarrow \left[ SNR = \frac{s}{\bar{\sigma}} \right] \quad (5.4)$$

where  $s$  is the signal,  $\overline{HU}_i$  is the mean HU for region  $i$ ;  $n$  is the number of regions and  $\bar{\sigma}$  is the mean standard deviation across all VOIs;  $\sigma_i$  is the standard deviation for region  $i$ .

A CT image of the Ge-68 phantom was used to calculate the NPS of the GE PET/CT Scanner at our institution. Ten axial slices; four regions per slice were selected in the uniform region of the phantom. A gain correction was applied by subtracting the mean value of the regions and the Fourier transform was applied to each region to create a two-dimensional noise power spectrum. Forthwith, a one-dimensional NPS was plotted from the two-dimensional noise power data (figure 5.2) by radial averaging.

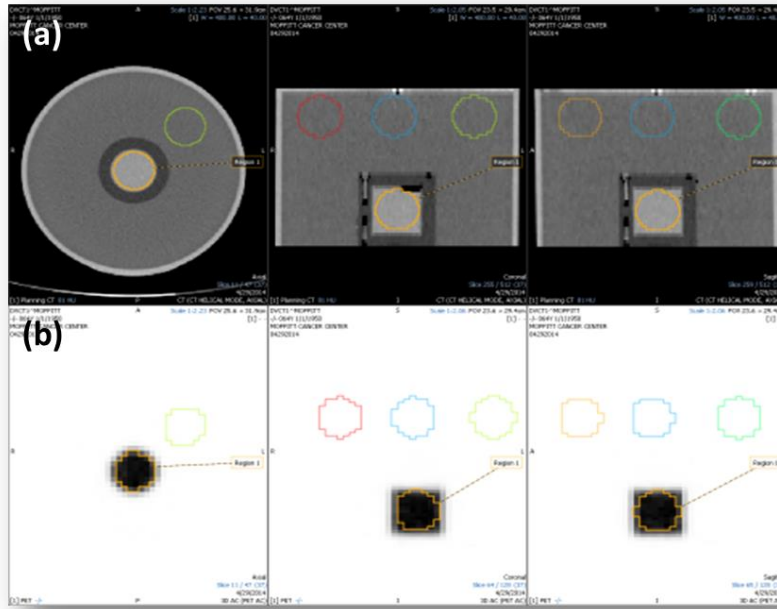


Figure 5.1 Method of SNR Calculation for (a) CT and (b) PET using Ge-68 solid epoxy phantom.

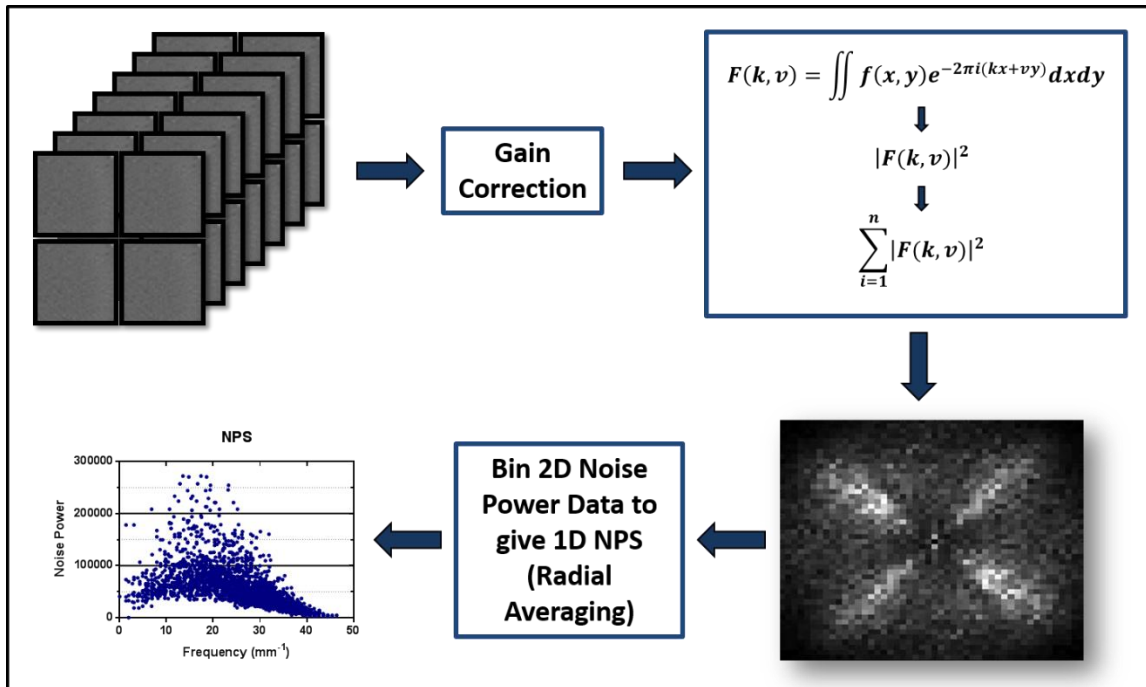


Figure 5.2 Method to measure CT noise power spectrum using Ge-68 solid epoxy phantom: 1) Gather 4D independent images. 2) Apply gain & offset correction and subtract mean. 4a) Take 2D Fourier Transform of each image. Take magnitude of results & square it. Calculate Average of 40 results. 5) Calculate 1D NPS by radial average of the 2D result.



### 5.2.2 Patient Study

Twenty-six non-small cell lung cancer (NSCLC) patients with 3D and RG PET/CT images were retrospectively selected for this study with ages from 47 to 83 years, eleven males and fifteen females. This study was approved with waived informed consent by the University of South Florida Institutional Review Board #105996. Standard of practice procedures at our institution were followed. Gaussian noise was applied to all 3D and one RG phase of PET and CT patient images according to the method described previously (equation 1). Resulting PET and CT datasets consisted of four separate image sets for each patient, an original image dataset and image datasets of low, medium, and high noise for 3D PET, 4D PET, 3D CT and 4D CT. Figure 5.3 demonstrates the noise levels for PET and CT for one case (coronal view).

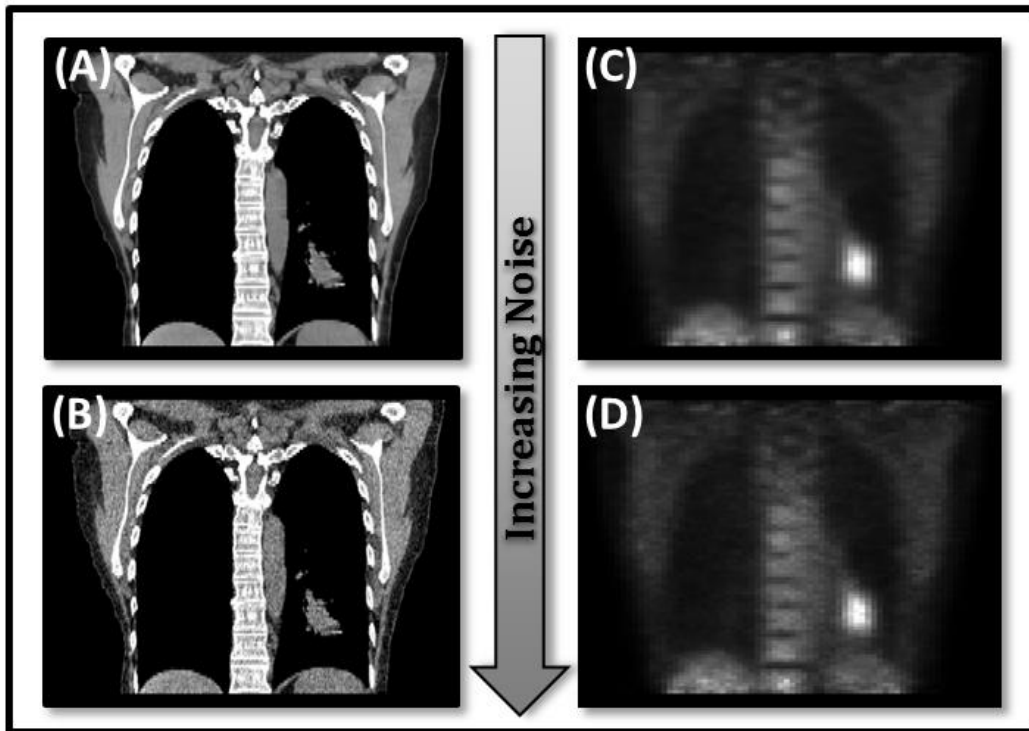


Figure 5.3 One coronal slice of a PET/CT image with and without noise. (A) Original image. (B) CT image with Gaussian noise ( $\sigma = 120$  HU). (C) PET image without noise. (D) PET image with Gaussian noise ( $\sigma = 0.06 \times \text{maximum intensity}$ ).

### 5.2.2.1 Feature Extraction

The original and noise-added image sets were imported, viewed, and contoured with Mirada Medical Software (RTx, Mirada Medical, Oxford, UK). Lung tumor contours were acquired separately for noise and original images. PET image tumors were contoured at 40% maximum intensity inside a defined volume of interest. On CT images, tumors were contoured with CT region segmentation. An in-house program extracted image features for the region represented inside each contour. Eighty-one image features were extracted: 11 shape features, 22 intensity features, 26 Gray Level Co-occurrence Matrix (GLCM) features, 11 Run Length Matrix (RLM) features and 11 Gray Level Size Zone Matrix (GLSZM) features (29, 63, 64, 111) (see Table 2.1 for a complete list.) The dimension of the co-occurrence matrices were 128x128, calculated based on the 3D images with a step size of 1 voxel in 13 directions. The gray levels were binned into 128 levels with equal intensity intervals for the run-length matrices. The run length was calculated with the 3D images in 13 directions. These 13 directions are defined by Xu *et. al* (110, 112).

### 5.2.2.2 Statistical Analysis on Patient Data

Conventional and 4D PET and CT image feature differences were evaluated separately resulting in four datasets: 3D PET, 3D CT, 4D PET, and 4D CT. Features from original images were compared to image sets with varying Gaussian noise levels for each case. Percent difference (Equation 5.5) was used to compare image features extracted from noise images (low, medium, and high noise) and original images.

$$\left[ \%Diff_{ij}^{noise} = 100 \times \left| \frac{NV_{ij} - OV_{ij}}{OV_{ij}} \right| \right] \quad (5.5)$$

where  $NV_{ij}$  is the value of feature  $j$  at noise level  $i$  and  $OV_{ij}$  is the value of feature  $j$  from the original image. The percent differences were averaged for each level of noise across all patients. Features varying on average by more than 100% were considered “non-robust” (those features that are not reliable or reproducible across noise). Those varying less than 10% were

considered “robust”. Features were classified into one of 11 categories for CT: R1, R2, R3, R4, R5, NR1, NR2, NR3, NR4, NR5, and B; and 7 categories for PET: R1, R2, R3, NR1, NR2, NR3, and B. These are defined in Table 5.1.

Table 5.1 CT and PET Feature Classifications.

Feature Classifications	
R1	$\%Diff < 10\%$ for CT noise level GN120 and PET highest noise level; Robust1
R2	$\%Diff < 10\%$ for CT noise level GN80 and PET mid noise level; Robust2
R3	$\%Diff < 10\%$ for CT noise level GN50 and PET lowest noise level; Robust3
R4	$\%Diff < 10\%$ for CT noise level GN20; Robust4
R5	$\%Diff < 10\%$ for CT noise level GN10; Robust5
NR1	$\%Diff > 100\%$ for CT noise level GN10 and PET lowest noise level; Non-robust1
NR2	$\%Diff > 100\%$ for CT noise level GN20 and PET mid noise level; Non-robust2
NR3	$\%Diff > 100\%$ for CT noise level GN50 and PET highest noise level; Non-robust3
NR4	$\%Diff > 100\%$ for CT noise level GN80; Non-robust4
NR5	$\%Diff > 100\%$ for CT noise level GN120; Non-robust5
B	$10\% < \%Diff < 100\%$ for CT noise level GN10 and PET lowest noise level

In addition to percent difference, the concordance correlation coefficient (CCC) was calculated for each feature between different levels of noise and the original image to assess whether feature values were consistent for different noise values. The strength of agreement classification is demonstrated in Table 5.2 (113). The mean CCC for each feature subtype was calculated and the median and range for each feature across noise levels were plotted.

Table 5.2 CCC Strength of Agreement Scale.

Strength of Agreement	CCC Score
High	>0.99
Substantial	0.95-0.99
Moderate	0.90-0.95
Poor	<0.90

\*This scale originated from McBride (113).

## 5.3 Results

### 5.3.1 Phantom Study

The SNR for PET and CT images behaved as expected by decreasing as noise level increased (Figure 5.4) indicating that uncorrelated Gaussian noise increased the image noise. The measured NPS (Figure 5.5) showed the noise texture associated with the scanner at our institution (using the phantom CT image). The NPS for the original CT image was spatial frequency dependent, indicating correlated noise texture (Figure 5.5a). Whereas, the noise power spectra of the CT images with high levels of added noise were spatial frequency independent indicating that the addition of Gaussian noise overwhelmed the correlated noise inherent to the image generation process (Figures 5.5d, 5.5e, 5.5f). The noise power spectra of the low noise level images, GN10 and GN20 (Figures 5.5b, 5.5c) were not completely spatial frequency independent demonstrating that the inherent scanner noise power was still represented in part at these levels.

### 5.3.2 Patient Study

The automatic contouring via intensity threshold in the lung was not significantly affected by the noise. The R1 ( $\%Diff < 10\%$  for highest added noise, Table 5.3) features with ( $\%Diff < 2\%$ ) for 3D CT were minimum intensity ( $<2\%$ ), peak intensity ( $<2\%$ ), mean intensity ( $<1\%$ ), root-mean-square ( $<2\%$ , RMS), I30 ( $<1\%$ , intensity ranging from lowest to 30% highest intensity volume), 1<sup>st</sup> order entropy ( $<1\%$ ) and inverse difference moment ( $<2\%$ ). The results were comparable for 4D CT (see Table 5.3). In addition to the 3D CT R1 features, the 4D CT R1 features included short axis, eccentricity, max intensity, V10-V90 (percentage volume with at least 10% intensity minus percentage volume with at least 90% intensity), and histogram entropy. However, the 4D R1 features did not include 1<sup>st</sup> order contrast and local homogeneity. Minimum intensity, peak intensity, mean intensity, RMS, I30, 1<sup>st</sup> order entropy, and inverse difference moment exhibited differences less than 1% for 4D CT. No features from the GLSZM were categorized as R1 features from 3D or 4D CT.

Non-robust features were defined as features that exhibited  $\%Diff > 100\%$  for the lowest level of added noise (NR1, Table 5.4). The 3D CT NR1 features included V70 (113%, percentage volume with at least 70% intensity) and V80 (278%, percentage volume with at least 80% intensity) from intensity features, as well as large-area emphasis (LAE, 105%), low-intensity emphasis (LIE, 375%), low-intensity large-area emphasis (LILAE, 410%), high-intensity large-area emphasis (HILAE, 184%), and intensity variability (IV, 135%) from the GLSZM. The CT NR1 features for 4D CT included V40 (138%, percentage volume with at least 40% intensity) and variance (182%) from the intensity features and SAE (121%, small-area emphasis), LIE (115%), LISAE (159%, low-intensity small-area emphasis), HISAE (245%, high-intensity small-area emphasis), LILAE (136%), and HILAE (930%) from the GLSZM.

For PET (3D and RG), shape features that depended solely on automatically drawn contours were the most stable. R1 features exhibiting  $\%Diff < 2\%$  included: surface/volume, sphericity, spherical disproportion, mean intensity, RMS, I30, 1<sup>st</sup> order entropy, 1<sup>st</sup> order local homogeneity, histogram entropy, entropy (<2%), inverse difference moment (<1%), inverse difference (<1%), sum average, sum entropy (<2%), information measure of correlation 2 (<2%), SRE (<0.5%). There was one GLSZM R1 feature: zone percentage (ZP). In 4D PET, the R1 shape and intensity features were the same as 3D PET, excluding contrast (1<sup>st</sup> order). GLCM R1 features were the same as R1 features. Entropy, inverse difference moment (<1%), inverse difference (<1%), sum entropy, and information measure of correlation 2 exhibited  $\%Diff < 2\%$ . RLM R1 features included SRE (<0.5%), LRE, high gray-level run emphasis (HGRE), short run high gray-level emphasis (SRHGE), long run high gray-level emphasis (LRHGE), gray-level non-uniformity (GLNU), RLNU, and RPC. The GLSZM R1 feature was ZP (the same as in 3D PET).

The non-robust features from the lowest level of noise (NR1) from 3D PET included LGRE (167%), short run low gray-level emphasis (SRLGE, 168%), and long run low gray-level emphasis (LRLGE, 164%) from the RLM and LIE (275%), LISAE (253%), LILAE (1437%) from

the GLSZM. There were fewer NR1 features from 4D PET. These features were from the GLSZM and included LIE (541%), LISAE (701%), and LILAE (610%).

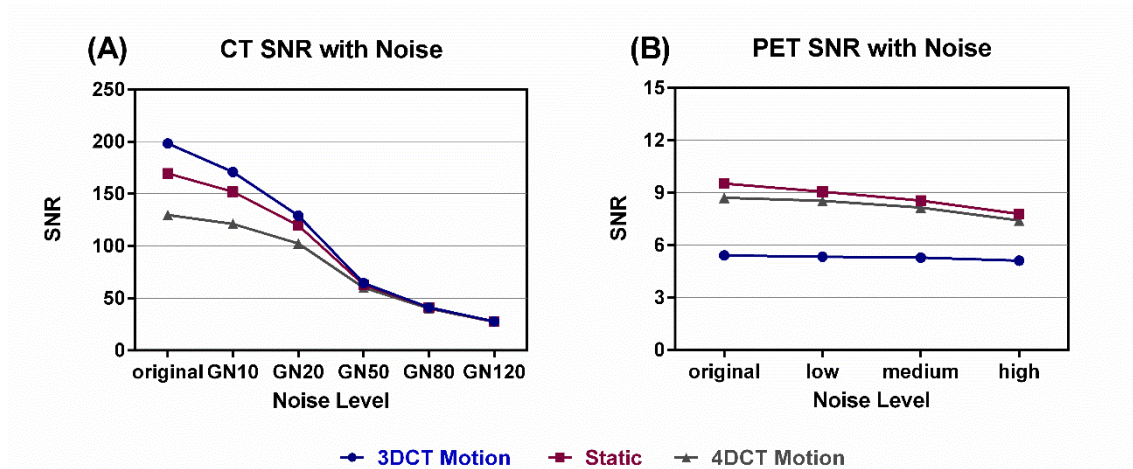


Figure 5.4 Signal to noise ratios for: (A) 3D and 4D CT (one phase) of phantom with motion table and static (3D only). (B) 3D and 4D PET phantom with motion table and static (3D only).

Figure 5.6 shows the trend between average percent differences for feature subgroups. For PET, shape, intensity, and GLCM features demonstrate an increase in difference with added noise. In CT this trend applied only to GLCM and GLSZM features, only with 3D CT. However, in both PET and CT, shape features exhibit the least change with uncorrelated noise (<4% average difference in PET and <11% average difference CT) and GLSZM features were the most sensitive to uncorrelated noise.

The CCC values further demonstrated that feature subtypes responded differently to added noise. GLSZM features demonstrated average CCCs below 0.90 for all modalities and all levels of noise (<0.70 for PET and <0.62 for CT). This demonstrated a poor agreement between the image features from noise and original images for GLSZM and supports our percent difference results. There was a discrepancy between PET and CT with the RLM CCC scores. In PET (3D and 4D), the RLM features demonstrated the highest CCC values across noise levels, followed by the shape descriptors, first order features, GLCM, and GLSZM (except for the medium level of noise where GLCM has a higher average CCC than first order features). Although the feature subtypes had an order, the distinction was not pronounced. Besides the

GLSZM features, all other average CCCs were greater than 0.95 and thus there was a substantial strength-to-agreement between these features derived from noise images and original images for feature subtypes excluding GLSZM.

In CT, the average CCCs were highest for shape descriptors, followed by the first-order features, GLCM, RLM, and GLSZM. Unlike PET, there was a clear distinction between the CCC values for different feature subtypes. Figures 5.7 and 5.8 demonstrate the median CCCs for each feature across noise with the ranges (min to max) for 3D CT, 4D CT, 3D PET, and 4D PET.

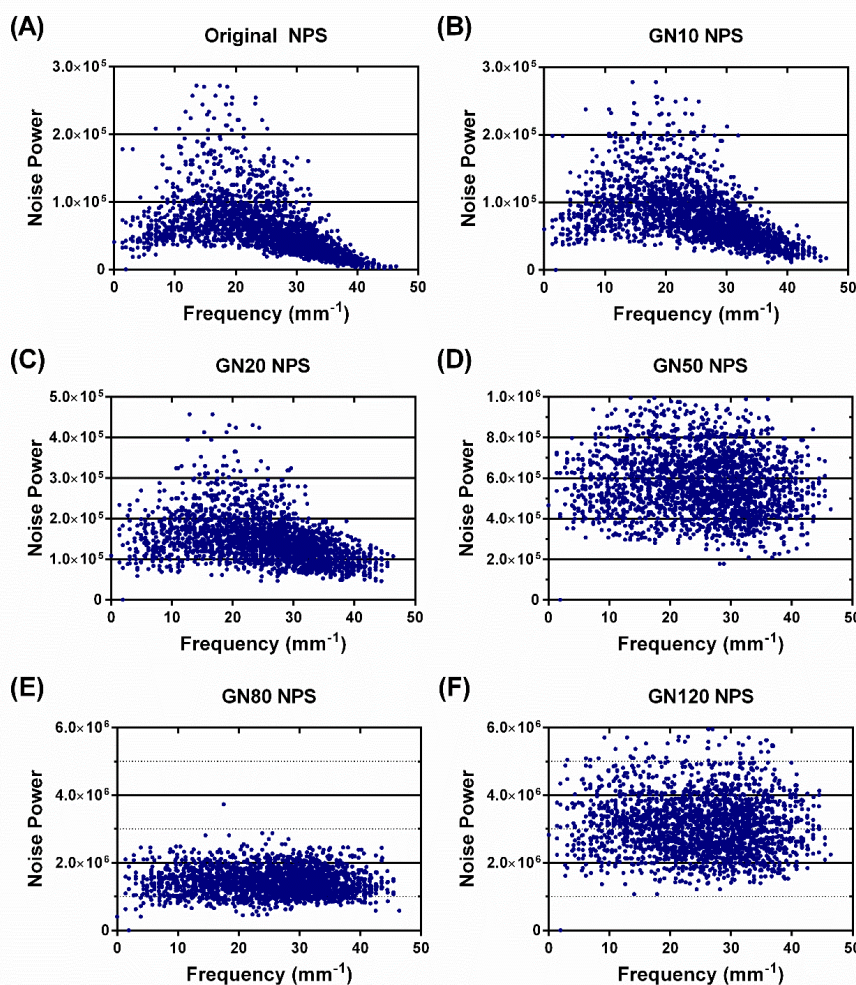


Figure 5.5 NPS of Ge-68 solid epoxy phantom for GE Discovery STE PET/CT scanner with increasing image noise. (A) Original NPS of CT image of phantom. (B) NPS of CT image with added Gaussian noise ( $\sigma=10$  HU). (C) NPS of CT image with added Gaussian noise ( $\sigma=20$  HU). (D) NPS of CT image with added Gaussian noise ( $\sigma=50$  HU). (E) NPS of CT image with added Gaussian noise ( $\sigma=80$  HU). (F) NPS of CT image with added Gaussian noise ( $\sigma=120$  HU). Note: Y-scale changes from (A) to (F).

Table 5.3 R1 Features (%Diff < 10%) for CT and PET.

Subtype	Feature	3D CT	4D CT	3D PET	4D PET
SHAPE	Volume	✓	✓	✓	✓
	Surface Area	✓	✓	✓	✓
	Surface Area/Volume	✓	✓	✓	✓
	Sphericity	✓	✓	✓	✓
	Compactness	✓	✓	✓	✓
	Spherical Disproportion	✓	✓	✓	✓
	Long Axis	✓	✓	✓	✓
	Short Axis		✓	✓	✓
	Eccentricity		✓	✓	✓
Convexity	✓	✓	✓	✓	
INTENSITY	Minimum Intensity	✓	✓	✓	✓
	Maximum Intensity	✓	✓	✓	✓
	Peak Intensity	✓	✓	✓	✓
	Mean Intensity	✓	✓	✓	✓
	Standard Deviation			✓	✓
	Skewness			✓	✓
	Coefficient of Variation			✓	✓
	TGV	✓	✓	✓	✓
	RMS	✓	✓	✓	✓
	I30	✓	✓	✓	✓
	I10-I90			✓	✓
	V10-V90		✓	✓	✓
	1st Order Energy	✓	✓	✓	✓
	1st Order Entropy	✓	✓	✓	✓
	1st Order Contrast	✓		✓	✓
1st Order Local Homogeneity	✓		✓	✓	
Histogram Entropy	✓	✓	✓	✓	
Uniformity			✓	✓	
GLCM	Homogeneity			✓	✓
	2nd Order Entropy			✓	✓
	Dissimilarity			✓	✓
	Co-occurrence Mean	✓	✓	✓	✓
	Inverse Difference Moment	✓	✓	✓	✓
	Inverse Difference	✓	✓	✓	✓
	Sum Average			✓	✓
	Sum Entropy			✓	✓
	Difference Average			✓	✓
	Difference Variance			✓	✓
	Difference Entropy			✓	✓
Info Correlation 1			✓	✓	
Info Correlation 2			✓	✓	
RLM	SRE	✓	✓	✓	✓
	LRE			✓	✓
	HGRE				✓
	SRHGE				✓
	LRHGE				✓
	GLNU				✓
	RLNU			✓	✓
GLSZM	RPC	✓	✓	✓	✓
	ZP			✓	✓

Abbreviations: GLCM = gray-level co-occurrence matrix; RLM = run length matrix; GLSZM = gray-level size zone matrix

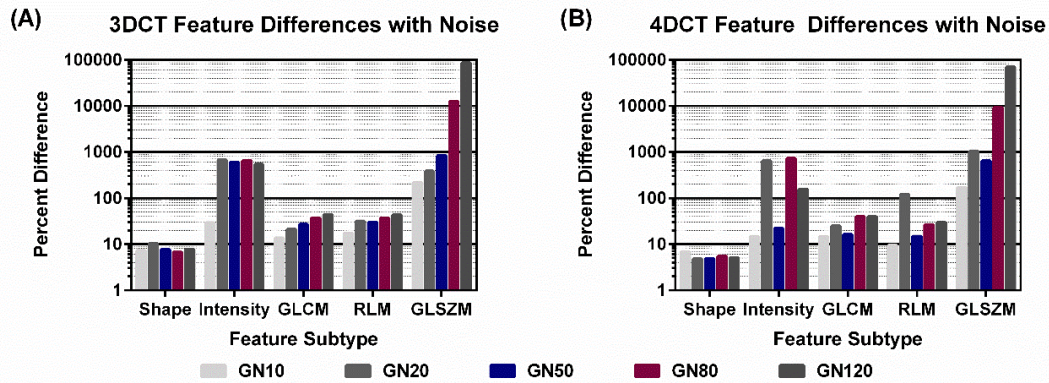


Table 5.4 NR1 Features (%Diff > 100%) for PET and CT.

Subtype	Feature	3D CT	4D CT	3D PET	4D PET
INTENSITY	V40		✓		
	V70	✓			
	V80	✓			
GLCM	Co-occurrence Variance		✓		
RLM	LGRE			✓	
	SRLGE			✓	
	LRLGE			✓	
GLSZM	SAE		✓		
	LAE	✓			
	LIE	✓	✓	✓	✓
	LISAE			✓	✓
	HISAE		✓		
	LILAE	✓	✓	✓	✓
	HILAE	✓	✓		
	IV	✓			

Abbreviations: GLCM = gray-level co-occurrence matrix; RLM = run length matrix; GLSZM = gray-level size zone matrix

### CT IMAGE FEATURE % DIFF WITH NOISE



### PET IMAGE FEATURE % DIFF WITH NOISE

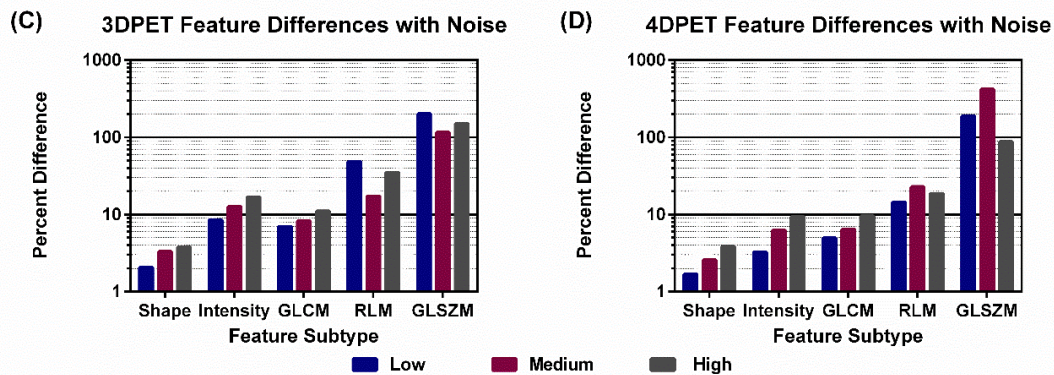


Figure 5.6 Average percent differences between noise and original images across feature sub-types for low, medium, and high noise in (A) 3D CT, (B) 4D CT, (C) 3D PET, and (D) 4D PET.

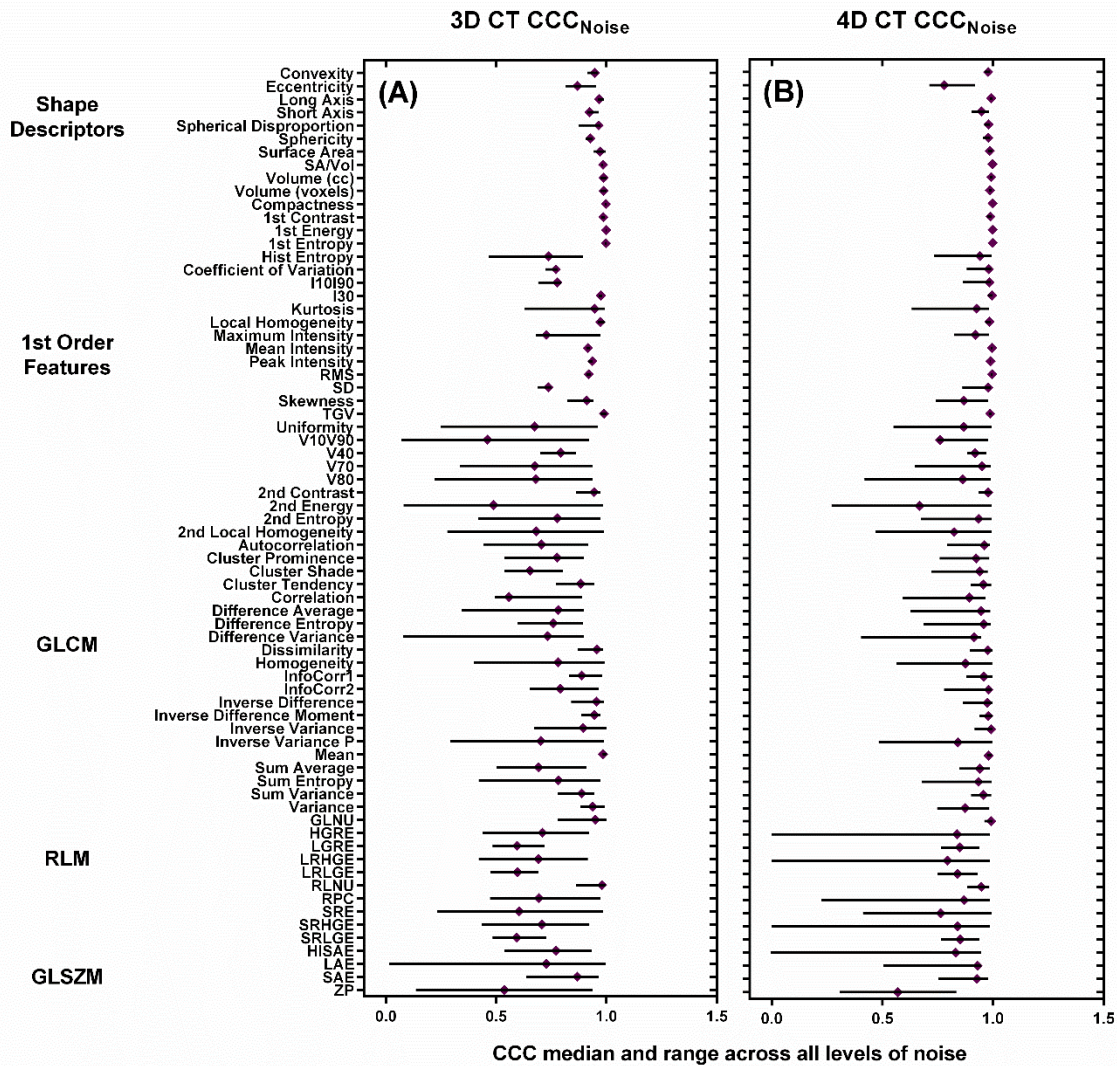


Figure 5.7 Median CCCs for all CT features – (A) 3D CT and (B) 4D CT – for the levels of noise including GN10, GN20, GN50, GN80, and GN120 range.

### 5.4 Discussion

We applied uncorrelated noise to phantom and patient images to analyze its effect on image features. We found that uncorrelated noise effects in GLCM, RLM, and GLSZM features were generally greater than those seen in shape features. Given what these texture features seek to measure, this finding is not surprising. Since GLCM, RLM, and GLSZM features measure the relationships between pixels and the addition of noise (correlated or uncorrelated) alters these relationships, these texture features would be affected more than shape features

which depend mainly on the contour defining the tumor volume (VOI). Specifically, the GLCM measures spatial relationships between pixel pairs and the RLM measures runs of the same gray level across an image. GLSZM, introduced by Thibault et al., is an advanced statistical matrix that measures homogeneity (41, 63). All matrices, except the GLSZM, were calculated along multiple directions. Shape features however, are based on the size, shape, and convexity of VOI's contour, which were essentially not affected by the addition of uncorrelated noise.

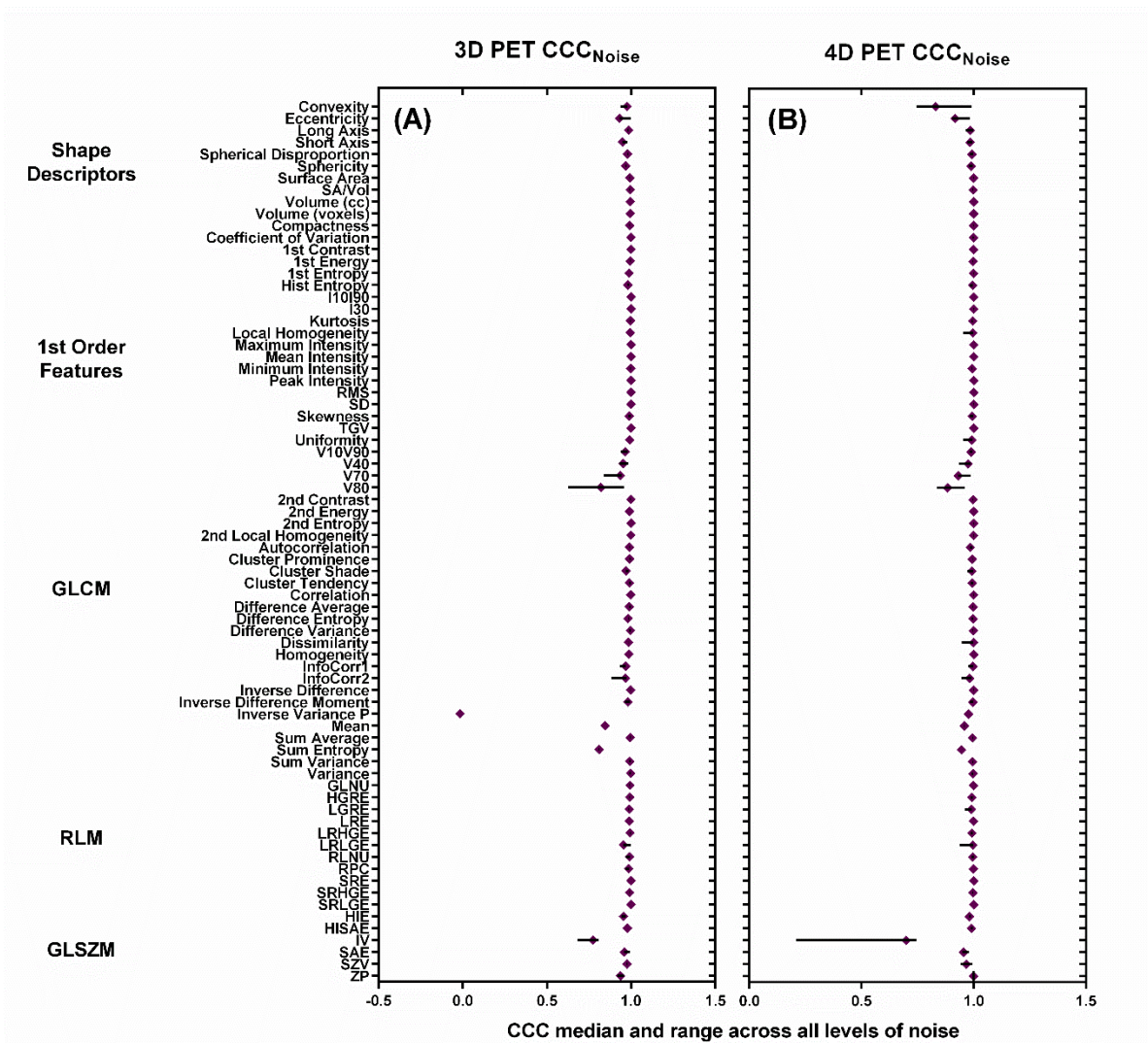


Figure 5.8 Median CCCs for all PET features – (A) 3D PET and (B) 4D PET – for the levels of noise including low, medium, and high levels of noise.

Adding uncorrelated noise to PET images with large areas of uptake in the tumor, brain or bladder resulted in less significant affect from added noise. The large areas of high uptake created a bigger dynamic window of intensities and thus the added noise appeared to be less significant. For this reason, in situations where there are large tumors with high uptake, the uncorrelated image noise may not be a significant problem in feature analysis.

We discovered that the effects of Gaussian added noise in CT were usually smaller in 4D images than 3D images. We believe that since the original image noise of 4D images was greater than 3D images, the difference between original image features and noise image features was not as prominent. This also affected the CT SNR. The visible trend for PET wherein feature differences increased as noise increased was less distinguishable in CT; especially in 4D CT (Figure 5.6). The added noise appears to have altered the CT SNR to a greater degree than PET SNR (see Figures 5.4a and 5.4b). It is clear that at higher noise levels, beginning at the GN50 noise level, the CT SNR converges implying that the Gaussian noise destroys the differences in SNR due to acquisition. Figure 5.4 demonstrates that at the GN50 noise level, the SNR for 4D CT had decreased by a factor of two. At the GN120 noise level, the SNR for CT decreased nearly 5 times compared to PET where the SNR decreased by a factor of one.

It is interesting that the SNR for 3D PET is lower than that of 4D PET especially when compared with the CT SNR. This could be due to motion effects. 4D PET accounts for motion. Since PET images are acquired over a rather long period of time (4 minutes per bed position), when motion is introduced into an image and not accounted for (as in 3D images) an averaging effect is introduced into the image and the true tumor location, size and shape is warped. We believe this is why the static PET image and 4D motion images have similar SNRs and 3D motion exhibits the lowest SNR. The SNR results in CT were drastically different from PET. To start, 3D motion had the highest SNR followed by static and finally 4D motion. Since CT images are acquired much faster than PET images, on the order of milliseconds and are much less

sensitive to motion, it makes sense that the static and 3D motion images had the highest SNRs; they received a higher number of counts than 4D CT. It is also interesting that the PET SNR is considerably lower (40x for 3D motion) than the CT SNR even for the original images.

In PET images, the 3D and 4D image feature differences were comparable. Although there were fewer features in the NR1 category for 4D PET, differences were not consistently larger for 3D or 4D PET across all feature subtypes. In addition, the percent difference in PET features did not always increase with respect to added noise. For instance, in the RLM and GLSZM features, average differences reached the maximum (4D PET) or minimum (3D PET) percent difference at the medium noise level (Figure 5.6). This could be caused by the large pixel size in PET, the high level of noise in the image due to decreased counts, or simply a saturation of the uncorrelated noise in the image at the low or medium noise levels. Figure 5.4 demonstrates that the PET SNR did not decrease sharply for 3D Motion, 4D Motion, or 3D Static PET indicating high levels of initial image noise.

The finding that shape descriptors were less affected by noise than GLCM, RLM, and GLSZM features is favorable for the field of radiation therapy. Increasingly common, the gross tumor volume (GTV) for radiation treatment planning is contoured using both PET and CT. The addition of PET as a diagnostic tool in radiation therapy has improved GTV definition and demonstrated a 21-100% change in tumor volumes (120). PET in radiation treatment planning improves the contouring accuracy of the GTV which promotes decreased toxicity to healthy tissue (105, 121).

The large differences in intensity, GLCM, RLM, and GLSZM features demonstrate that uncorrelated image noise affects image feature analysis. GLSZM features are highly unstable, particularly in 3D CT, with average values nearing 100,000% in some cases but as low as 0.4% in others. The full extent of this effect requires further investigation but it is clear that image features, especially those from intensity, GLCM, RLM, and GLSZM are affected by uncorrelated noise. Investigators that are using large numbers of images from multiple scanners should be

aware of the effects of image noise on image feature analysis. This was recently illustrated nicely for correlated noise by the work of Nyflot et al. (48). Although we did not compare results from multiple scanners, quantitative accuracy in PET/CT is still being established (48). Multi-center PET/CT trials testing the stability and repeatability of PET data from different sites demonstrated that the quantitative PET measurement, standardized uptake values (SUV), were within the PET Response Criteria in Solid Tumors (RECIST) limitations, but were higher than in a previous study conducted in smaller single-center studies (48). Even inside a single institution, patients imaged on the same scanner demonstrated SUV differences approaching 50% on test and retest analysis (2). It is clear that studies involving multiple scanners should be aware of the effects of image noise on their features.

The noise texture, defined by the measured NPS was uniform in shape for the highest levels of added noise demonstrating that we had indeed added uncorrelated noise to the images. This uncorrelated noise is commonly termed “white noise” and in this study is considered random noise of a Gaussian distribution. A distinct difference existed between noise phantom NPS and original phantom NPS due to the shift from the inherent correlated noise in the image to uncorrelated Gaussian noise.

Uncorrelated noise is not the only criteria that affect PET/CT image features. Nyflot et al. tested the effect of correlated (stochastic) noise on image features, as well as patient size, lesion size and image reconstruction method. They determined how stochastic noise have various effects on different feature subtypes – what they term “classes of metrics” - concluding that additional standards are warranted for prospective PET image feature analysis studies towards predicting clinical outcome or treatment response (48). Other studies have shown that motion, bin width, and SUV discretization, acquisition modes and reconstruction parameters also affect image features and in some cases the extent of these effects are feature dependent (45, 47, 48, 122). A common conclusion of these studies was that standardization of image feature analysis in Radiomics is needed. We join in agreement with these studies that

standardization is crucial as we look towards the application of Radiomics in Radiotherapy and other fields. We are also advocating for standardization of image feature analysis, especially in PET/CT, to promote accuracy and patient safety (if features are applied prospectively) when measuring image features for clinical purposes and to encourage accurate image feature study comparisons between scanners, institutions, and manufacturers.

There were limitations to this study. The major limitation of this study was that we did not have access to the pre-reconstruction PET/CT data due to proprietary reasons. This was a retrospective study; however, we wanted to determine the effect of noise on archived patient studies. These limitations affected the approach in which noise was added to the PET/CT images as typically done (48, 123). In typical imaging systems PET and CT image noise is integrated into the reconstruction method and not necessarily additive, except in the case of electronic noise. Nevertheless, our method still allowed us to measure the sensitivity and degradation of Radiomic features due to noise.

## **5.5 Conclusions**

Uncorrelated noise was added to PET and CT images. Shape, intensity, GLCM, RLM, and GLSZM image features were extracted from VOIs and image features that were non-robust with respect to the additional noise were identified. Many intensity, GLCM, RLM, and GLSZM features varied significantly with noise. Percent change between original and noise image features increased as noise level increased for intensity and GLCM features in PET, and GLCM, RLM, and GLSZM features in CT. GLSZM features were the most sensitive to noise both for CT and PET. A good understanding of features sensitivity to noise is essential for image features analysis and Radiomics studies involving a large number of images from multiple scanners as would be the case in multi-institutional clinical trials. This study adds support to the proposal for standardization of clinical processes and analysis involved in Radiomics.

## Chapter Six: Fiducials vs. 18F FDG PET/CT for Esophageal Cancer GTV Delineation for Radiotherapy Treatment Planning<sup>6</sup>

### 6.1 Introduction

The five year survival rate for esophageal cancer patients is 18% and there are approximately 17,000 estimated new esophageal cancer cases and nearly 16,000 estimated deaths for 2016 (124). This is perpetuated because most esophageal patients present with locally advanced or metastatic disease (54). To improve outcomes, neoadjuvant therapy is recommended for patients with locally advanced disease prior to surgery (125). At present, data on individual sensitivity to chemotherapy and radiation therapy is lacking, thus patients are advised to undergo standard of care chemoradiation based on their clinical, rather than molecular, factors. Clinicians rely on radiographic indicators to assess response but, in the absence of progression at restaging, patients proceed to esophageal resection; the quality of life implications and medical cost of this are profound if patients have a pathologic complete response and yet have undergone removal of their esophagus. Conversely, if patients are found at the time of surgery to have had no response, their outcomes are no better than if they went directly to surgery upfront (57). In fact, in the case of the pathologic non-responder, there is also the consideration of the potential acute neoadjuvant toxicity incurred for no demonstrable benefit at a delay of at least 12 weeks from diagnosis until definitive surgery.

3D <sup>18</sup>F-Fluorodeoxyglucose (FDG) PET/CT is obtained routinely for the initial staging of esophageal cancer and has been shown in several sites to alter the GTV delineation (126-129). Specifically in the esophagus, 3D PET/CT has been shown to improve staging (130). Theoretically, PET/CT could eliminate the need for additional staging methods in patients with distant metastatic disease (125). This would expedite treatment and avoid potentially ineffective

---

<sup>6</sup> Portions of this chapter have been submitted for publication.



treatment methods. Not only does 3D PET/CT identify the primary tumor location, it is an early assessment tool for treatment response, outcome prediction and therapy modification (131-133). FDG PET is able to detect most primary tumors and lymph nodes. Studies have demonstrated a sensitivity of 30-93% and a specificity of 79-100% for lymph node detection (134).

Endoscopically-placed fiducial markers have facilitated determination of respiratory associated tumor motion in the treatment of esophageal cancer as well as strategies of abdominal compression to decrease such motion, which has enhanced a conformal approach, particularly when used in conjunction with image-guided radiation therapy (IGRT) (135). It is common, when planning esophageal cancer radiotherapy treatment, that accounting for microscopic disease, nodal involvement, and tumor motion is associated with larger planning target volumes (PTV) and consequently, increased concern about the amount of normal tissue irradiated (136). Studies have shown that increased areas of irradiated tissues can result in harmful effects such as radiation pneumonitis, pericardial effusion, and pleural effusion (136). At our institution, all trimodality esophageal cancer patients receive 3D PET/CT imaging prior to treatment. In a study on 81 esophageal patients, respiratory associated superior-inferior tumor movement of 1.25 cm for proximal and middle esophageal tumors and 1.75 cm for those in the distal esophagus were noted (136).

The role of 3D PET/CT in esophageal tumors that move with respiration and have the potential for significant mucosal inflammation is unclear. The GI research group at Moffitt Cancer Center previously reported the stability data of esophageal fiducial markers endoscopically implanted under ultrasound guidance within 1 cm from the superior and inferior edges of the tumor (135). However, the correlation between gross tumor volumes derived from 3D PET/CT vs. endoscopically placed fiducial markers has not yet been reported. This work tested the correlation between metabolic tumor volumes (MTV) derived from 3D PET/CT and endoscopically placed fiducial markers using ultrasonography.

## 6.2 Materials and Methods

### 6.2.1 Patient Population

Sixty-two patients with esophageal cancer were selected for this retrospective, IRB-approved analysis. Twenty-one patients were deemed unfit for the study for various reasons. Each patient underwent placement of a VISICOIL™ (RadioMed Corporation, an IBA Company, Bartlett, TN) 10 mm x 0.75 mm gold fiducial marker at the inferior and superior borders of the tumor and received 3D PET/CT prior to Radiotherapy (RTx). Patients receiving only one fiducial marker were not included in this study. Refer to Table 6.1 for patient statistics and characteristics. Twenty patients received fiducials pre-PET/CT and 21 patients received fiducials post-PET/CT. In cases where patients underwent fiducial placement post-PET/CT, planning CTs were used to delineate the location of the fiducial marker. 3D planning CTs were imported from the Pinnacle treatment planning system (TPS; Version 9.8 Philips Medical System™, Fitchburg, WI). 3D PET/CT and planning CT images were imported into an image analysis software system (Mirada RTx, Mirada Medical, Oxford, UK) for measurements.

Table 6.1 Patient Characteristics.

<b>Characteristic</b>	<b>n (%)</b>
Age (median, years)	66
Gender	
Male	32 (78.0)
Female	9 (19.5)
Location of Tumor	
Upper/Middle	1 (2.4)
Middle	3 (7.3)
Middle/Lower	2 (4.9)
Lower	30 (73.2)
GEJ/Lower	4 (9.8)
GEJ	1 (2.4)
MTV (median, cm <sup>3</sup> )	22.4

### 6.2.2 Measurement of MTV

Patients were imaged with a GE Discovery STE PET/CT Scanner (GE Medical Systems) at our institution. A tumor threshold was created using a background uptake method. A 3 cm

spherical volume-of-interest (VOI) was dropped onto a homogenous uptake region in the liver. The mean and standard deviation of the standardized uptake value (SUV) was extracted to calculate a threshold for the tumor volume as shown in Equation 6.1.

$$Threshold = [Liver_{\mu} + 2Liver_{\sigma}] \quad (6.1)$$

where  $\mu$  is mean and  $\sigma$  is standard deviation. In cases where contours extended into the stomach or the heart, a Boolean tool was used to create a conformal MTV. These difficult contours were then physician-verified and/or edited. The fiducial was delineated on CT via an absolute threshold for HU greater than 350. The centroid was determined as the center of mass of the fiducial contour.

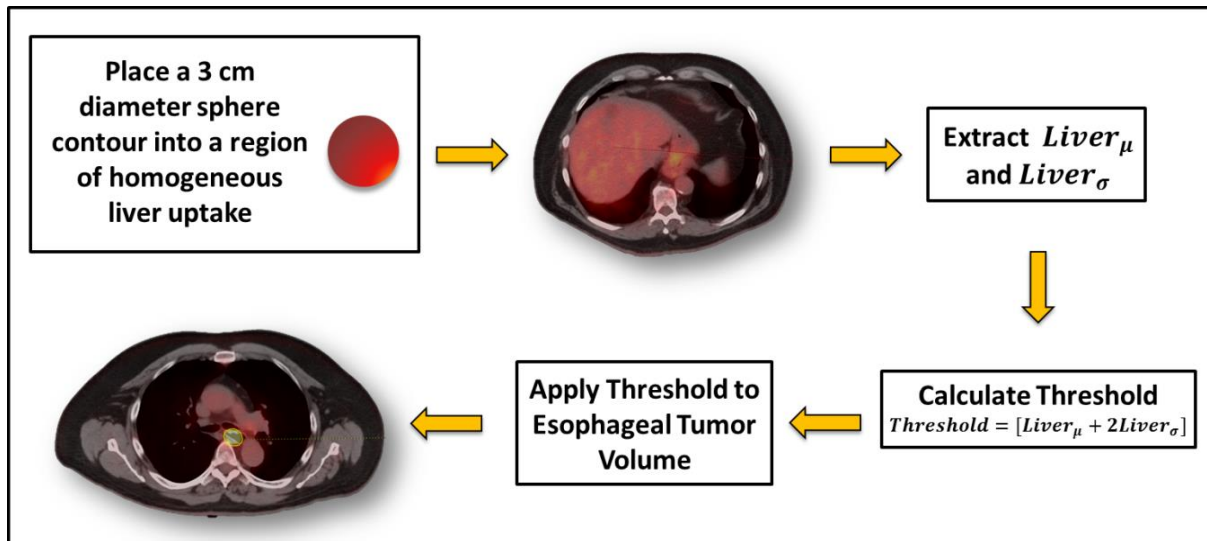


Figure 6.1 Method of determining MTV threshold for each esophageal tumor. On the fused PET/CT image, a 3-cm spherical region is placed in center of liver to account for background uptake.

### 6.2.3 Measurement of Fiducial Distance

The MTV contour was specified using the above defined liver threshold method and the axial slices were used to measure the distance between each centroid and corresponding tumor border. This distance was defined as the number of slices between the centroid of the fiducial and the first axial PET slice that included the MTV contour. The number of axial slices was then multiplied by slice thickness (3.27 mm for PET/CT and 3.0 mm for planning CT) to provide the

distance in centimeters. The distance and absolute values of the distances were recorded. Negative values described distances where fiducials were located inferior to the MTV border for both the superior and inferior margins of the tumor. Descriptive statistics such as: mean, median, standard deviation, maximum, and minimum were calculated. Cases involving large distances were investigated. The Concordance Correlation Coefficient (CCC) was used to determine correlation between the MTV threshold and superior fiducial distance (SFD), MTV threshold and inferior fiducial distance (IFD), tumor site and IFD, tumor site and SFD, time between fiducial and PET/CT, patient age and IFD, patient age and SFD. Tumor site represents the location of the tumor in the esophagus (upper, mid or distal/GE junction). The strength-of-agreement scale was as follows: CCC > 0.99: high; CCC 0.95-0.99: substantial; CCC 0.90-0.95: moderate; CCC <0.90: poor (113).

### 6.3 Results

The median MTV threshold was 2.51 SUV (1.6-3.6) for all patients. For patients receiving fiducials before undergoing PET/CT (PrePF), the median MTV threshold was 2.45 SUV (1.6-3.6). For patients receiving fiducials after undergoing PET/CT (PostPF), the median MTV threshold was 2.6 SUV (1.8-3.4). There was not much difference in MTV thresholds between the two cohorts. A two-tailed t-test demonstrated a p-value of 0.58 between the two cohorts demonstrating they were not significantly different. The median relative uptake for the liver contour was 24% (5%-79%). Refer to Table 6.1-6.3. The median distance between MTV and fiducials was -0.3 cm (-3.90 cm – 2.70 cm) and 1.3 cm (-2.1 cm – 6.87 cm) for inferior and superior tumor borders, respectively (Table 6.2-6.4). These values were comparable to those from the two groups. PrePF patients (Table 6.3) demonstrated a median distance between MTV and fiducials of -0.82 cm (-2.62 cm - 2.62 cm) and 1.64 cm (-0.33 cm – 6.87 cm) for inferior and superior borders, respectively. PostPF (Table 6.4) patients demonstrated a median distance between MTV and fiducials of -0.30 cm (-3.90 cm – 2.70 cm) inferiorly and 0.60 cm (-4.20 cm – 3.90 cm) superiorly. A poor strength-of-agreement (CCC < 0.90) was calculated between MTV

threshold and superior fiducial distance (SFD), MTV threshold and inferior fiducial distance (IFD), tumor site and IFD, tumor site and SFD, time between fiducial and PET/CT, patient age and IFD, patient age and SFD.

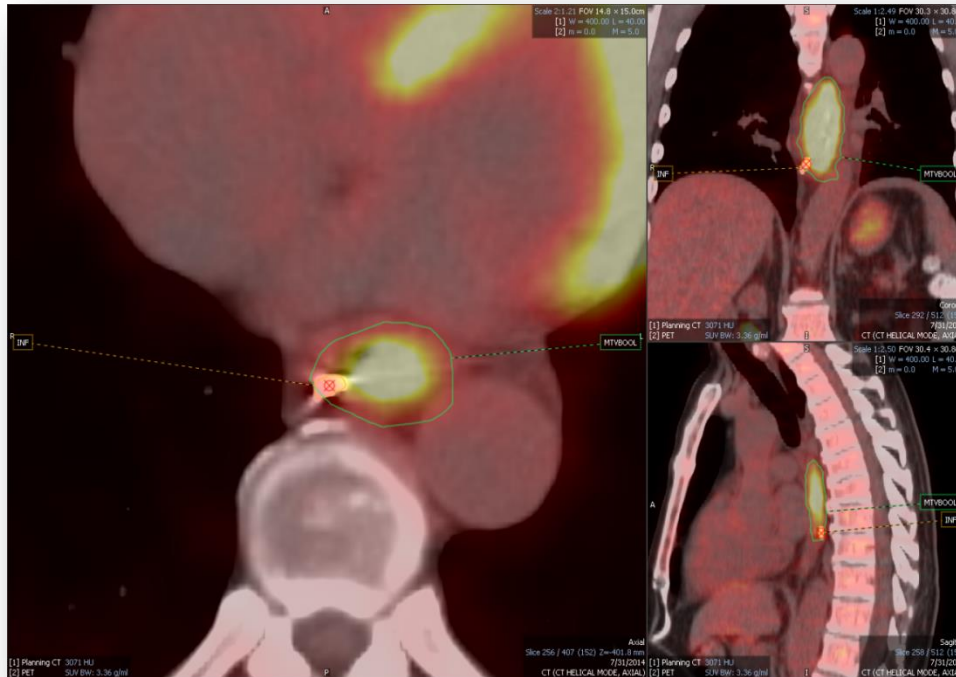


Figure 6.2 Method of identifying the fiducial and marking the centroid at the inferior border of the MTV.

24 of 42 (58.5%) cases had inferior fiducials located superior to the MTV border. In 13 cases (31.7%), inferior fiducials were below the MTV border. 4 cases (9.5%) demonstrated perfect agreement between the inferior fiducial and MTV border (all PrePF). The superior fiducial and MTV border did not have perfect agreement in any case. In 34 cases (82.9%) as shown in Figure 6.3b, the superior fiducial was located inferior to the MTV border. In 7 cases (17.1%) the superior fiducial was located superior to the MTV border. Of these cases, 3 of 34 (8.8%) distances were less than 0.5 cm for the superior location and 7 of 24 (29.2%) distances were less than 0.5 cm for the inferior location (Figure 6.3).

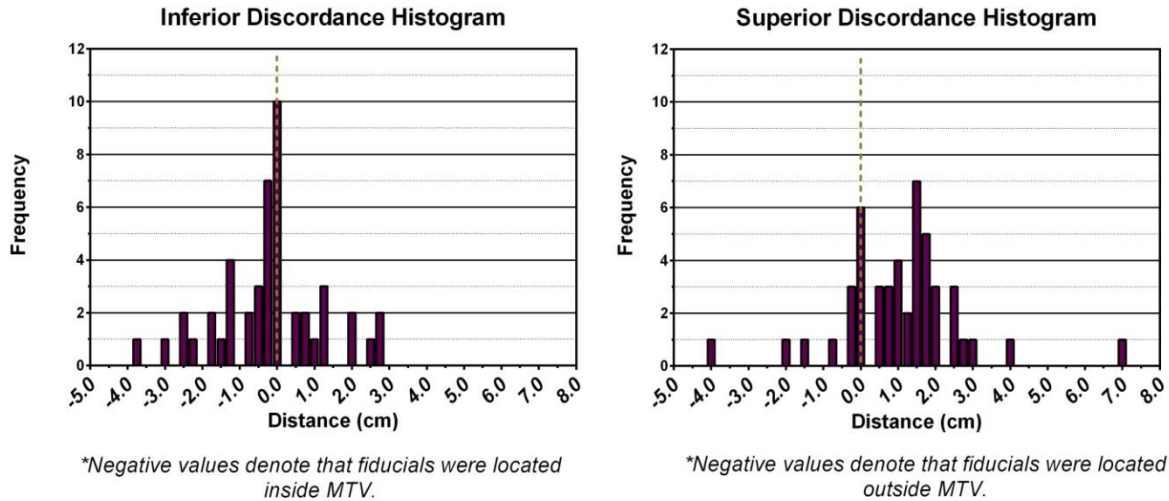


Figure 6.3 Inferior discordance and superior discordance histograms. Note: For inferior discordance, negative values denote that fiducials were located inside MTV. For superior discordance, negative values denote that fiducials were located outside MTV.

### 6.3.1 Large Discordances

In 8 cases, the superior fiducial-MTV discordance was greater than 2.0 cm. Of these cases, time between fiducials and PET ranged from 2-27 days. There was only 1 occurrence where the superior fiducial was inferior to the MTV border (negative distance). In the case of the patient with 6.87 cm discord (see Figure 6.4), the patient was diagnosed with extensive esophagitis and several nodules at the gastroesophageal (GE) junction. This discord was attributed to esophagitis. Of the 1 occurrence where the superior fiducial was superior to the MTV border, it was the only patient with medically inoperable stage 1 cancer (2 patients had stage 2 cancer in the cohort). The patient was diagnosed with adenocarcinoma of the distal esophagus although the uptake was in the mid esophagus. This patient had Barrett's esophagus from the mid esophagus to the GE junction which may have influenced the uptake in the mid esophageal region (Figure 6.5). There was no correlation between MTV-to-fiducial distances greater than 2 cm and the gastroenterologist that performed the fiducial implantation.

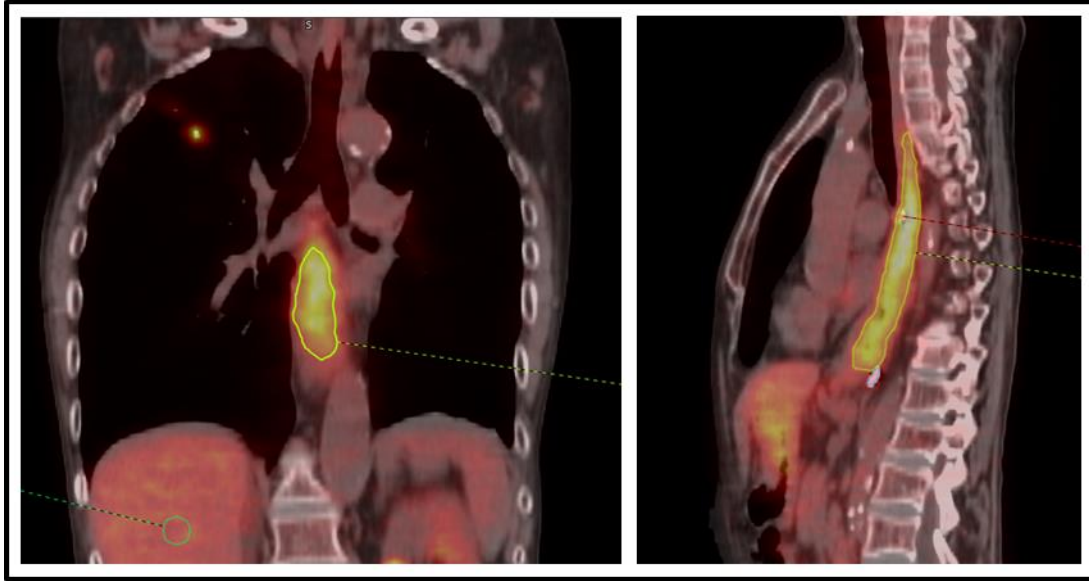


Figure 6.4 Case with largest discord (6.87 cm) between superior fiducial and superior MTV border.

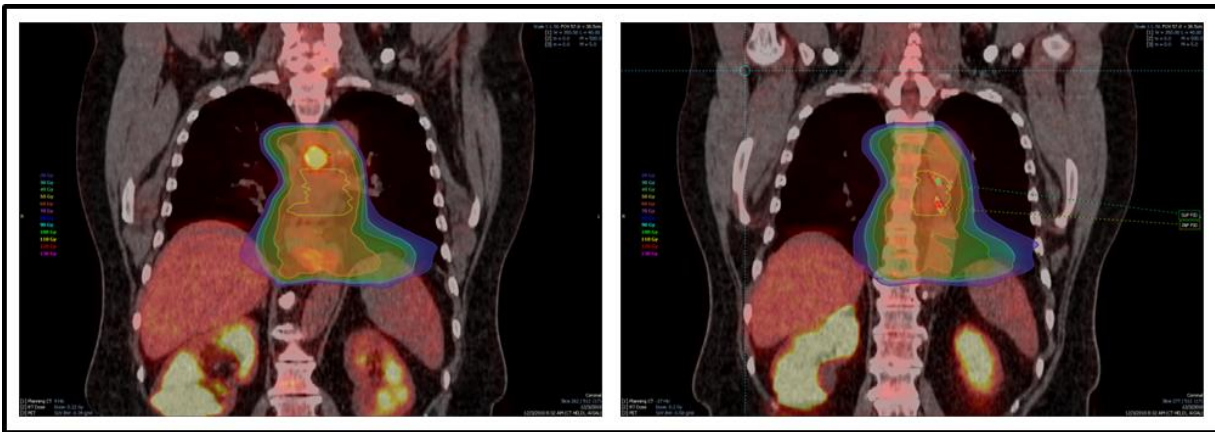


Figure 6.5 Example of discordance between fiducials and MTV at the inferior and superior location (-4.20 cm discord between superior fiducial and superior MTV border.) The red contour represents MTV threshold. The green dot represents the superior fiducial. The red dot is the inferior fiducial.

Table 6.2 Fiducial vs MTV Border

Total	Inferior Fiducial Distance (cm)	Superior Fiducial Distance (cm)	MTV Threshold (SUV)	Relative MTV (%)	MTV (cm <sup>3</sup> )
Mean	-0.27	1.28	2.51	28.97%	30.53
Median	-0.30	1.31	2.50	24.00%	23.55
Min	-3.90	-2.10	1.60	5.00%	1.80
Max	2.70	6.87	3.60	79.00%	107.10
SD	1.50	1.52	0.42	18.19%	28.21

Table 6.3 FID Pre PET Fiducial vs MTV Border

FID Pre PET	Inferior Fiducial Distance (cm)	Superior Fiducial Distance (cm)	MTV Threshold (SUV)	Relative MTV (%)	MTV (cm <sup>3</sup> )
Mean	-0.64	1.68	2.44	23.50%	40.13
Median	-0.82	1.64	2.45	18.50%	34.70
Min	-2.62	-0.33	1.60	5.00%	3.00
Max	2.62	6.87	3.60	74.00%	107.10
SD	1.31	1.43	0.42	16.09%	33.47

Table 6.4 FID Post PET Fiducial vs MTV Border

FID Post PET	Inferior Fiducial Distance (cm)	Superior Fiducial Distance (cm)	MTV Threshold (SUV)	Relative MTV (%)	MTV (cm <sup>3</sup> )
Mean	-0.03	0.60	2.56	34.10%	21.91
Median	-0.30	0.60	2.60	31.00%	20.60
Min	-3.90	-4.20	1.80	11.00%	1.80
Max	2.70	3.90	3.40	79.00%	74.80
SD	1.68	1.77	0.44	18.54%	18.50

## 6.4 Discussion

Esophageal tumors can have significant respiration-induced tumor motion. A study by Jin *et al.* measured the peak-to-peak magnitudes of the motion (137). The greatest motion was found for the distal esophagus in the cranial-caudal direction with a median distance of 5.4 mm. Median displacements for the proximal and middle esophagus were 2.9 mm and 3.7 mm, respectively. Interestingly, motion in the cranial-caudal direction was shown to have the strongest correlation with respiratory curves (138). In that particular study, motion in the cranial-caudal direction reached 13.8 mm in the lower thoracic esophagus, 7.4 mm in the middle esophagus, and 4.3 mm in the upper esophagus. Investigators are beginning to realize the benefit of fiducials for radiation treatment planning for cancers of the esophagus (139) and studies have demonstrated that implantation of esophageal fiducial markers are both safe and feasible for target volume delineation purposes on CT (139). However, to our knowledge, the



discordance between endoscopically-placed fiducial markers and PET MTV in esophageal cancer has not been investigated.

A retrospective study of esophageal cancer patients treated with preoperative or definitive chemoradiation at our institution between 2000 and 2012 demonstrated 3-year overall survival rates of 44.8% for 3-D conformal radiation therapy (3DCRT) and 41.5% for intensity-modulated radiation therapy (IMRT) (140). The rates for that study agree well with the national 5-year relative survival rate of 17% between 2010-2012 (141-143). In the current study, patients were treated with radiotherapy between 2009 and 2014.

It is common practice at our institution for patients with esophageal tumors that move with respiration to undergo fiducial placement in addition to  $^{18}\text{F}$ -FDG PET/CT to facilitate conformal delivery of a simultaneous integrated boost to the gross tumor volume (144). Our data recently reported a 55% complete pathologic response rate, significantly higher than the 29% reported with the current standard of care CROSS regimen (145), which may be secondary to integration of both fiducials and MTV. Indeed, it is perhaps the combination of fiducials and MTV that may allow for precise dose painting – the treatment of tumor areas that are more metabolically active with higher doses of radiation. This incorporation has facilitated our confidence in dose painting the gross tumor volume to a total dose of 56 Gy in 28 fractions simultaneously with the clinical volume dosed to 50.4 Gy while ensuring the reproducibility of our daily image guided delivery. Focal dose escalation is of particular concern in the region of the gastroesophageal junction where stomach filling can cause additional motion (146). The fiducials delimit the visible endoscopic mucosal tumor burden which improves target volume delineation in precise conjunction with daily dose delivery. MTV, on the other hand, identifies metabolically active tumor regions and submucosal microscopic spread of disease that may not be visible endoscopically or on a CT image.

Although our hypothesis in this study was neither confirmed nor disproven, a robust correlation was found between the inferior fiducial location and the border of the MTV and the

absolute SUV values were reasonably close to commonly used values of 2.5. However, relative SUV values were lower than the typical values of 40% of SUVmax. A study by Zhong *et al.* found that an SUV cutoff of 2.5 was best to estimate gross tumor length in squamous cell carcinoma of the esophagus using 18F-FDG PET (147). This agrees with our median MTV threshold of 2.51 SUV.

Some cases demonstrated large areas of uptake that extended well beyond the implanted fiducial (-4.2 - 6.9 cm). In a few instances, tumors had small MTV with large distances between the MTV border and fiducial location. These could represent tumors that were not FDG avid. The results demonstrated that in 81% of cases, the superior fiducial was located below the MTV threshold. Thus, in these cases, the MTV extended above the superior fiducial placement. This discordance could be due to inflammation or esophagitis or potentially to disease that was not endoscopically visible. In comparison, 57% of cases the MTV extended below the inferior fiducial. Of these cases, 3 of 34 (8.8%) were less than 0.5 cm at the superior tumor border and 7 of 24 (29.2%) were less than 0.5 cm at the inferior tumor border. The majority of these cases were diagnosed as distal or GE junction tumors. Thus, accurate fiducial placement may not have been possible in these cases given the proximity of the stomach. There was a clear indication that in most cases the MTV extended beyond the superior fiducial and that timing of the PET before or after fiducial placement was not a significant factor.

It is unclear, however, the etiology of the discordance superiorly, with the PET/CT showing high uptake at and above the endoscopically placed marker, potentially representing benign secondary esophagitis such as in the setting of luminal obstruction, the presence of malignant nodes, inflammation caused by the technical aspects of the fiducial placement itself, or potential submucosal disease. In the case of the largest discordance, the patient had been diagnosed prior to treatment with extensive esophagitis. According to these findings, the incorporation of a fiducial marker inferiorly into the routine management of locally advanced

esophageal cancer may offer accurate target volume delineation when compared with 3D PET/CT alone.

One limitation of this study is that about half of the patients underwent PET/CT imaging prior to fiducial implantation. Thus, the PET/CT and planning CT had to be fused, and the fiducials and distances between fiducials and the MTV borders had to be identified and measured. Image fusion may have led to some uncertainty (148).

## **6.5 Conclusion**

The inferior fiducial location and MTV border for esophageal cancer had a robust correlation. Thus, it may be prudent to incorporate an inferior fiducial in the routine management of locally advanced esophageal cancer. The etiology of the discordance between the superior fiducial location and MTV border could be caused by inflammation from the fiducial placement itself, submucosal disease, or benign secondary esophagitis. Regardless of the discordance, having both fiducials is important for image guidance, especially in cases of focal dose escalation in tumors involving the GE junction. The factors confounding FDG uptake superiorly need further investigation to optimize MTV delineation.

## Chapter Seven: Conclusions and Future Work

### 7.1 Research Outcomes

This research explored PET/CT diagnostics and treatment capabilities in lung and esophageal cancer. The main research outcomes of this work were that:

- Image Features are affected by motion, especially in tumors that move greatly (ranging ~1-2 cm) - such as those that move with the respiratory cycle (lung tumors located near the diaphragm).
- Certain features are less affected by motion and noise. GLSZM are highly sensitive and varied greatly with the addition of uncorrelated Gaussian noise. This is not acceptable as an additional decision factor for clinical image analysis for PET/CT systems with high levels of noise but it could be an advantage in low noise situations. GLSZM features may vary significantly (more than 100 % for CT data) with change of image texture pattern. Shape features were the least affected by the addition of uncorrelated Gaussian noise. GLCM and RLM features were highly sensitive to image noise (change in noise level caused changes >100%), although not as much as for the GLSZM features.
- There was discordance between endoscopically placed fiducial markers and MTV. This demonstrates a need for both techniques synergistically in the clinic to provide a more complete view of the tumor location.

Investigators must be aware of these effects of noise and motion on image feature analysis and account for them to avoid false positives. Chalkidou *et al.* found that published studies using image feature analysis to predict clinical outcomes had an average probability of type-I error (false positive) of 76% (82). These false positives could be caused by tumor motion and image noise as well as other factors that affect image features such as SUV bin

discretization (where SUV bins are divided into equally spaced bins) and reconstruction parameters (47, 48, 81-83). Other studies have demonstrated that particular features are correlated with tumor volume (82). Thus, the predictive abilities of these image features are characteristics of the tumor volume and not the features themselves (71). It cannot be emphasized enough that investigators involved in image feature analysis, especially in a clinical setting, must be aware of how factors such as noise and motion affect their image features. This work highly recommends that image feature analysis for PET and CT imaging modalities be standardized and a protocol developed for reproducibility and accuracy across institutions.

## **7.2 Future Work**

Currently, investigators are pursuing active research correlating treatment outcomes to image features (54, 80). One goal of image feature analysis in Radiotherapy is to identify predictive and prognostic features and to use such features to identify optimal treatment regimen for patients (personalized medicine). This technique will allow the Radiation Oncologist to select the optimal treatment regimen for a patient prior to treatment thus providing the best chance of survival and quality of life (tumor control and reduced toxicity). Esophageal cancers will especially benefit from this technique because they are time-sensitive. An ineffective treatment could lead to an early death. A patient's chance of survival decreases each day an ineffective treatment regimen is used.

The next step of our research would be to extend the noise study from Chapter 4 and assess the sensitivity of correlated noise effects on image features. This would require access to raw PET/CT data which was unavailable during this study due to proprietary reasons. Noise would be added directly to the sinograms in the case of PET and then reconstructed with the noise inherent in the image, thus providing a more accurate noise image because it more accurately represents the stochastic effect of the random variations in photon counting for CT and radioactive decay in PET (21, 48). The ASIM PET simulator, open-access software developed by the University of Washington (<https://depts.washington.edu/asimuw/index.html>) to

simulate emission data, attenuation correction, and noise propagation would be used to apply correlated noise to the raw PET data (sinograms) at our institution (H. Lee Moffitt Cancer Center).

Another interesting and potentially promising idea would look at ways to account for motion and noise effects in image feature analysis. Vaidya et al accounted for motion in their study of 2 NSCLC patients by applying a deconvolution algorithm – an inverse filter compensating for blurring through a motion kernel (38). Aerts et al accounted for noise by using data that was in its raw form, without pre-processing or normalization (84). Although a 2D study, Ganeshan *et al.* removed the effects of noise in CT using image filtration, using features larger than 4 voxels (72).

Our current research serves as a warning to investigators that image features are not independent. They are affected by many factors that require acquisition protocol standardization across institutions.

Another future step would be standardization of Radiomic features. This would be a collaborative effort between many institutions and research groups. Buvat stated that feature descriptor names, definitions, and equations vary between studies (71) which warrants the standardization of features a necessary step toward clinical applications.

## List of References

1. The Role of PET/CT in Radiation Treatment Planning for Cancer Patient Treatment. IAEA, Vienna: International Atomic Energy Agency, 2008 Contract No.: IAEA-TECDOC-1603.
2. Kumar V, Nath K, Berman CG, Kim J, Tanvetyanon T, Chiappori AA, Gatenby RA, Gillies RJ, Eikman EA. Variance of Standardized Uptake Values for FDG-PET/CT Greater in Clinical Practice than Under Ideal Study Settings. Clin Nucl Med. 2013;38(3):175-82.
3. Cancer Facts & Figures 2015. American Cancer Society, 2015.
4. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-Small Cell Lung Cancer: Epidemiology, Risk Factors, Treatment, and Survivorship. Mayo Clin Proc. 2008;83(5):584-894.
5. Eschrich S, Zhang H, Zhao H, Boulware D, Lee JH, Bloom G, Torres-Roca JF. Systems biology modeling of the radiation sensitivity network: a biomarker discovery platform. International journal of radiation oncology, biology, physics. 2009;75(2):497-505.
6. Wiener RS, Schwartz LM, Woloshin S, Welch HG. Population-Based Risk for Complications After Transthoracic Needle Lung Biopsy of a Pulmonary Nodule: An Analysis of Discharge Records. Annals of Internal Medicine. 2011;155(3):137-44.
7. Cook GJ, Yip C, Siddique M, Goh V, Chicklore S, Roy A, Marsden P, Ahmad S, Landau D. Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? Journal of Nuclear Medicine. 2013;54(1):19-26.
8. Jeraj R, Bowen S, Jallow N, Nyflot M, Vanderhoek M. Molecular Imaging in Radiation Oncology. The Modern Technology of Radiation Oncology: A compendium for Medical Physicists and Radiation Oncologists. 3. Madison, Wisconsin: Medical Physics Publishing; 2013.

9. Bushberg JT, Seibert JA, Edwin M. Leidholdt J, Boone JM. The Essential Physics of Medical Imaging. Third Edition ed. Philadelphia, PA: LIPPINCOTT WILLIAMS & WILKINS, a WOLTERS KLUWER business; 2012.
10. Nestle U, Kremp S, Grosu AL. Practical integration of [18F]-FDG-PET and PET-CT in the planning of radiotherapy for non-small cell lung cancer (NSCLC): the technical basis, ICRU-target volumes, problems, perspectives. *Radiotherapy and Oncology*. 2006;81(2):209-25.
11. Moses WW. Fundamental Limits of Spatial Resolution in PET. *Nucl Instrum Methods Phys Res A*. 2011;648 Supplement 1:S236-S40.
12. Kumar V, Nath K, Berman CG, Kim J, Tanvetyanon T, Chiappori AA, Gatenby RA, Gillies RJ, Eikman EA. Variance of SUVs for FDG-PET/CT is greater in clinical practice than under ideal study settings. *Clin Nucl Med*. 2013;38(3):175-82.
13. Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJ. Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. *European journal of nuclear medicine and molecular imaging*. 2013;40(1):133-40.
14. El Naqa I, Grigsby P, Apte A, Kidd E, Donnelly E, Khullar D, Chaudhari S, Yang D, Schmitt M, Laforest R, Thorstad W, Deasy JO. Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern recognition*. 2009;42(6):1162-71.
15. Borst GR, Belderbos JS, Boellaard R, Comans EF, De Jaeger K, Lammertsma AA, Lebesque JV. Standardised FDG uptake: a prognostic factor for inoperable non-small cell lung cancer. *European journal of cancer*. 2005;41(11):1533-41.
16. Fahey FH, Kinahan PE, Doot RK, Kocak M, Thurston H, Poussaint TY. Variability in PET quantitation within a multicenter consortium. *Medical Physics*. 2010;37(7):3660.
17. Alessio AM, Kohlmyer S, Branch K, Chen G, Caldwell J, Kinahan P. Cince CT for attenuation correction in cardiac PET/CT. *J Nucl Med*. 2007;48:794-801.
18. Hounsfield GN. Historical notes on computerized axial tomography. *J Can Assoc Radiol*. 1976;27(3):135-42.



19. Kahn FM. The Physics of Radiation Therapy. 4th Edition ed. Philadelphia, PA: Lippincott Williams & Wilkins, a Wolters Kluwer business; 2010.
20. Xu Q, Yuan K, Ye D. Respiratory motion blur identification and reduction in ungated thoracic PET imaging. *Physics in medicine and biology*. 2011;56(14):4481-98.
21. Hasegawa BH. The Physics of Medical X-ray Imaging: (or the Photon and Me, how I Saw the Light): Medical Physics Pub.; 1991.
22. Hanson KM. Noise and contrast discrimination in computed tomography. In: Newton TH, Potts DG, editors. *Radiology of the Skull and Brain: Technical Aspects of Computed Tomography*. 5. St. Louis: C.V. Mosby; 1981. p. 3941-55.
23. Prince JL, Links JM. *Medical Imaging Signals and Systems*. Upper Saddle River, NJ: Pearson Education, Inc.; 2006.
24. Saha GB. *Basics of PET Imaging: Physics, Chemistry, and Regulations*. New York, NY: Springer Science+Business Media, Inc.; 2005.
25. Park SJ, Ionascu D, Killoran J, Mamede M, Gerbaudo VH, Chin L, Berbeco R. Evaluation of the combined effects of target size, respiratory motion and background activity on 3D and 4D PET/CT images. *Physics in medicine and biology*. 2008;53(13):3661-79.
26. Razifar P, Sandstrom M, Schnieder H, Langstrom B, Maripuu E, Bengtsson E, Bergstrom M. Noise correlation in PET, CT, SPECT and PET/CT data evaluated using autocorrelation function: a phantom study on data, reconstructed using FBP and OSEM. *BMC medical imaging*. 2005;5:5.
27. Zhao W, Rowlands JA. Digital radiology using active matrix readout of amorphous selenium: Theoretical analysis of detective quantum efficiency. *Med Phys*. 1997;24(12):1819.
28. Zucker SW, Terzopoulos D. Finding structure in Co-occurrence matrices for texture analysis. *Computer Graphics and Image Processing*. 1980;12(3):286-308.
29. Haralick RM, Shanmugam K, Dinstein Ih. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*. 1973;SMC-3(6):610-21.

30. Aggarwal N. First and Second Order Statistics Features for Classification of Magnetic Resonance Brain Images. *Journal of Signal and Information Processing*. 2012;03(02):146-53.
31. Duncan JS, Ayache N. Medical image analysis: progress over two decades and the challenges ahead. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000;22(1):85-106.
32. Ganeshan B, Abaleke S, Young RC, Chatwin CR, Miles KA. 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*. 2010;10:137-43.
33. Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K. Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. *European radiology*. 2012;22(4):796-802.
34. Al-Kadi OS, Watson D. Texture analysis of aggressive and nonaggressive lung tumor CE CT images. *IEEE Trans Biomed Eng*. 2008;55(7):1822-30.
35. Hunter LA, Krafft S, Stingo F, Choi H, Martel MK, Kry SF, Court LE. High quality machine-robust image features: identification in nonsmall cell lung cancer computed tomography images. *Med Phys*. 2013;40(12):121916.
36. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, Zegers CM, Gillies R, Boellard R, Dekker A, Aerts HJ. Radiomics: extracting more information from medical images using advanced feature analysis. *European journal of cancer*. 2012;48(4):441-6.
37. Willaime JM, Turkheimer FE, Kenny LM, Aboagye EO. Quantification of intra-tumour cell proliferation heterogeneity using imaging descriptors of <sup>18</sup>F fluorothymidine-positron emission tomography. *Physics in medicine and biology*. 2013;58(2):187-203.
38. Vaidya M, Creach KM, Frye J, Dehdashti F, Bradley JD, El Naqa I. Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer. *Radiother Oncol*. 2012;102(2):239-45.

39. Leijenaar RT, Carvalho S, Velazquez ER, van Elmp J, Parmar C, Hoekstra OS, Hoekstra CJ, Boellaard R, Dekker AL, Gillies RJ, Aerts HJ, Lambin P. Stability of FDG-PET Radiomics features: an integrated analysis of test-retest and inter-observer variability. *Acta oncologica*. 2013;52(7):1391-7.
40. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges JP, Corcos L, Visvikis D. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *Journal of Nuclear Medicine*. 2011;52(3):369-78.
41. Tixier F, Hatt M, Le Rest CC, Le Pogam A, Corcos L, Visvikis D. Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. *Journal of Nuclear Medicine*. 2012;53(5):693-700.
42. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D, Goldgof DB, Hall LO, Lambin P, Balagurunathan Y, Gatenby RA, Gillies RJ. Radiomics: the process and the challenges. *Magnetic resonance imaging*. 2012;30(9):1234-48.
43. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RT, Hermann G, Lambin P, Haibe-Kains B, Mak RH, Aerts HJ. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother Oncol*. 2015;114(3):345-50.
44. Cunliffe A, Armato SG, 3rd, Castillo R, Pham N, Guerrero T, Al-Hallaq HA. Lung texture in serial thoracic computed tomography scans: correlation of radiomics-based features with radiation therapy dose and radiation pneumonitis development. *International journal of radiation oncology, biology, physics*. 2015;91(5):1048-56.
45. Galavis PE, Hollensen C, Jallow N, Paliwal B, Jeraj R. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta oncologica*. 2010;49(7):1012-6.

46. Cheng NM, Fang YH, Yen TC. The promise and limits of PET texture analysis. *Annals of nuclear medicine*. 2013;27(9):867-9.
47. Leijenaar RT, Nalbantov G, Carvalho S, van Elmpt WJ, Troost EG, Boellaard R, Aerts HJ, Gillies RJ, Lambin P. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep*. 2015;5(11075):11075.
48. Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. *Journal of Medical Imaging*. 2015;2(4).
49. Foster B, Bagci U, Mansoor A, Xu Z, Mollura DJ. A review on segmentation of positron emission tomography images. *Comput Biol Med*. 2014;50:76-96.
50. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, Forster K, Aerts HJ, Dekker A, Fenstermacher D, Goldgof DB, Hall LO, Lambin P, Balagurunathan Y, Gatenby RA, Gillies RJ. Radiomics: the process and the challenges. *Magnetic resonance imaging*. 2012;30(9):1234-48.
51. Ganeshan B, Panayiotou E, Burnand K, Dizdarevic S, Miles K. Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. *European radiology*. 2012;22(4):796-802.
52. Chicklore S, Goh V, Siddique M, Roy A, Marsden PK, Cook GJ. Quantifying tumour heterogeneity in  $^{18}\text{F}$ -FDG PET/CT imaging by texture analysis. *European journal of nuclear medicine and molecular imaging*. 2013;40(1):133-40.
53. Cook GJR, Yip C, Siddique M, Goh V, Chicklore S, Roy A, Marsden P, Ahmad S, Landau D. Are Pretreatment  $^{18}\text{F}$ -FDG PET Tumor Textural Features in Non-Small Cell Lung Cancer Associated with Response and Survival After Chemoradiotherapy. *J Nucl Med*. 2013;54:19-26.

54. Tixier F, Le Rest CC, Hatt M, Albarghach N, Pradier O, Metges JP, Corcos L, Visvikis D. Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med.* 2011;52(3):369-78.
55. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebbers F, Rietbergen MM, Leemans CR, Dekker A, Quackenbush J, Gillies RJ, Lambin P. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014;5:4006.
56. Cancer Facts & Figures 2016. 2016.
57. Dittrick GW, Weber JM, Shridhar R, Hoffe S, Melis M, Almhanna K, Barthel J, McLoughlin J, Karl RC, Meredith KL. Pathologic Nonresponders after Neoadjuvant Chemoradiation for Esophageal Cancer Demonstrate no Survival Benefit Compared with Patients Treated with Primary Esophagectomy. *Annals of Surgical Oncology.* 2012;19(5):1678-84.
58. van Hagen P, Hulshof MC, van Lanschot JJ, Steyerberg EW, van Berge Henegouwen MI, Wijnhoven BP, Richel DJ, Nieuwenhuijzen GA, Hospers GA, Bonenkamp JJ, Cuesta MA, Blaisse RJ, Busch OR, ten Kate FJ, Creemers GJ, Punt CJ, Plukker JT, Verheul HM, Spillenaar Bilgen EJ, van Dekken H, van der Slangen MJ, Rozema T, Biermann K, Beukema JC, Piet AH, van Rij CM, Reinders JG, Tilanus HW, van der Gaast A. Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med.* 2012;366(22):2074-84.
59. Cooper JS, Guo MD, Herskovic A, Macdonald JS, Martenson JA, Jr., Al-Sarraf M, Byhardt R, Russell AH, Beitler JJ, Spencer S, Asbell SO, Graham MV, Leichman LL. Chemoradiotherapy of locally advanced esophageal cancer: long-term follow-up of a prospective randomized trial (RTOG 85-01). Radiation Therapy Oncology Group. *JAMA.* 1999;281(17):1623-7.

60. Cheedella NK, Suzuki A, Xiao L, Hofstetter WL, Maru DM, Taketa T, Sudo K, Blum MA, Lin SH, Welch J, Lee JH, Bhutani MS, Rice DC, Vaporciyan AA, Swisher SG, Ajani JA. Association between clinical complete response and pathological complete response after preoperative chemoradiation in patients with gastroesophageal cancer: analysis in a large cohort. *Ann Oncol.* 2013;24(5):1262-6.
61. Burton EC, Troxclair DA, Newman lii WP. Autopsy Diagnoses of Malignant Neoplasms. *Jama.* 1998;280(14):1245.
62. Lu W, Wang J, Zhang HH. Computerized PET/CT image analysis in the evaluation of tumour response to therapy. *Br J Radiol.* 2015;88(1048):20140625.
63. Thibault G, Fertil B, Navarro C, Pereira S, Cau P, Levy N, Sequeira J, Mari J-L. Texture indexes and gray level size zone matrix. Application to cell nuclei classification. *Pattern Recognition and Information Processing.* 2009:140-5.
64. Galloway MM. Texture Analysis Using Gray Level Run Lengths. *Computer Graphics and Image Processing.* 1975;4:172-9.
65. Tustison NJ, Gee JC. Run Length Matrices For Texture Analysis. 2008.
66. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They are Data. *Radiology.* 2016;278(2):563-77.
67. Hatt M, Tixier F, Rest CCL, Pradier O, Visvikis D. Robustness of intratumour <sup>18</sup>F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *European journal of nuclear medicine and molecular imaging.* 2013;40:1662-71.
68. Brunetti J. PET-CT in Radiation Treatment Planning. 2011. In: *Clinical PET-CT in Radiology: Integrated Imaging in Oncology* [Internet]. New York: Springer; [121-9].
69. Adams MC, Turkington TG, Wilson JM, Wong TZ. A systematic review of the factors affecting accuracy of SUV measurements. *AJR Am J Roentgenol.* 2010;195(2):310-20.

70. Osman MM, Wall A, Miller MD, Nguyen N, Oliver DA, Bucholz RD, Walz BJ.  $^{18}\text{F}$ FDG-PET/CT for Treatment Planning and Follow-Up. 2007. In: Robotic Radiosurgery Treating Tumors that Move with Respiration [Internet]. Springer; [55-60].
71. Buvat I, Orhac F, Soussan M. Tumor Texture Analysis in PET: Where Do We Stand? *J Nucl Med*. 2015;56(11):1642-4.
72. Ganeshan B, Skogen K, Pressney I, Coutroubis D, Miles K. Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clin Radiol*. 2012;67(2):157-64.
73. Ganeshan B, Goh V, Mandeville HC, Ng QS, Hoskin PJ, Miles KA. Non-Small Cell Lung Cancer: Histopathologic Correlates for Texture Parameters at CT. *Radiology*. 2013;266(1):326-36.
74. Bagci U, Yao J, Miller-Jaster K, Chen X, Mollura DJ. Predicting future morphological changes of lesions from radiotracer uptake in  $^{18}\text{F}$ -FDG-PET images. *PLoS One*. 2013;8(2):e57105.
75. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *European journal of cancer*. 2009;45(2):228-47.
76. Dong X, Xing L, Wu P, Fu Z, Wan H, Li D, Yin Y, Sun X, Yu J. Three-dimensional positron emission tomography image texture analysis of esophageal squamous cell carcinoma: relationship between tumor  $^{18}\text{F}$ -fluorodeoxyglucose uptake heterogeneity, maximum standardized uptake value, and tumor stage. *Nuclear Medicine Communications*. 2012;34:40-6.

77. Hatt M, Majdoub M, Vallieres M, Tixier F, Rest CCL, Groheux D, Hindie E, Martineau A, Pradier O, Hustinx R, Perdrisot R, Guillevin R, Naqa IE, Visvikis D. 18F-FDG PET Uptake Characterization Through Texture Analysis: Investigating the Complementary Nature of Heterogeneity and Functional Tumor Volume in a Multi-Cancer Site Patient Cohort. *J Nucl Med.* 2014;56(1):38-44.
78. Brooks FJ, Grigsby PW. The Effect of Small Tumor Volumes on Studies of Intratumoral Heterogeneity of Tracer Uptake. *J Nucl Med.* 2015;55(1):37-42.
79. Orhac F, Michael S, Maisonobe J-A, Garcia CA, Vanderlinden B, Buvat I. Tumor Texture Analysis in <sup>18</sup>F-FDG PET: Relationships Between Texture Parameters, Histogram Indices, Standardized Uptake Values, Metabolic Volumes, and Total Lesion Glycolysis. *J Nucl Med.* 2014;55(3):414-22.
80. Tan S, Kligerman S, Chen W, Lu M, Kim G, Feigenberg S, D'Souza WD, Suntharalingam M. Spatial-Temporal [<sup>18</sup>F]FDG-PET Features for Predicting Pathologic Response of Esophageal Cancer to Neoadjuvant Chemoradiation Therapy. *International journal of radiation oncology, biology, physics.* 2012;85(5):1375-82.
81. Yan J, Chu-Shern JL, Loi HY, Khor LK, Sinha AK, Quek ST, Tham IW, Townsend D. Impact of Image Reconstruction Settings on Texture Features in <sup>18</sup>F-FDG PET. *J Nucl Med.* 2015;56(11):1667-73.
82. Chalkidou A, O'Doherty MJ, Marsden PK. False Discovery Rates in PET and CT Studies with Texture Features: A Systematic Review. *PLoS One.* 2015;10(5):e0124165.
83. Oliver JA, Budzevich M, Zhang GG, Dilling TJ, Latifi K, Moros EG. Variability of Image Features Computed from Conventional and Respiratory-Gated PET/CT Images of Lung Cancer. *Transl Oncol.* 2015;8(6):524-34.



84. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, Bussink J, Monshouwer R, Haibe-Kains B, Rietveld D, Hoebbers F, Rietbergen MM, Leemans CR, Dekker A, Quackenbush J, Gillies RJ, Lambin P. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*. 2014;5:4006.
85. Kumar V, Nath K, Berman CG, Kim J, Tanvetyanon T, Chiappori AA, Gatenby RA, Gillies RJ, Eikman EA. Variance of SUVs for FDG-PET/CT is Greater in Clinical Practice Than Under Ideal Study Settings. *Clinical Nuclear Medicine*. 2013;38(3):175-82.
86. Ravanelli M, Farina D, Morassi M, Roca E, Cavalleri G, Tassi G, Maroldi R. Texture analysis of advanced non-small cell lung cancer (NSCLC) on contrast-enhanced computed tomography: prediction of the response to the first-line chemotherapy. *European radiology*. 2013;23(12):3450-5.
87. De Ruyscher D, Sharifi H, Defraene G, Kerns SL, Christiaens M, De Ruyck K, Peeters S, Vansteenkiste J, Jeraj R, Van Den Heuvel F, van Elmpt W. Quantification of radiation-induced lung damage with CT scans: the possible benefit for radiogenomics. *Acta oncologica*. 2013;52(7):1405-10.
88. Balagurunathan Y, Kumar V, Gu Y, Kim J, Wang H, Liu Y, Goldgof DB, Hall LO, Korn R, Zhao B, Schwartz LH, Basu S, Eschrich S, Gatenby RA, Gillies RJ. Test-retest reproducibility analysis of lung CT image features. *J Digit Imaging*. 2014;27(6):805-23.
89. Grove O, Berglund AE, Schabath MB, Aerts HJWL, Dekker A, Wang H, Velazquez ER, Lambin P, Gu Y, Balagurunathan Y, Eikman E, Gatenby RA, Eschrich S, Gillies RJ. Quantitative Computed Tomographic Descriptors Associate Tumor Shape Complexity and Intratumor Heterogeneity with Prognosis in Lung Adenocarcinoma. *PLoS One*. 2015;10(3):e0118261.
90. QIBA Profile. FDG-PET/CT as an Imaging Biomarker Measuring Response to Cancer Therapy. Quantitative Imaging Biomarkers Alliance, 2013.

91. Kalpathy-Cramer J, Freymann JB, Prior FW. Quantitative Imaging Network: Data Sharing and Competitive Algorithm Validation Leveraging The Cancer Imaging Archive. *Transl Oncol.* 2014;7(1):147-52.
92. Dholakia AS, Chaudhry M, Leal JP, Chang DT, Raman SP, Hacker-Prietz A, Su Z, Pai J, Oteiza KE, Griffith ME, Wahl RL, Tryggestad E, Pawlik T, Laheru DA, Wolfgang CL, Koong AC, Herman JM. Baseline metabolic tumor volume and total lesion glycolysis are associated with survival outcomes in patients with locally advanced pancreatic cancer receiving stereotactic body radiation therapy. *International Journal of Radiation Oncology\*Biophysics\*Physics.* 2014;89(3):539-46.
93. Ion AL. Methods for Knowledge Discovery in Images. *Inf Technol Control.* 2009;38(1):43-50.
94. Gadkari D. Image quality analysis using GLCM: University of Central Florida; 2004.
95. Wallis KF. A note on the calculation of entropy from histograms. Department of Economics: University of Warwick, 2006.
96. Liang M. 3D co-occurrence matrix based texture analysis applied to cervical cancer screening: Uppsala Universitet; 2012.
97. Xiuhua G, Tao S, huan W, Zhigang L. Prediction Models for Malignant Pulmonary Nodules Based-on Texture Features of CT Image. 2011. In: *Theory and Applications of CT Imaging and Analysis [Internet]. Available from: <http://www.intechopen.com/books/theory-and-applications-of-ct-imaging-and-analysis/prediction-models-for-malignant-pulmonary-nodules-based-on-texture-features-of-ct-image>.*
98. Haralick RM. Statistical and Structural Approaches to Texture. *Proceedings of the IEEE.* 1979;67(5):786-803.
99. Chu A, Sehgal CM, Greenleaf JF. Use of gray value distribution of run lengths for texture analysis. *Pattern Recognition Letters.* 1990;11(6):415-20.

100. Dasarathy BV, Holder EB. Image characterizations based on joint gray-level run-length distributions. *Pattern Recognition Letters*. 1991;12(8):497-502.
101. Thibault G, Angulo J, Meyer F, editors. Advanced statistical matrices for texture characterization: Application to DNA chromatin and microtubule network classification. 18th IEEE International Conference on Image Processing (ICIP); 2011; Brussels, Germany: IEEE.
102. Lardinois D, Weder W, Hany TF, Kamel EM, Korom S, Seifert B, von Schulthess GK, Steinert HC. Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. *N Engl J Med*. 2003;348(25):2500-7.
103. Nestle U, Kremp S, Schaefer-Schuler A, Sebastian-Welsch C, Hellwig D, Rube C, Kirsch CM. Comparison of different methods for delineation of <sup>18</sup>F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-Small cell lung cancer. *J Nucl Med*. 2005;46(8):1342-8.
104. Erdi YE, Macapinlac H, Rosenzweig KE, Humm JL, Larson SM, Erdi AK, Yorke ED. Use of PET to monitor the response of lung cancer to radiation treatment. *Eur J Nucl Med*. 2000;27(7):861-6.
105. Bradley J, Thorstad WL, Mutic S, Miller TR, Dehdashti F, Siegel BA, Bosch W, Bertrand RJ. Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer. *International journal of radiation oncology, biology, physics*. 2004;59(1):78-86.
106. Vicente AMG, Castrejon AS, Martin AAL, Garcia BG, Woll JPP, Munoz AP. Value of 4-Dimensional <sup>18</sup>F-FDG PET/CT in the Classification of Pulmonary Lesions. *J Nucl Med*. 2011;39(2):91-9.
107. Yip S, McCall K, Aristophanous M, Chen AB, Aerts HJ, Berbeco R. Comparison of texture features derived from static and respiratory-gated PET images in non-small cell lung cancer. *PLoS One*. 2014;9(12):e115510.

108. Kuykendall CC, Budzevich MM, Latifi K, Moros EG, Hoffe SE, Dilling TJ, Zhang GG, Montilla-Soler JL, Eikman EA. 4D PET/CT: Radiology Imaging to Radiation Therapy. Practical Radiation Oncology. 2013;3(2 Suppl 1):S28.
109. Boellaard R, Oyen WJ, Hoekstra CJ, Hoekstra OS, Visser EP, Willemsen AT, Arends B, Verzijlbergen FJ, Zijlstra J, Paans AM, Comans EF, Pruim J. The Netherlands protocol for standardisation and quantification of FDG whole body PET studies in multi-centre trials. European journal of nuclear medicine and molecular imaging. 2008;35(12):2320-33.
110. Kurani AS, Xu D-H, Furst JD, Raicu DS, editors. Co-occurrence matrices for volumetric data. The 7<sup>th</sup> IASTED International Conference on Computer Graphics and Imaging; CGIM; 2004 August 17-19; Kauai, Hawaii, USA.
111. Tustison NJ, Gee JC. Run Length Matrices For Texture Analysis. The Insight Journal. 2008(January-June 2008).
112. Xu D-H, Kurani AS, Furst JD, Raicu DS. Run-length encoding for volumetric texture. The 4<sup>th</sup> IASTED International Conference on Visualization, Imaging and Image Processing; VIIP; September 6-8; Marbella, Spain 2004.
113. McBride GB. Equivalence measures for comparing the performance of alternative methods for the analysis of water quality variables. New Zealand: National Institute of Water & Atmospheric Research Ltd 2007.
114. Low DA, White BM, Lee PP, Thomas DH, Gaudio S, Jani SS, Wu X, Lamb JM. A novel CT acquisition and analysis technique for breathing motion modeling. Physics in medicine and biology. 2013;58(11):L31-6.
115. Teo BK, Saboury B, Munbodh R, Scheuermann J, Torigian DA, Zaidi H, Alavi A. The effect of breathing irregularities on quantitative accuracy of respiratory gated PETCT. Med Phys. 2012;39(12):7390-7.

116. Cook GJR, Yip C, Siddique M, Goh V, Chicklore S, Roy A, Marsden P, Ahmad S, Landau D. Are Pretreatment  $^{18}\text{F}$ -FDG PET Tumor Textural Features in Non-Small Cell Lung Cancer Associated with Response and Survival After Chemoradiotherapy. *Journal of Nuclear Medicine*. 2013;54:19-26.
117. Boellaard R. Standards for PET image acquisition and quantitative data analysis. *Journal of Nuclear Medicine*. 2009;50(5 (Suppl)):11S-20S.
118. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. *Journal of Nuclear Medicine*. 2009;50 Suppl 1(5 (Suppl)):122S-50S.
119. Latifi K, Huang TC, Feygelman V, Budzevich MM, Moros EG, Dilling TJ, Stevens CW, van Elmpt W, Dekker A, Zhang GG. Effects of quantum noise in 4D-CT on deformable image registration and derived ventilation data. *Physics in medicine and biology*. 2013;58(21):7661-72.
120. Nestle U, Kremp S, Grosu A-L. Practical integration of [ $^{18}\text{F}$ ]-FDG-PET and PET-CT in the planning of radiotherapy for non-small cell lung cancer (NSCLC): The technical basis, ICRU-target volumes, problems, perspectives. *Radiotherapy and Oncology*. 2006;81(2):209-25.
121. Brunetti J. PET-CT in Radiation Treatment Planning. 2011:121-9.
122. Doumou G, Siddique M, Tsoumpas C, Goh V, Cook GJ. The precision of textural analysis in  $^{18}\text{F}$ -FDG-PET scans of oesophageal cancer. *European radiology*. 2015.
123. Harrison RL, Elston Bf, Doot RK, Lewellen TK, Mankoff DA, Kinahan PE. A Virtual Clinical Trial of FDG-PET Imaging of Breast Cancer: Effect of Variability on Response Assessment. *Translational Oncology*. 2014;7(1):138-46.
124. Cancer Facts & Figures 2015. 2015.
125. Shridhar R, Almhanna K, Meredith KL, Biagoli MC, Chuong MD, Cruz A, Hoffe SE. Radiation Therapy and Esophageal Cancer. *Cancer Control*. 2013;20(2):97-110.

126. Spratt DE, Diaz R, McElmurray J, Csiki I, Duggan D, Lu B, Delbeke D. Impact of FDG PET/CT on delineation of the gross tumor volume for radiation planning in non-small-cell lung cancer. *Clin Nucl Med.* 2010;35(4):237-43.
127. Zheng Y, Sun X, Wang J, Zhang L, Di X, Xu Y. FDG-PET/CT imaging for tumor staging and definition of tumor volumes in radiation treatment planning in non-small cell lung cancer. *Oncol Lett.* 2014;7(4):1015-20.
128. Gregoire V, Haustermans K, Geets X, Roels S, Lonneux M. PET-based treatment planning in radiotherapy: a new standard? *J Nucl Med.* 2007;48 Suppl 1(1):68S-77S.
129. Joiner MC, van der Kogel A. PET Image Segmentation. *Basic Clinical Radiobiology Fourth Edition.* 4 ed: CRC Press; 2009.
130. Wang Y-C, HSIEH T-C, Yu C-Y, Yen K-Y, Chen S-W, Yang S-N, Chien C-R, HSU S-M, Pan T, Kao C-H, Liang J-A. The clinical application of 4D <sup>18</sup>F-FDG PET/CT on gross tumor volume delineation for radiotherapy planning in esophageal squamous cell cancer. *Journal of Radiation Research.* 2012;53:594-600.
131. Westerterp M, Westreenen HLv, Reitsma JB, Hoekstra OS, Stoker J, Fockens P, Jager PL, Eck-Smit BLFV, Plukker JTM, Lanschot JJBv, Sloof GW. Esophageal Cancer: CT, Endoscopic US, and FDG PET for Assessment of Response to Neoadjuvant Therapy—Systematic Review. *Radiology.* 2005;236(3):841-51.
132. Wieder HA, Ott K, Lordick F, Becker K, Stahl A, Herrmann K, Fink U, Siewert JR, Schwaiger M, Weber WA. Prediction of tumor response by FDG-PET: comparison of the accuracy of single and sequential studies in patients with adenocarcinomas of the esophagogastric junction. *European journal of nuclear medicine and molecular imaging.* 2007;34(12):1925-32.

133. Lordick F, Ott K, Krause BJ, Weber WA, Becker K, Stein HJ, Lorenzen S, Schuster T, Wieder H, Herrmann K, Bredenkamp R, Hofler H, Fink U, Peschel C, Schwaiger M, Siewert JR. PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial. *The Lancet Oncology*. 2007;8(9):797-805.
134. Muijs CT, Beukema JC, Pruim J, Mul VE, Groen H, Plukker JT, Langendijk JA. A systematic review on the role of FDG-PET/CT in tumour delineation and radiotherapy planning in patients with esophageal cancer. *Radiother Oncol*. 2010;97(2):165-71.
135. Fernandez DC, Hoffe SE, Barthel JS, Vignesh S, Klapman JB, Harris C, Almhanna K, Biagioli MC, Meredith KL, Feygelman V, Rao NG, Shridhar R. Stability of endoscopic ultrasound-guided fiducial marker placement for esophageal cancer target delineation and image-guided radiation therapy. *Practical Radiation Oncology*. 2013;3(1):32-9.
136. Fukada J, Hanada T, Kawaguchi O, Ohashi T, Takeuchi H, Kitagawa Y, Seki S, Shiraishi Y, Ogato H, Shigematsu N. Detection of Esophageal Fiducial Marker Displacement During Radiation Therapy With a 2-dimensional On-board Imager: Analysis of Internal Margin for Esophageal Cancer. *International journal of radiation oncology, biology, physics*. 2013;85(4):991-8.
137. Jin P, Hulshof MC, de Jong R, van Hooft JE, Bel A, Alderliesten T. Quantification of respiration-induced esophageal tumor motion using fiducial markers and four-dimensional computed tomography. *Radiother Oncol*. 2016.
138. Yamashita H, Kida S, Sakumi A, Haga A, Ito S, Onoe T, Okuma K, Ino K, Akahane M, Ohtomo K, Nakagawa K. Four-dimensional measurement of the displacement of internal fiducial markers during g320-multislice computed tomography scanning of thoracic esophageal cancer. *International journal of radiation oncology, biology, physics*. 2011;79(2):588-95.

139. Machiels M, Van Hooft J, Jin P, van Berge Henegouwen MI, van Laarhoven HM, Alderliesten T, Hulshof MC. Endoscopy/EUS-guided fiducial marker placement in patients with esophageal cancer: a comparative analysis of 3 types of markers. *Gastrointest Endosc.* 2015;82(4):641-9.
140. Freilich J, Hoffe SE, Almhanna K, Dinwoodie W, Yue B, Fulp W, Meredith KL, Shridhar R. Comparative outcomes for three-dimensional conformal versus intensity-modulated radiation therapy for esophageal cancer. *Diseases of the Esophagus.* 2015;28(4):352-7.
141. *Cancer Facts & Figures 2010.* 2010.
142. *Cancer Facts & Figures 2011.* 2011.
143. *Cancer Facts & Figures 2012.* 2012.
144. Venkat P, Oliver JA, Jin W, Dault J, Frakes JM, Hoffe SE, Pimiento JM, Almhanna K, Fontaine J, Pena L, Latifi K. Prognostic value of 18F-FDG PET/CT metabolic tumor volume for complete pathologic response and clinical outcomes after neoadjuvant chemoradiation therapy for locally advanced esophageal cancer. *J Clin Oncol.* 2016;34(suppl 4S).
145. Shapiro J, van Lanschot JJ, Hulshof MC, van Hagen P, van Berge Henegouwen MI, Wijnhoven BP, van Laarhoven HW, Nieuwenhuijzen GA, Hospers GA, Bonenkamp JJ, Cuesta MA, Blaisse RJ, Busch OR, Ten Kate FJ, Creemers GJ, Punt CJ, Plukker JT, Verheul HM, Bilgen EJ, van Dekken H, van der Slangen MJ, Rozema T, Biermann K, Beukema JC, Piet AH, van Rij CM, Reinders JG, Tilanus HW, Steyerberg EW, van der Gaast A. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. *The Lancet Oncology.* 2015;16(9):1090-8.
146. Bouchard M, McAleer MF, Starkschall G. Impact of gastric filling on radiation dose delivered to gastroesophageal junction tumors. *International journal of radiation oncology, biology, physics.* 2010;77(1):292-300.



147. Zhong X, Yu J, Zhang B, Mu D, Zhang W, Li D, Han A, Song P, Li H, Yang G, Kong FM, Fu Z. Using 18F-fluorodeoxyglucose positron emission tomography to estimate the length of gross tumor in patients with squamous cell carcinoma of the esophagus. International journal of radiation oncology, biology, physics. 2009;73(1):136-41.
148. Latifi K, Zhang G, Stawicki M, van Elmpt W, Dekker A, Forster K. Validation of three deformable image registration algorithms for the thorax. J Appl Clin Med Phys. 2013;14(1):3834.

## Appendix A: List of Publications

**J.A. Oliver**, M. Budzevich, G.G. Zhang, T.J. Dilling, K. Latifi, E.G. Moros. Variability of Image Features Computed From Conventional and Respiratory-Gated PET/CT Images of Lung Cancer. *Translational Oncology*. 2015; 8(6):524-34.  
<http://dx.doi.org/10.1016/j.tranon.2015.11.013>

Latifi, K., **Oliver, J.**, Baker, R., Dilling, T.J., Stevens, C., Kim, J., Yue, B., DeMarco, M., Zhang, G., Moros, E., Feygelman, V. 2014. Study of 201 non small-cell lung cancer (NSCLC) SABR patients shows local control dependence on dose calculation algorithm. *Int J Rad Onc Biol Phys*. 2014 Apr 1;88(5):1108-13. doi: 10.1016/j.ijrobp.2013.12.047. Epub 2014 Feb 12. PMID: 24529716.

UNDER REVIEW:

**J.A. Oliver**, M. Budzevich, D. Hunt, E.G. Moros, K. Latifi, T.J. Dilling, V. Feygelman, G.G. Zhang. Sensitivity of Image Features to Noise in Conventional and Respiratory-Gated PET/CT Images of Lung Cancer: Uncorrelated Noise Effects.

Venkat, P., **Oliver, J.A.**, Montilla-Soler, J., Klapman, J., Dhadham, G.C., Harris, C., Emanuel, K., Werner, J., Frakes, J.M., Shridhar, R., Dilling, T.J., Zhang, G.G., Moros, E.G., Hoffe, S.E., Latifi, K. Prognostic Value of <sup>18</sup>F-FDG PET/CT Metabolic Tumor Volume for Complete Pathologic Response and Clinical Outcomes after Neoadjuvant Chemoradiation Therapy for Locally Advanced Esophageal Cancer.

**Oliver, J.A.**, Venkat, P., Montilla-Soler, J., Klapman, J., Dhadham, G.C., Harris, C., Emanuel, K., Werner, J., Frakes, J.M., Shridhar, R., Dilling, T.J., Zhang, G.G., Moros, E.G., Hoffe, S.E., Latifi, K. Fiducials vs. 18F-FDG PET/CT for Esophageal Cancer GTV Delineation for Radiotherapy Treatment Planning.

UNPUBLISHED MANUSCRIPTS:

**J.A. Oliver, G. G. Zhang**, M. Budzevich, K. Latifi, S.E. Hoffe, T.J. Dilling, R.J. Gillies, E.G. Moros. PET/CT Image Feature Analysis: A Review.

M. Budzevich, **J.A. Oliver**, G.G. Zhang, K. Latifi, K. Kuykendall, T.J. Dilling, E.G. Moros. Determining the Optimal Number of Bins for 4D PET/CT using Phantom.

ABSTRACTS:

**J.A. Oliver**, M.M. Budzevich, D.C. Hunt, E.G. Moros, G.G. Zhang. Effect of Uncorrelated Noise Texture On Computed Tomography Quantitative Image Features. July 2015. Abstract for ePoster Presentation, American Association of Medical Physicists Annual Meeting, Anaheim, CA.

**Oliver, J.A.**, Montilla-Soler, J., Klapman, J., Dhadham, G.C., Harris, C., Emanuel, K., Werner, J., Frakes, J.M., Shridhar, R., Dilling, T.J., Zhang, G.G., Moros, E.G., Hoffe, S.E., Latifi, K. Fiducials vs. 18F-FDG PET/CT for Esophageal Cancer GTV Delineation for Radiotherapy Treatment Planning. May 2015. Abstract for Poster Presentation, Moffitt Scientific Symposium, Tampa, FL

**J. Oliver**, M. Budzevich, G. Zhang, C. Kuykendall, T. Dilling, E. Moros. Texture Analysis of Conventional and Respiratory-Gated PET/CT Images in Lung Cancer. February 2015. Oral Presentation. Emerging Researchers National Conference, Washington, D.C.

**J. Oliver**, M. Budzevich, G. Zhang, K. Latifi, T. Dilling, Y. Balagurnathan, Y. Gu, O. Grove, V. Feygelman, R. Gillies, E. Moros. Dependence of 3D/4D PET Quantitative Image Features On Noise. Med. Phys. 41, 380 (2014). <http://dx.doi.org/10.1118/1.4888997>.

**J. Oliver**, M. Budzevich, G. Zhang, K. Latifi, T. Dilling, Y. Balagurnathan, Y. Gu, O. Grove, V. Feygelman, R. Gillies, E. Moros. Dependence of Computerized Tomography Image Features on Noise. May 2014. Abstract for Poster Presentation, Moffitt Scientific Symposium, Tampa FL.

**J.A. Oliver**, M. Budzevich, G. Zhang, K. Latifi, C. Kuykendall, S. Hoffe, J. Montilla-Soler, E. Eikman, T. Dilling, E. G. Moros. Texture Analysis of 3D and 4D Positron Emission Tomography Images in Adenocarcinoma of the Lung. August 2013. Abstract for Poster Presentation, American Association of Medical Physicists Annual Meeting, Indianapolis, IN

K Latifi, **J Oliver**, T Dilling, C Stevens, M DeMarco, G Zhang, E Moros, V Feygelman. Local Control Differences for SBRT Lung Patients Planned with Pencil Beam Vs. Collapsed Cone Convolution Algorithms. August 2013. Abstract for Oral Presentation, American Association of Medical Physicists Annual Meeting, Indianapolis, IN

**J.A. Oliver**, M. Budzevich, G. Zhang, K. Latifi, C. Kuykendall, S. Hoffe, J. Montilla-Soler, E. Eikman, T. Dilling, E. G. Moros. Texture Analysis of 3D and 4D Positron Emission Tomography Images in Adenocarcinoma of the Lung. May 2013. Abstract for Poster Presentation, Moffitt Scientific Symposium, Tampa, FL

**J.A. Oliver**, M. Budzevich, G. Zhang, K. Latifi, C. Kuykendall, S. Hoffe, J. Montilla-Soler, E. Eikman, T. Dilling, E. G. Moros. Positron Emission Tomography Texture Analysis of Necrosis in Lung Cancer Tumors. February 2013. Abstract for Poster Presentation, Cancer Imaging and Radiation Therapy Symposium, Orlando, FL

**J.A. Oliver**, M. Budzevich, G. Zhang, K. Latifi, C. Kuykendall, S. Hoffe, J. Montilla-Soler, E. Eikman, T. Dilling, E. G. Moros. Positron Emission Tomography Texture Analysis of Necrosis in Lung Cancer Tumors. October 2012. Abstract for Poster Presentation, HENAAC Great Minds in Stem, Orlando, FL

K. Forster, J. Oliver, S. Hoffe, R. Gillies, E. Eikman, T. Dilling. The optimization of 4D PET acquisition parameters for radiation oncology. June 2012. Abstract for Poster Presentation, Society of Nuclear Medicine Annual Meeting, Miami Beach, FL

K. Forster, J. Oliver, S. Hoffe, R. Gillies, E. Eikman, T. Dilling. The effect of PET reconstruction algorithms on tumor delineation for radiation oncologists. June 2012. Abstract for Poster Presentation, Society of Nuclear Medicine Annual Meeting, Miami Beach, FL

#### TALKS:

**J.A. Oliver**, M.M. Budzevich, D.C. Hunt, E.G. Moros, G.G. Zhang. Dependence of Image Features on Noise in Conventional and Respiratory-Gated PET/CT Images of Lung Cancer. Oral Presentation. Radiomics Retreat, Clearwater FL.

**J.A. Oliver**, M.M. Budzevich, D.C. Hunt, E.G. Moros, G.G. Zhang. Impact of Noise on Radiomics Features Extracted from 3D & 4D CT Images of Lung Cancer. April 2015. Oral Presentation. FLAAPM 2015 Spring Meeting, Orlando, FL.

**J. Oliver**, M. Budzevich, G. Zhang, C. Kuykendall, T. Dilling, E. Moros. Texture Analysis of Conventional and Respiratory-Gated PET/CT Images in Lung Cancer. February 2015. Oral Presentation. Emerging Researchers National Conference, Washington, D.C.

**J.A. Oliver**, M. Budzevich, G. Zhang, T. Dilling, E.G. Moros. Texture Variability Between 3D and 4D PET/CT modalities. October 2013. Oral Presentation, Moffitt Cancer Center Radiomics Retreat, Clearwater, FL.

## Appendix B: Copyright Permissions

**J.A. Oliver**, M. Budzevich, G.G. Zhang, T.J. Dilling, K. Latifi, E.G. Moros. *Variability of Image Features Computed From Conventional and Respiratory-Gated PET/CT Images of Lung Cancer. Translational Oncology*. 2015; 8(6):524-34.

<http://www.sciencedirect.com/science/article/pii/S1936523315000832>



RightsLink®

Home

Account Info

Help



**Title:** Variability of Image Features Computed from Conventional and Respiratory-Gated PET/CT Images of Lung Cancer

**Author:** Jasmine A. Oliver, Mikalai Budzevich, Geoffrey G. Zhang, Thomas J. Dilling, Kujtim Latifi, Eduardo G. Moros

**Publication:** Translational Oncology

**Publisher:** Elsevier

**Date:** December 2015

Copyright © 2015 The Authors. Published by Elsevier Inc.

Logged in as:  
Jasmine Oliver

LOGOUT

### Creative Commons Attribution-NonCommercial-No Derivatives License (CC BY NC ND)

This article is published under the terms of the [Creative Commons Attribution-NonCommercial-No Derivatives License \(CC BY NC ND\)](#).

For non-commercial purposes you may copy and distribute the article, use portions or extracts from the article in other works, and text or data mine the article, provided you do not alter or modify the article without permission from Elsevier. You may also create adaptations of the article for your own personal use only, but not distribute these to others. You must give appropriate credit to the original work, together with a link to the formal publication through the relevant DOI, and a link to the Creative Commons user license above. If changes are permitted, you must indicate if any changes are made but not in any way that suggests the licensor endorses you or your use of the work.

Permission is not required for this non-commercial use. For commercial use please continue to request permission via Rightslink.

BACK

CLOSE WINDOW

Copyright © 2015 [Copyright Clearance Center, Inc.](#) All Rights Reserved. [Privacy statement](#). [Terms and Conditions](#). Comments? We would like to hear from you. E-mail us at [customercare@copyright.com](mailto:customercare@copyright.com)

## Appendix C: IRB MCC 16784 Determination Letter



August 8, 2007

Thomas Dilling, M.D.  
Department of Interdisciplinary Oncology  
MDC 44 / MCC-MOD 9

Attn: Margaret Lomartire

RE: **Expedited Approval** for Initial Review including your Application for Waiver/Alteration of HIPAA Authorization  
IRB#: 105996d  
Title: *Retrospective Analysis of Image-Guided Radiation Treatment Delivery to the Lung on a Novalis Treatment Machine MCC# 15269*  
Study Approval Period: 07/31/07 to 07/29/08

Dear Dr. Dilling:

On 07/31/07, Institutional Review Board (IRB) reviewed and **APPROVED** the above protocol **for the period indicated above**. It was the determination of the IRB that your study qualified for expedited review based on the federal expedited category number 5. This approval includes a Waiver of Informed Consent.

The IRB also reviewed and **APPROVED** your Application for Waiver/Alteration of HIPAA Authorization for the above protocol as outlined below:

**Your HIPAA Waiver/Alteration application has been approved for you to access IMPAC and CernerPowerchart systems at H. Lee Moffitt Cancer Center (MCC) in order to conduct a retrospective review of medical records involving approximately fifty patients who were diagnosed with lung cancer and who were treated with image guided radiation therapy (IGRT) on the Novalis treatment machine at MCC, during May 2006 to July 10, 2007.**

Please note: approval of your waiver/alteration of HIPAA authorization is based upon your having satisfied the following HIPAA Privacy Rule mandates:

1. The use or disclosure of protected health information (PHI) involves no more than minimal risk to the privacy of the individual, based on the presence of: a) an adequate plan to protect the identifiers from improper use and disclosure; b) an adequate plan to destroy the identifiers at the earliest opportunity consistent with the conduct of the research unless there is health or research justification for retaining the identifiers or such retention is otherwise required by law; and c) adequate written assurance that the PHI will not be reused or disclosed to any other person or entity, EXCEPT 1) as required by law, 2) for authorized oversight of the research study, or 3) for other research for which the use or disclosure of PHI would be permitted by the HIPAA Privacy Regulations.

DIVISION OF RESEARCH INTEGRITY & COMPLIANCE  
University of South Florida • 3702 Spectrum Blvd., Suite 155 • Tampa, FL 33612-9445  
(813) 974-5638 • Fax (813) 974-7091



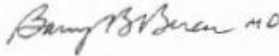
2. The research could not practicably be conducted without the waiver or alteration; and
3. The research could not practicably be conducted without access to and use of the PHI.

Please note, if applicable, the **enclosed informed consent/assent documents are valid during the period indicated by the official, IRB-Approval stamp located on page one of the form.** Valid consent must be documented on a copy of the most recently IRB-approved consent form. Make copies from the enclosed original.

**Please reference the above IRB protocol number in all correspondence** regarding this protocol with the IRB or the Division of Research Integrity and Compliance. In addition, we have enclosed an Institutional Review Board (IRB) Quick Reference Guide providing guidelines and resources to assist you in meeting your responsibilities in the conduction of human subjects research. Please read this guide carefully. It is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to the Human Research Protections Program. If you have any questions regarding this matter, please call 813-974-9343, or, for HIPAA-specific questions, please contact Vinita Witanachchi J.D., Research Privacy Officer, at 974-5478.

Sincerely,



Barry B. Bercu, M.D., Chairperson  
USF Institutional Review Board

Enclosures: (If applicable) IRB-Approved, Stamped Informed Consent/Assent Document(s)  
IRB Quick Reference Guide

Cc: Sandra Partap, USF IRB Professional Staff  
Vinita Witanachchi, J.D., Research Privacy Officer  
MCC  
FAO

IRB Approval-EXPEDITED+Waiver-0601



RESEARCH INTEGRITY AND COMPLIANCE  
Institutional Review Boards, FWA No. 00001669  
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799  
(813) 974-5638 • FAX (813) 974-7091

6/26/2013

Thomas Dilling, M.D.  
H Lee Moffitt Cancer Center  
Moffitt Cancer Ctr - Radiation Oncology  
12902 Magnolia Drive  
Tampa, FL 33606

RE: **Expedited Approval for Amendment**

IRB#: Ame1\_Pro00006748

Title: Retrospective Assessment of Patients Treated with Chemoradiotherapy (CRT) to 70 Gy  
for Locally-Advanced Non-Small Cell Lung Cancer (NSCLC)

Dear Dr. Dilling:

On 6/25/2013, the Institutional Review Board (IRB) reviewed and **APPROVED** your Amendment. The submitted request has been approved for the following:

- Deletion of Dr. Craig Stevens, Andrew Vivas, and Ryan Baker from study staff
- Addition of Olya Grove and Jasmine Oliver to study staff

Approved Item(s):

Protocol Document(s):

[16784 Protocol 70 Gy Chemorads Ver 3 dated 06/25/2013](#)

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

E. Verena Jorgensen, M.D., Chairperson  
USF Institutional Review Board

## Appendix D: IRB MCC 16567 Determination Letter

**From:** [rsch-arc@usf.edu](mailto:rsch-arc@usf.edu)  
**To:** [Lomartire, Margaret M.](#)  
**Subject:** eIRB: Personnel Change Request Approved  
**Date:** Tuesday, June 30, 2015 11:14:13 AM

---



### IRB Personnel Change Request Approved

**To:** Margaret Lomartire  
**RE:** Amendment 19 for IRB Study #Pro00003382  
Retrospective analysis of all patients with gastrointestinal malignancies (primary and metastatic) treated with external beam radiation, brachytherapy, or radioembolization at MCC  
**PI:** Sarah Hoffe  
**Link:** [Ame19\\_Pro00003382](#)

On 6/30/2015, the Institutional Review Board (IRB) reviewed and **APPROVED** your Personnel Change Request. The submitted request has been approved for the following:

**Description of Changes**  
Key Personnel added: Jasmine Oliver

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call the Office of Research Integrity & Compliance at 813-974-5638.

---

*DO NOT REPLY: To ensure a timely response, please direct correspondence to Research Integrity & Compliance either through your project's workspace or the contact information below.*

**Research Integrity & Compliance**  
University of South Florida - Research and Innovation  
ARC Help Desk (eIRB, eIACUC, eCOI): (813) 974-2880  
Email: [rsch-arc@usf.edu](mailto:rsch-arc@usf.edu)  
Mail: 12901 Bruce B. Downs Blvd, MDC 35, Tampa, FL 33612-4799

Template:\_000 - IRB Amendment: Personnel Change Approved

Please consider the environment before printing this email.

### **About the Author**

Jasmine A. Oliver received her Bachelor's degree in physics from South Carolina State University in 2011 and her Master's degree in physics from the University of South Florida in 2014. Jasmine is a native of Valrico, FL. She enjoys spending time with her parents, sister, grandmother, and her dog Ninja. Jasmine also enjoys reading, singing and playing the piano. She is a classically trained pianist, loves music and is a member of Sigma Alpha Iota International Music Fraternity. Jasmine is excited about the future and all that God has in store for her life. Her life motto and favorite scripture is Philippians 4:13 (KJV), "I can do all things through Christ that strengthens me."