LINEAR SYSTEM ANALYSIS FOR PHOTON-COUNTING COMPUTED TOMOGRAPHY PROTOCOL OPTIMIZATION

By

Celeste I. Leary

A DISSERTATION

Presented to the Department of Medical Physics
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

May 2020

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the PhD dissertation of

Celeste I. Leary

has been approved

_____
Mentor/Advisor

_____
Member

_____
Member

_____
Member

_____
Member

_____
Member

TABLE OF CONTENTS

List of Figures

List of Tables

Abbreviations

CI Areas – Confidence Interval Areas

CNR – Contrast-to-Noise Ratio

CRLB – Cramer Rao Lower Bound

CT – Computed Tomography

CZT - Cadmium Zinc Telluride

DECT – Dual Energy CT

E-bin – Energy-bin

EIDs – Energy Integrating Detectors

FOM – Figure-of-Merit

Gd - Gadolinium

GUI – Graphical User Interface

Hf - Hafnium

I - Iodine

ICP-MS - Inductively Coupled Plasma Mass Spectrometry

MR – Magnetic Resonance

PCCT – Photon-Counting CT

PCDs – Photon Counting Detectors

QC – Quality Control

SD – Standard Deviation

SDNR – Signal-Difference-to-Signal-Noise Ratio

SNR – Signal-to-Noise Ratio

VNC – Virtual Non-Contrast

Acknowledgements

I would first like to express my very great appreciation and thanks to my advisor, Dr. Thomas Griglock, for everything he has done for me throughout my graduate student career. I have been (figuratively and literally) all over the place in the six years we have worked together, and he has always remained even-keeled. The opportunities and guidance he has given me these past six years have been extraordinary, and his kindness and encouragement helped me persevere. I am looking forward to continuing to work with Dr. Griglock at OHSU for another two years for my Residency.

Next, I would like to acknowledge my committee members, whose valuable input helped to shape this research. Dr. Conroy Sun pointed me in the direction of ICP-MS for chemical testing and allowed me to use his lab to prepare my samples. Dr. Richard Crilly suggested the addition of the error analysis, which lead us on a wild statistics journey that was absolutely vital to this project. Dr. Anthony Butler helped me with his expertise in photon-counting CT, and it is because of him and the collaboration with Mars Bioimaging that I was able to do this research. Dr. Lindsay DeWeese helped me navigate administrative issues that were a near miss. Dr. Anna Mench joined my Oral Exam Committee as the final member, saving me a great deal of stress. I am grateful to have mentors in my life who I can rely on for support and guidance.

I could not have completed this work without my research colleagues. My special thanks are extended to the researchers in the Mars Bioimaging team. I am truly blessed to have spent five months with them throughout my graduate career. They are all such wonderful, intelligent people, and have been an absolute joy to work with. In particular, I would like to express my very great appreciation to Dr. Aamir Raja. His helpfulness and kindness cannot be overstated. Last but not

least, I need to extend another huge thank you to Aysouda and Mackenzie. They completed my final five scans in New Zealand, and Aysouda went above and beyond for me; she was fast and effective and great communicator. I am grateful for their assistance.

There are many others whose help was vital to my research. I would like to thank Jack Wiedrick of OHSU's biostatistics program, for guidance on statistics that was instrumental in this research. Assistance provided by Drs. Megan Neufeld and Anna Brown in Dr. Sun's lab was greatly appreciated. I would like to acknowledge the service, guidance, and follow-up information from Dr. Martilla Ralle for ICP-MS sample testing.

I would like to express my gratitude for my family. I am lucky to have the unwavering support of family members, each of whom provided me with unique and appreciated love and guidance.

Finally, I simply cannot thank my fiancée, Dylan Winters, enough for all of his love and support. He has shared every success and frustration with me, and his partnership has made me a better person and a better scientist.

Abstract

Contrast-enhanced computed tomography (CT) has found widespread use as a diagnostic imaging tool, with contrast used in approximately 60% of all CT scans. Dual-Energy CT (DECT) offers improvements to standard CT contrast imaging, particularly through material differentiation which improves diagnostic criteria and makes virtual non-contrast imaging possible. Photon-Counting CT (PCCT) improves upon DECT with the use of multiple adjustable energy-thresholds. This technology promises better material differentiation, multi-contrast imaging, increased CNR, reduced dose for maintained image quality, and material-specific K-edge imaging, but optimization of PCCT protocols is needed in order to maximize these benefits.

Optimizing PCCT energy-threshold placement is a difficult problem because it is highly task-dependent and the parameters are interrelated. There are many promising PCCT protocol optimization methods circulating in research today, primarily based on the inherent tradeoff between energy-resolution and SNR or CNR. Both theoretical and empirical optimization models exist, and both have benefits and drawbacks. A discrepancy exists between research and practice for K-edge protocol optimization. Theoretical optimization research commonly concludes that optimal placement of an energy-threshold is at the K-edge, while a popular method in practice involves placing thresholds a few keV off of the K-edge energy.

This research presents an empirical method for protocol optimization based on analyzing the material-basis linear system. An automated analysis GUI in MATLAB was created, which allows energy-images to be loaded, viewed, and processed, and produces a report summarizing the results. The analysis metrics were designed to measure accuracy and precision of the material-basis linear

system. Linear system accuracy, Δ, was calculated as the distance in material-concentration space between the known and calculated concentration (equivalently, the residual sum of squares). Energy-bin accuracy, $\Delta_{\perp}$(E-bin i), was calculated as the distance in material-concentration space between the known concentration and energy-bin-line i. And the linear system precision, 68% and 95% confidence interval areas, were calculated using a Monte Carlo error propagation.

Five protocols for scanning mixtures of iodine and gadolinium were tested to determine the optimal placement of the energy-thresholds. The statistical analysis showed that only the confidence interval areas had statistically significant differences for different protocols, with the "$-2\ keV$" and "$-4\ keV$" protocols resulting in larger average area than all other protocols. Additionally, differences in sample concentration produced statistically significant differences for all metrics (Δ, $\Delta_{\perp}$, and CI Areas). These results indicated that linear-system accuracy was relatively insensitive to energy-threshold adjustment of $\leq$ 4keV, but system precision was slightly worse for protocols with the lowest minimum energy-bin photon count. This relative insensitivity is promising for clinical implementation as protocols need to be robust for intra-scan changes in X-ray beam intensity.

There are many methods of protocol optimization in PCCT, and this research presents a new way of directly analyzing the material-basis linear system. The use of empirical data and automation with a MATLAB GUI makes this a promising tool for future clinical use in PCCT QC, allowing users to verify theoretical optimization results, and adjust clinical protocols for different patient sizes or contrast materials.

1. Introduction

1.1. Current Clinical Contrast CT

Contrast-enhanced computed tomography (CT) has found widespread use as a diagnostic imaging tool, with contrast used in approximately 60% of all CT scans [1]. Dual-Energy CT (DECT) offers some improvements to standard CT contrast imaging, particularly through material differentiation, which improves diagnostic criteria and makes virtual non-contrast imaging possible. However, DECT captures data in only two energy-bins which limits material identification to, at most, three materials.

Additionally, most DECT scanners use energy integrating detectors (EIDs), which cannot differentiate between photon energy. Because of this, DECT spectral separation is accomplished by varying the X-ray spectrum (fast kVp switching, dual beam), or by using dual-layer detectors. This constrains energy-binning to two stationary bins, which makes optimizing protocols for different materials impossible. Also, spectral overlap is inherent in all of these methods, and this introduces noise which degrades material differentiation accuracy.

1.2. Photon-Counting CT Prospects

Photon-Counting CT (PCCT) accomplishes detector-level spectral separation with photon-counting detectors (PCDs) instead of EIDs. PCDs allow for better spectral separation and more energy-bins with adjustable energy-thresholds (2 - 4 are energy-bins typically used in literature). This creates the possibility for optimizing spectral imaging protocols for differentiating specific materials. Reported benefits of PCCT include: better material differentiation, increased contrast-

to-noise ratio (CNR), reduced dose for maintained image quality, and material-specific K-edge imaging. Background information about Photon Counting CT is discussed further in Chapter 2.3.

1.3. Need for PCCT Optimization

The possible advantages of PCCT over DECT will only be realized with appropriate energy-bin thresholds, and should be maximized with optimal energy-bin settings [2], [3]. PCCT will be available in the future for clinical imaging, therefore protocol optimization techniques are needed to optimize different imaging tasks.

Optimizing energy-bin thresholds is a difficult problem because it is highly task-dependent and the parameters are interrelated. Bin-placement depends on many factors: the materials being imaged, the number of bins used, the subject thickness, the incident photon spectrum, and the detector response [4], [5]. Additionally, protocol optimization is further complicated by factors that degrade energy information (K-escape and charge sharing).

1.4. Current State of PCCT Optimization

There are many promising PCCT protocol optimization methods circulating in research today. Optimization methods are primarily based on the inherent tradeoff between energy-resolution and signal-to-noise ratio (SNR) or CNR, and most employ computer simulations or analytic models. There is a wide range in complexity of these models, and many seem oversimplified (two energy-bins, fixed-width bins, disregard pulse pile-up, ideal detector system). In all cases, it would be beneficial to test any theoretical results with empirical data.

A few empirical methods have also been proposed. For example, a method by Panta et al. uses energy-images to compare the performance of different PCCT protocols by calculating K-factor ($\frac{\mu(\text{E-bin after K-edge})}{\mu(\text{E-bin before K-edge})}$) and Relative X-ray Intensity ($\frac{\text{Intensity(E-bin i)}}{\text{Intensity(all E-bins)}}$) [6]. Another empirical method by Raja et al. directly tests the end effect of protocols on the material decomposition by calculating misidentification and misquantification in the material images [7]. Additionally, a popular method used in practice involves checking the linear-attenuation vs density, and linear attenuation vs energy-bin of calibration materials. Empirical methods like these are straightforward and produce good results. It would be beneficial to create a user-friendly program which heavily automates protocol analysis and gives the user information to guide the placement of energy-thresholds.

An interesting K-edge-threshold discrepancy between research and practice exists that ought to be explored. The overwhelming majority of theoretical optimization routines conclude that, when optimizing a PCCT protocol for imaging a K-edge contrast material, the optimal placement of an energy-threshold is at the K-edge energy (e.g. K-edge of iodine is 33.2 keV, on-K-edge energy-threshold = 33 keV). In practice however, thresholds are often placed a few keV off of the K-edge energy, with researchers reporting better results for these off-K-edge protocols in some cases [7].

1.5. Specific Aims

The primary objective of this study is to develop a user-friendly empirical method for multi-contrast protocol optimization, based on analyzing the material-basis linear system. We aim to use the method developed in this research to find the optimal protocol, of those tested, for scanning iodine and gadolinium mixtures.

The following work was performed to meet this goal:

1. Creation of calibration and test samples: Omnipaque (iodine contrast) and Omniscan (gadolinium contrast) were diluted with ultrapure water to make a series of calibration vials. Mixtures of these materials were also made to use as test samples. Oregon Health & Science University's (OHSU) Elemental Analysis Core performed Inductively Coupled Plasma Mass Spectrometry (ICP-MS) analysis of these vials for accurate concentration measurements. The materials were shipped to the University of Otago, Christchurch for PCCT imaging.

2. Testing K-edge protocols: Five protocols were designed to test the K-edge threshold placement. The "on-K-edge" protocol (20 – 33 – 50 – 80 – 118 keV) was repeated five times for statistical analysis. For the other "off-K-edge" protocols, thresholds were adjusted by $\pm$ 2 and $\pm$ 4 keV:

    - Protocol 1: [20 – 33 – 50 – 80 – 118 keV]
    - Protocol 2: [18 – 31 – 48 – 78 – 118 keV]
    - Protocol 3: [22 – 35 – 52 – 82 – 118 keV]
    - Protocol 4: [16 – 29 – 46 – 76 – 118 keV]
    - Protocol 5: [24 – 37 – 54 – 84 – 118 keV]

3. Analysis Graphical User Interface (GUI) created: MATLAB App Designer was used to create an analysis GUI that allows the user to load and view images, select cylindrical volumes of interest, enter materials and their respective concentrations, and generate a report and save report metrics.

4. Linear system analysis: Metrics were designed to assess:

- Accuracy of linear system as a whole: Δ = distance between the known and calculated concentration, $\Delta = \sqrt{(\mathrm{I_{known}} - \mathrm{I_{calc}})^2 + (\mathrm{Gd_{known}} - \mathrm{Gd_{calc}})^2}$
- Accuracy of each energy-bin: $\Delta_{\perp}$(E-bin i) = distance from the known concentration point to Energy-bin i.
- Precision of the linear system as a whole: Confidence Interval Areas = Area of the 68% and 95% confidence ellipses. The 68% and 95% confidence ellipses were calculated with a Monte Carlo error calculation which accounted for error in the calibration and error in the sample measurement.

2. Background

2.1. CT

2.1.1. CT History

The first computed tomography device was developed in 1967 by Godfrey Hounsfield. By mathematically combining 180 degrees of X-ray projections through the body, unobstructed axial slices of anatomy could be viewed which solved the problem of overlapping anatomy and poor depth information that is intrinsic to planar radiography. This new technology proved so revolutionary to diagnostic medicine that Hounsfield, and his colleague, A.M. Cormack were awarded the Nobel Prize in Medicine for its invention in 1979 [8].

Driven by goals of faster scan times and better image quality, CT technology progressed through four generations of scanner geometry from 1971 to 1976 [9]. Continuous helical CT acquisition became possible in 1987 with the development of slip-ring technology and Willi Kalender's interpolation algorithms [9], [10]. Then in 1992, increasing the number of detector rows from one to two rows in the z-direction initiated modern multi-slice CT (MSCT), and by 2005 CT scanners were available with up to 320 detector rows [10]. Modern CT designs are based on the third-generation detector geometry with a high-power X-ray tube rigidly attached to the gantry opposite a multi-slice detector array. A full history of the development of CT is beyond the scope of this dissertation, but aspects of CT imaging relevant to this research are discussed in the following sections.

2.1.2. CT Imaging

Figure 1 and Figure 2 illustrate the concepts of CT X-ray attenuation and detection which are the foundation of image formation, and provide essential background for understanding the differences between CT, DECT and PCCT.

A simplified model of a CT scan is shown in Figure 1. An X-ray beam (Figure 1A) travels through the study subject (Figure 1B) and the attenuated X-ray spectra (Figure 1C) is detected (Figure 1D). The beam and detectors rotate continuously around the subject with rotation time of <0.5 seconds (~0.2 seconds), and 1000 - 3000 projections are acquired per rotation. The signal for each projection is acquired over a projection time of 0.2 – 0.5 ms [11]. In this way, data are captured from many angles in a sinogram, then transformed into a brightness signal in each pixel representing the material X-ray attenuation from the corresponding volume of material in the study subject.

Figure 2 represents the mathematical concepts behind each simplified CT stage shown in Figure 1. The incident X-ray beam contains a range of X-ray energies depending on the peak X-ray tube energy (kVp), tube current (mA), target material, and filtration. Typical CT X-ray beams have 80 - 140 kVp, tungsten targets, and filtration of 5 – 10 mm Al, resulting in X-ray spectra similar to that shown in Figure 2A. Figure 2B shows the energy dependence of attenuation of common materials in the body (water, fat, calcium). Contrast between materials in CT is formed due to differences in material X-ray attenuation, i.e., low-attenuating materials appear dark and highly attenuating materials appear bright.

Current clinical CT scanners use Energy Integrating Detectors (Figure 2D) which collect the total signal in each projection. All X-rays detected during each projection time contribute to the signal, resulting in a total attenuation signal that is made from a sampling of all X-ray energies in the X-ray spectrum. As a consequence, integrating detectors average out the energy dependence of material attenuation, producing a final attenuation signal near the average energy.

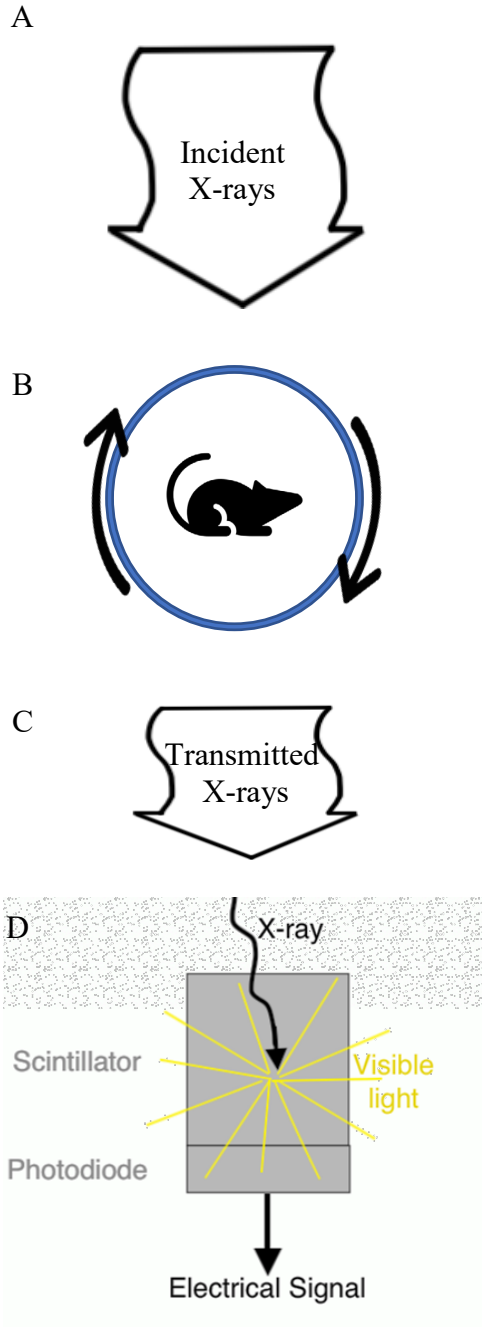


Figure 1: CT Imaging Chain: A) Incident X-ray spectrum. B) CT beam and detector rotate around the subject. C) Transmitted X-ray spectrum. D) X-ray detection.

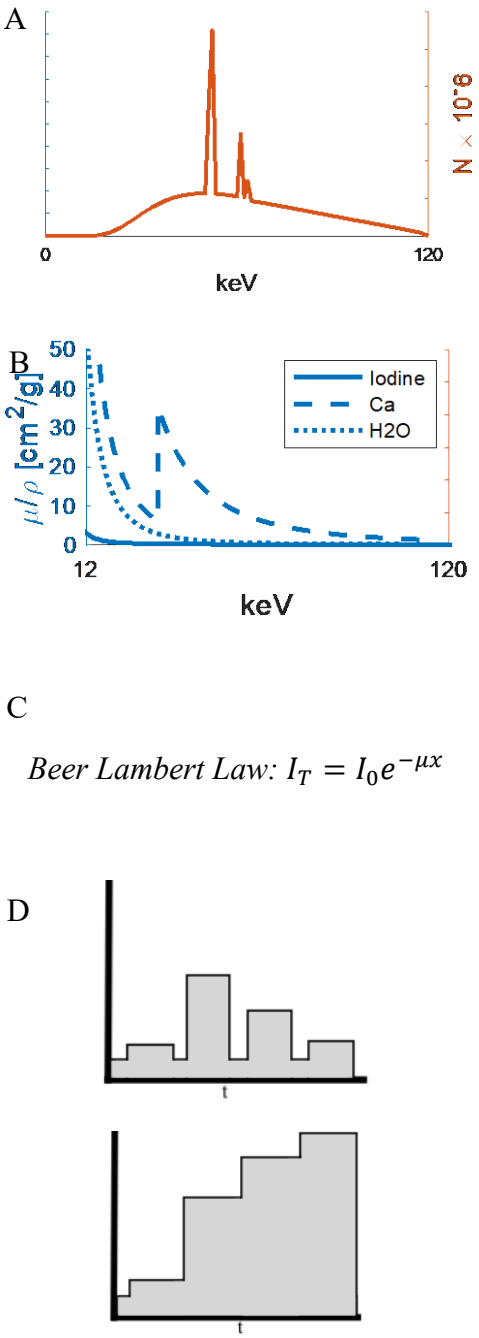


Figure 2: CT imaging chain physics: A) Histogram of approximate X-ray spectra used in CT. B) The scanned subject is made of materials with different attenuating properties which are energy dependent, $\mu/\rho = F(E)$. C) The incident X-ray beam is attenuated based on the Beer Lambert Law D) Integrating detectors capture the transmitted X-ray signal. The total signal is captured, regardless of photon energy.

Effective energy is used to approximate average beam energy because it describes a monoenergetic beam with the similar attenuation behavior. Effective energy is defined as the energy of a monoenergetic X-ray beam with the same half-value layer as the polyenergetic beam. For diagnostic imaging, this is approximately 1/3 – 1/2 of the kVp [11]. Attenuation for a polyenergetic beam can be estimated as the attenuation for a monoenergetic beam with E = effective energy, assuming no beam hardening effects [12]. In reality, average energy changes with position in the gantry and subject thickness due to differences in filtration and beam hardening.

The X-ray beam passing through the subject is attenuated based on the Beer-Lambert Law (Figure 2C) , Equation 1, where *I* is the intensity of the attenuated beam, I_o is the intensity of the incident beam, *x* is the thickness of the material through which the beam travels, and μ is the average linear attenuation of the material [11].

$$I = I_0 e^{-\mu x} \tag{1}$$

By acquiring >180 degrees of data, each voxel's attenuation can be computed. The attenuation measured in each voxel results from summing the attenuation of the materials in the voxel [13]:

$$\mu = \frac{\mu}{\rho_1} \cdot \rho_1 + \frac{\mu}{\rho_2} \cdot \rho_2 + \cdots + \frac{\mu}{\rho_N} \cdot \rho_N \tag{2}$$

In a CT image, this information is displayed with voxel values of Hounsfield units (HU) which scales the displayed attenuation information so the HU of water is 0 and the HU of air is -1000, Equation [3].

$$HU = 1000 \cdot \frac{\mu - \mu_{water}}{\mu_{water} - \mu_{air}} \tag{3}$$

2.1.3. CT Contrast Imaging

Image contrast is defined as the percent difference in signal between an area (A) and the background (B), Equation 4 [11], and signal in CT is due to X-ray attenuation. The primary attenuation interactions in the diagnostic energy range are photoelectric absorption and Compton scattering, while Rayleigh scattering makes up approximately 5 – 10% of interactions in the diagnostic energy range [11], [14]. The total linear attenuation of a material is the sum of the attenuation due to these interactions, Equation 5. In absence of a K-edge, the relationship between photon energy (E), absorbing material density and atomic number (ρ, Z), and the interaction cross sections of Rayleigh scattering (R), the photoelectric effect (τ) and Compton scattering (σ) are given by Equations 6, 7 and 8 [11], [14]. The probability of Compton scattering is proportional to electron density, which is proportional to density for materials in the body with the exception of hydrogen [11].

Contrast: $$C = \frac{A - B}{A} \quad (4)$$

Linear Attenuation: $$\mu = \mu_R + \mu_\tau + \mu_\sigma \quad (5)$$

Rayleigh: $$\mu_R \propto \frac{1}{E^{1.2}} \quad (6)$$

Photoelectric Effect: $$\mu_\tau \propto \frac{Z^3}{E^3} \quad (7)$$

Compton Scattering:

$$\mu_\sigma \propto \rho, \text{ in range } 10 - 100 \text{ keV}$$
$$\mu_\sigma \propto \frac{\rho}{E}, E_\gamma > 100\text{keV} \quad (8)$$

Contrast between soft tissues is poor because soft tissues – blood, fat, muscle, organs – have similar effective atomic number and density, thus their X-ray attenuation is similar. Equations 6 - 8 show that Rayleigh, photoelectric, and Compton effect from tissues of similar atomic number and density would be similar, thus the linear attenuation (Equation 5) would be similar, resulting in poor contrast (Equation 4) between such materials.

Contrast agents containing higher Z elements can be added to structures of interest to increase their visibility. The most commonly used CT contrast elements are iodine and barium. Iodinated agents are usually used intravenously to increase the contrast of organs and vasculature, while barium-containing agents are used to increase gastrointestinal contrast. Gadolinium-based agents are not typically used in CT, but can be used instead of iodine in cases where iodine is contraindicated, e.g., for those with iodine allergy or reduced kidney function [15]. However, the fact that gadolinium has a K-edge at 50.2 keV and is already used safely in patients for MRI contrast makes it a low-hanging fruit for translation to multi-contrast imaging in conjunction with iodine.

Along with having properties that allow for safe human use, the K-edges of I, Ba, and Gd are well-suited for diagnostic CT energies, shown in Figure 3. The K-edge is the sudden increase in attenuation at the K-shell electron binding energy. This discontinuous increase in attenuation is due to photoelectric absorption – absorption of a photon and emission of an electron – which can

only occur at photon energies at or above the electron binding energy. As atomic number increases, so does the X-ray energy of the K-edge, shown in Table 1.

Table 1: K-edge increases with atomic number

Element	Z	K-edge [keV]
Calcium (Ca)	20	4.0381
Iodine (I)	53	33.1694
Barium (Ba)	56	37.4406
Gadolinium (Gd)	64	50.2391

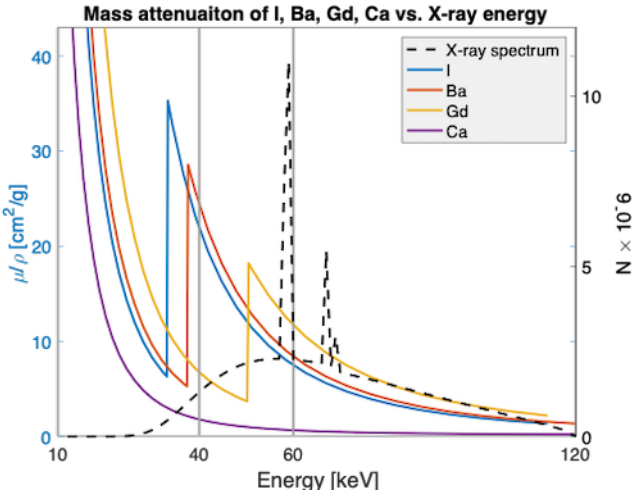


Figure 3: Mass attenuation of water, fat, calcium, iodine, barium, and gadolinium. Note the increase in attenuation in the range of diagnostic effective energies for I, Ba, and Gd. X-ray spectrum data from SpekCalc [16]. Attenuation data from NIST X-ray Form Factor, Attenuation, and Scattering Tables [17].

Materials with a K-edge near the average X-ray beam energy result in the most significant increase in attenuation, which is important for keeping the contrast agent dose low while still increasing target contrast. Contrast doses range based on body size and age group, with a typical iodinated contrast agent dose in the range of 45 – 150 mL of 320 mg/ml iodine solution [15], [18]. After injection, scans are timed for maximum target enhancement. Target iodine concentrations of 2-15 mg/ml generally agreed upon in literature [1], [3], [18].

2.1.4. Clinical CT Imaging

CT use has continued to rise since its introduction in 1973, reaching approximately 70 million annual scans in the United States in 2015 [19]. CT is used throughout healthcare for diagnosis,

treatment (e.g. CT fluoroscopy), triage in the emergency department, cancer treatment planning, etc. In 2017, it was reported that contrast enhancement was used in more than 60% of CT examinations [1].

Due to the significant rise in CT radiation exposure to the population as a whole, dose reduction and monitoring campaigns have been launched to optimize the use of radiation in CT exams. These programs include: Image Wisely, Image Gently, and the American College of Radiology Dose Registry Index. The purpose of CT dose optimization is to use only the radiation necessary to achieve the appropriate image quality required for each study.

2.1.5. CT Imaging Optimization

Protocol optimization in CT is based on balancing the tradeoff between image quality and dose. Dose should be reduced as much as possible while still being sufficient enough to produce a diagnostic quality image for the given study. Protocol parameters affecting image quality and dose include: tube current-time product (mAs), kilo-voltage-peak (kVp), detector configuration, slice thickness, pitch, patient positioning, and reconstruction algorithm [20]. Of these, dose is directly affected by all except reconstruction algorithm.

SNR and CNR are parameters used to discuss image quality because image quality is highly dependent on the appearance of noise in the image. Noise arises in a CT image from many sources: electronic, structural, anatomical, and quantum. Of these sources, quantum noise is affected by kVp, mAs, slice thickness, and pitch. Quantum noise arises due to the stochastic nature of X-ray counting statistics, therefore obeys Poisson statistics, i.e., for N recorded counts in a pixel the noise

is given by $\sqrt{N}$. Signal-to-noise ratio (SNR = $\frac{N}{\sqrt{N}}$) is used to describe the appearance of noise in an image, and contrast-to-noise ratio (CNR = $\frac{C}{\sqrt{N}}$) is used to describe the ability to resolve areas through noise.

X-ray techniques refer to the tube potential (kVp) and the current-time product (mAs). The X-ray tube output ($\propto$ N) is linearly related to the mA, ms, and mAs, and is related to the kVp^2 [21]. Although both kVp and mAs affect N, in practice, the kVp is set to provide appropriate penetration and contrast, then the mAs is adjusted to provide appropriate SNR [22], [23]. Since N is linearly related to mAs, doubling the mAs would double the dose and would increase the SNR by a factor of $\sqrt{2}$. In modern clinical CT, automatic tube current modulation is used to equalize noise throughout the image and reduce patient dose for the overall study [23].

Protocol Optimization for iodine contrast scans:

Low-kVp protocols have been shown to be beneficial for lowering dose in contrast scans of thin patients [23]. Iodine contrast is improved with lower kVp spectra due to the increased spectral overlap of the iodine K-edge. X-ray beam output (and therefore dose) changes rapidly with kVp, so a reduction in kVp must also be accompanied by an increase in mAs to maintain the same noise level in an image. For small patients, lower kVp scans can lower the dose while maintaining contrast enhancement (CNR) or have equivalent dose while increasing contrast enhancement. However, for larger patients, lower-energy beams may not provide enough penetration and result in noisier images.

2.2. Dual Energy CT

2.2.1. DECT Summary

Dual-energy CT systems collect information at low and high X-ray energies, then use the energy dependence of attenuation to add information to CT images. Both DECT and PCCT are called "Spectral CT" because they make use of the energy information from the incident X-ray spectrum. The concept of DECT was first described by Hounsfield in 1973; later, a technique for energy-selective reconstructions was proposed by Alvarez and Macovski in 1976 [24], [25]. There are a handful of different methods for achieving dual-energy scans, including slow kVp switching, dual source, fast kVp switching, and layered detectors [2], [26]. These methods are discussed in Section 2.2.3.

2.2.2. Spectral CT: Dual Energy CT or Photon-Counting CT?

Both Dual Energy CT and Photon-Counting CT are called Spectral CT in literature because they use photon energy information. However, DECT measures attenuation for only two different average energies and current commercial DECT scanners primarily use energy integrating detectors. PCCT always uses photon-counting detectors, usually using more than two energy bins to sample attenuation at multiple average energy points. As a result, DECT CT is restricted to two stationary energy ranges, while PCCT can have >2 bins with bin placement set by the user.

Photon-Counting CT is also commonly called: Spectral Imaging, Spectral Molecular Imaging, Spectral Photon-Counting CT, and Spectral CT. Throughout this paper, the naming convention "Photon-Counting CT, PCCT" will be used to prevent confusion between Spectral Photon-Counting CT and Spectral Dual-Energy CT.

2.2.3. Types of DECT

2.2.3.1. Slow kVp Switching

In slow kVp switching DECT, low and high kVp scans of the entire volume or slice are acquired consecutively, shown in Figure 4. The disadvantages of this include patient movement between energies which can result in artifacts, and use of the same tube (with the same amount of filtration) which results in spectral overlap. Movement artifacts can be reduced, though not eliminated, by switching source energy after acquiring data for 180°+ fan angle.

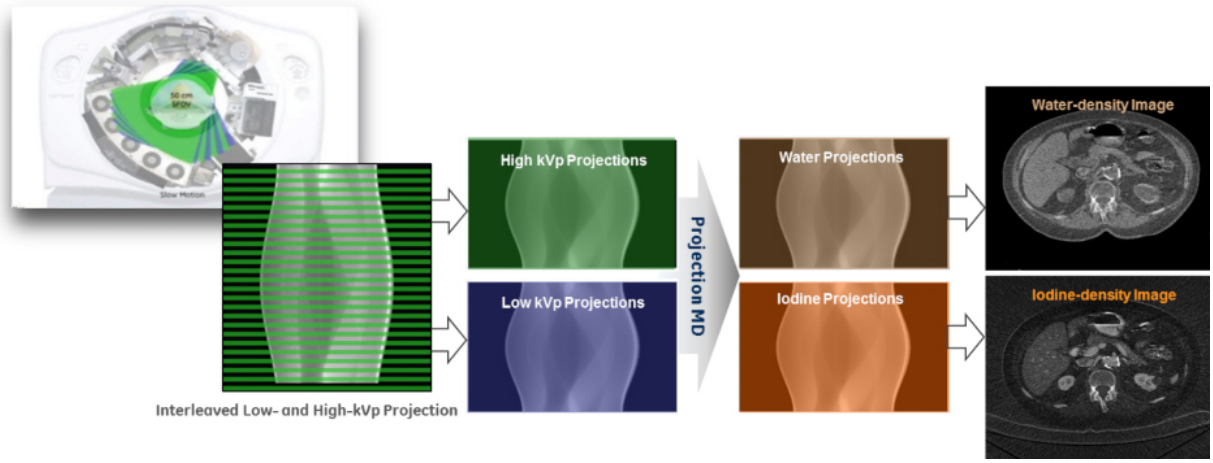


Figure 4: Slow kVp switching DECT: low-energy and high-energy slices are acquired consecutively. Low- and high- energy data sets are used to make material images (water and iodine shown). [27]

2.2.3.2. Dual Source DECT

Dual-source CT was introduced in 2005 for ultrafast cardiac scans and has the added benefit of being operated as a dual energy CT [28]. In dual-source DECT, two X-ray tubes and their detectors are attached to the same gantry at 90 degrees from each other, shown in Figure 5. The tubes are typically operated at 80 and 140 kVp. Benefits of dual-source DECT include reduced artifacts due

to patient motion, fast acquisition time, and increased spectral separation resulting from a more heavily filtered high-energy beam. A disadvantage is noise due to scatter; scatter origination from one tube can be detected by the other detector, which reduces spectral separation.

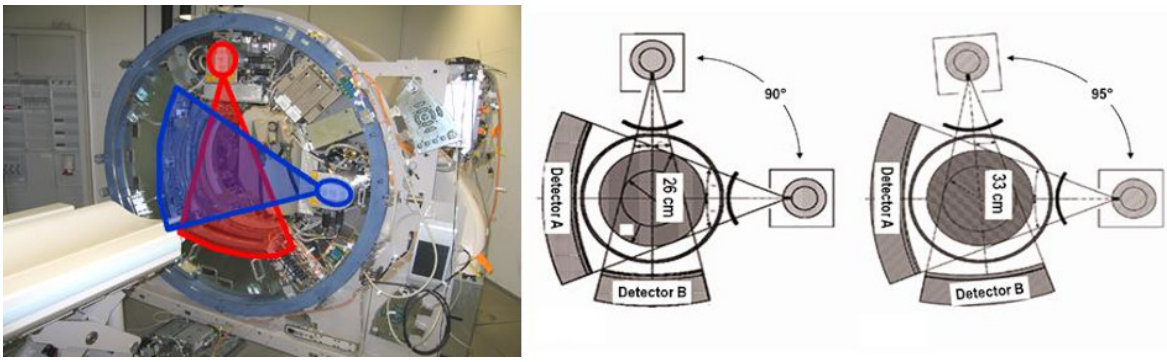


Figure 5: Dual-source DECT: a low-energy and a high-energy X-ray tube at 90 degrees from each other acquire images at the same time. [27]

2.2.3.3. Fast kVp Switching

In fast kVp switching DECT, a single source alternates between low and high energies for consecutive projections. Because kVp is modulated at sub-millisecond times, asymmetric sampling – shorter ms for high kVp, longer ms for low kVp – is used to compensate for tube output difference between the low and high energy exposures. Fast kVp switching allows for near-simultaneous acquisition to further reduce motion misregistration artifacts. However, spectral overlap resulting from using one tube (i.e. one filter) at two potentials is still a source of noise.

2.2.3.4. Detector-based DECT

In detector-based DECT, layered detectors are used to separate low-energy and high-energy photons originating from a single high kVp X-ray source. Low-energy photons are preferentially absorbed in the first detector layer, and higher-energy photons are detected in the second layer. Benefits of dual-layered DECT include simultaneous acquisition (no misregistration artifacts),

increased spectral separation, and clinical implementation as a single energy CT with the option to retrospectively reconstruct dual energy information [29].

2.2.4. DECT Advantages and Disadvantages

Clinical DECT scanners primarily use energy integrating detectors to collect CT information from two X-ray spectra [2], [26], [30]. The dual-energy information can be used to perform material decomposition, reconstruct effective-Z images or monoenergetic images, create virtual non-contrast images (VNC), and increase diagnostic specificity.

Advantages of DECT include fast scan times due to high flux operation of integrating detectors, more diagnostic information than CT, dose savings (particularly from VNC scans), and reduction of beam hardening artifacts from monoenergetic reconstructions.

Issues encountered with DECT include spectral overlap (better spectral separation means better SNR), limitation to two relatively inflexible energy-bins, increased computation times and data storage compared to conventional CT, increased dose as a result of some vendor's DECT protocols, and difficulty with introducing a new modality into the radiologist workflow.

2.2.5. DECT Material Decomposition

DECT can produce material-specific images of two or three materials through a process called material decomposition. Material decomposition involves solving a system of two equations (attenuation at low and high energies) and two unknowns (concentration of two materials). Solving

for three materials is an under-solved system, but three-material decomposition can be achieved by adding an additional volume constraint (material volume fractions must sum to 1) [31].

There are two standard methods for material decomposition: "basis set" decomposition (also called ρZ decomposition), and "basis material" decomposition. Both types of material decomposition can be done pre-reconstruction or post-reconstruction (image domain) and these methods are based on the energy dependence of attenuation (μ). The basis material metric can be derived from the ρZ metric, although the reverse is not true [13]. The basis material method models μ as a sum of linear attenuations from each material component contribution (refer to Figure 6). The basis set method models μ as the sum of attenuation due to particle interaction cross-sections (Compton scattering and photoelectric effect); refer to Figure 7 and Figure 8.

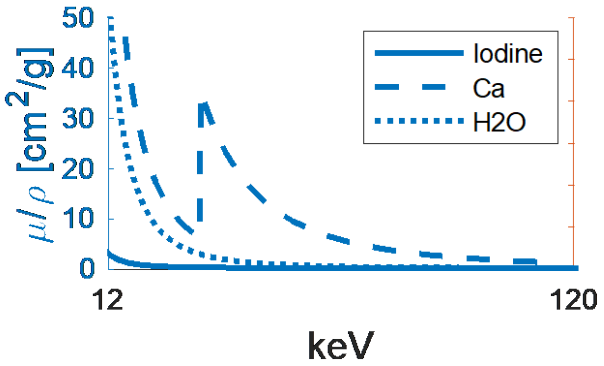


Figure 6: This figure shows μ/ρ of various materials. In the basis-material method, μ in each voxel is modeled as the density-averaged of μ from all material in the voxel. Data from NIST X-ray Form Factor, Attenuation, and Scattering Tables [17].

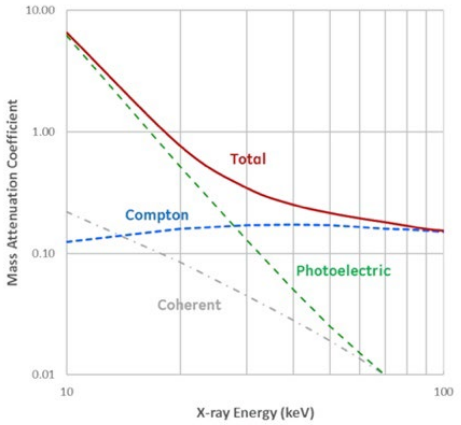


Figure 7: This figure shows total attenuation coefficient of muscle as sum of attenuation due to different interactions. [27]

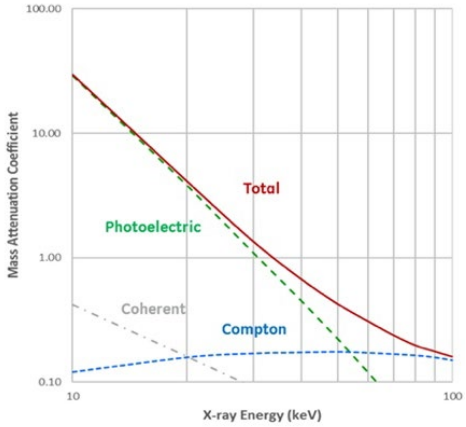


Figure 8: This figure shows total attenuation coefficient of cortical bone as sum of attenuation due to different interactions. [27]

Basis-set material decomposition models attenuation based on particle interaction probabilities vs energy. This version of material decomposition was not used in this research. Assuming (1) there is no K-edge present, and (2) an energy range where photoelectric absorption and Compton scattering are the dominant interactions, the linear attenuation along the beam path can be described using Equation 9. [32]

$$\mu = photoelectric + Compton$$
$$\int \mu(\vec{r}, E)ds = \alpha \int \rho_{eff} \frac{Z_{eff}^{k}}{E^{n}} ds + \beta \int \rho_{eff} f_{KN}(E) ds \tag{9}$$

Where $k \approx 3-4, n \approx 3-3.5$, α and β are constants and $f_{KN}(E)$ is the Klein-Nishina formula. (Equation 10) [14].

$$\frac{E_s}{E_0} = \frac{1}{1 + \frac{E_0}{511\ keV}(1 - \cos(\theta))} \tag{10}$$

Basis material decomposition for DECT models all materials as a combination of two or three materials [33]. The attenuation along each projection for a two-material decomposition can be described using Equation 11. [32]

$$\int \mu(\vec{r}, E)ds = \int \rho_{eff}\, m_1 \mu_{m1}(\vec{r}, E)ds + \int \rho_{eff}\, m_2 \mu_{m2}(\vec{r}, E)ds \tag{11}$$

For reference to these material decomposition techniques, the following papers are recommended:

Pre-reconstruction: (Liu [32], Lehmann [33], Kalender [34], Williamson [13])

Post-reconstruction: (Liu [32], Heismann [35], Macovski [25], Williamson [13], Mendonça [31])

Post-reconstruction basis material decomposition is relevant to this research therefore outlined in greater detail below. Equation 12 writes the linear attenuation of each energy-image as a density weighted sum of its constituent parts. Equation 13 shows the application of Equation 12 to two energies. Equation 14 rewrites Equation 13 as a linear system.

$$\mu(E_\gamma) = \sum_{i=1}^{Nmaterials} \frac{\mu}{\rho}(E_\gamma)_i \cdot \rho_i \tag{12}$$

2 materials,

2 energies:

$$\mu(E_{Low}) = \frac{\mu}{\rho}(E_{Low})_1 \cdot \rho_1 + \frac{\mu}{\rho}(E_{Low})_2 \cdot \rho_2$$

$$\mu(E_{High}) = \frac{\mu}{\rho}(E_{High})_1 \cdot \rho_1 + \frac{\mu}{\rho}(E_{High})_2 \cdot \rho_2 \tag{13}$$

Linear system: (14)

$$Ax = b$$

$$\begin{bmatrix} \frac{\mu}{\rho}(E_{Low})_1 & \frac{\mu}{\rho}(E_{Low})_2 \\ \frac{\mu}{\rho}(E_{High})_1 & \frac{\mu}{\rho}(E_{High})_2 \end{bmatrix} \cdot \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

2.2.6. Clinical DECT Applications

DECT can be used to identify materials, remove highly attenuating structures, and reprocess images for more optimal viewing as a virtual monoenergetic image. Examples of material identification use include iodine identification in perfusion blood pool imaging, uric acid identification for urinary stone characterization and gout diagnosis, and detecting silicone leaks in breast implants [2]. Uses of removing high attenuating structures include automatic bone removal in CT angiography, virtual non contrast images, atherosclerotic plaque removal, and virtual non-calcium images [2].

2.2.7. DECT Contrast Imaging

Contrast imaging in DECT is improved over standard contrast CT. Material-only images improve CNR because the background signal is processed out, and better CNR improves the detectability of small or low-uptake lesions in iodine contrast scans. Additionally, material identification can help discriminate between contrast enhanced and non-contrast enhanced lesions that would normally have similar CT attenuation.

DECT can also create VNC images which are single acquisitions in which the signal from iodine is subtracted out, creating an anatomy-only image. In standard CT, multiple acquisitions are needed to acquire non-contrast and contrast scans. VNC images reduce the number of phases required for some contrast studies, providing significant dose savings. [2]

Intrinsic contrast can also be exploited to aid in diagnosis. For example, identification of uric acid through the dual-energy index ($DEI = \frac{HU(low\ E) - HU(high\ E)}{HU(low\ E) + HU(high\ E) + 2000}$) provides a non-invasive gout diagnosis. Kidney stones can also be detected and categorized by their material makeup [36], [37].

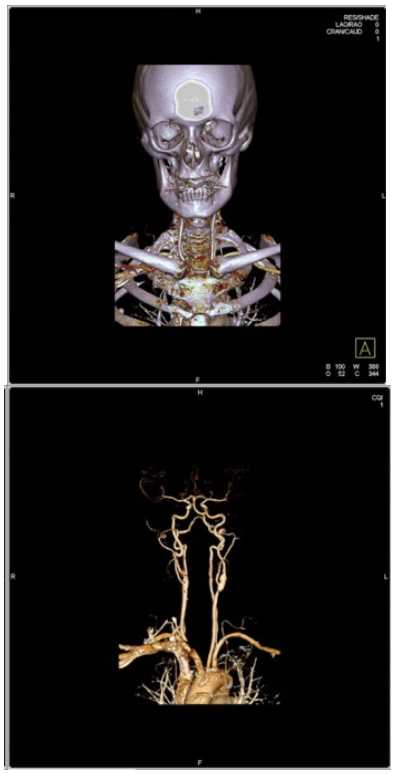


Figure 9: DECT angiogram can separate bone and vasculature from the rest of the anatomy. [27]

Figure 10: DECT images show uric acid crystals, which indicate gout. [27]

Improved CNR and material differentiation with DECT has increased interest in creating new contrast agents. Iodine and barium work best at low energies, providing a 70-90% greater CT number at 80 kVp than at 140 kVp [3]. However, iodine (Z = 53) and barium (Z = 56) cannot be separated from each other with DECT material differentiation because they have nearly the same atomic number, providing similar attenuation signals. Higher Z materials could compliment the use of iodine and barium in DECT scans by providing a greater high-energy attenuation signal and allowing material differentiation of two contrast agents at once. New contrast agents could add diagnostic information without impacting patient dose, although the any new contrast material must be extremely well-studied to ensure its safety [3].

2.2.8. DECT Optimization

DECT optimization focuses on improving spectral-imaging abilities. For example, increasing spectral separation between low-energy and high-energy spectra has been shown to increase SNR in VNC images [38]. Spectral separation can be improved by using greater filtration and a higher kVp for the high-energy X-ray beam and a lower kVp for the low-energy beam. Figure 11 shows how spectral overlap could be decreased with changes to filtration. Using a lower low-kVp spectrum could also increase spectral separation. However, penetrability issues in the low-kVp beam may be an issue for larger patients and high currents would wear out tubes quickly [39].

Optimizing the contrast agent used with DECT has also been proposed, although significant research into safety is an obstacle for any new contrast agent. Because DECT has two relatively stationary energy-spectra, having one contrast agent with a K-edge in the lower-energy range and a second with a K-edge in the higher-energy range would be optimal for material decomposition [3].

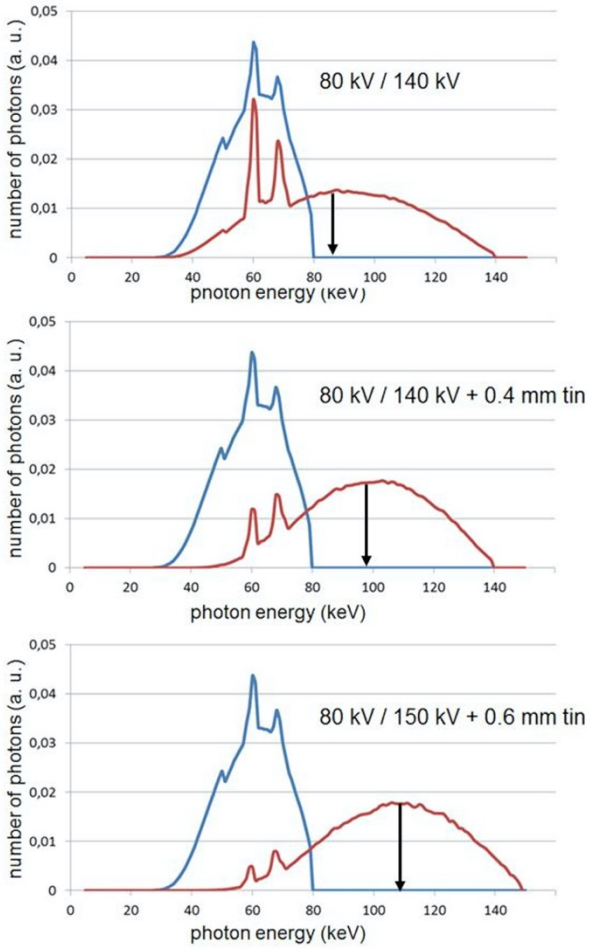


Figure 11: 80 kVp and 140 kVp X-ray spectra with varying filters for the high-energy spectrum. Illustrates spectral overlap in typical DECT spectra [27].

2.3. Photon Counting CT

2.3.1. Photon Counting Detectors

Photon-counting CT pairs one X-ray tube with photon-counting detectors to expand and improve the use of photon energy information. The significant differentiating factor between CT, DECT, and PCCT is the use of PCDs instead of EIDs. PCDs are direct-conversion semiconductor detectors, usually made of CZT, CdTe, or Si, that are combined with a fast-counting ASIC rather than a signal-integrating ASIC. This enables pulse height discrimination of single-photon events [40], [41]. The user sets energy thresholds above which photons are counted, then the number of

photons in each energy-bin are computed by subtraction. More than two energy-bins can be used due to the energy discriminating capabilities of PCDs, opening up the possibility for multi-contrast imaging in the future.

2.3.2. PCCT History

PCDs were developed for particle tracking in high energy physics experiments at CERN. Development started in the 1980's and the first "hybrid pixel detectors" were installed in 1995 [42]. Pixel hybridization describes combining a sensor and readout chip for on-chip signal amplification, discrimination, and digitization [43]. By 2006, multiple LHC experiments included early PCDs including ATLAS, CMS, and ALICE.

Different groups have worked in collaboration with CERN and individually to develop PCDs for specialized uses. Versions of PCDs include MEDIPIX, XPAD, PILATUS, and PIXIE. Some targeted uses for these pixels are synchrotron experiments, measuring X-ray polarization in space, diffraction imaging, and biomedical imaging**.** Both Medipix and Xpad were designed for biomedical imaging applications, and Medipix detectors are installed in the PCCT scanner used for this research. [40, 41]

The PCCT scanner used in this research was produced by Mars Bioimaging Ltd.; Mars Bioimaging was founded in 2007 and they collaborate with CERN, the Universities of Otago and Canterbury, and others for research and development of PCCT for medical imaging [44]. Currently they sell the MARS Small Bore Spectral Scanner, pictured in Figure 12, for small-animal imaging and pre-clinical PCCT research. They are currently developing a human scanner, and the first human scans of the wrist and ankle of Dr. Phil Butler, CEO of Mars Bioimaging Ltd., were published in 2018

(Figure 13 and Figure 14) [45]. The MARS Small Bore Scanner was used in this research, and details about this scanner can be found in Chapter 3.

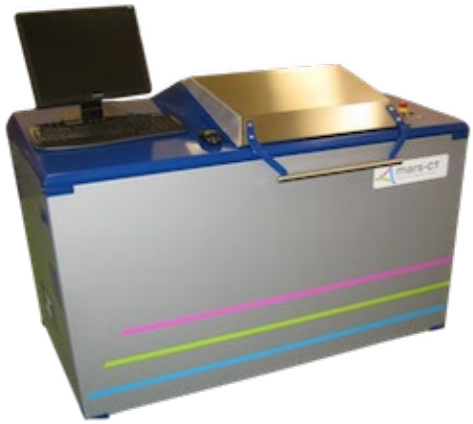


Figure 12: The MARS Small Bore Scanner, produced and sold by MARS Bioimaging Ltd. for photon-counting CT research.
(Figure reproduced with permission from https://www.marsbioimaging.com)

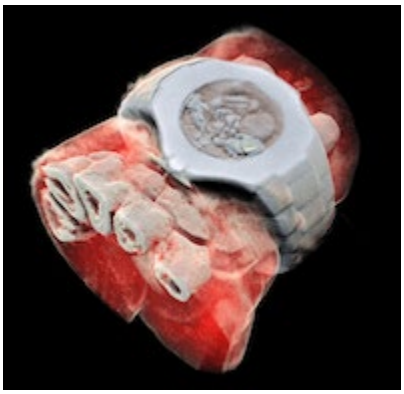


Figure 13: 3D Material image of Dr. Phil Butler's wrist and watch. (Figure reproduced with permission from https://www.marsbioimaging.com)

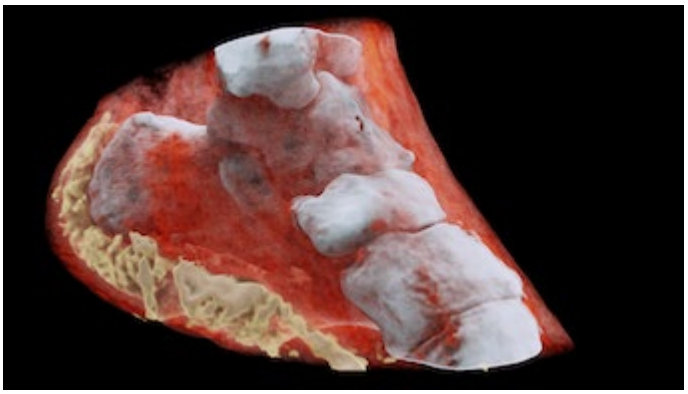


Figure 14: 3D material image of Dr. Phil Butler's Ankle. (Figure reproduced with permission from https://www.marsbioimaging.com)

2.3.3. PCCT Benefits and Current Limitations

There are many benefits of PCCT. The use of more energy-bins makes material differentiation of more materials possible, and user-set energy thresholds allow for optimization of the imaging task. Better spectral separation due to discrete energy bins (energy resolution reported to be 1 - 3.7 keV)

can provide better image quality and material differentiation than DECT [46], [47]. Photon-counting detectors can reject electronic noise by setting a low-energy threshold, below which pulses from electronic noise are not counted, and they have approximately 30% higher geometric efficiency than energy integrating detectors [2]. Photon-counting detectors also have straight response to energy rather than an energy-weighted response, which can be used to optimally weight CT images [48], [49].

Current limitations of PCCT will need to be overcome before all proposed clinical uses are possible. Energy detection can be degraded by charge sharing and K-escape, and counts can be lost due to pulse pile-up. K-escape results in the partial loss of photon energy due to a characteristic X-ray escaping detection. Charge sharing results when one photon is counted as multiple lower-energy photons, although these effects can be improved with charge-summing mode, where simultaneous events detected in neighboring pixels are summed and assigned to the pixel with the largest signal [2].

The largest barrier to clinical use is pulse pile-up, which causes loss of counts as mA is increased above what the detector dead-time can handle [2]. This results in scan speeds that are longer than what is currently possible with integrating detectors due to necessary low mA. However, pileup has been shown to be controlled for $\leq$550 mA at 140 kVp for 0.5 second rotation time ($\sim 2.5 \times 10^{11}$ photons per cm^2 per second) with the use of small sub-pixels (255 μm pitch), the idea being that increasing the number of pixels per mm decreases the chance of coincident photon detection in the same pixel [49]. In 2016, the first human scans comparing two-energy-bin PCCT and dual-layer DECT of the abdomen with iodine contrast showed good performance of PCCT [47].

2.3.4. Proposed Clinical Applications of PCCT

Improved spectral separation translates easily into use as an improved DECT, e.g., the 2016 human trial. In the future, PCCT should be able to do anything DECT can do, but with better noise statistics due to better spectral separation. Applications of PCCT will also go beyond DECT because multiple energy bins can be set according to the imaging task. Some proposed applications are contrast imaging, bone imaging, cancer imaging, and atherosclerotic plaque imaging.

Contrast imaging in PCCT focuses on K-edge imaging. Specific imaging of high-Z contrast elements with PCCT is most often called K-edge imaging because an element's unique K-edge can be identified by placing energy bins around the K-edge [50]. Multiple and mixed contrast agents can also be identified and quantified because PCCT has multiple energy bins. This presents opportunities for new future contrast agents.

Quantitative bone imaging has already been proven useful for monitoring bone health with Dual-Energy X-ray Absorptiometry (DXA) and Quantitative Computed Tomography (QCT). Both of these modalities compute bone mineral density (BMD) in an area or volume – a task easily accomplished by PCCT. It is possible that PCCT could add to the bone imaging toolkit in the future with improved image quality and more accurate 3D BMD measurement.

PCCT could play a role in early detection or treatment monitoring of cancer; early detection of could be accomplished with PCCT by visualizing biomarkers of cancer while it is still at a treatable

stage. Monitoring treatment with PCCT could also be beneficial for treatment-resistant tumors (e.g. regions of hypoxia).

Atherosclerotic plaque imaging may help prevent stroke and heart attack in the future. Biomarkers of plaque in danger of rupturing and causing a stroke or heart attack could be identified and used for initiating treatment. PCCT could also be used to simultaneously monitor drug delivery and the response of both the plaque and the patient to treatment.

2.3.5. PCCT Material Decomposition

PCCT material decomposition is based on the same principle as DECT material decomposition: attenuation is related to material and energy. Unlike DECT, which is limited to sampling the photon spectrum at two energies, and therefor can only differentiate two or three materials, PCCT often employs more than two energy bins, and therefore should be able to differentiate at least as many materials as the number of energy-bins.

As discussed in Section 2.2.5, material decomposition can be accomplished with either the basis set or basis material method in the pre-reconstruction or post-reconstruction domain. However, the basis set material decomposition must be modified in the presence of K-edge materials because the K-edge introduces a discontinuity in attenuation which is not described by Compton and photoelectric components alone. The next section (Section 2.3.6) explains this and its relation to basis material decomposition.

There are many different mathematical methods for accomplishing material differentiation with more than two energy-bins. Some methods include maximum likelihood estimator (MLE) [4], [12], [51], least squares [52], maximum *a posteriori* MAP estimator [53], primary component analysis (PCA)[54], and deep learning [55]. Issues in PCCT material decomposition include computation time, beam hardening through different sized subjects, energy signal deterioration through k-escape and cross talk, and energy-bin threshold optimization.

2.3.6. PCCT Contrast Imaging

Contrast imaging will likely be improved and expanded with the use of PCCT. Increased spectral separation will improve CNR, which will improve detection in contrast-enhanced scans.
Increased number of bins will make multi-contrast imaging possible [3], [47], and having more flexible energy bins creates opportunities for new higher-Z contrast elements. Multi-contrast imaging could be used in the future to image physiological processes, for example, the body's response to drug delivery. The full potential of this technological advance in contrast imaging is not yet known.

K-edge imaging is heavily researched for PCCT. This technique identifies contrast materials by placing energy thresholds around their unique K-edge, and is best accomplished with energy bins both before and after the K-edge discontinuity [50]. Iodine, with a relatively low K-edge at 33.2 keV, is therefore not an ideal candidate for K-edge imaging in humans because X-rays below this energy are filtered or attenuated out [51].

Material decomposition of K-edge imaging is usually described by extending the ρZ method. Above the K-edge, attenuation can be described as a linear combination of photoelectric and Compton cross sections [33], [56].

$$\mu(E,\vec{x}) = a_{ph}(\vec{x})\frac{1}{E^3} + a_{Co}(\vec{x})f_{\mathrm{KN}}(E/E_e) \tag{15}$$

Here $\vec{x}$ describes the dependence on position, f_{KN} is the Klein-Nishina formula, and $E_e \approx 511\ keV$ (electron rest-mass energy).

In the presence of K-edges, one term for each K-edge present is added to Equation 15. Iodine and gadolinium are used as examples below (Equation 16), but any number of K-edge elements could be included.

$$\mu(E,\vec{x}) = a_{ph}(\vec{x})\frac{1}{E^3} + a_{Co}(\vec{x})f_{\mathrm{KN}}(E/E_e) + a_{Gd}(\vec{x})f_{Gd}(E) + a_I(\vec{x})f_I(E) \tag{16}$$

The terms $a_\alpha(\vec{x})$ and $f_\alpha(E)$ (for K-edge material α) are the local density and the mass attenuation coefficient, respectively. The mass attenuation coefficient encompasses all attenuation contributions: photoelectric, Compton, and K-edge.

Equation 15 and 16 can be transformed from the basis-set (ρZ) decomposition method to the basis-material decomposition. Lehmann et al. explains how the basis-set method and the basis-material method are related [33]:

1) *Any non-K-edge material can be described as a linear combination of a_{ph} and a_{co} in Equation 16.*

2) *Any two linearly independent sums of two basis functions span the space and can therefore be used as basis functions.*

This also explains why having energy-bins below and above the K-edge is needed for K-edge material decomposition. The jump in attenuation from the K-edge distinguishes it from being identified as a combination of other materials. For example, bones (calcium) can be identified as a mixture of iodine and water when an iodine-water-gold material basis is used. Calcium (no K-edge) can be described as a linear combination of iodine and water [50]. Cormode et al. also reported that images were noisy when a material basis of iodine-water-calcium-gold was used [48]. Without having a bin before the iodine K-edge, the water-iodine-calcium material basis is a linearly dependent set, and material misregistration/noise arises from this collinearity. This highlights the importance of developing higher Z contrast elements for clinical PCCT K-edge imaging.

2.3.7. PCCT Optimization

Optimizing PCCT energy-threshold placement is important for PCCT imaging because threshold placement impacts contrast and noise in PCCT images [57]. Most PCCT protocol optimization techniques are based on the tradeoff between energy resolution and SNR in energy-bins, illustrated in Figure 15 and Figure 16. Large energy-bins (Figure 15) capture large amount of photons, resulting in and high SNR at the cost of averaging the attenuation signal over wider portions of the energy spectrum, i.e., poor energy resolution. Conversely, small energy-bins (Figure 16) allow for better energy resolution at the cost of SNR.

There is also a relationship between protocol robustness and number of energy-bins; a study with photon-counting projection imaging showed that ten optimized bins had the same material differentiation noise level as an un-optimized protocol with sixty-four energy-bins [58]. This observation that a greater number of energy bins result in more robust protocols may be due to the fact that more energy thresholds result in more constraints on the material differentiation results. Additionally, greater bin density reduces freedom of threshold placement which may also limit variation in material differentiation results. If the opposite is true for fewer bins, then protocol optimization is especially important for protocols with few energy-bins.

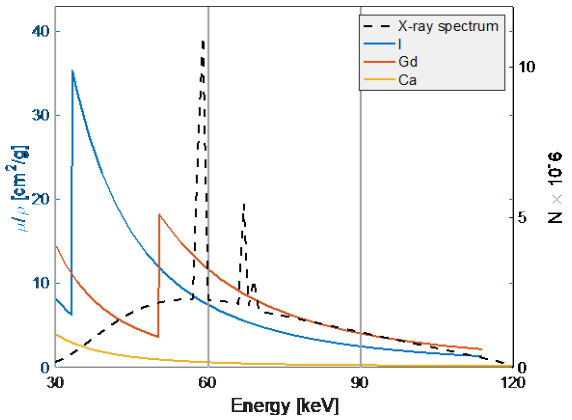


Figure 15: Model of a 3-bin PCCT protocol for scanning iodine, gadolinium, and calcium with a 120 kVp X-ray spectrum. [16], [17]

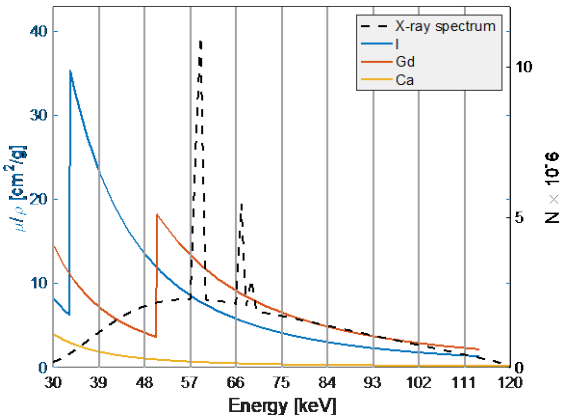


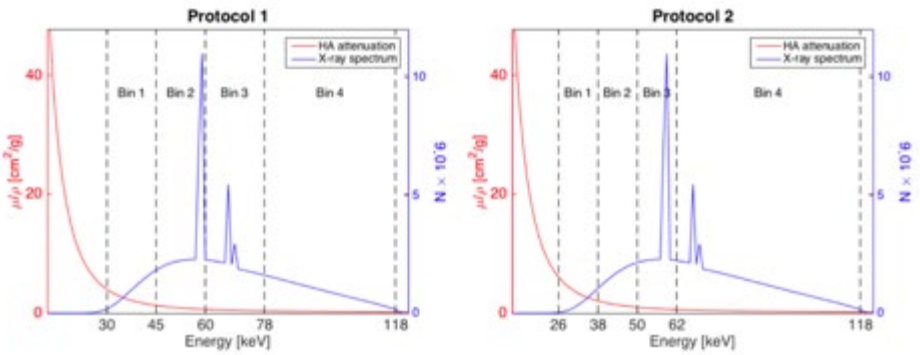
Figure 16: Model of a 10-bin PCCT protocol for scanning iodine, gadolinium, and calcium with a 120 kVp X-ray spectrum. [16], [17]

In a protocol with four energy-bins the user must place four thresholds. In absence of a K-edge, it might make sense to place the thresholds so an approximately equal number of photons are in each bin. But with the addition of K-edge materials, maximizing K-edge signal also becomes a concern; for example, in Figure 15, both the gadolinium and iodine K-edges are in Energy-bin 1. This doesn't make sense from a K-edge imaging perspective because having an energy-bin before the

K-edge is valuable. Additionally, the K-edge of Gadolinium is approximately halfway along the photon-weighted spectrum, which would effectively average out the entire K-edge signal.

Lack of photons in the low-energy and high-energy range are also a concern. The shape of the X-ray spectra shown in Figures 25 and 26 were modeled using a program called SpekCalc. In reality, the spectrum pre- and post-attenuation vary with X-ray tube, position in scanner, and attenuator thickness. Because of this, placing bins too close to the tail ends of the spectrum could result in adverse effects due to lack of signal. An example of this is given in Figure 17; here, two energy-bin schemes were tested to distinguish Calcium, Fat, and Water, shown in Figure 17 A: Protocol 1: 30, 45, 60, 75, 118 keV, and Protocol 2: 26, 38, 50, 62, 118 keV. The attenuation signal for calcium, graphed in Figure 17 B, did not behave as expected in Protocol 2, Bin 1. Attenuation for calcium should decrease with energy, but the attenuation in Bin 1 was lower than the attenuation in Bin 2. This is likely due to the low energy-bin being too low or too small, and therefore having poor SNR. Based on this result, we selected Protocol 1 to scan bones.

A



(a) Protocol 1: Energy bins: [30, 45, 60, 78, 118] (b) Protocol 2: Energy bins: [26, 38, 50, 62, 118]

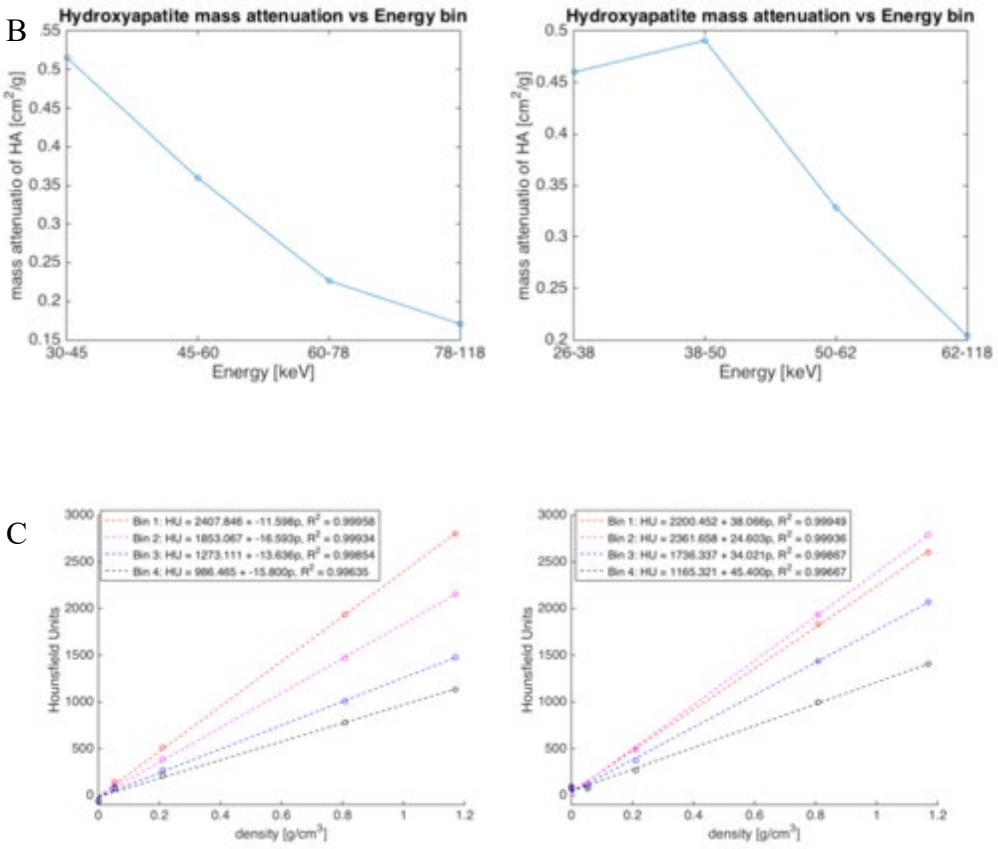


Figure 17: Two calcium scanning protocols compared. (A) Model of the two PCCT protocols, [16], [17] (B) The mas attenuation vs energy bin was graphed – results for protocol 1 behaved as expected. Energy-bin 1 and 2 in protocol 2 did not present the decreasing attenuation signal that was expected. (C) All bins showed good linearity with density, but bin 2 of protocol 2 had higher slope (i.e. attenuation vs concentration) than bin 1.

2.3.8. PCCT Protocol Optimization Techniques

Many protocol optimization techniques exist for PCCT imaging, and this section provides a review of the existing literature. However, PCCT protocol optimization is an active area of research so it is likely that this does not cover all existing techniques.

For the most part, optimization methods define a Figure-of-Merit (FOM), usually based on the bin size tradeoff between energy resolution and noise, then maximize or minimize the FOM to find the ideal energy-bin configuration. Most methods model the attenuation which can be arbitrarily complex in an effort to represent the physical system. The benefit of a computer models is that,

given enough time, they can test all possible bin configurations. However, as number of bins increases, so does computation time. Another limitation of computer models is that the results may only be as good as the model.

Methods 5 and 7 covered in this section are based on empirical data rather than models. Benefits of these methods include feedback on performance of the system as a whole, ability to be done on-the-fly with as many measurements as is practical, and less requirements of computing power or models. Testing all possible energy-bin configurations would be extremely time intensive, but it is possible that not all energy-bin configurations need to be tested.

Technique covered in this section include:

1) Cramér Rao lower bound
2) CNR vs bin width
3) Signal-difference-to-signal-noise ratio (SDNR)
4) $SDNR^2$ and SNR^2
5) K-factor ranking method
6) Confidence interval minimization
7) On-demand protocol testing (based on experience, not published)

2.3.8.1. Cramér Rao Lower Bound (CRLB):

Cramer Rao lower bound (CRLB) is a mathematical technique used to find the lower bound of the variance. In PCCT protocol optimization it has been used to compute the minimum theoretical noise in the basis images as a function of the energy-bin width. The CRLB

technique defines a minimization problem that can be used to find the energy-bin thresholds that would produce the least noise in the basis images (photoelectric, Compton, K-edge).

This method of PCCT bin optimization was first used by Roessl and Proska (2006), who applied the CRLB method to two basis functions: photoelectric and Compton [59]. They used SNR^2/exposure as a FOM and ran simulations with up to four E-bins, looked at variation of CRLB with changing energy thresholds for different filtration thicknesses and compared the FOM result with equidistant energy thresholding (evenly spaced bins).

The 2009 paper by Roessl and Herrmann, "Cramér-Rao Lower Bound of Basis Image Noise in Multiple-Energy X-Ray Imaging," expands the application of the CRLB method to an arbitrary number of basis functions [4]. This paper also describes the analytic computation of the CRLB basis noise. The result is that the basis noise is bounded with a minimum possible value computed by the inverse of the Fischer information matrix. This paper maximized SNR as a FOM, but states that there is nothing special about using SNR, rather, this technique is applicable to any FOM that is based on noise (i.e. variance). The authors also note that the detector model can be made as simple or complex as wanted, with an "ideal" situation giving the lower limit of noise. The example problem in this paper uses three basis functions (photoelectric, Compton, and K-edge of gadolinium), and five comparators defining four energy-bins, with only the middle comparator adjusted (10, 30, X, 70, 90). The result showed that SNR in the gadolinium basis image is optimized when threshold three is at the gadolinium K-edge.

More examples of use include Wang and Pelc "Sufficient Statistics as a Generalization of Binning in Spectral X-Ray Imaging" (2011) and Roessl et al. "Sensitivity of Photon-Counting Cased K-Edge Imaging in X-Ray Computed Tomography" (2011).

2.3.8.2. CNR as a Function of Bin-width

He et al. "Material Discrimination Based on K-edge Characteristics" (2013) defines CNR as a function of energy bin width, and in this way, finds the optimal bin width to maximize CNR for contrast material decomposition. Their model calculates Expected Image and variance as a function of bin width from a model of the photon spectrum and Filtered Back Projection. This model only deals with one energy bin starting at the K-edge and extending for a width *w*. Limitations of this model include the following: only one bin-width modeled, models can rarely take all effects into account, and CNR of a contrast element relies on concentration.

2.3.8.3. SDNR

The signal difference to noise ratio (SDNR) was used by He et al. in the 2012 paper "Optimization of K-edge Imaging with Spectral CT" [60]. In the simulation study, two energy-bins were placed on either side of the gadolinium K-edge and energy-bin width was optimized using SDNR as a FOM.

SDNR is defined as:

$$SDNR = \frac{\overline{\mu}_R - \overline{\mu}_L}{\sqrt{\sigma_R^2 + \sigma_L^2}}$$

Here $\bar{\mu}$ is the average attenuation, σ is the standard deviation of the attenuation, the subscripts R and L indicate the right and left side of the K-edge – i.e. the attenuation after and before the K-edge. SDNR as a FOM was introduced due to the tradeoff between precise attenuation signal with a narrow energy window, and low noise with a wide energy window (energy resolution and SNR).

He et al. developed an equation for SDNR as a function of bin width based on a Filtered Back Projection numerical simulation, then maximized SDNR with respect to the bin width. Limitations of this method include its assumption of same energy-bin width on either side of the K-edge and10,000 photons count per 1 keV, and results are w = ~4 keV, which is approximately the energy resolution of photon counting detectors.

2.3.8.4. SDNR^2 and SNR^2

A paper by Zheng et al. (2018) [5], "Robustness of Optimal Energy Thresholds in Photon-counting Spectral CT," uses signal-difference-to-noise ratio squared (SDNR^2) and signal-to-noise ratio squared (SNR^2) as a FOM to optimize energy-bins for two different detector models (silicon strip and CZT PCCT). SDNR^2 is used for optimizing the thresholds of an optimally weighted image (not material decomposed image).

$$SDNR^2 = \frac{(w\Delta \mathrm{g})^2}{(K^b + K^t)w^T}$$

Where $\Delta g = g^b - g^t$ is the difference between background counts and target counts ($g = (I_1, \dots, I_N)^T$ where I_N is the expected number of counts in bin N calculated from this model. K^b

and K^t are the covariance matrices of g, and w is the optimal weighting factor (Barrette & Meyers 2013). In their detector model, pulse pile-up and charge sharing are considered.

SNR^2 is used to optimize thresholds for a material decomposed image. The example material used in this equation is gadolinium. In their choice of FOM, Zheng et al. cites the Roessl 2011 paper on Cramer Rao lower bound, explaining that maximizing SNR^2 is synonymous with minimizing the CRLB lower bound noise.

$$SNR^2 = \frac{A_{Gd}^2}{\sigma_{A_{Gd}}^2} \leq \frac{A_{Gd}^2}{CRLB_{Gd,Gd}}$$

Here A_{Gd} is the line integral of the Gd basis material, and σ is the variance.

Maximizing two FOM at once is solved using the Global Search method (Matlab) run twenty times to ensure stable maxima. Furthermore, to create FOM independent of photon flux, $SDNR^2$ and SNR^2 are normalized by ideal $SDNR^2$ and SNR^2, which are then defined as relative $SDNR^2$ and relative SNR^2, respectively. Ideal $SDNR^2$ and SNR^2 are calculated by covering the energy spectrum with 1 keV energy bins. The model consisted of 1 mg/ml iodine and 10 mg/ml iodine, bone, and tumor, and 2 – 8 energy bins were investigated with 15 cm, 30 cm, and 50 cm phantom sizes.

Results indicated optimal threshold placement is highly related to the detected spectrum, which was influenced by the attenuator and the detector response. Bins for the 30 cm phantom had the least variable performance across all phantom ranges – though this could be because 10 cm – 60

cm were tested, and 30 cm is near the center of this range. Increasing the number of bins improved the FOM, but also increased the amount of data to process (tradeoff: number of bins vs. computation time). Near optimal $SDNR^2$ and SNR^2 were observed for 6 – 8 bins. More bins might have shown marginal increases in $SDNR^2$ and SNR^2 at the cost of computation time.

2.3.8.5. K-factor

The K-factor is a metric used to rank protocols for scanning multiple high-Z materials; this ranking method was introduced by Panta et al. in a 2018 paper "Element-Specific Spectral Imaging of Multiple Contrast Agents: a Phantom Study" [6].

K-factor is the percent increase in attenuation caused by the K-edge.

$$\text{K-Factor} = \frac{\text{X-ray attenuation in K-edge containing range}}{\text{X-ray attenuation in the preceding energy range}} \times 100\%$$

A large K-factor would indicate a better protocol for identifying the high-Z material. However, there is a tradeoff between K-factor and SNR, due to the tradeoff between SNR and energy resolution. Narrow energy-bins placed before and after the K-edge would give a large K-factor but would capture few photons. On the other hand, wide energy-bins would provide better SNR, but would average out the energy-dependent attenuation information, resulting in lower K-factor.

Instead of directly using SNR, Panta uses "Relative X-ray Intensity" in the K-edge containing energy range, where relative X-ray intensity is defined as:

$$\text{Relative X-ray intensity} = \frac{\text{X-ray intensity in a particular energy range}}{\text{X-ray intensity in all energy ranges}} \times 100\%$$

Relative X-ray intensity and SNR are both directly related to the number of photons in a particular energy range. This method only takes into account relative X-ray intensity in the K-edge containing range for each high Z-element, which can be interpreted as "signal strength for the elements K-edge." However, signal strength also depends on the energy-bin preceding the K-edge. In this paper, four bins are used and the first bin does not contain a K-edge and therefore is not taken into account in the relative intensity scoring.

2.3.8.6. Mahalanobis Distance Mapping/ Confidence Interval Minimization

Nik, Meyer, and Watts 2011 [61] "Optimal Material Discrimination Using Spectral X-Ray Imaging" presents a model to optimize energy-bins by minimizing the confidence region of thicknesses making up the material basis.

Given Poisson statistics (mean λ, and noise $= \sqrt{\lambda}$) the 63% confidence interval for the null hypothesis that the measurement, x, is equal to the mean, λ, is defined as:

$$z = \frac{x - \lambda}{\sigma} = 1 \text{ for 63\% confidence interval}$$

This can be expanded into higher dimensions (more energy bins) with the Mahalanobis distance:

$$z = \left\{\sum_{k=1}^{n}[(x_k - \lambda_k)]^2 \times \frac{1}{n}\right\}^{\frac{1}{2}} = \left\{\left[\sum_{k=1}^{n}\left(x_k - \lambda_k(\vec{t})\right)^2 \times \frac{1}{\lambda_k(\vec{t})}\right] \times \frac{1}{n}\right\}^{\frac{1}{2}} = 1$$

where $\vec{t}$ is the set of thicknesses, and the factor $\frac{1}{n}$ is introduced to negate z-dependence on number of energy-bins. Mapping z in thickness space represents a distribution of all thicknesses that could make up the signal. The thickness set for $z = 0$ is the combination of material thicknesses that are most consistent with the measurement x. For two bins, thickness sets for $z \leq 1$ lie within the 63% confidence region. For three bins, z = 1 defines a confidence volume of 61%.

The FOM introduced in this paper represents the size of the confidence interval. Instead of computing the area or volume of the confidence region, Nik et al. uses the size of a rectangular box circumscribing the confidence region. This method avoids computing the Fischer information matrix.

$$FOM = \left(\sum_{i=1}^{m}(\sigma_{t_i}/\tau_i)\right)^{-\frac{1}{2}}$$

Here σ_{t_i} is the standard deviation (SD) of the thicknesses t_i inside the confidence region. Minimizing this FOM is used to optimize protocols; minimization is done by doing an exhaustive search through all possible protocols, which is feasible for low number or bins ($n \leq 3$). For higher number of bins, Nik recommends a simulated annealing algorithm or other such optimization model.

Results for separating iodine and calcium with two energy-bins indicate that placing the bin border at the K-edge of iodine is optimal (also indicated by Roessl and Herrmann 2009 CRLB optimization). Separation of iodine and water was best with the bin border at 60 keV – this achieves greater average attenuation ratios between energy-bins. It is also shown that in calcium-water-fat discrimination, calcium can be resolved but water and fat have an elongated confidence region, suggesting that only their sum can be resolved.

2.3.8.7. On-demand Protocol Testing

Published protocol optimization techniques rely on calculating Figures-of-Merit then optimizing the protocol by minimizing or maximizing the FOM. Experience has shown the use of a graphical method to pick between a few protocols. The steps of the graphical method are outlined as follows:

1) Choose a few protocols to test based on materials being scanned.
2) Test the protocols
    a. Plot:
        i. HU vs. Energy (Figure 18)
        ii. HU vs. Concentration (Figure 19)
    b. Analyze:
        i. HU vs. Energy: should decrease if there is no K-edge, increase at K-edge
        ii. HU vs. Concentration: should be linear
3) Choose a protocol based on the result.

As written above, the analysis focuses on confirming that the protocol is behaving as expected. The K-edge should be apparent in the appropriate energy-bin, and the attenuation (HU) should

linearly increase with concentration. Figure 18 and Figure 19 show a protocol that is behaving as expected and could therefore be used in practice. Figure 20 shows a calcium protocol that is not behaving as expected; the attenuation should be monotonically decreasing with energy but the second energy-bin shows higher attenuation than the first. This protocol did not pass the protocol testing and was not used.

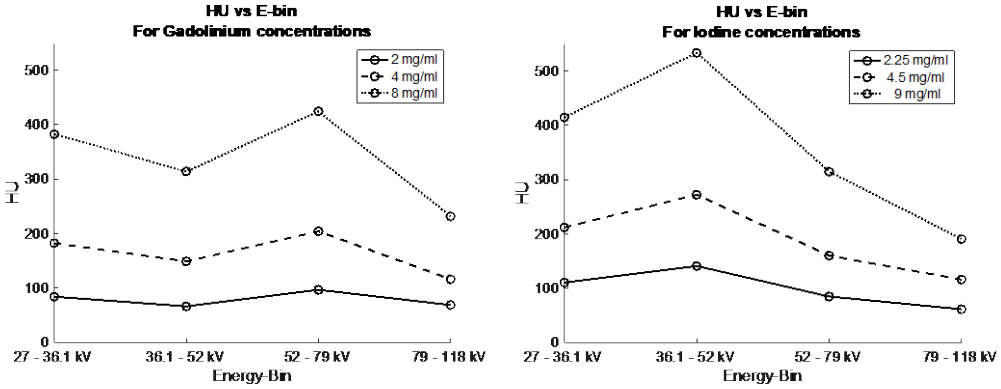


Figure 18: HU vs E-bin plotted for gadolinium (left) and iodine (right), shows the increase in attenuation due to the K-edge's (50.2 keV and 33.2 keV, respectively), appear in the expected energy-bin. Figures from personal research.

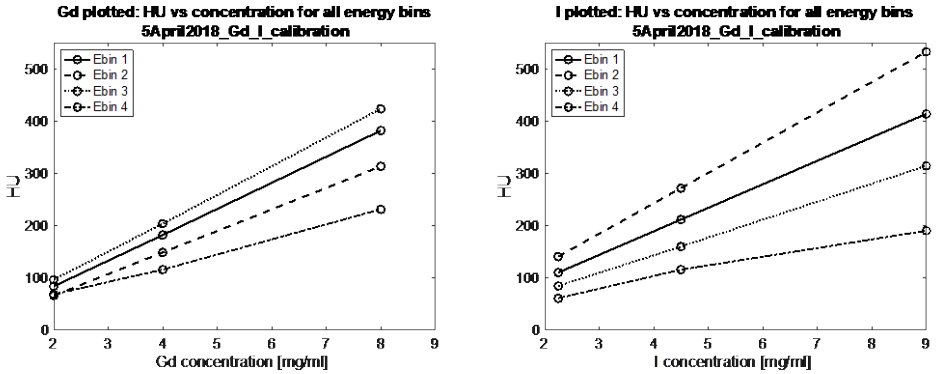


Figure 19: HU vs concentration in each energy bin plotted for gadolinium (left) and iodine (right), shows the expected linear behavior of attenuation with concentration. Figures from personal research.

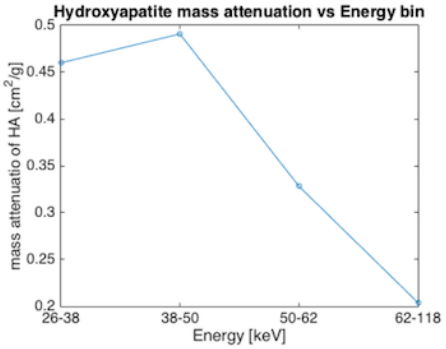


Figure 20: A bad protocol for calcium hydroxyapatite. The mass attenuation is expected to decrease with energy but increases from bin 1 to bin 2. Figure from personal research.

2.3.9. Implications

Similarities between FOMs used to optimize protocols indicate that both noise and K-edge signal are known to be important for protocol success. This is highlighted by the fact that both optimization research and practice use noise-based and K-edge-based tests. Noise-based FOMs used in research include CRLB (used to find the theoretical lower bound of noise for any FOM involving noise, e.g. SNR and SNR^2/exposure), SNR^2, size of confidence intervals, and relative X-ray intensity. K-edge signal metrics used in research include CNR as function of bin width, SDNR, and K-factor. In practice, both noise and K-edge signal factor into protocol acceptance in the "on-demand" empirical protocol testing method, where graphs of attenuation vs. energy and attenuation vs. density are analyzed for unexpected results. Protocols rejected with this test are usually thought behave unexpectedly due to noise.

Although protocol optimization research and practice agree on the importance of noise and K-edge signal, there is still a disconnect regarding placement of K-edge thresholds. Excluding one result from Nik et al.'s two-bin iodine-water protocol, optimized with an energy-threshold of 60 keV, optimization research indicates that energy thresholds should be positioned on the K-edge when imaging K-edge contrast materials. However, in practice, Mars Bioimaging's research group advises that K-edge thresholds should be skewed a few keV above or below the K-edge. The rationale behind this is that the system has imperfect energy resolution, so when a threshold is placed on the K-edge, there is some signal sharing between pre- and post-K-edge bins. It appears that skewing the threshold preserves the K-edge better than placing it directly on the K-edge. However, the magnitude and direction which the threshold should be moved away from the K-edge are not known, and the impact of this decision on PCCT protocols needs to be quantified.

Out of seven protocol optimization methods discussed, five were based on models and two were based on empirical data. Again, this highlights a disconnect between research and practice; research trends toward models, which can be as complex as desired, but still may not cover machine-to-machine variation. In practice, the Mars Bioimaging researchers are using a simple empirical method to choose protocols. Although this "educated guess and check" method seems simplistic, Mars Bioimaging produces excellent results and their many publications are a testament to the validity of this method.

The effectiveness of the empirical guess and check method also begs the question: how optimized does a protocol need to be? The fact that a protocol can still be effective even if it is imperfect is promising for clinical implementation, where varying patient thickness cause variations in beam hardening and would therefore make threshold optimization a moving target within a single scan. For future clinical implementation, it is still important to optimize protocols to obtain the best material decomposition for the least dose, and it is also important to understand how the protocol affects the accuracy of material measurement.

This research presents an empirical protocol optimization routine that attempts to add information to bridge the gap between research and practice. Choice of using an empirical method is twofold: (1) it avoids many limitations of computer models and (2) this builds on the method used in practice. Unlike a computer model, empirical models take into account all complexities of the system they are testing without sacrificing computation time or being only relevant for one machine. The proposed method also goes beyond the currently employed empirical testing method by adding information based the true goal of PCCT protocols: accurate material quantification.

Furthermore, collecting empirical protocol data could eventually be used to investigate current unknowns, such as K-edge thresholds, impact of SNR vs. Energy Resolution, and thresholds adjustment for different-sized subjects.

3. Materials & Methods

3.1. MARS System

The PCCT scanner used in this study was the mars small-animal scanner (Mars Bioimaging Ltd., Christchurch, NZ). The detectors in this scanner are Medipix3RX detectors, made from Cadmium Zinc Telluride (CZT). The particular scanner we used was equipped with a 7-chip camera. Each chip has a 128 × 128 array of pixels with 110 um pixel pitch. Information about the scan parameters used in this research is provided in Section 3.3.2.

Medipix3RX detectors have eight energy thresholds: one arbitration threshold used for noise elimination at 7 keV, three thresholds which can be used in single-pixel mode, and four thresholds which can be used in charge-summing mode. Charge-summing mode was developed to improve energy resolution; in events where a photon signal is split between multiple pixels (charge sharing), the energy signal is summed and assigned to the pixel with the highest signal. Single-pixel mode treats every pixel as a separate address, resulting in an increase in low-energy counts due to charge sharing events. Charge-summing mode was used in this study.

3.2. Calibration and Test Materials

The proposed optimization routine relies on having previous knowledge of calibration and test material concentrations. The analysis is set up to test two materials in water and must be calibrated with single-material concentration series. Materials in this research were iodine in the form of Omnipaque and gadolinium in the form of Onmiscan. Omipaque, chemical name Iohexol ($C_{19}H_{26}I_3N_3O_9$), is a water-soluble iodine-based radiocontrast agent with an iodine concentration of 350mg/ml [62]. Omniscan, chemical name gadodiamide ($C_{16}H_{28}GdN_5O_9$), is an aqueous

gadolinium-based Magnetic Resonance (MR) contrast agent with a concentration of 287 mg/ml gadodiamide, equating to 76.28 mgGd/ml [63].

These contrast agents were diluted with Milli-Q® water (ultrapure water [64]) and mixed with the goal of making an iodine and a gadolinium calibration series of approximately 2.5, 5, 10, and 15 [mg/ml], and test mixtures of 1/9, 2.5/7.5, 5/5, 7.5/7.5, 7.5/2.5, and 9/1 [$\frac{mgI}{ml}/\frac{mgGd}{ml}$]. These dilutions and mixtures split into three sets: one set of samples was sent to the ICP-MS lab at OHSU to be tested for more accurate iodine and gadolinium concentrations, the second set of samples was shipped to the MARS Bioimaging group at the University of Otago, Christchurch, and the third was stored in case either of these steps needed to be repeated.

ICP-MS measurements were performed in the OHSU Elemental Analysis Core with partial support from the NIH instrumentation grant S10RR025512. Table 2 provides the concentration results as reported by ICP-MS. Following the advice of the director of the Elemental Analysis Core, Dr. Martina Ralle, we used one significant figure for these concentrations in the analysis.

Table 2: ICP-MS Results

ID	I [MG/ML]	I COV [MG/ML]	GD [MG/ML]	GD COV [MG/ML]
I/GD 7.5/7.5	7.389	0.297	7.639	0.054
I/GD 7.5/2.5	7.512	0.302	2.596	0.019
I/GD 9/1	8.925	0.359	1.107	0.008
I/GD 1/9	1.270	0.051	9.302	0.066
I/GD 2.5/7.5	2.676	0.108	7.386	0.053
I/GD 5/5	4.884	0.197	5.285	0.038
I 2.5	2.745	0.110	<0.003	N/A
I 5	5.239	0.211	<0.002	N/A
I 10	11.095	0.447	<0.002	N/A
I 15	15.041	0.605	<0.002	N/A

GD 2.5	< 0.06	N/A	2.735	0.019
GD 5	< 0.06	N/A	5.341	0.038
GD 10	< 0.06	N/A	11.195	0.080
GD 15	< 0.15	N/A	16.704	0.119
BLANK	< 0.04	N/A	<0.006	N/A

3.3. Optimization Routine

The proposed optimization routine is an iterative procedure that allows protocols to be analyzed graphically and with figures-of-merit based on material quantification accuracy for each energy-bin and for the system as a whole. Additionally, the routine is built into a GUI that progresses the user through each step in the analysis. Some benefits of this GUI include:

- Calibration and test samples can be scanned either together or separately. Using in-scan calibration reduces the number of scans needed.
- Images can be viewed inside the GUI; there is no need to open a separate program (ImageJ) to check images.
- The analysis is split into tabs in the GUI, which makes the process forgiving to user error. If an error is made, the user can start again from the last correct tab to re-analyze instead of starting over completely.

3.3.1. GUI

The Protocol Optimization GUI was built using MATLAB App Designer. It is built in two modules which are separated onto two main tabs on the GUI window. Figure 21 shows module 1, “New Scan,” which allows the user to input the DICOM energy-image reconstructions and results in a saved calibration file and sample data file. Figure 22 shows module 2, “Load Data & Make Report,” which allows the user to load calibration and sample files, then generates a report PDF

and saves the data from the report into a report data file. Details of the data collection and processing are discussed in the following sections: 3.3.1.1 and 3.3.1.2.

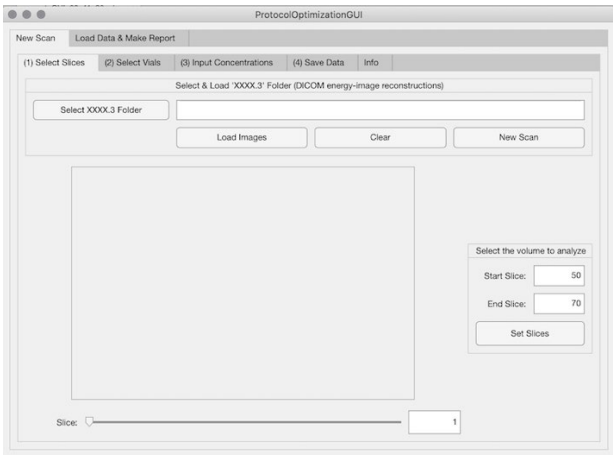


Figure 21: GUI module 1," New Scan."

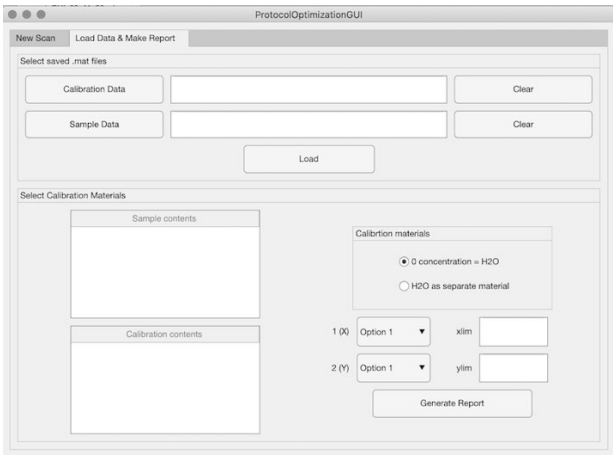


Figure 22: GUI module 2, "Load Data & Make Report."

3.3.1.1. "New Scan" Module

The "New Scan" module has multiple tabs which the user is automatically progressed through after the completion of each stage; sections 0 - 0 describe the function of each tab. All figures are placed at the end of Section 3.3.1.1 and would ideally be viewed alongside the sections as they are mentioned. The MATLAB App is published for download with a supplementary image folder (10 slices) at https://github.com/celesteleary/PCCT_Analysis_GUI.

Select Slices Tab

On the Select Slices tab, the user first selects the folder containing the DICOM energy-image reconstructions, then loads the images. The Load button loads the images and DICOM information into Matlab while displaying a progress bar (Figure 23). The image files are organized as a 4-

dimensional matrix (pixel rows × pixel columns × number of slices × number of energy-bins) and are rescaled by the appropriate values stored in the DICOM information:

images = images*info.RescaleSlope + info.RescaleIntercept.

The stack of images from the first energy-bin are displayed on the UIAxes, and the user can scroll through the images to pick the set of slices that are appropriate for analysis. Slices used for the analysis in this research were 235 – 275, as shown in Figure 24.

Pressing the Set Slices button first disables buttons on this tab, then loads the selected slices onto the next tab figure, and finally sets the next tab as active (user is moved automatically to the next tab). The center slice from the selected slices is displayed on the next tab's figure axes, sliders are made available for window/leveling and scrolling through slices, and an invisible green pixel mask is initialized on top of the figure for later use in displaying sample selections.

Select Vials Tab

On the Select Vials tab, shown in Figure 25, the user enters the number of samples to analyze and the radius (in pixels) of the circular sample areas, then presses Start. In all of the scans in this research, fifteen vials were selected with a sample radius of 17 pixels.

Pressing Start opens a new window displaying the center slice of the selected image series. Again, an invisible green pixel mask is initialized over this image. When the user clicks on the centers of the samples, green circles showing the sample boundary and sample number appear. MATLAB's ginput function is used to identify and store the pixel coordinates that are clicked, and circles are

drawn by turning the opacity of the green mask from 0 to 1 in a pixel range from $radius - 2$ to $radius$ from the selected center pixel coordinate (e.g., pixels that are 15-17 pixels away from the center coordinate are made opaque green). Text indicating the sample number is also displayed on the selected pixel. Figure 26 shows the image during sample selection. MATLAB displays crosshairs for pixel selection, but they become inactive during a screen capture so are not shown here.

After the last sample is selected, the figure automatically closes, and the green circles are displayed using the green mask over the figure on the GUI. The user can then scroll through the images and window/level to check the efficacy of their sample selection (Figure 27).

After vials are selected and checked, the user inputs a list of materials (other than air and water) in the scan and presses Enter. The Enter button displays a progress bar while data for the next page is initialized. At this point the pixel means and standard deviations in each sample volume are calculated, saved, and displayed on the next tab, and tables are initialized for the user to input the sample concentrations. The next tab is activated when this process is complete.

Input Concentrations Tab

The plot on the Input Concentrations tab shows linear attenuation vs. energy-bin for each sample volume (Figure 28 and Figure 29). Pixel values in the images represent the linear attenuation of each voxel, so the mean and standard deviation of the linear attenuation if measured by taking the mean and standard deviation of the pixel values in each sample volume. The sample numbers from the previous tab correspond to the line numbers, and the lines can be toggled on and off with the

"plot on/off" check boxes for better visualization of sample subsets. To illustrate this use, Figure 28 shows the linear attenuation for the series of iodine concentration vials, and Figure 29 shows the linear attenuation for the series of gadolinium concentration vials.

The user inputs the known concentration in each sample into the "Materials" table, and checks boxes indicating water and air vials. A few shortcut options are given for inputting the same concentrations for multiple analyses: "Save new file concentrations.mat," "Load from file concentrations.mat," and "Load coded concentrations." The "Save new file" button will save the table contents to a file named "concentrations.mat" in the parent directory of the DICOM image folder. The "Load from file" button will open a file selection dialog box, from which one selects a previously saved "concentrations.mat" file. The "Load" button is useful for repeatedly entering the same concentrations quickly, but in order to use it the user must type a properly formatted table of concentration into the "Loadcodedconcentrations" function in the App code.

When the concentration table is finalized, the Enter button initializes the tables on the next tab, then automatically moves the user to the Save Data tab.

Save Data Tab

The Save Data tab is pictured in Figure 30 and has two sections which are numbered and written in bold. In Section 1, the user checks boxes indicating which samples to use in the calibration matrix (A) and sample vector (b). The buttons at the top are shortcuts for auto-selecting all calibration or sample checkboxes. In Section 2, options for the saved data must be specified, then the data can be saved, and the calibration plots can be displayed if desired. Option 1 adds the water

sample as the y-intercept of the calibration curve. If there is more than one water sample, the pooled average and error are used. Option 2 allows the user to select energy-bins. This study used the water-intercept for both I and Gd and did not use the lowest energy-bin (bin 7-X keV).

The Save Data button processes the data based on the selections and saves information for the calibration matrix (A) and sample measurement vector (b). This information is saved as "calibration.mat" and "samples.mat" files in the directory containing the DICOM image folder. Descriptions of the file contents are in Table 3 and Table 4. Details of the calculations of A, SD(A), b, and SD(b) are provided in Section 3.4.

The Calibration Plots button opens and saves figures that show the saved calibration information. Examples of these are given in Figure 31, Figure 32, and Figure 33. For each material and energy-bin, the mean linear attenuation vs. concentration data is plotted with error bars showing the standard deviation of the linear attenuation, and the average linear fit and the standard deviation of the linear fit. For water and air, the mean and standard error of the linear attenuation is plotted for each energy-bin.

Table 3: The file calibration.mat contains a structure, "Cal," with fields: m, m_SD, concentration, and EbinString.

Cal.m:	a table of mean slope for each material and energy-bin, used to create the calibration matrix, A.
Cal.m_SD:	a table of slope SDs for each material and energy-bin, used in the error propagation, discussed in Section 3.4.3.
Cal.concentration:	a table storing material concentrations that were used to create the calibration curve, used for reference in the "Load Data & Make Report" module.
Cal.EbinString:	a cell array containing which energy bins were used in the calibration. This is used to check that the calibration and sample are compatible in the "Load Data & Make Report" module.

Table 4: The file samples.mat contains a structure, "Samples," with fields: averages: SDs, concentrations, info, EbinString, calculated_concentrations.

Samples.averages:	a matrix of mean linear attenuation, Samples × Energy-bin.
Samples.SDs:	a matrix of linear attenuation SD, Samples × Energy-bin.
Samples.concentrations:	a table of known material concentrations that are in each sample, organized as Samples × Material.
Samples.info:	the DICOM info from the sample scan.
Samples.EbinString:	a cell array containing which energy-bins were used in the calibration.
Samples.calculated_concentrations:	a table of calculated material concentrations that are in each sample, organized as Samples × Material.

Info Tab

The Info tab has information about the resulting files, "calibration.mat" and "samples.mat," see Figure 34.

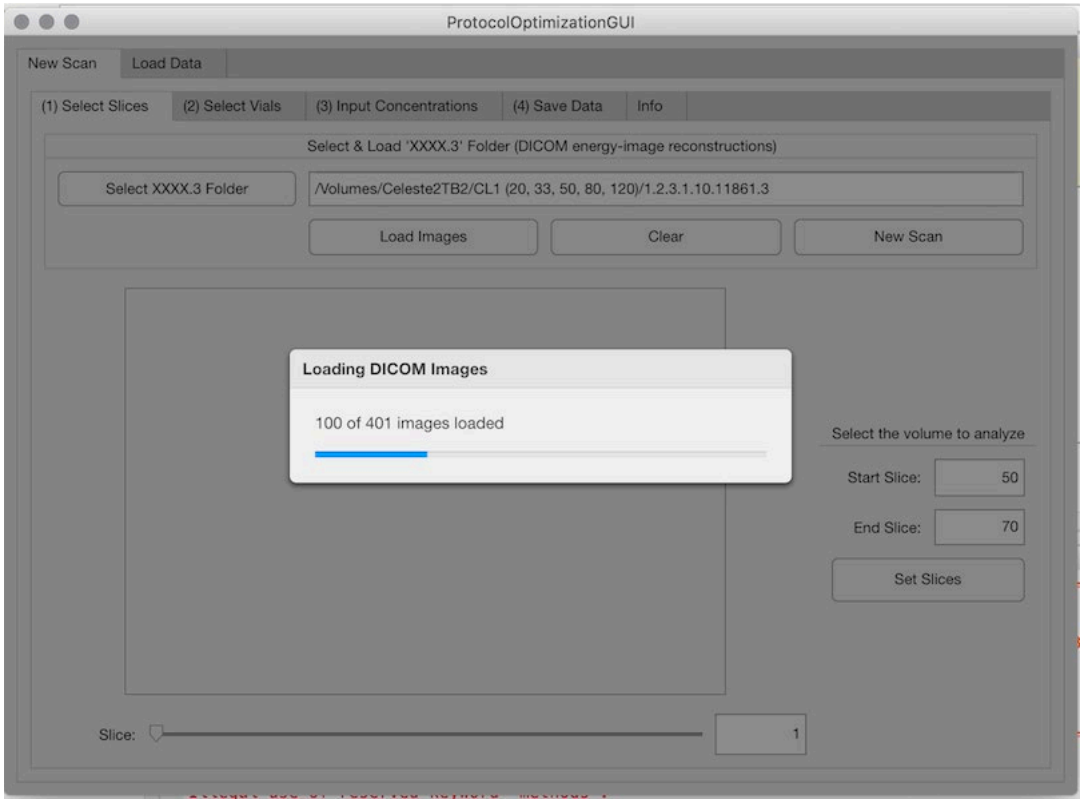


Figure 23: Select DICOM energy-image file & load images

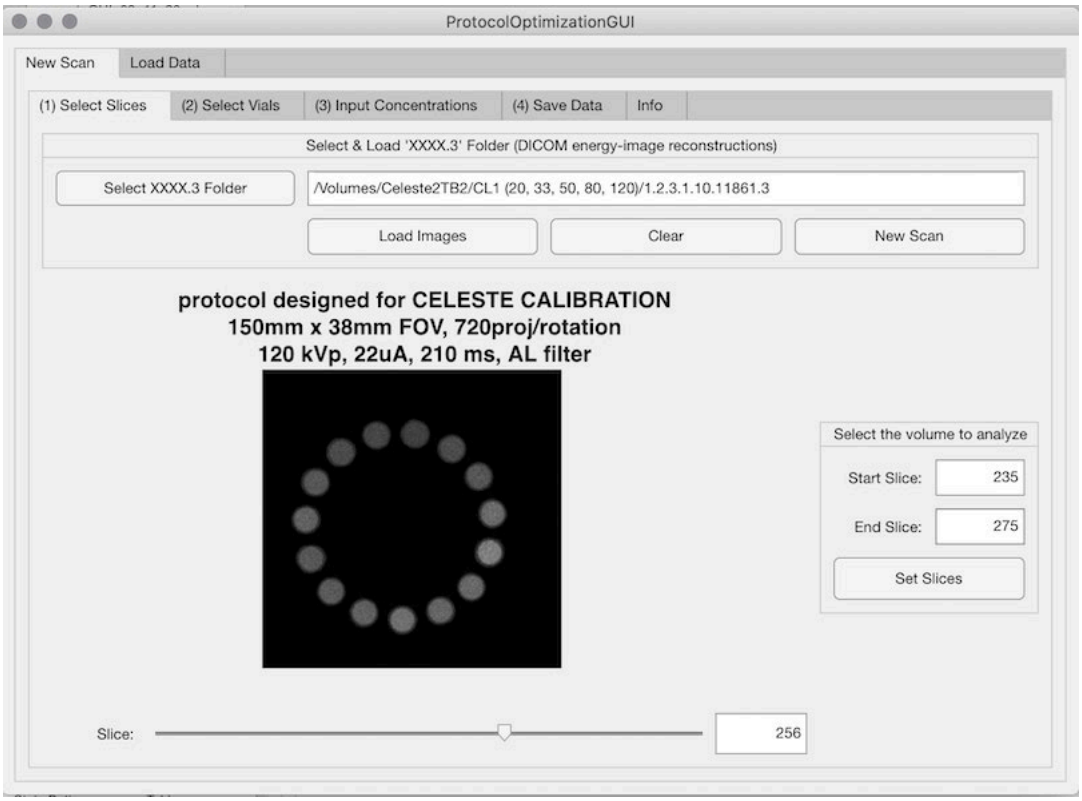


Figure 24: Scroll through the images and select the series of slices to analyze.

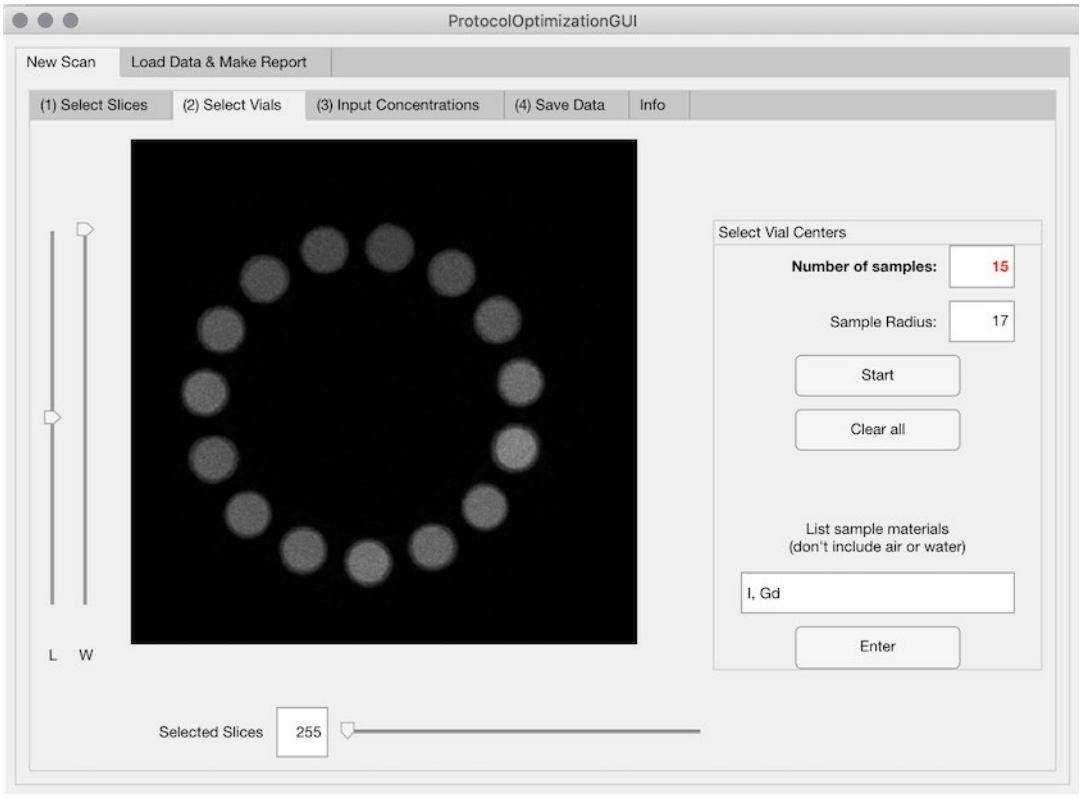


Figure 25: Input the number of samples in the scan and the sample radius (in pixels).

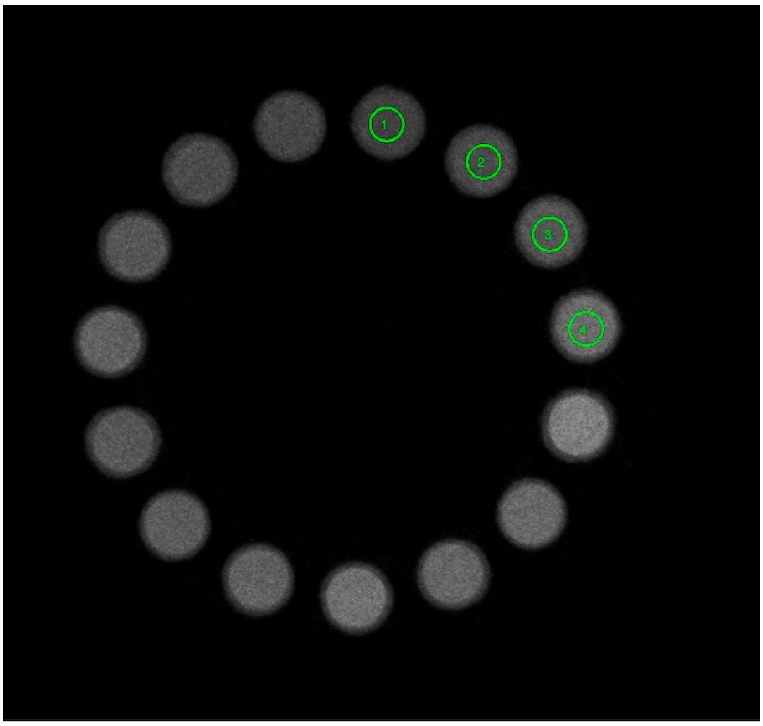


Figure 26: A new window will open, allowing you to resize the image and select the center of each sample. Click on the centers of each sample vial (crosshairs in the program, not apparent in screen capture).

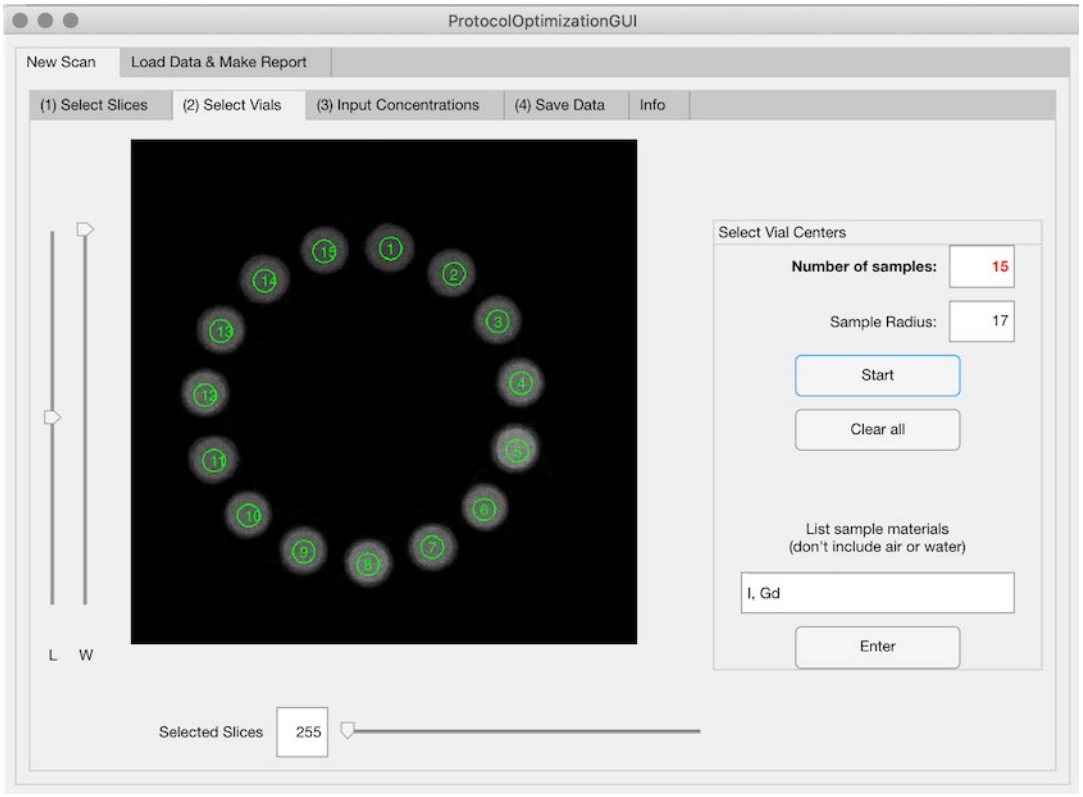


Figure 27: Scroll through the images and use Window/Level to ensure the selections are valid.

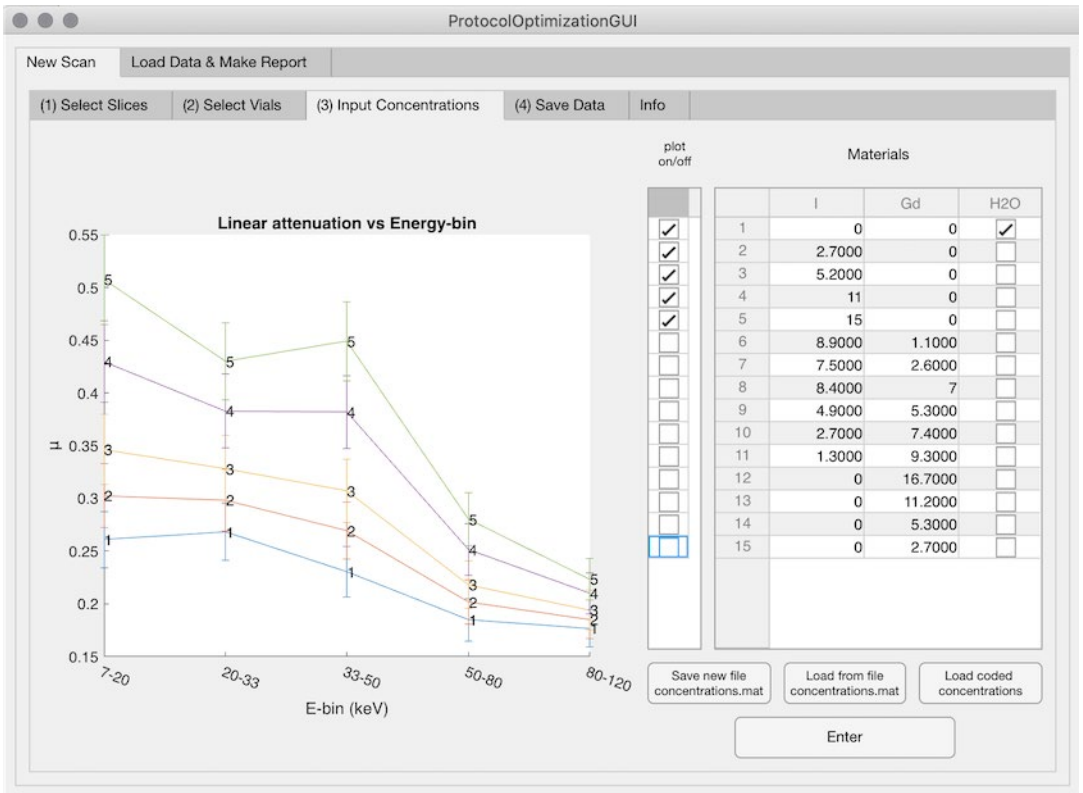


Figure 28: Enter the known concentrations into the table. Linear attenuation vs Energy-bin is displayed in the figure axis.

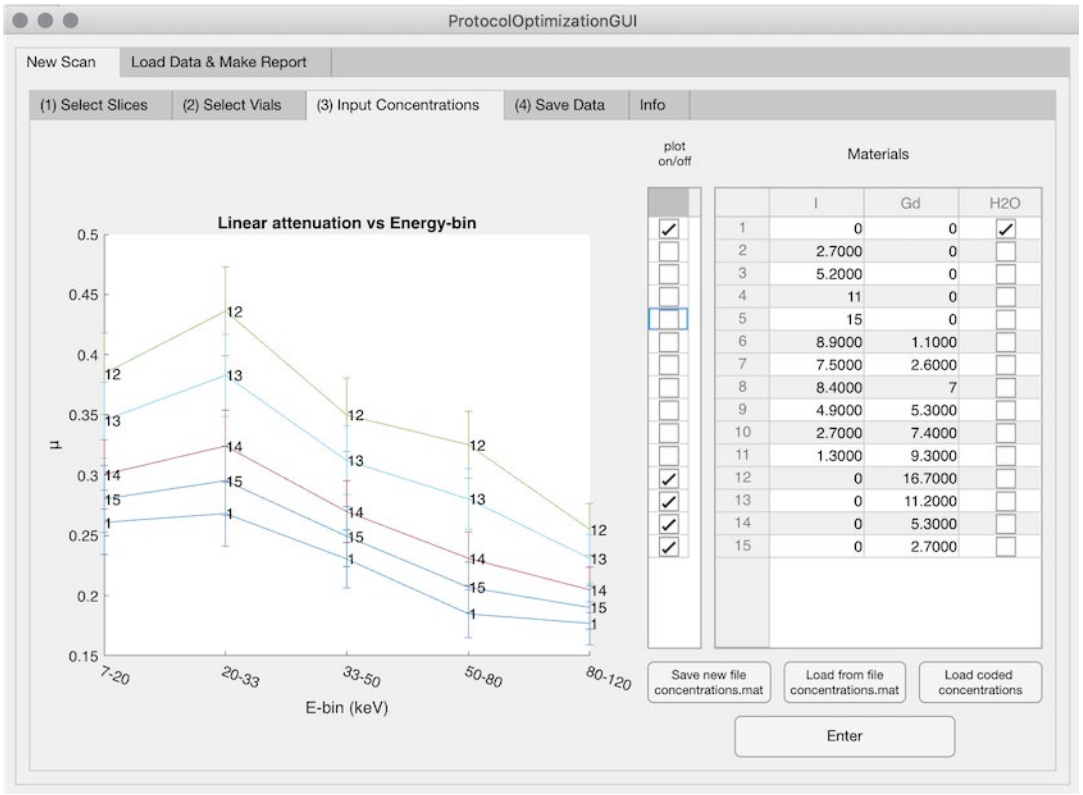


Figure 29: Linear attenuation vs E-bin for each sample can be toggled on/off.

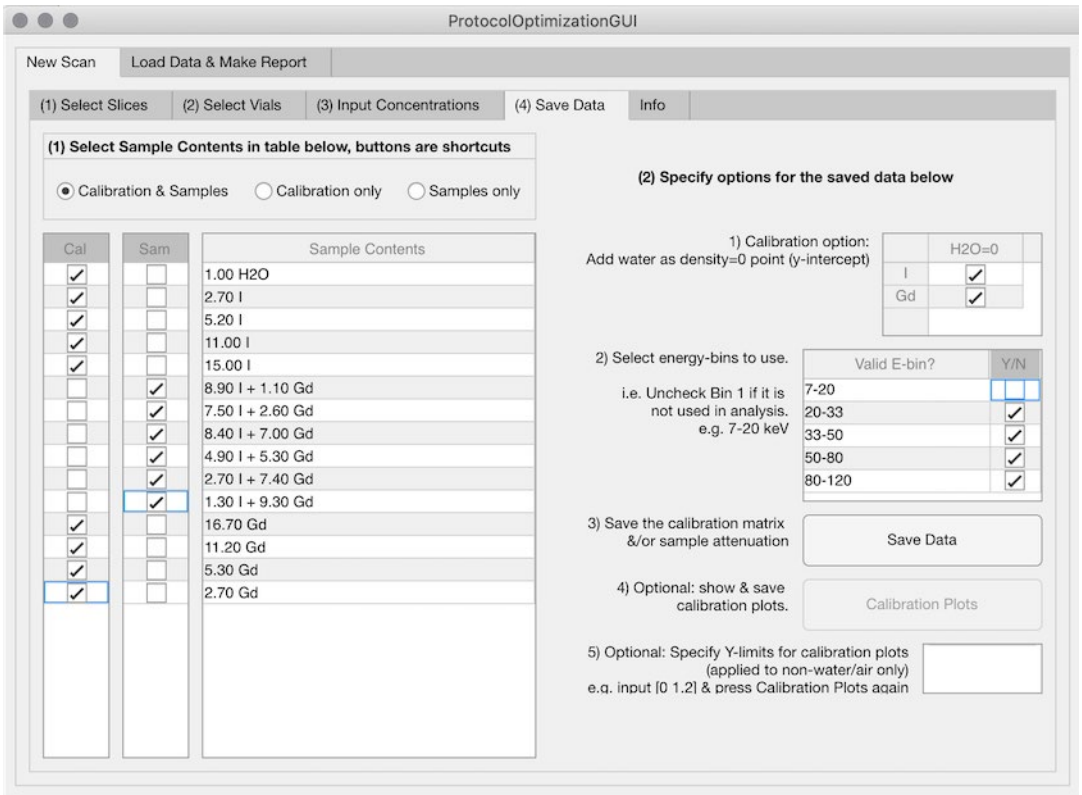


Figure 30: Numbered directions: (1) select the samples to use in the calibration and the optimization test. (2) Data options, Save Data, Show/Save Calibration plots.

I

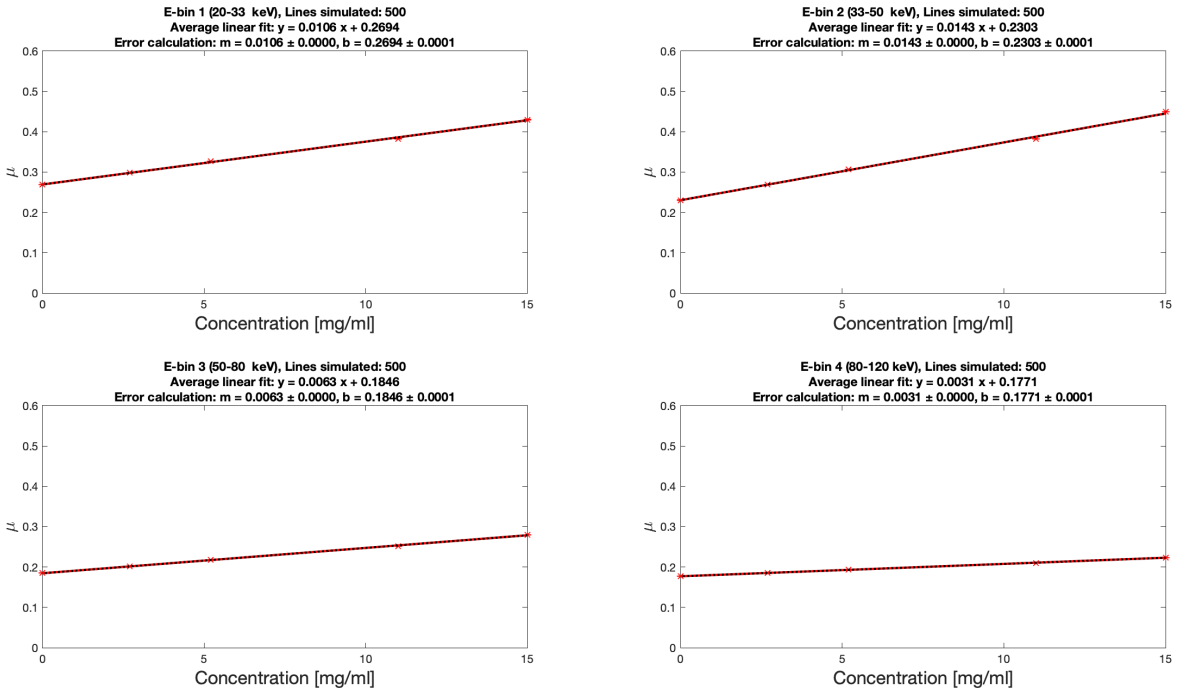


Figure 31: Iodine calibration plots.

Gd

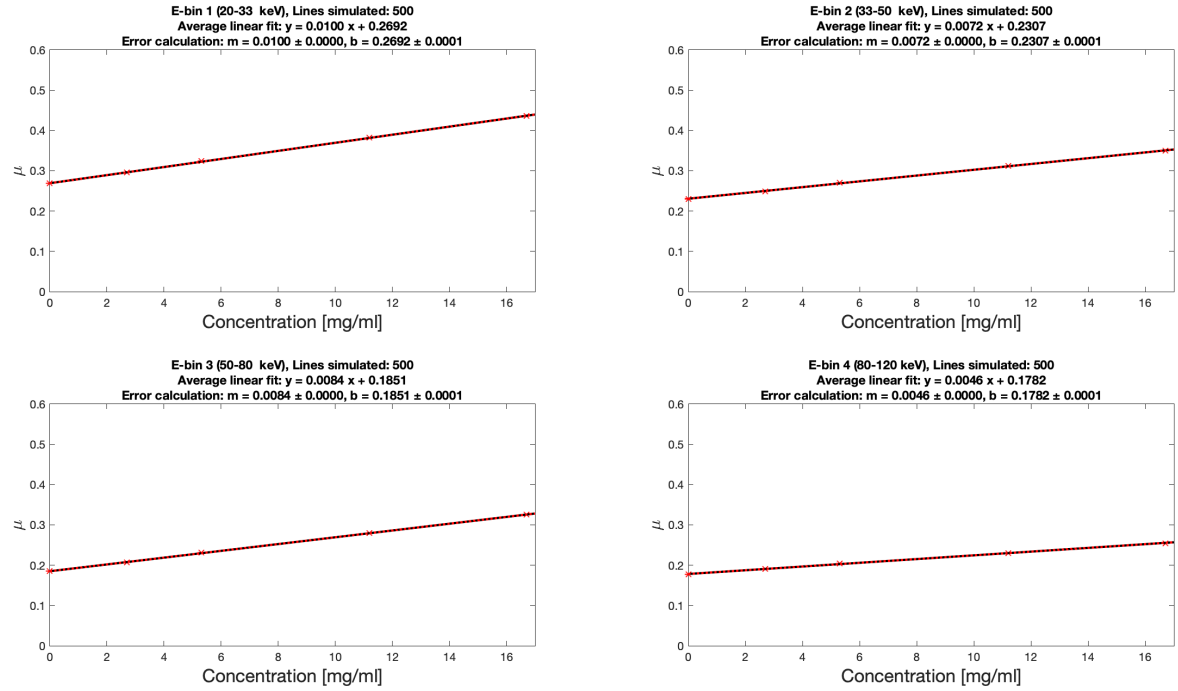


Figure 32: Gadolinium calibration plots.

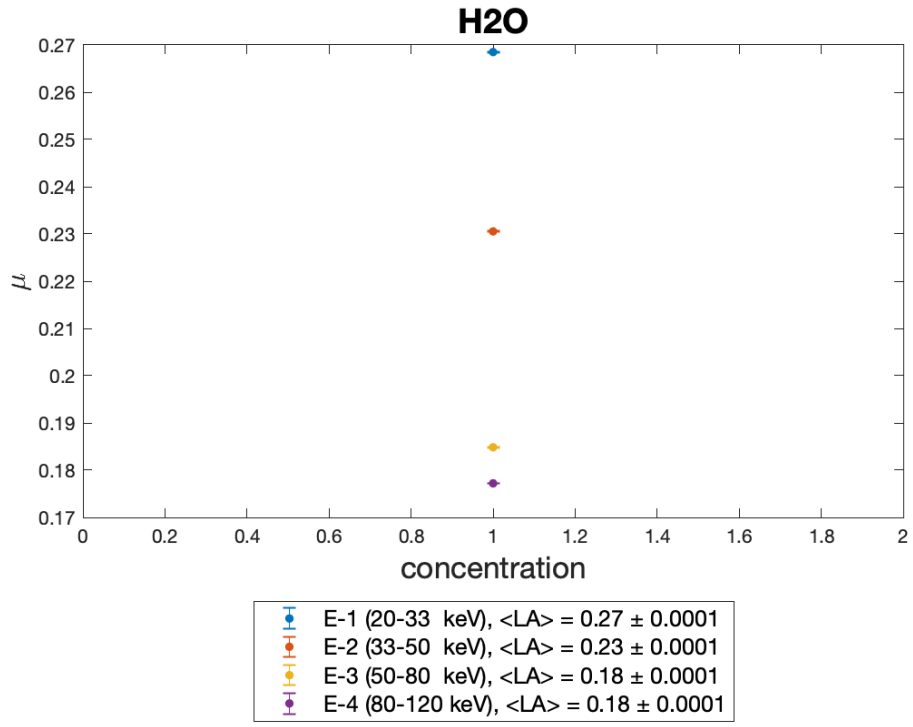


Figure 33: Linear attenuation for water in each energy- bin.

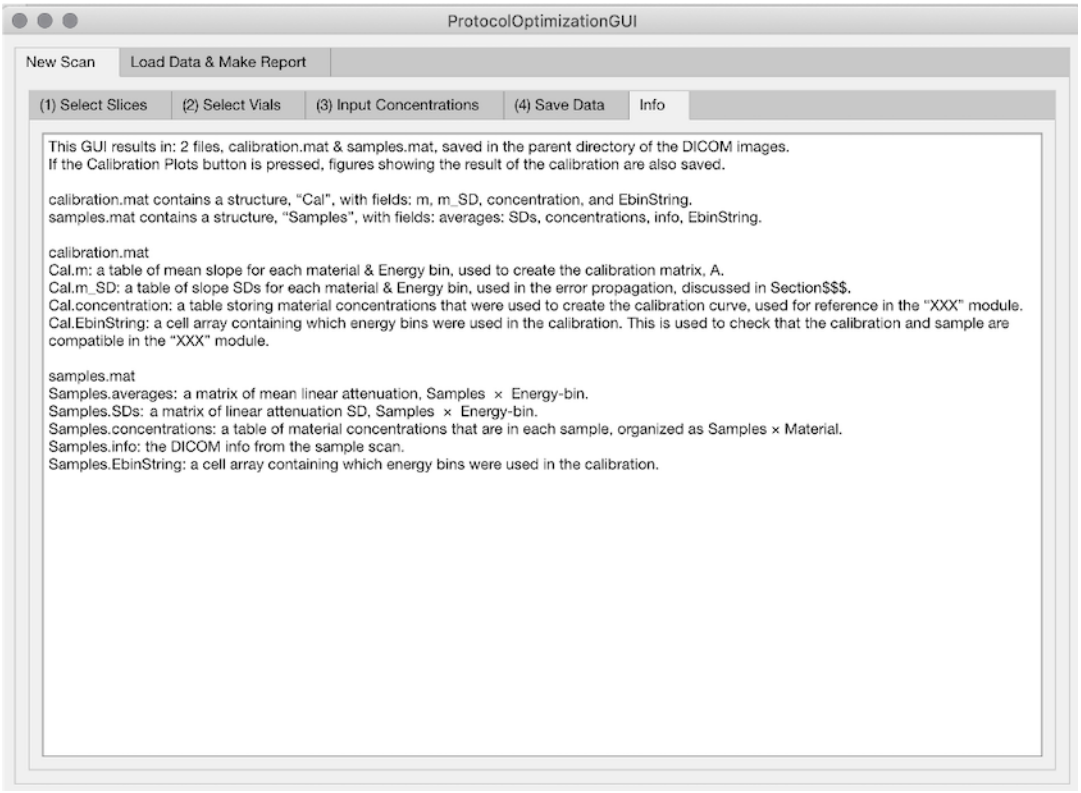


Figure 34: Info tab displays details of results of the "New Scan" module.

3.3.1.2. "Load Data & Make Report" Module

The Load Data & Make Report module allows the user to load the saved calibration.mat and sample.mat files, select the material basis, and generate the protocol report. This module is pictured in Figure 35, and an example of a report is shown in Figure 36.

When the user selects the calibration.mat file after pressing the Calibration Data button, the Sample Data field is automatically filled with a path to a sample.mat file in the same directory as the calibration.mat file. This is useful for loading data made from the same scan. If data from a different scan is needed, the Sample Data button can be pressed to select any sample.mat file.

The Load button initializes the Sample Contents and Calibration Contents tables. The Sample Contents table lists the materials concentrations in each sample. The Calibration Contents table lists each material saved in calibration.mat and shows the concentrations that were used to make the calibration curves.

The options on the right side of the panel are used to select the material basis. This GUI is set up to display 2D graphs of a three-material system of water plus two materials. This is accomplished by subtracting the water linear attenuation from the sample mixture attenuation. The assumption behind this subtraction and the effect on noise are discussed in Section 3.4.2.

The Generate Report function performs all calculations and makes the figures that are in the report, then saves and opens the generated report. There are multiple steps to this process: making and

saving figures, calculating metrics (Δ, $\Delta_\perp$, confidence interval area), saving the report data, and generating and saving the report (ProtocolReport.pdf).

When the Generate Report button is pressed, the fields Cal.EbinString and Samples.EbinString are first checked to make sure the calibration is correct for the sample scan. If they do not have the same energy-bins, the process does not run. Instead, a warning message pops up telling the user that the calibration and sample scan energy-bins do not match, and the energy bins stored in each file are printed to the warning message. If the energy-bins of the calibration and sample files match, a progress bar is displayed while the data is processed, and the report is generated. A folder named "Report" is created in the parent directory of the DICOM image folder, and it is used to store the report (ProtocolReport.pdf), the data calculated in the report (ReportMetrics.mat), and the figures generated for the report (saved in a folder called "Figs").

The materials in the X and Y dropdowns are used to make the calibration matrix and initialize the X and Y axes. For each sample, a figure is made that shows the basis-material linear system, Ax=b, and report metrics. Each plot shows a line for each energy-bin, a red X for the known concentration of the sample, a red point for the least squares solution to the linear system, and ellipses defining the 68% and 95% confidence intervals. The calculation of the energy-bin lines and the confidence intervals are given in Section 3.4.

Metrics are calculated for each plot: Δ(known, calculated), $\Delta_\perp$(E-bin), and the confidence interval areas. $\Delta(\text{known, calculated})$ is the Euclidean distance between the known and calculated concentration, equivalently, the residual sum of squares. $\Delta_\perp(\text{E-bin } i)$ is the Euclidean distance

between Energy-bin line *i* and the known concentration. The confidence interval areas are the area of each confidence ellipse. Two figures are made to compare Δ and $\Delta_{\perp}$ across samples. These figures are explained in the Results (Section 4).

The report, ProtocolReport.pdf, is made using MATLAB's Publish function [65]. Formatting and adding figures using *Publish* is accomplished through Publishing Markup, specifically formatted MATLAB comments [66]. For example, including a saved picture in a report is done by typing the comment:

```
%
% <<FILENAME.PNG>>
%
```

, where FILENAME includes the full path of the image file. In order to create a PDF report that has comments that change to pull the correct images into the report, a function was written that uses sprintf() to print a Markup file specific to the calibration and sample files that are being analyzed. The customized Markup file is Published, then deleted. The resulting PDF is saved in the Report folder and opened for the user to view.

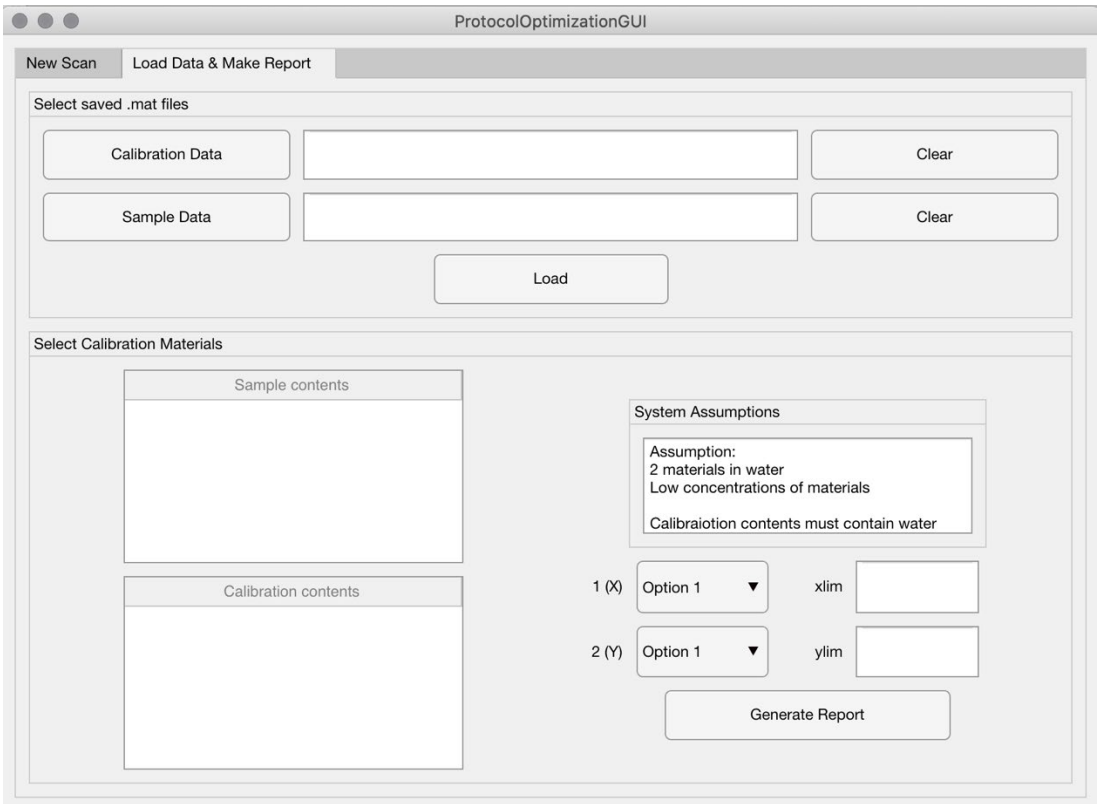
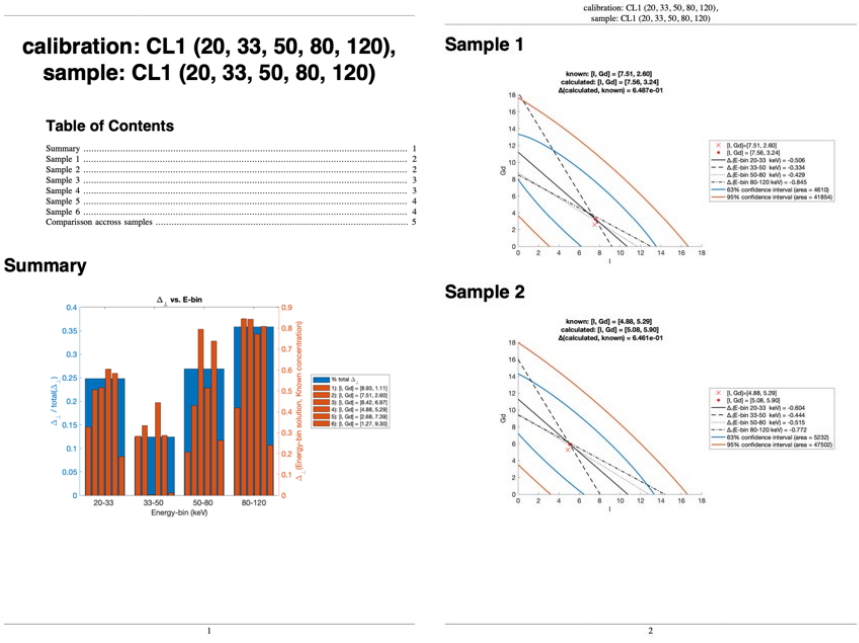


Figure 35: "Load Data & Make Report tab." The files calibration.mat and samples.mat are loaded, the material basis is selected, and the report is generated and opened.



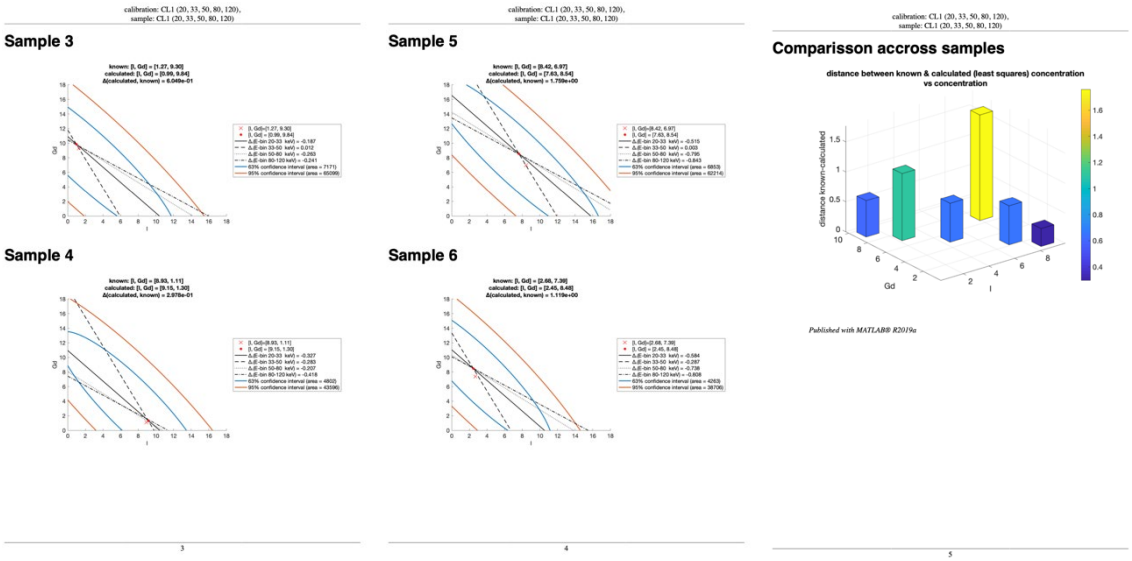


Figure 36: ProtocolOptimization.pdf, report generated with MATLSB's Publish function.

3.3.2. Scans

The empirical protocol optimization routine proposed in this research was set up as an iterative procedure. The initial plan was to complete three series of scans: the first series would test the K-edge thresholds and quantify the scan-to-scan variability, pictured in Figure 37, then, the second and third series of scans would find the low-energy threshold and the high-energy threshold.

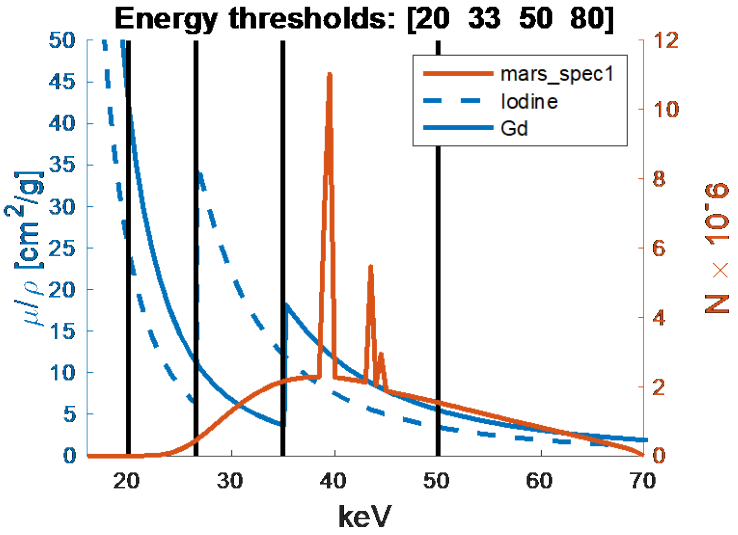


Figure 37: Scan used to test the on-K-edge thresholds and scan repeatability.

However, in the analysis of the first series of scans it was discovered that protocols were relatively insensitive to moving the K-edge thresholds by a few keV. These results were explored for this research. The scans that were completed are given in Table 5, and the "Future Studies" section proposes how this research could be continued given our discoveries.

Table 5: Completed Scans

#	SCAN ID	E1 [KEV]	E2 [KEV]	E3 [KEV]	E4 [KEV]
1	CL1	20	33	50	80
2	CL2	20	33	50	80
3	CL3	20	33	50	80
4	CL3_2	20	33	50	80
5	CL3_3	20	33	50	80
6	CL4	18	31	48	78
7	CL5	16	29	46	76
8	CL6	22	35	52	82
9	CL7	24	37	54	84

Scan parameters were kept consistent across scans while energy-thresholds were varied. Table 6 shows the scan parameters used.

Table 6: Scan Parameters

kVp	120
μA	22
ms	21
FOV [mm^2]	150×38
Slices	401
Slice Thickness [mm]	0.09
Projections/Rotation	720
Filter	Al
Source-Detector Distance	271.9
Source-Patient Distance	211.9

3.4. Linear System Plots and Error Analysis

The main idea of this research was to obtain information about the effect of energy-threshold placement on accuracy and precision of mixed-contrast PCCT scans by analyzing the linear

system, $Ax = b$, of a water plus two contrast-material basis. Each energy-bin was visualized as a line in material concentration space, and metrics quantifying the material decomposition accuracy of each line and the system as a whole were calculated. This is explained in the following section, 3.4.1.

To assess the precision of the final material concentration calculation, the error from the calibration matrix (A) and the measurement vector (b) was first calculated, then a Monte Carlo method was used to propagate the error through to the solution vector (x). Sections 3.4.3, 3.4.2, and 3.4.4 explain how error in each component, A, b, and x was calculated.

3.4.1. Linear System Plots

In this research, a three-material basis of water, iodine, and gadolinium was used. This linear system can be written as:

$$Ax = b$$

$$\begin{bmatrix} \frac{\mu}{\rho}(E_1)_{Gd} & \frac{\mu}{\rho}(E_1)_{I} & \frac{\mu}{\rho}(E_1)_{H2O} \\ \frac{\mu}{\rho}(E_2)_{Gd} & \frac{\mu}{\rho}(E_2)_{I} & \frac{\mu}{\rho}(E_2)_{H2O} \\ \vdots & \vdots & \vdots \\ \frac{\mu}{\rho}(E_N)_{Gd} & \frac{\mu}{\rho}(E_N)_{I} & \frac{\mu}{\rho}(E_N)_{H2O} \end{bmatrix} \cdot \begin{bmatrix} \rho_{Gd} \\ \rho_{I} \\ \rho_{H2O} \end{bmatrix} = \begin{bmatrix} \mu(E_1) \\ \mu(E_2) \\ \vdots \\ \mu(E_N) \end{bmatrix} \quad (17)$$

To visualize the linear system and confidence intervals in the iodine-gadolinium concentration plane, the attenuation contribution of water was subtracted from both sides of Equation 17 to reduce the three-material basis to a two-material basis (Equation 18). This assumes that the concentration of water in each voxel does not change with contrast concentration, which is a good

approximation for small concentrations of iodine and gadolinium. The effect of subtraction on standard deviation was accounted for (see Section 3.4.2).

$$\begin{bmatrix} \frac{\mu}{\rho}(E_1)_{Gd} & \frac{\mu}{\rho}(E_1)_I \\ \frac{\mu}{\rho}(E_2)_{Gd} & \frac{\mu}{\rho}(E_2)_I \\ \vdots & \vdots \\ \frac{\mu}{\rho}(E_N)_{Gd} & \frac{\mu}{\rho}(E_N)_I \end{bmatrix} \cdot \begin{bmatrix} \rho_{Gd} \\ \rho_I \end{bmatrix} = \begin{bmatrix} \mu(E_1) - \mu_{H2O}(E_1) \\ \mu(E_2) - \mu_{H2O}(E_2) \\ \vdots \\ \mu(E_N) - \mu_{H2O}(E_N) \end{bmatrix} \tag{18}$$

Equation 18 can be plotted as *N* energy-bin lines. To illustrate this, the expanded equation for Energy-bin *n* is given in Equation 19. This is an equation of a line in the form $ax + by = c$. Where a and b are from the calibration, and c is (sample linear attenuation − water linear attenuation). Each line in the system is plotted as shown in Figure 38.

$$\frac{\mu}{\rho}(E_n)_{Gd} \cdot \rho_{Gd} + \frac{\mu}{\rho}(E_n)_I \cdot \rho_I = \mu(E_n) - \mu_{H2O}(E_n) \tag{19}$$

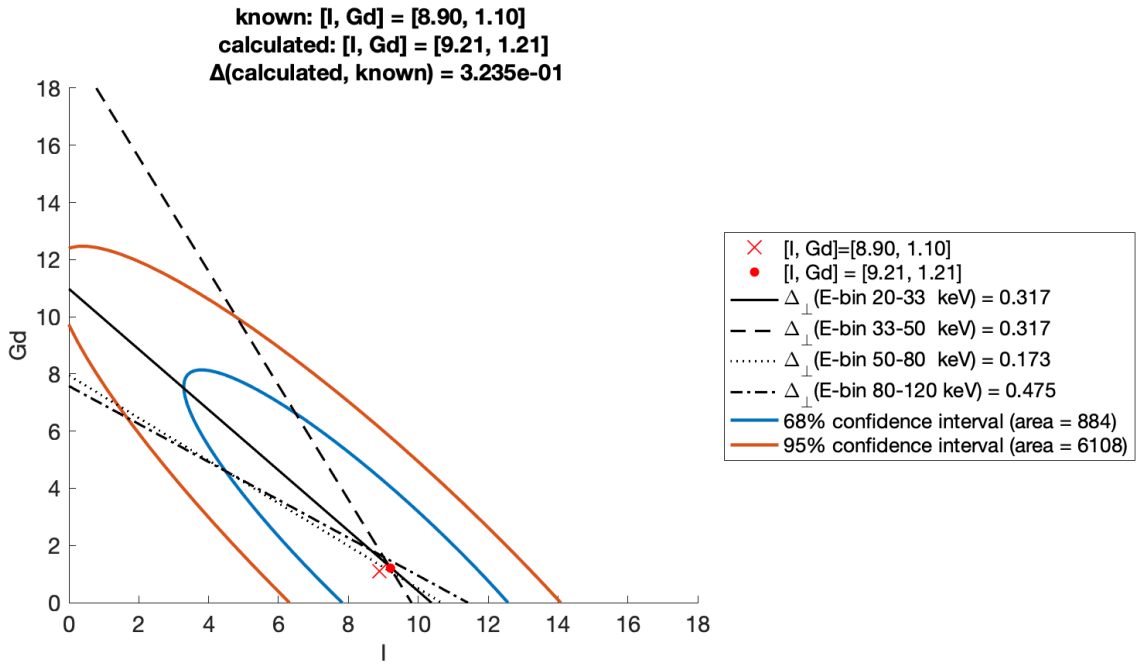


Figure 38: Linear system plot shows energy-bin lines, known solution, calculated solution, confidence intervals, and plot metrics.

The linear system plots also include known sample concentration; red X, calculated sample concentration; red dot, Δ(known, calculated); in title, $\Delta_{\perp}$(E-bin i); in legend, and 68% and 95% confidence intervals; blue and orange ellipses.

The calculated sample concentration is computed using MATLAB's mldivide function, which solves the linear equation $Ax = b$, coded as x = A\b. This function is appropriate for solving overdetermined systems, like one in this research. For an overdetermined system, it returns the least-squares solution to the linear system [67]–[69].

Δ(known, calculated) is defined as the Euclidean distance between the known and calculated sample concentration. The coordinates, (x_{known}, y_{known}) and $(x_{calculated}, y_{calculated})$, are known, so the distance can be calculated with Equation 20.

$$\Delta(\text{known, calculated}) = \sqrt{(x_{known} - x_{calculated})^2 + (y_{known} - y_{calculated})^2} \quad (20)$$

$\Delta_{\perp}$(E-bin i) is defined as the perpendicular distance from Energy-bin line i to the known sample concentration. This is calculated using the Hessian normal form of a plane and computing the point-plane distance (in this case, point-line distance). With a plane in the form $ax + by + cz + d = 0$, and a point (x_0, y_0, z_0), the plane can be described in Hessian normal form as $\hat{\boldsymbol{n}} \cdot \boldsymbol{x} = -p$. The point-plane distance is computed as $D = \hat{\boldsymbol{n}} \cdot \boldsymbol{x_0} + p$, where $\hat{\boldsymbol{n}}$ is the unit normal vector of the plane and p is the distance from the origin to the plane. In this research $\Delta_{\perp}$(E-bin i) = $|D|$.

$a = \frac{\mu}{\rho}(E_n)_{Gd}$, $b = \frac{\mu}{\rho}(E_n)_I$, $c = 0$, $d = -(\mu(E_n) - \mu_{H2O}(E_n))$, and $(x_0, y_0, z_0) = (\frac{I\ mg}{ml}_{known}, \frac{Gd\ mg}{ml}_{known}, 0)$. Equations 21, 22, and 23 are used to calculate $\hat{\boldsymbol{n}}$ and p [70], [71].

$$n_x = \frac{a}{\sqrt{a^2 + b^2}} \quad (21)$$

$$n_y = \frac{b}{\sqrt{a^2 + b^2}} \quad (22)$$

$$p = \frac{d}{\sqrt{a^2 + b^2}} \quad (23)$$

This calculation was performed by writing the function:

```
function D = dist_from_line(app, A, b, x0)
   % perpendicular distance to hyperplane (hessian normal distance)
   % e.g. 3D plane:
   % Ax + By + Cz = b
   % A = [A B C] %calibration coefficients
   % b = signal measured in bin
   % point x0 = [two vectors] % known concentration
   n = A./rssq(A);
   p = -b/rssq(A);
   D = n(1)*x0(:,1)+n(2)*x0(:,2)+p;
end
```

The 68% and 95% confidence intervals (1 SD and 2 SD) for the calculated solution were plotted by calculating error in *A* and *b* separately, then using a Monte Carlo method to simulate 500 possible concentration solutions. The distribution of solutions was used to calculate the confidence regions. This is further explained in the following sections, 3.4.2, 3.4.3, and 3.4.4.

3.4.2. Error Part 1: b, Measurement Vector

The measurement vector, *b*, is calculated as the average linear attenuation of the sample minus the average linear attenuation of water (Equation 24). The pixel value mean and standard deviation were stored as the linear attenuation mean and standard deviation, and error propagation for subtraction was used to calculate error in *b* (Equation 25).

$$\vec{b} = \begin{bmatrix} \mu(E_1) - \mu_{H2O}(E_1) \\ \mu(E_2) - \mu_{H2O}(E_2) \\ \vdots \\ \mu(E_N) - \mu_{H2O}(E_N) \end{bmatrix} \tag{24}$$

$$b_n = \mu(E_n) - \mu_{H2O}(E_n)$$

$$\sigma_{b_n} = \sqrt{\sigma_{\mu_n}^2 + \sigma_{\mu_n,H2O}^2} \tag{25}$$

In the Equations 24 and 25, μ and σ_μ are the mean and standard deviation of pixel values in the test sample, and μ and $\sigma_{\mu,H2O}$ are the mean and standard deviation of pixel values in the water sample. Normality of the pixel value distributions was verified against the expected shape of a gaussian with mean and standard deviation equal to the pixel mean and standard deviation (Figure 39).

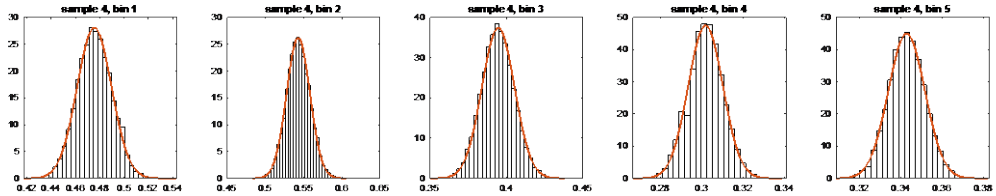


Figure 39: Histograms of pixel values in samples (blue) were checked against Normal distribution (red) to confirm gaussian distribution.

3.4.3. Error Part 2: A, Calibration Matrix

Recall the calibration matrix, A:

$$\overline{A} = \begin{bmatrix} \frac{\mu}{\rho}(E_1)_{Gd} & \frac{\mu}{\rho}(E_1)_I & \frac{\mu}{\rho}(E_1)_{H2O} \\ \frac{\mu}{\rho}(E_2)_{Gd} & \frac{\mu}{\rho}(E_2)_I & \frac{\mu}{\rho}(E_2)_{H2O} \\ \vdots & \vdots & \vdots \\ \frac{\mu}{\rho}(E_N)_{Gd} & \frac{\mu}{\rho}(E_N)_I & \frac{\mu}{\rho}(E_N)_{H2O} \end{bmatrix}$$

Using the linear relationship between μ and ρ (Equation 2), each entry in A can be calculated by plotting the calibration series, μ vs. ρ, and taking the slope of the line fitted to the points (Equation 26).

$$\begin{gathered} \mu = \frac{\mu}{\rho} \cdot \rho + \mu_0 \\ y = mx + b \end{gathered} \tag{26}$$

However, the average linear attenuation measured in the calibration vials, $\bar{\mu}$, has associated error, $\sigma_{\bar{\mu}}$, which must be propagated through to error in each calibration entry, $\frac{\mu}{\rho}$, $\sigma_{\mu/\rho}$. The standard error of the mean linear attenuation was calculated using Equation 27.

$$\begin{gathered} \sigma_{\bar{\mu}} = \frac{\sigma}{\sqrt{n_{pixels}}} \\ n_{pixels} = floor(\pi \cdot r_{selection}^2 \times n_{slices}) \end{gathered} \tag{27}$$

In this research, $r_{selection}$= 17, and $n_{slices} = 275 - 235 = 40$. A Monte Carlo method was used to calculate both $\frac{\mu}{\rho}$ and $\sigma_{\mu/\rho}$. Linear attenuation for each sample density was randomly sampled from the appropriate gaussian distribution using the function normrnd($\bar{\mu}, \sigma_{\bar{\mu}}$). Then, the polyfit function was used to calculate the slope and the intercept of the least-squares fitted line to the points (used as: polyfit($\vec{\rho}$, normrnd($\bar{\mu}, \sigma_{\bar{\mu}}$), 1) [72]. This was repeated 500 times, storing the slope and the intercept for each run. The average slope and standard deviation of the slopes were calculated and used as the estimate of μ/ρ and $\sigma_{\mu/\rho}$. Figure 41 displays the distribution of lines used in an example of this Monte Carlo solution, and Figure 41 shows one plot that is generated by the Calibration Plots button.

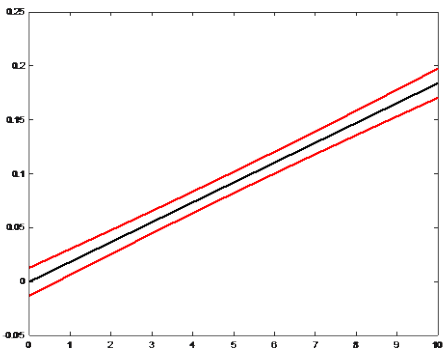


Figure 40: 500 lines were simulated to find the mean and standard deviation of the slope, used for the calibration matrix.

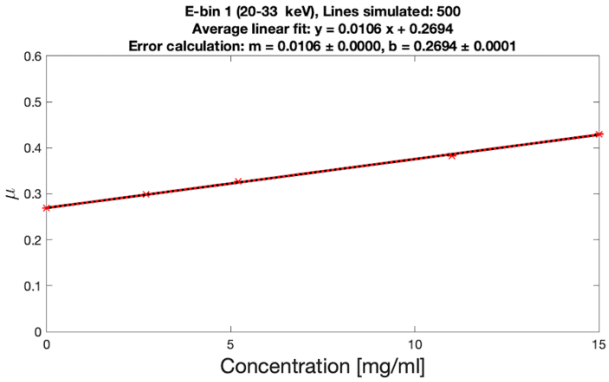


Figure 41: Sample averages and standard deviations are shown with red X's and error bars, average linear fit shown in black, and red dotted lines show the ± 1 SD of the 500 simulated lines.

3.4.4. Error Part 3: x, Concentration Solution Vector

A Monte Carlo method was again used to calculate the confidence intervals of the final calculated solution. At this point, $A, \sigma_A, b,$ and σ_b have been calculated, and by randomly sampling each entry from its respective gaussian distribution, a distribution of solutions can be computed and stored.

This is accomplished by using the command x = normrnd(A,SD_A)\normrnd(b,SD_b). Figure 42 shows the distribution of solutions stored with this method.

Confidence intervals can be plotted by approximating the distribution as a multivariate gaussian, and using the mean, covariance, and cumulative chi-squared distribution [73]. Figure 43 shows confidence intervals drawn over the solution distribution. The implementation of this given below for p = [.68, .95] (1 SD and 2 SD). In this code, confidence intervals are plotted, and the area of each confidence ellipse is calculated and stored as well.

```
%plot confidence intervals here
Cov = cov(sim_x(:, 1),sim_x(:, 2));
mu = mean(sim_x);
l_conf_int=cell(1, length(p)); %legend initialize
for idx = 1:length(p)
        %source, explanation: https://www.xarg.org/2018/04/how-to-plot-a-covariance-error-ellipse/
        s = -2 * log(1 - p(idx));
        [V, D] = eig(Cov * s); %[eigenvalues, eigenvectors]
        t = linspace(0, 2 * pi);
        a = (V * sqrt(D)) * [cos(t(:))'; sin(t(:))']; %coordinates of ellipse
        c_int_area(samp_idx, idx) = pi*D(1, 1)*D(2, 2); %area of ellipse
        l_conf_int{:, idx} = sprintf('%i%% confidence interval (area = %.0f)', p(idx)*100, c_int_area(samp_idx,
        idx)); %legend entry
        hold on;
        h_conf_int(idx) = plot(ax1, a(1, :) + mu(1), a(2, :) + mu(2), 'LineWidth', 2);
end
```

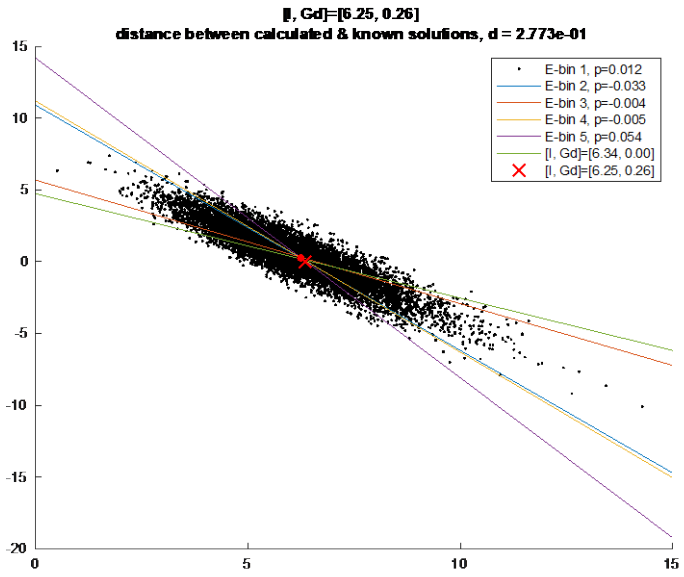


Figure 42: Black points show 500 material concentration solutions (x) to Ax=b, with pixel noise in A and b taken into account.

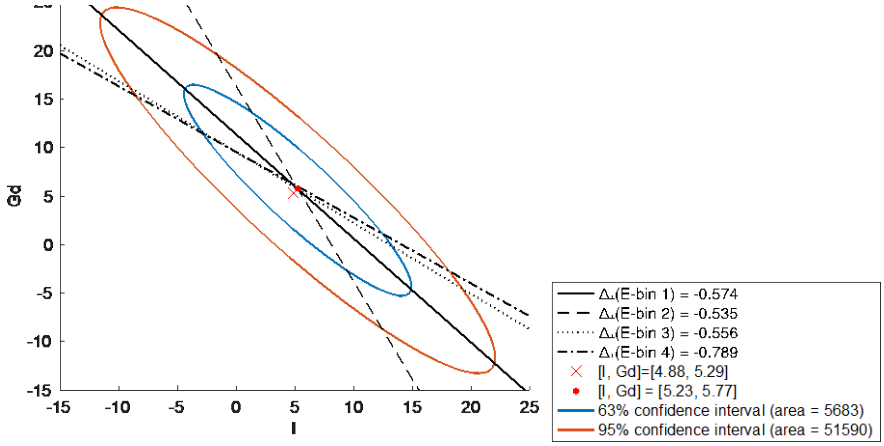


Figure 43: Confidence intervals were plotted by fitting the solution point-cloud to a multivariate gaussian. Here the 68%, and 95% confidence intervals are shown.

4. Results

Reports were generated for the nine different scans and the report metrics were analyzed. A reference table for scan names and their respective spectral protocols (energy-thresholds) are given in Table 7, and a reference table for sample numbers and their respective material contents is given in Table 8. For aid in the discussion, Figure 44 shows a histogram comparing the number of photons in each energy-bin for each scan, raw data for which is given in Table 12. A sample report from scan CL1 is included in Section 4.2. Page 1 of the report shows the calibration and test scan information and a figure summarizing the $\Delta_\perp$ results. Pages 2 - 4 show the linear system graphs and report metrics for each sample. The last page includes a figure comparing Δ for each tested sample composition.

The calculated report metrics are Δ, $\Delta_\perp$, and Confidence Interval Areas (CI Areas). Each metric was analyzed for statistically significant differences, and the GUI was run an additional 20 times for scan "CL1" to analyze the reproducibility of each measurement. Section 4.3 presents results for Δ, Section 4.4 presents results for CI Areas, and Section 4.5 presets results for $\Delta_\perp$. Implications of all results are discussed Section 5.

Table 7: The nine scans used in this research, their reference names, and their spectral protocols (energy-thresholds).

#	SCAN ID	E1 [KEV]	E2 [KEV]	E3 [KEV]	E4 [KEV]
Repeat Scans	CL1	20	33	50	80
	CL2	20	33	50	80
	CL3	20	33	50	80
	CL3_2	20	33	50	80
	CL3_3	20	33	50	80
2	CL4	18	31	48	78
3	CL5	16	29	46	76
4	CL6	22	35	52	82
5	CL7	24	37	54	84

Table 8: Reference table for sample composition.

SAMPLE	COMPOSITION
1	8.9 I + 1.1 Gd
2	7.5 I + 2.6 Gd
3	8.4 I + 7.0 Gd
4	4.9 I + 5.3 Gd
5	2.7 I + 7.4 Gd
6	1.3 I + 9.3 Gd

Table 9: Iodine concentration. Known concentration compared to calculated concentration from all scans.

I (MG/ML)	**CALCULATED IODINE CONCENTRATION**										
KNOWN	CL1	CL2	CL3	CL3-2	CL3-3	CL4	CL5	CL6	CL7	**Mean**	**SD**
8.9	9.21	9.24	9.26	9.19	9.16	9.26	9.39	9.25	9.12	**9.23**	**0.08**
7.5	7.70	7.74	7.64	7.69	7.65	7.54	7.75	7.77	7.74	**7.69**	**0.07**
8.4	7.79	7.78	7.74	7.77	7.72	7.69	7.91	7.81	7.79	**7.78**	**0.06**
4.9	5.21	5.27	5.16	5.14	5.08	5.12	5.33	5.17	5.18	**5.18**	**0.08**
2.7	2.56	2.60	2.56	2.53	2.60	2.43	2.62	2.49	2.50	**2.54**	**0.06**
1.3	1.09	0.97	1.03	1.01	1.00	0.93	1.11	1.08	1.06	**1.03**	**0.06**

Table 10: Gadolinium concentration. Known concentration compared to calculated concentration from all scans.

GD (MG/ML)	**CALCULATED GADOLINIUM CONCENTRATION**										
KNOWN	CL1	CL2	CL3	CL3-2	CL3-3	CL4	CL5	CL6	CL7	**Mean**	**SD**
1.1	1.21	1.13	1.13	1.35	1.03	1.07	1.08	1.19	1.19	**1.15**	**0.10**
2.6	3.12	3.08	2.91	3.11	2.92	3.11	3.04	3.07	3.04	**3.05**	**0.08**
7	8.40	8.62	8.33	8.60	8.52	8.40	8.58	8.47	8.39	**8.48**	**0.10**
5.3	5.78	5.70	5.53	5.61	5.39	5.62	5.47	5.71	5.69	**5.61**	**0.13**
7.4	8.40	8.45	8.56	8.51	8.34	8.68	8.41	8.39	8.51	**8.47**	**0.10**
9.3	9.78	9.82	9.81	9.91	9.90	9.92	9.77	9.74	9.81	**9.83**	**0.06**

Table 11: Average error = |(known concentration) – (average calculated concentration)| , $\%\ error = \frac{average\ error}{konwn\ concentration}$

Average error

	I		**Gd**	
	Average error [mg/ml]	% error	Average error [mg/ml]	% error
SAMPLE 1	0.33	4%	0.05	5%
SAMPLE 2	0.19	3%	0.45	17%

SAMPLE 3	0.62	7%	1.48	21%
SAMPLE 4	0.28	6%	0.31	6%
SAMPLE 5	0.16	6%	1.07	14%
SAMPLE 6	0.27	21%	0.53	6%

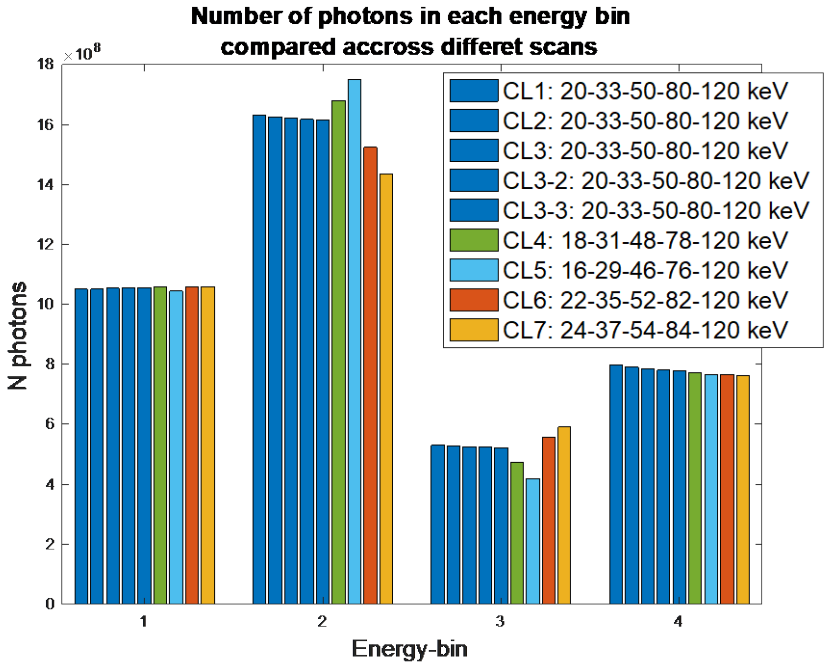


Figure 44: Histogram showing the number of photons recorded in each energy-bin in each scan.

Table 12: Number of photons recorded in each energy-bin.

	E-BIN 1	**E-BIN 2**	**E-BIN 3**	**E-BIN 4**
CL1	1050351390	1630068169	530500870	797136224
CL2	1050061143	1624558556	527428713	791110282
CL3	1053788249	1622506524	525352704	785633388
CL3-2	1053671034	1617278365	522721514	780152697
CL3-3	1054841565	1615150861	521315096	776882647
CL4	1057706993	1680538433	472395907	771907551
CL5	1044198950	1749501213	418959347	765364090
CL6	1059453084	1523426648	557116434	766122350
CL7	1059344356	1433319893	590612065	762655704

4.1. Statistical Analysis

A similar approach was used to test each metric. This involved answering the questions: Is there a difference in measurements? If there is a difference, where do the differences occur? These questions are commonly answered with an analysis of variance (ANOVA) omnibus test, followed by a multi-comparison post hoc test for significant differences.

An ANOVA tests the null hypothesis that the group means are equal, against the alternative hypothesis that there is a difference in at least one pair of group means:

H0: $\mu_1 = \mu_2 = \cdots = \mu_n$

HA: $\mu_i \neq \mu_j$, for some i, j.

Five assumptions need to be met in order to conduct a classic ANOVA [74]:

1) Continuous dependent variable ($\Delta, \Delta_\perp$, CI Area).
2) Independent variables consisting of independent, categorical groups (Scans, Samples).
3) Independent observations, e.g. data from Scan 1 is not included in Scan 4 data.
4) No significant outliers. This was assumed to be true for all data due to the small sample size available to us; we could not justify excluding any measurements.
5) Normality of the dependent variable for each combination of independent variables, tested using the Anderson-Darling test (null hypothesis: data comes from a population with a normal distribution).
6) Homogeneity of variance for each combination of independent variables, tested with Bartlett's test for homogeneity of variance.

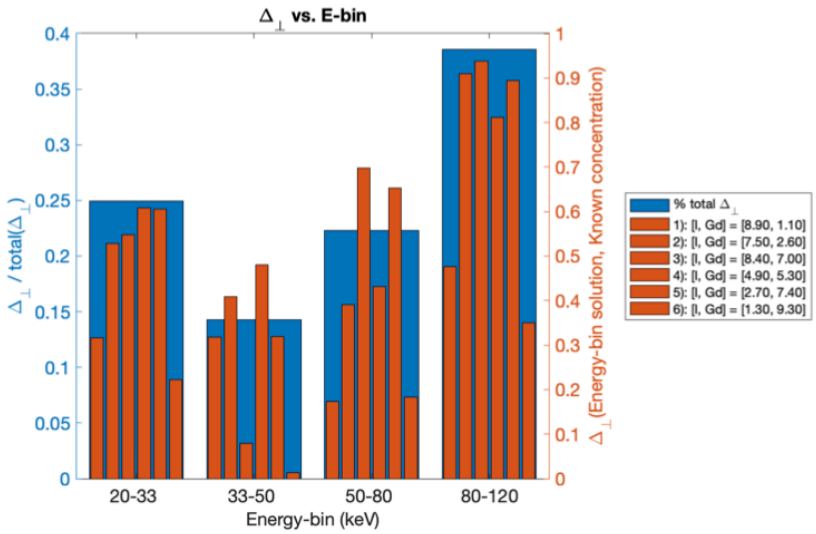
If all assumptions are met, the data was tested with a classic ANOVA and statistically significant findings were followed up with a Tukey's Honestly Significant Difference (Tukey's HSD) post hoc test.

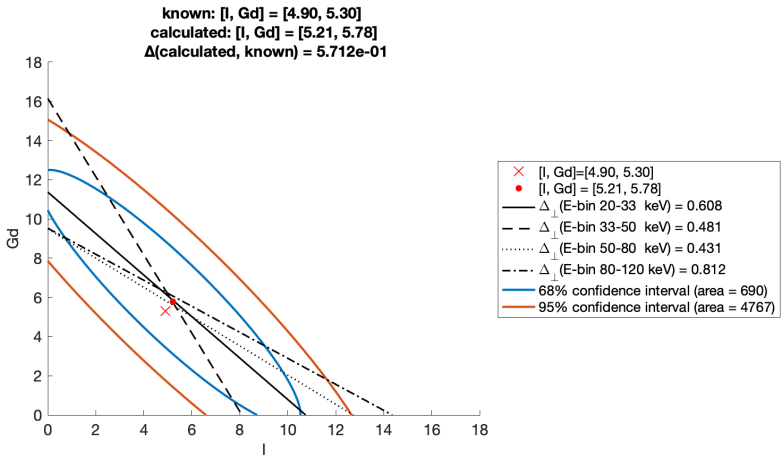
4.2. Reports

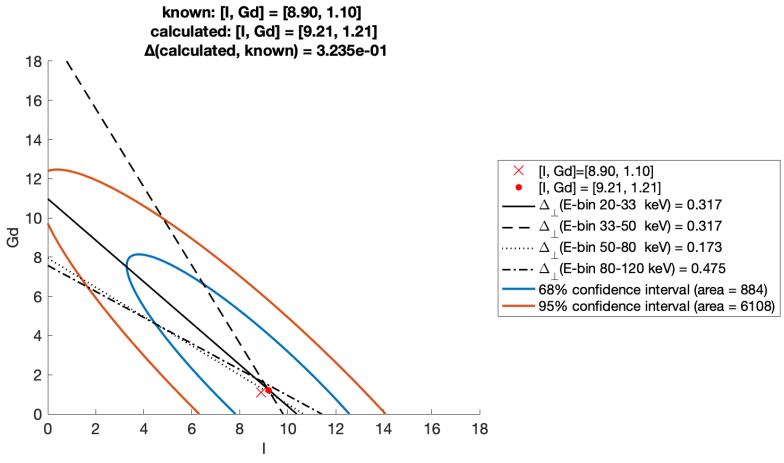
calibration: CL1 (20, 33, 50, 80, 120), sample: CL1 (20, 33, 50, 80, 120)

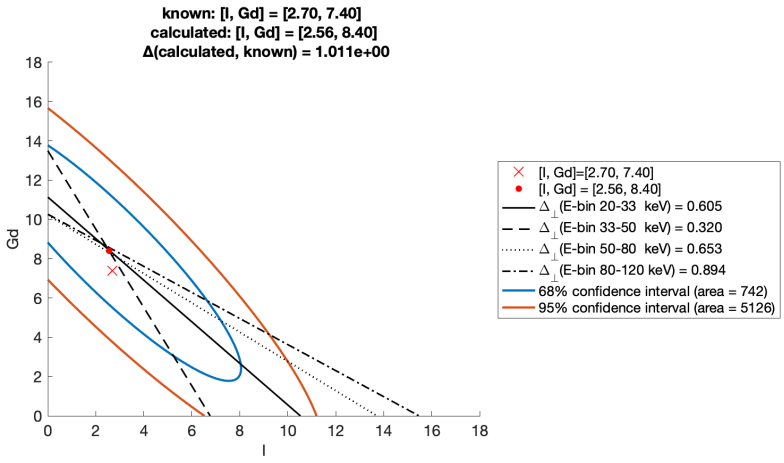
Table of Contents

Summary



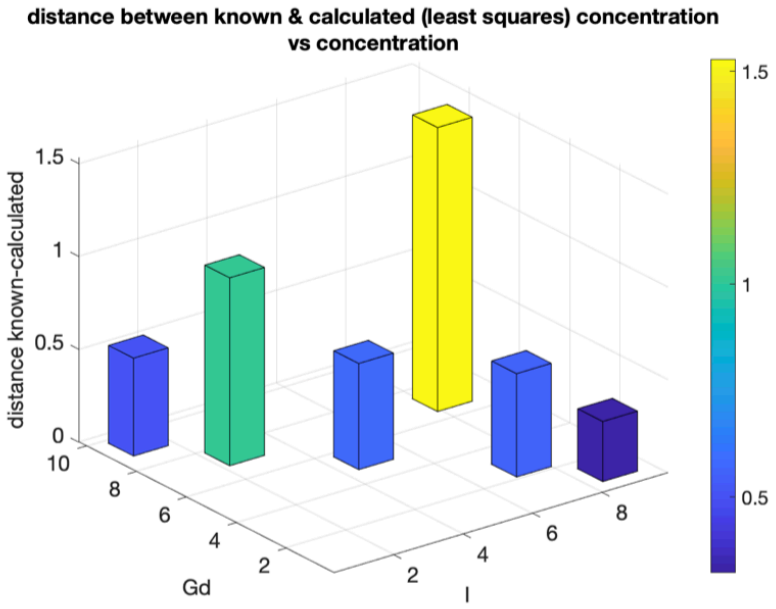

known: [I, Gd] = [4.90, 5.30]
calculated: [I, Gd] = [5.21, 5.78]
Δ(calculated, known) = 5.712e-01
Gd
I
0
2
4
6
8
10
12
14
16
18
[I, Gd]=[4.90, 5.30]
[I, Gd] = [5.21, 5.78]
Δ⊥(E-bin 20-33 keV) = 0.608
Δ⊥(E-bin 33-50 keV) = 0.481
Δ⊥(E-bin 50-80 keV) = 0.431
Δ⊥(E-bin 80-120 keV) = 0.812
68% confidence interval (area = 690)
95% confidence interval (area = 4767)


known: [I, Gd] = [8.90, 1.10]
calculated: [I, Gd] = [9.21, 1.21]
Δ(calculated, known) = 3.235e-01
Gd
I
0
2
4
6
8
10
12
14
16
18
[I, Gd]=[8.90, 1.10]
[I, Gd] = [9.21, 1.21]
Δ⊥(E-bin 20-33 keV) = 0.317
Δ⊥(E-bin 33-50 keV) = 0.317
Δ⊥(E-bin 50-80 keV) = 0.173
Δ⊥(E-bin 80-120 keV) = 0.475
68% confidence interval (area = 884)
95% confidence interval (area = 6108)


known: [I, Gd] = [2.70, 7.40]
calculated: [I, Gd] = [2.56, 8.40]
Δ(calculated, known) = 1.011e+00
Gd
I
[I, Gd]=[2.70, 7.40]
[I, Gd] = [2.56, 8.40]
Δ⊥(E-bin 20-33 keV) = 0.605
Δ⊥(E-bin 33-50 keV) = 0.320
Δ⊥(E-bin 50-80 keV) = 0.653
Δ⊥(E-bin 80-120 keV) = 0.894
68% confidence interval (area = 742)
95% confidence interval (area = 5126)

calibration: CL1 (20, 33, 50, 80, 120),
sample: CL1 (20, 33, 50, 80, 120)

Comparison across samples



Published with MATLAB® R2019a

4.3. Δ, known-calculated distance

Δ is defined as the Euclidean distance between the known concentration and the least-squares concentration solution. Δ was calculated for each can protocol and sample, and the raw data is organized in Table 13. Each scan report also includes a figure showing Δ vs. Iodine and Gadolinium concentrations to compare Δ across all tested samples, and these figures are shown below (Figure 45 - Figure 53). The axes for these figures are represent distance on a material-concentration plane, the units that would be most appropriate for the axes would be mg/ml of the basis materials.

In this analysis we wanted to test: (1) whether Δ is affected by scan protocol, (2) whether sample composition has an effect on Δ, and (3) whether the effect of protocol is different depending on the sample composition. These tests were accomplished by a two-way ANOVA. Preliminary analyses of the ANOVA assumptions are presented in Section 4.3.1 and results of the two-way ANOVA are reported in Section 4.3.2.

Table 13: Δ, for each scan and sample.

SCAN:	CL1	CL2	CL3	CL3-2	CL3-3	CL4	CL5	CL6	CL7
SAMPLE 1	0.3235	0.3396	0.3614	0.3812	0.2701	0.3638	0.487	0.3596	0.2389
SAMPLE 2	0.5563	0.5383	0.3434	0.5441	0.3493	0.5163	0.5022	0.5441	0.5031
SAMPLE 3	1.5288	1.7323	1.4871	1.7172	1.6643	1.5744	1.6512	1.5856	1.5148
SAMPLE 4	0.5712	0.5445	0.3458	0.3918	0.207	0.386	0.4602	0.4923	0.4831
SAMPLE 5	1.0114	1.0565	1.1682	1.1262	0.9419	1.3053	1.0123	1.0107	1.1236
SAMPLE 6	0.5268	0.6127	0.578	0.6764	0.6716	0.7265	0.5121	0.4969	0.5682
	---------- Repeat protocols ----------								

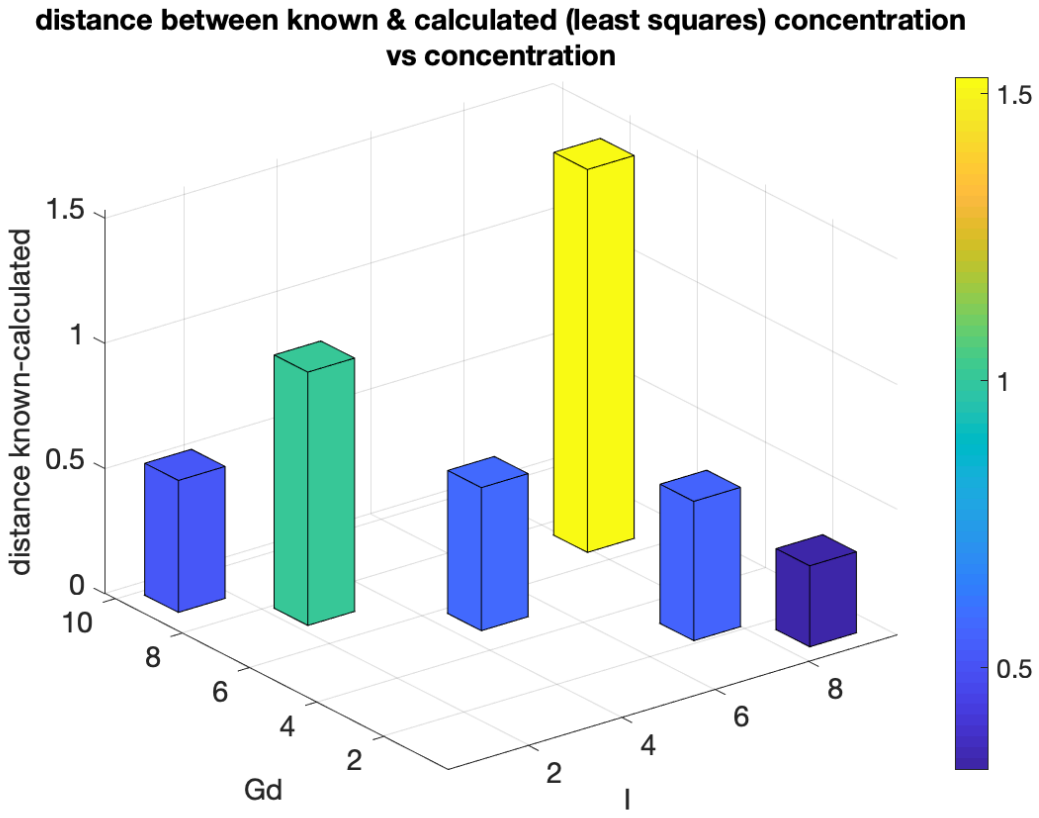


Figure 45: Δ vs I & Gd concentration for scan CL1, Axes units for these figures represent distance on a material-concentration plane, the units that would be most appropriate for the axes would be mg/ml of the basis materials.

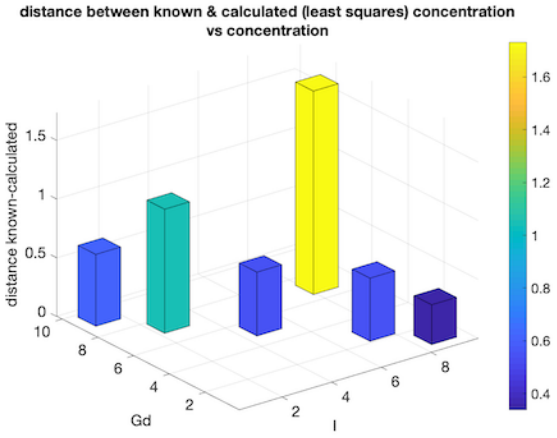


Figure 46: Δ vs. I & Gd concentration for scan CL2

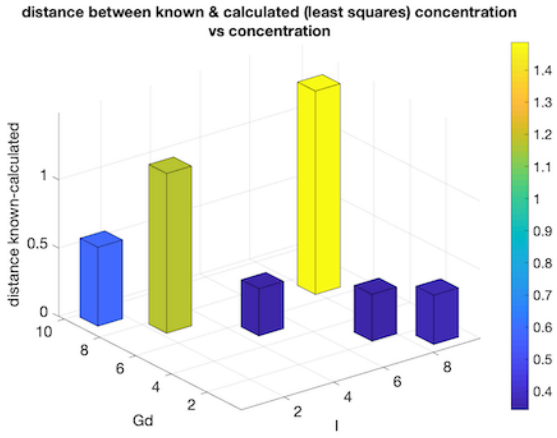


Figure 47: Δ vs. I & Gd concentration for scan CL3

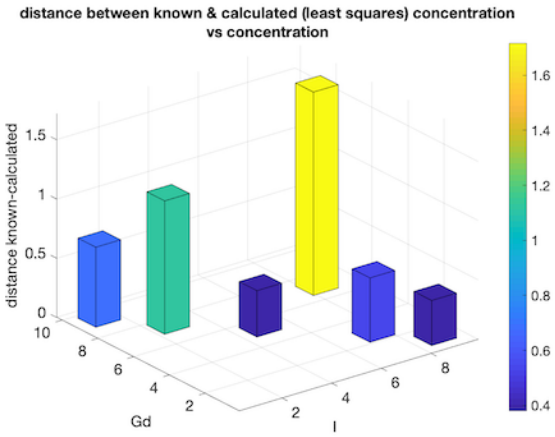


Figure 48: Δ vs. I & Gd concentration for scan CL3-2

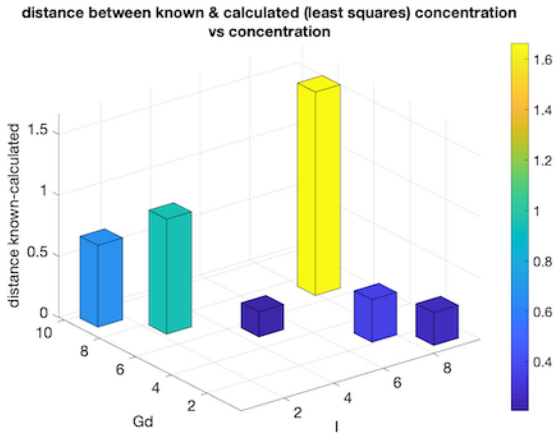


Figure 49: Δ vs. I & Gd concentration for scan CL3-3

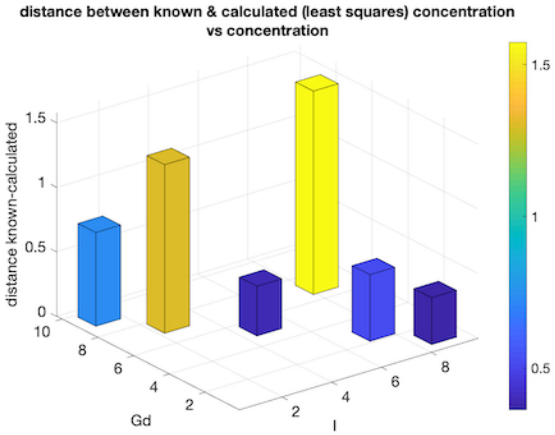


Figure 50: Δ vs. I & Gd concentration for scan CL4

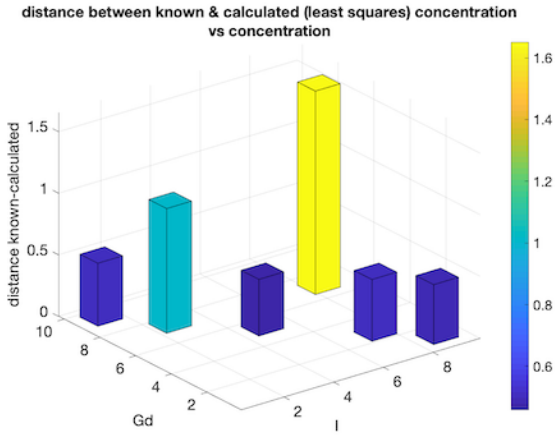


Figure 51: Δ vs. I & Gd concentration for scan CL5

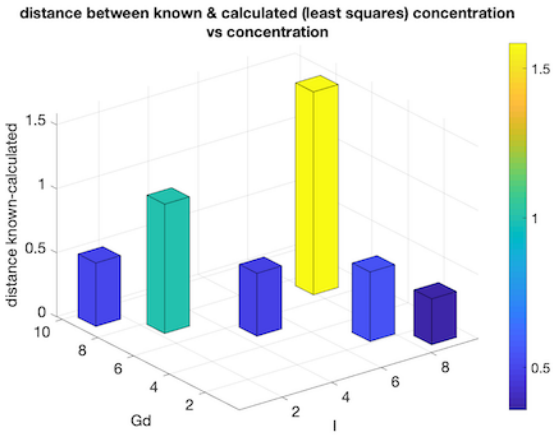


Figure 52: Δ vs. I & Gd concentration for scan CL6

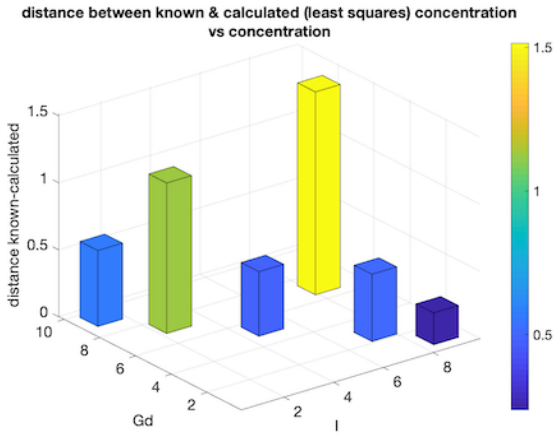


Figure 53: Δ vs. I & Gd concentration for scan CL7

4.3.1. Preliminary Analysis

4.3.1.1. Normality

The Anderson-Darling test was used to check for normality of Δ with respect to Scans, Samples and the intersection of the Repeat Protocol× Samples. Results for each are given in Table 14, Table 15, and Table 16, respectively.

Three distributions, Δ(Scan CL5), Δ(Sample 2), and Δ(Sample 2 × Repeat Protocol) have significant p-values of 0.0118, 0.0025, and 0.0283, respectively. Because most groups did not reject the null hypothesis, and the fact that an ANOVA is robust for violations of normality, we concluded that the assumption of normality is sufficient for executing a two-way ANOVA [75].

Table 14: Anderson-Darling test results for Δ vs Scan.

Δ(SCAN, ALL SAMPLES)	
Scan	p-value
CL1	0.1237
CL2	0.0956
CL3	0.0604
CL3-2	0.1741
CL3-3	0.2504
CL4	0.2013
CL5	**0.0118**
CL6	0.0568
CL7	0.1997

Table 15: Anderson-Darling test results for Δ vs Sample.

Δ(SAMPLE, ALL SCANS)	
Sample	p-value
1	0.3217
2	**0.0025**
3	0.6833
4	0.6816
5	0.3851
6	0.6515

Table 16: Anderson-Darling test results for Δ vs Repeat Protocol × Sample.

Δ(SAMPLE, REPEAT PROTOCOL)	
Sample	p-value
1	0.7964
2	**0.0283**
3	0.319
4	0.7047
5	0.9623
6	0.6609

4.3.1.2. Homogeneity of Variance

Bartlett's test was used to verify the assumption of homogeneity of variance of Δ with respect to Scans, Samples, and the intersection of the Repeat Protocol × Samples. Results in Table 17 indicate that this assumption was not violated for any group.

Table 17: Bartlett's test for homogeneity of variance: Δ(Protocol, all samples), Δ(Sample, all protocols), and Δ(Sample, Repeat Protocol).

	Δ(SCAN)	Δ(SAMPLE)	Δ(SAMPLE, REPEAT PROTOCOL)
BARTLETT'S STATISTIC	0.34022	2.55446	6.22461
DEGREES OF FREEDOM	8	5	5
P-VALUE	0.99997	0.76827	0.28497

4.3.2. ANOVA Results

Table 18 reports results of the two-way ANOVA for protocol and sample effect on Δ. There was no statistically significant interaction found between Protocol and Sample for Δ ($p = 0.7908$), which indicates that the effect of sample composition is the same for each protocol in our study. The null hypothesis was also rejected for Δ vs Protocol ($p = 0.7012$). However, there was a statistically significant result for Δ vs Sample ($p = 8.15$ E-17). Therefore, a post hoc Tukey's HSD test for Δ vs Sample was completed (Section 4.3.2.1). Box plots showing distributions of Δ for groupings of Scan and Sample are shown in Figure 54 and Figure 55.

Table 18: Two-way ANOVA for the effects of scan protocol and sample on Δ.

SOURCE	SUM SQ.	D.F.	MEAN SQ.	F	PROB>F
PROTOCOL	0.0223	4	0.00557	0.55	0.7012
SAMPLE	6.922	5	1.38441	136.52	**8.15E-17**
PROTOCOL*SAMPLE	0.1417	20	0.00708	0.7	0.7908
ERROR	0.2434	24	0.01014		
TOTAL	11.2055	53			

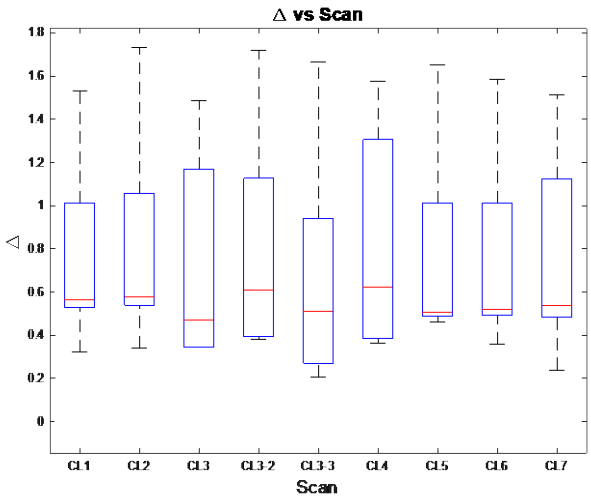


Figure 54: Box plots showing distribution of Δ vs Scan.

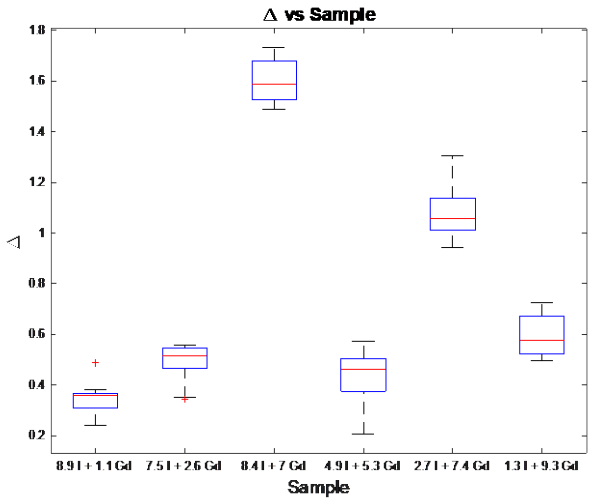


Figure 55: Box plots showing distribution of Δ vs Sample.

4.3.2.1. Samples

Tukey's HSD post hoc test (Table 19) indicated that Samples 3 and 5 had statistically significant differences in Δ from all other samples, and the pairwise difference between Sample 1 vs Sample 6 was also statistically significant. The multiple comparison graph (Figure 56) illustrates that Sample 1 has the smallest average $\overline{\Delta}$ value, and Samples 3 and 5 have statistically significantly larger $\overline{\Delta}$ than other samples, with Sample 3 having the largest $\overline{\Delta}$ measurement.

Table 19: Tukey HSD results for the main effect of sample on Δ.

SAMPLES		P-VALUE
1	2	0.14605
1	3	**2.07E-08**
1	4	0.64342
1	5	**2.07E-08**
1	6	**0.0082591**
2	3	**2.07E-08**
2	4	0.90592
2	5	**2.53E-08**
2	6	0.77229
3	4	**2.07E-08**
3	5	**2.23E-07**

3	6	**2.07E-08**
4	5	**2.14E-08**
4	6	0.21721
5	6	**8.55E-08**

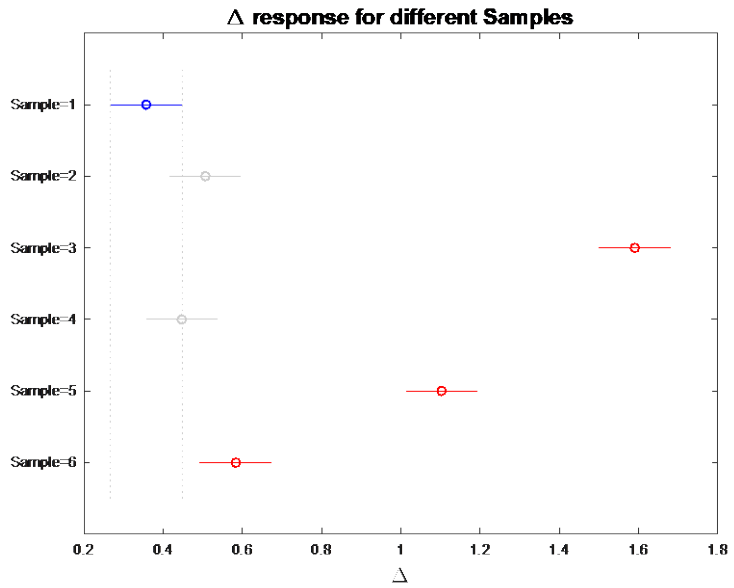


Figure 56: Multiple comparison of Δ for different samples.

4.3.3. Repeat GUI Test of Δ

Values of Δ for repeated runs of the GUI had coefficients of variation (CoV) of 4% - 11% depending on the sample, and a pooled CoV of 6%. The raw repeat data of Δ vs Sample for each analysis of Scan CL1 with the GUI is given in Table 20, and the mean, standard deviation, and CoV for each sample is summarized in Table 21. Values of Δ for all tested scans (CL1 – CL7) had coefficients of variation of 6% - 26% depending on the sample, and a pooled CoV of 12% (Table 22).

Table 20: Δ from the 20 repeated GUI runs of scan CL1.

REPEAT#	SAMPLE 1	SAMPLE 2	SAMPLE 3	SAMPLE 4	SAMPLE 5	SAMPLE 6
1	0.3676	0.5619	1.4475	0.5513	0.989	0.449
2	0.4438	0.5403	1.5612	0.5234	0.9802	0.6092

3	0.4134	0.5625	1.5823	0.5775	1.0264	0.4307
4	0.3976	0.5355	1.4734	0.5729	1.0263	0.5114
5	0.414	0.4574	1.3336	0.5116	1.0415	0.4337
6	0.3725	0.7025	1.4961	0.5925	1.0725	0.5417
7	0.386	0.4984	1.4959	0.5802	1.0381	0.4263
8	0.3559	0.4939	1.5252	0.5493	0.9948	0.4674
9	0.3478	0.557	1.548	0.5377	1.0297	0.464
10	0.4535	0.5512	1.3542	0.5132	0.9385	0.445
11	0.3996	0.5091	1.383	0.5127	1.0113	0.5312
12	0.407	0.5302	1.4702	0.4876	0.9152	0.4256
13	0.4263	0.5648	1.4715	0.4574	0.9822	0.3949
14	0.3762	0.5867	1.5211	0.6063	1.0398	0.4889
15	0.3978	0.5084	1.4328	0.5876	1.0352	0.4874
16	0.3677	0.4973	1.4767	0.5736	0.9965	0.4519
17	0.4404	0.5782	1.5316	0.616	1.0536	0.5371
18	0.4037	0.5962	1.4977	0.5563	1.0275	0.451
19	0.4186	0.5358	1.5008	0.5625	0.9718	0.4634
20	0.4175	0.5357	1.4327	0.5576	1.0035	0.4623

.

Table 21: Δ: mean, standard deviation, and coefficient of variation for repeat test of GUI.

REPEAT TEST OF GUI, SCAN CL1

SAMPLE	Count	Mean	Std	CoV
1	20	0.40034	0.02945	7%
2	20	0.54515	0.05094	9%
3	20	1.47677	0.06551	4%
4	20	0.55136	0.04056	7%
5	20	1.00869	0.03867	4%
6	20	0.47359	0.05087	11%
POOLED	120	0.74265	0.0474	6%

Table 22:Δ: mean, standard deviation, and coefficient of variation of all protocol scans.

ALL PROTOCOL SCANS

SAMPLE	Count	Mean	Std	CoV
1	9	0.34722	0.07037	20%
2	9	0.48857	0.08282	17%
3	9	1.60618	0.08914	6%
4	9	0.43133	0.11229	26%
5	9	1.08401	0.10962	10%
6	9	0.59658	0.08082	14%
POOLED	54	0.75898	0.09212	12%

4.4. Confidence Interval Area

The 68% and 95% confidence interval areas represent the area on the concentration plane that is 1 and 2 standard deviations around the mean concentration, as estimated by 500 simulations of the linear system. Data for the 68% CI Area for each sample and scan can be found in Table 23, and data for the 95% CI Area is in Table 24.

In this analysis we used a two-way ANOVA to test: (1) whether the scan protocol affects the CI Area, (2) whether the sample composition affects the CI Area, and (3) whether the effect of protocol is different depending on the sample composition. Preliminary analyses are presented in Section 4.4.1, results of the ANOVA are in Section 4.4.2, and Section 4.4.3 contains the GUI repeatability analysis for CI Areas.

Table 23: 68% confidence interval area, for each scan and sample

	CL1	CL2	CL3	CL3-2	CL3-3	CL4	CL5	CL6	CL7
SAMPLE 1	883.58	894.75	895.69	707.84	699.7	1246.1	1653.3	787.29	807.49
SAMPLE 2	847.97	936.08	840.39	951.33	902.26	1076.2	1361.4	910.99	902.96
SAMPLE 3	984.42	1241.6	1101.6	1175.4	1130.7	1459.6	1708.1	1121.2	1144.6
SAMPLE 4	689.7	694.09	853.15	943.21	880.32	1068	1241.8	784.29	821.76
SAMPLE 5	741.54	787.81	631.75	870.63	750.85	943.16	1153.1	635.42	815.67
SAMPLE 6	688.78	785.7	803.39	702.53	727.78	1009.8	1424.4	620.4	795.64

Table 24: 95% confidence interval area, for each scan and sample

	CL1	CL2	CL3	CL3-2	CL3-3	CL4	CL5	CL6	CL7
SAMPLE 1	6107.6	6184.9	6191.4	4892.8	4836.6	8613.2	11428	5442	5581.7
SAMPLE 2	5861.5	6470.5	5809.1	6576	6236.8	7439.4	9410.7	6297.1	6241.6
SAMPLE 3	6804.7	8582.3	7615	8124.5	7815.8	10089	11807	7750.2	7911.8
SAMPLE 4	4767.4	4797.8	5897.3	6519.8	6085.1	7382.1	8583.7	5421.3	5680.3
SAMPLE 5	5125.8	5445.7	4366.9	6018.1	5190.2	6519.5	7970.6	4392.2	5638.2
SAMPLE 6	4761.1	5431	5553.3	4856.2	5030.7	6980.1	9845.6	4288.5	5499.8

4.4.1. Preliminary Analysis

4.4.1.1. Normality

The Anderson-Darling test was used to check normality of CI Areas with respect to scans and samples. Results are given in Table 25.

For CI Areas vs Scan, Scan CL7 had a statistically significant value of 0.0071. For CI Areas vs Sample, four distributions had p-values below 0.05 (two of which were below 0.01). Despite these results we chose to proceed with a two-way ANOVA.

Table 25: Anderson-Darling test results for: 68%CIArea(scan), 68%CIArea(sample), 95%CIArea(scan), 95%CIArea(sample).

68% CI AREA	
Scan	p-value
CL1	0.5099
CL2	0.2395
CL3	0.4305
CL3-2	0.5022
CL3-3	0.2834
CL4	0.3067
CL5	0.7268
CL6	0.5078
CL7	**0.0071**

68% CI AREA	
Sample	p-value
1	**0.01**
2	**0.0049**
3	**0.0314**
4	0.4814
5	0.4514
6	**0.0083**

95% CI AREA	
Scan	p-value
CL1	0.5099
CL2	0.2395
CL3	0.4305
CL3-2	0.5022
CL3-3	0.2834
CL4	0.3067
CL5	0.7268
CL6	0.5078
CL7	**0.0071**

95% CI AREA	
Sample	p-value
1	**0.01**
2	**0.0049**
3	**0.0314**
4	0.4814
5	0.4514
6	**0.0083**

4.4.1.2. Homogeneity of Variance

Bartlett's test was used to verify the assumption of homogeneity of variance of CI Areas with respect to Scan, Sample and the intersection of the Repeat Protocol × Sample. Results in Table 17 indicate that this assumption was not violated.

Table 26: Bartlett's test results for: 68%CIArea(scan), 68%CIArea(sample), 95%CIArea(scan), 95%CIArea(sample).

	68%, CI(SCAN)	**68%, CI(SAMPLE)**	**95%, CI(SCAN)**	**95%, CI(SAMPLE)**
BARTLETT'S STATISTIC	2.62755	5.33454	2.62755	5.33454
DEGREES OF FREEDOM	8	5	8	5
P-VALUE	0.95552	0.37643	0.95552	0.37643

Table contd.	**68%, CI(SAMPLE X REPEAT PROTOCOL)**	**95%, CI(SAMPLE X REPEAT PROTOCOL)**
BARTLETT'S STATISTIC	4.06758	4.06758
DEGREES OF FREEDOM	5	5
P-VALUE	0.53973	0.53973

4.4.2. ANOVA Results

Table 27 and Table 28 report the two-way ANOVA results for the 68% and 95% CI Areas grouped by Protocol, Sample, and the Protocol × Sample interaction. The null hypothesis that sample means were equal was rejected for CI Area vs Protocol (p = 1.754E-12) and CI Area vs Sample (p = 6.75E-8). Tukey's HSD post hoc tests for these results are reported in Section 4.4.2.1 and Section 4.4.2.2, respectively.

Table 27: Two-way ANOVA results for 68% confidence interval areas.

SOURCE	SUM SQ.	D.F.	MEAN SQ.	F	PROB>F
PROTOCOL	1947903.7	4	486975.9	64.42	**1.754E-12**
SAMPLE	764089.4	5	152817.9	20.21	**6.75E-08**
PROTOCOL*SAMPLE	174732.2	20	8736.6	1.16	0.364
ERROR	181433.1	24	7559.7		
TOTAL	3322942.5	53			

Table 28: Two-way ANOVA results for 95% confidence interval areas.

SOURCE	SUM SQ.	D.F.	MEAN SQ.	F	PROB>F
PROTOCOL	93072700	4	23268177.1	64.42	**1.754E-12**
SAMPLE	36508900	5	7301784.6	20.21	**6.75E-08**
PROTOCOL*SAMPLE	8348870	20	417443.6	1.16	0.364
ERROR	8669050	24	361210.4		
TOTAL	158773000	53			

4.4.2.1. Protocol

Tukey's HSD post hoc test for CI Area vs Protocol for the 68% CI Area and the 95% CI Area are reported in Table 29 and Table 30, respectively. Results indicate that the scan protocols used in CL4 and CL5 resulted in statistically significantly different CI Areas than all other protocols (and each other). The multiple comparison figures (Figure 63, Figure 64) show that the average CI Area

for Scan CL5 was the largest, and that from Scan CL4 was the next largest. Box-whisker plots for 68% CI Area vs Scan and 95% CI Area vs Scan (Figure 59 and Figure 60) also show that CL4 and CL5 have larger CI Areas than the other scans.

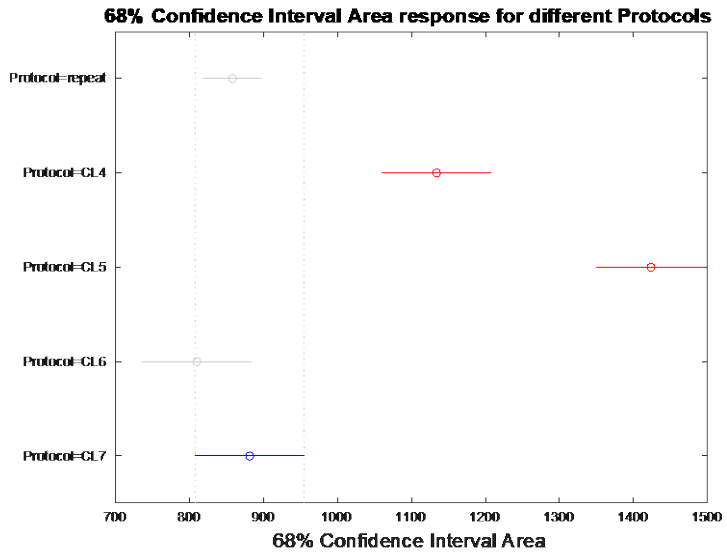


Figure 57: Multiple comparison for 68% CI Area vs Protocol

Table 29: Pairwise Tukey HSD results for 68% CI Area vs Protocol

PROTOCOLS		P-VALUE
1	2	**2.35E-06**
1	3	**9.94E-09**
1	4	0.73
1	5	0.97
2	3	**5.42E-05**
2	4	**1.05E-05**
2	5	**3.42E-04**
3	4	**1.00E-08**
3	5	**1.10E-08**
4	5	0.62

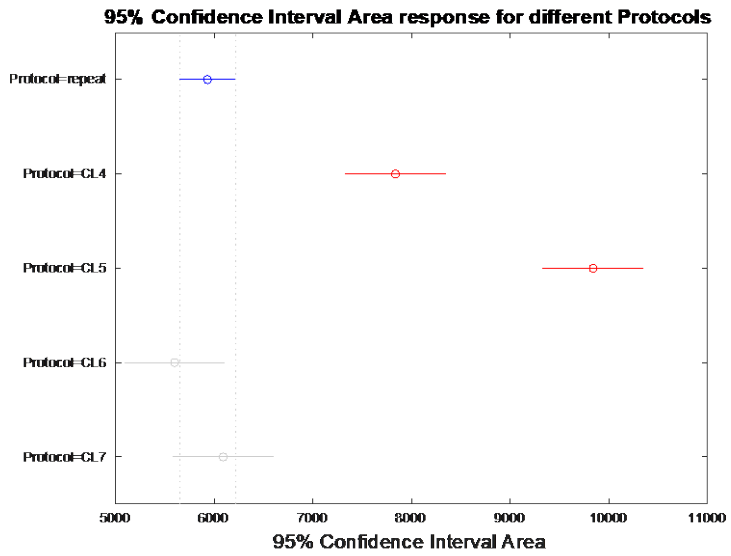


Figure 58: Multiple comparison for 95% CI Area vs Protocol

Table 30: Pairwise Tukey HSD results for 95% CI Area vs Protocol

PROTOCOLS		P-VALUE
1	2	**2.35E-06**
1	3	**9.94E-09**
1	4	0.73
1	5	0.97
2	3	**5.42E-05**
2	4	**1.05E-05**
2	5	**3.42E-04**
3	4	**1.00E-08**
3	5	**1.10E-08**
4	5	0.62

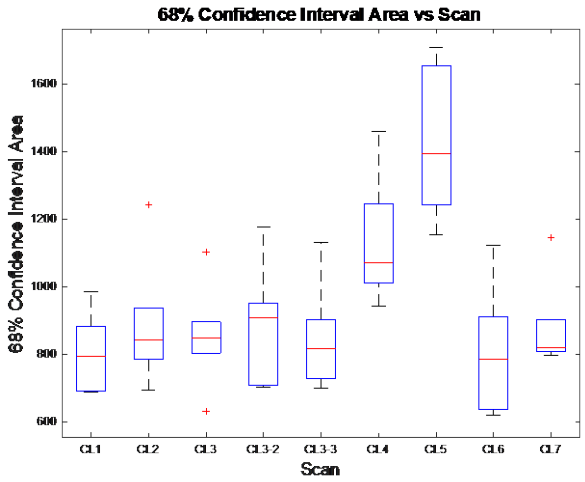


Figure 59: 68% CI Area vs Scan. Scans CL1 - CL3-3 are from the "repeat protocol".

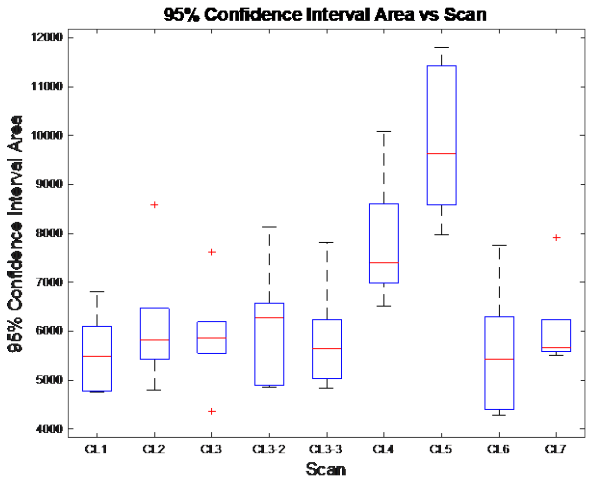


Figure 60: 95% CI Area vs Scan. Scans CL1 - CL3-3 are from the "repeat protocol".

4.4.2.2. Sample

Tukey's HSD post hoc test for CI Area vs Sample for the 68% CI Area and the 95% CI Area are reported in Table 31 and Table 32, and their respective multiple comparison figures are shown in

Figure 63 and Figure 64. Box plots for CI Area vs Sample are shown in Figure 61 and Figure 62 for better representation of the distribution of data.

Results show that Sample 3 (8.4 I + 7.0 Gd) had statistically significantly larger average CI Area than all other samples. Sample 5 (2.7 I + 7.4 Gd) had the smallest average CI Area, and this group was statistically significantly different than CI Areas from Sample 1 (8.9 I + 1.1 Gd) in addition to Sample 3.

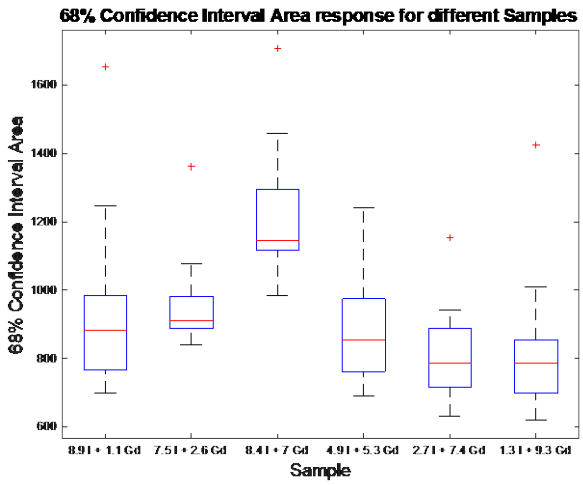


Figure 61: 68% CI Area vs Sample.

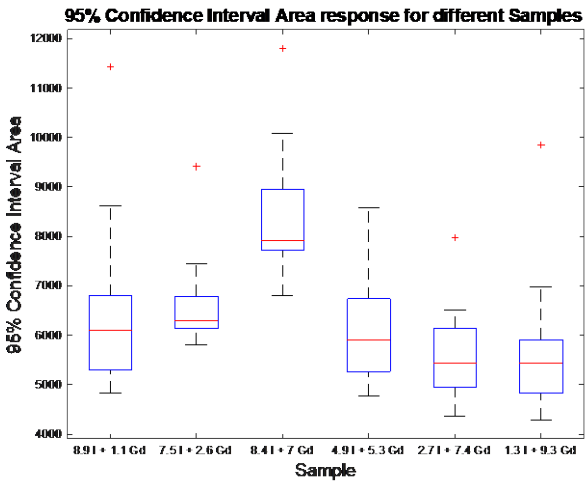


Figure 62: 68% CI Area vs Sample.

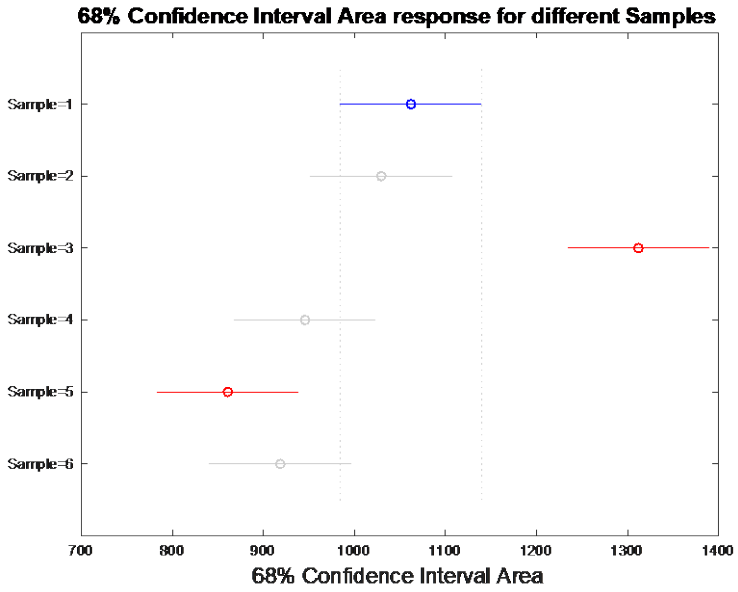


Figure 63: Multiple comparison for 68% CI Area vs Sample

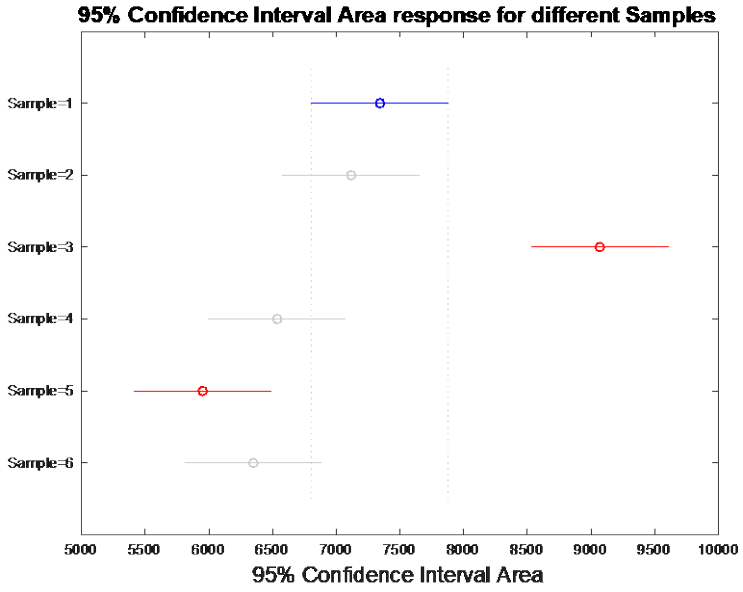


Figure 64: Multiple comparison for 95% CI Area vs Sample

Table 31: Pairwise Tukey HSD results for 68% CI Area vs Sample

SAMPLES		P-VALUE
1	2	0.99
1	3	5.89E-04
1	4	0.23
1	5	0.01
1	6	0.08
2	3	1.19E-04
2	4	0.57
2	5	0.03
2	6	0.27
3	4	2.28E-06
3	5	7.79E-08
3	6	6.89E-07
4	5	0.56
4	6	0.99
5	6	0.86

Table 32: Pairwise Tukey HSD results for 95% CI Area vs Sample

SAMPLES		P-VALUE
1	2	0.99
1	3	5.89E-04
1	4	0.23
1	5	0.01
1	6	0.08
2	3	1.19E-04
2	4	0.57
2	5	0.03
2	6	0.27
3	4	2.28E-06
3	5	7.79E-08
3	6	6.89E-07
4	5	0.56
4	6	0.99
5	6	0.86

4.4.3. Repeat Test of GUI

Raw data for the repeated GUI study is provided in Table 33, and mean, SD, and CoV of these results are reported in Table 34 and Table 35. The mean, SD, and CoV for CI Area vs Sample of all scans (CL1 – CL7) are reported in Table 36 and Table 37. The CoV for the repeat GUI test for Samples 1 – 6 was: 8%, 10%, 10%, 12%, 6%, and 10%, while the CoV of all scans for Samples 1 – 6 was: 32%, 17%, 18%, 20%, 20%, and 29%

Table 33: Confidence Interval Area vs. Sample for 20 GUI analyses of scan CL1.

	68% CONFIDENCE INTERVAL AREA						95% CONFIDENCE INTERVAL AREA					
#	1	2	3	4	5	6	1	2	3	4	5	6
1	839.70	729.14	1093.90	744.76	782.06	680.17	5804.30	5040.10	7561.50	5148.10	5405.90	4701.60
2	934.22	726.41	1147.80	820.70	737.34	698.21	6457.70	5021.20	7934.10	5673.00	5096.80	4826.30
3	1003.70	842.77	1192.50	611.83	740.32	740.07	6937.80	5825.60	8242.70	4229.20	5117.40	5115.70
4	1001.60	884.81	1083.20	752.42	722.39	618.33	6923.80	6116.10	7487.80	5201.00	4993.40	4274.10
5	854.75	674.27	1423.20	938.43	787.89	653.39	5908.40	4660.80	9837.90	6486.80	5446.20	4516.50
6	860.95	800.56	1089.50	774.64	859.41	658.67	5951.20	5533.80	7531.00	5354.60	5940.60	4553.00
7	863.23	739.88	1274.40	792.15	710.28	689.04	5967.00	5114.40	8809.20	5475.70	4909.70	4762.90
8	968.26	950.93	1359.50	669.69	726.51	671.94	6693.00	6573.20	9397.50	4629.20	5021.90	4644.70
9	888.17	828.52	1257.00	802.18	711.67	662.93	6139.30	5727.00	8688.60	5545.00	4919.30	4582.40
10	849.08	793.57	1207.20	913.00	726.33	714.48	5869.20	5485.50	8344.50	6311.00	5020.70	4938.70
11	877.00	878.65	1379.20	868.49	837.07	713.45	6062.10	6073.60	9533.40	6003.30	5786.20	4931.60
12	1025.30	663.34	1395.10	767.48	723.70	832.31	7087.50	4585.30	9643.30	5305.10	5002.50	5753.30
13	937.91	751.76	1239.90	788.84	720.99	643.15	6483.20	5196.40	8570.90	5452.80	4983.70	4445.70
14	809.80	847.42	1313.00	684.61	801.40	704.79	5597.70	5857.70	9076.00	4732.30	5539.60	4871.80
15	790.70	779.16	1122.10	897.61	660.91	764.84	5465.60	5385.90	7756.60	6204.60	4568.50	5286.90
16	881.43	844.84	1126.90	898.52	746.75	756.48	6092.80	5839.90	7789.60	6210.90	5161.80	5229.10
17	908.53	919.41	1148.60	868.00	811.70	851.74	6280.10	6355.30	7939.40	5999.90	5610.80	5887.60
18	817.80	744.81	1299.40	710.52	726.70	769.05	5652.90	5148.40	8981.90	4911.40	5023.30	5316.00
19	792.85	843.57	936.51	689.90	770.96	847.43	5480.50	5831.10	6473.50	4768.90	5329.20	5857.70
20	889.31	776.55	1134.30	703.54	747.51	872.50	6147.20	5367.80	7840.80	4863.10	5167.10	6031.10

Table 34: Repeat GUI study, 68% CI area vs Sample

68% CI AREA VS SAMPLE, REPEAT

GROUP	Count	Mean	Std	CoV
1	20	889.72	69.818	8%
2	20	801.02	77.504	10%
3	20	1211.16	126.427	10%
4	20	784.87	91.383	12%
5	20	752.6	48.108	6%
6	20	727.15	75.368	10%

Table 35: Repeat GUI study, 95% CI area vs Sample

95% CI AREA VS SAMPLE, REPEAT

GROUP	Count	Mean	Std	CoV
1	20	6150.06	482.608	8%
2	20	5536.94	535.738	10%
3	20	8372.01	873.913	10%
4	20	5425.29	631.677	12%
5	20	5202.22	332.544	6%
6	20	5026.33	520.97	10%

Table 36: 68% CI Area vs Sample: mean, standard deviation, and coefficient of variation of all scans

68% CI AREA VS SAMPLE, ALL SCANS

GROUP	Count	Mean	Std	CoV
1	9	952.85	308.352	32%
2	9	969.96	162.115	17%
3	9	1229.69	220.855	18%
4	9	886.25	178.155	20%
5	9	814.43	161.982	20%
6	9	839.82	244.728	29%

Table 37: 95% CI Area vs Sample: mean, standard deviation, and coefficient of variation of all scans

95% CI AREA VS SAMPLE, ALL SCANS

GROUP	Count	Mean	Std	CoV
1	9	6586.47	2131.45	32%
2	9	6704.74	1120.6	17%
3	9	8500.09	1526.63	18%
4	9	6126.09	1231.47	20%
5	9	5629.68	1119.68	20%
6	9	5805.14	1691.65	29%

4.5. $\Delta_{\perp}$(E-bin), known-line distance

$\Delta_{\perp}$(E-bin i) is defined as the perpendicular distance from Energy-bin-line i to the known concentration point. $\Delta_{\perp}$ was calculated for all energy-bins of each scan and sample. Raw data for $\Delta_{\perp}$ is reported in Table 38, and the summary figures of $\Delta_{\perp}$ for each scan are shown below (Figure 65 and Figure 71).

In this analysis we tested whether $\Delta_{\perp}$(E-bin) was affected by scan protocol or sample composition by using a two-way ANOVA for each energy-bin. Preliminary analyses to assess the assumptions of normality and homogeneity of variance are given in Section 4.5.1, and the results of the ANOVA are explained in Section 4.5.2.

Table 38: $\Delta_{\perp}$ for all scans

SCAN CL1	**$\Delta_{\perp}$(E1)**	**$\Delta_{\perp}$(E2)**	**$\Delta_{\perp}$(E4)**	**$\Delta_{\perp}$(E4)**
SAMPLE 1	0.31675	0.31703	0.17325	0.47548
SAMPLE 2	0.52843	0.40919	0.39016	0.90996
SAMPLE 3	0.54749	0.079053	0.69814	0.93736
SAMPLE 4	0.60788	0.48077	0.43095	0.81174
SAMPLE 5	0.60542	0.31952	0.65317	0.89416
SAMPLE 6	0.22299	0.013204	0.18348	0.34957

SCAN CL2	**$\Delta_{\perp}$(E1)**	**$\Delta_{\perp}$(E2)**	**$\Delta_{\perp}$(E4)**	**$\Delta_{\perp}$(E4)**
SAMPLE 1	0.37029	0.28266	0.081009	0.2424
SAMPLE 2	0.55778	0.41491	0.36675	0.85008
SAMPLE 3	0.69951	0.16648	0.81688	1.2413
SAMPLE 4	0.59775	0.49277	0.43095	0.6799
SAMPLE 5	0.66159	0.38171	0.71786	1.0013
SAMPLE 6	0.13424	0.062708	0.15431	0.40714

SCAN CL3	**$\Delta_{\perp}$(E1)**	**$\Delta_{\perp}$(E2)**	**$\Delta_{\perp}$(E4)**	**$\Delta_{\perp}$(E4)**
SAMPLE 1	0.30914	0.33036	0.15137	0.41834
SAMPLE 2	0.33927	0.25871	0.2255	0.60427
SAMPLE 3	0.41514	0.011308	0.61635	0.99841
SAMPLE 4	0.39047	0.3226	0.2464	0.44855
SAMPLE 5	0.69568	0.39237	0.75324	1.1737
SAMPLE 6	0.14287	0.0099438	0.21962	0.40793

SCAN CL3-2	**$\Delta_{\perp}$(E1)**	**$\Delta_{\perp}$(E2)**	**$\Delta_{\perp}$(E4)**	**$\Delta_{\perp}$(E4)**
SAMPLE 1	0.45312	0.34629	0.24857	0.48023
SAMPLE 2	0.58158	0.36619	0.36951	0.63838
SAMPLE 3	0.69236	0.11442	0.79669	1.0512
SAMPLE 4	0.44913	0.33022	0.28515	0.4797
SAMPLE 5	0.69214	0.31615	0.6836	0.92924
SAMPLE 6	0.24186	0.002112	0.24582	0.42896

SCAN CL3-3	**$\Delta_{\perp}$(E1)**	**$\Delta_{\perp}$(E2)**	**$\Delta_{\perp}$(E4)**	**$\Delta_{\perp}$(E4)**
SAMPLE 1	0.16275	0.20204	0.046631	0.22791
SAMPLE 2	0.32761	0.27158	0.25947	0.5993
SAMPLE 3	0.51928	0.06498	0.75225	1.0719
SAMPLE 4	0.20717	0.20507	0.15746	0.24528

SAMPLE 5	0.57589	0.32332	0.64757	0.80508
SAMPLE 6	0.18704	0.0055418	0.23006	0.53358

SCAN CL4	**$\Delta_{\perp}$(E1)**	**$\Delta_{\perp}$(E2)**	**$\Delta_{\perp}$(E4)**	**$\Delta_{\perp}$(E4)**
SAMPLE 1	0.3351	0.27193	0.11201	0.14696
SAMPLE 2	0.39082	0.27475	0.32011	0.71296
SAMPLE 3	0.41123	0.028019	0.62328	0.76632
SAMPLE 4	0.40975	0.33325	0.31898	0.47371
SAMPLE 5	0.6671	0.35678	0.72735	1.1144
SAMPLE 6	0.15074	0.039993	0.18578	0.44277

SCAN CL5	**$\Delta_{\perp}$(E1)**	**$\Delta_{\perp}$(E2)**	**$\Delta_{\perp}$(E4)**	**$\Delta_{\perp}$(E4)**
SAMPLE 1	0.50534	0.34054	0.15373	0.33097
SAMPLE 2	0.5537	0.40466	0.35015	0.70864
SAMPLE 3	0.66821	0.36821	0.74305	1.2783
SAMPLE 4	0.52966	0.4141	0.31437	0.43826
SAMPLE 5	0.64836	0.41128	0.62114	0.8645
SAMPLE 6	0.21799	0.04196	0.16258	0.31411

SCAN CL6	**$\Delta_{\perp}$(E1)**	**$\Delta_{\perp}$(E2)**	**$\Delta_{\perp}$(E4)**	**$\Delta_{\perp}$(E4)**
SAMPLE 1	0.32589	0.18632	0.028833	0.3441
SAMPLE 2	0.51047	0.40089	0.36781	0.79041
SAMPLE 3	0.51337	0.069961	0.74347	0.82246
SAMPLE 4	0.54719	0.39244	0.36712	0.55565
SAMPLE 5	0.60684	0.32485	0.74253	0.94567
SAMPLE 6	0.17881	0.014943	0.26356	0.32148

SCAN CL7	**$\Delta_{\perp}$(E1)**	**$\Delta_{\perp}$(E2)**	**$\Delta_{\perp}$(E4)**	**$\Delta_{\perp}$(E4)**
SAMPLE 1	0.32589	0.18632	0.028833	0.3441
SAMPLE 2	0.51047	0.40089	0.36781	0.79041
SAMPLE 3	0.51337	0.069961	0.74347	0.82246
SAMPLE 4	0.54719	0.39244	0.36712	0.55565
SAMPLE 5	0.60684	0.32485	0.74253	0.94567
SAMPLE 6	0.17881	0.014943	0.26356	0.32148

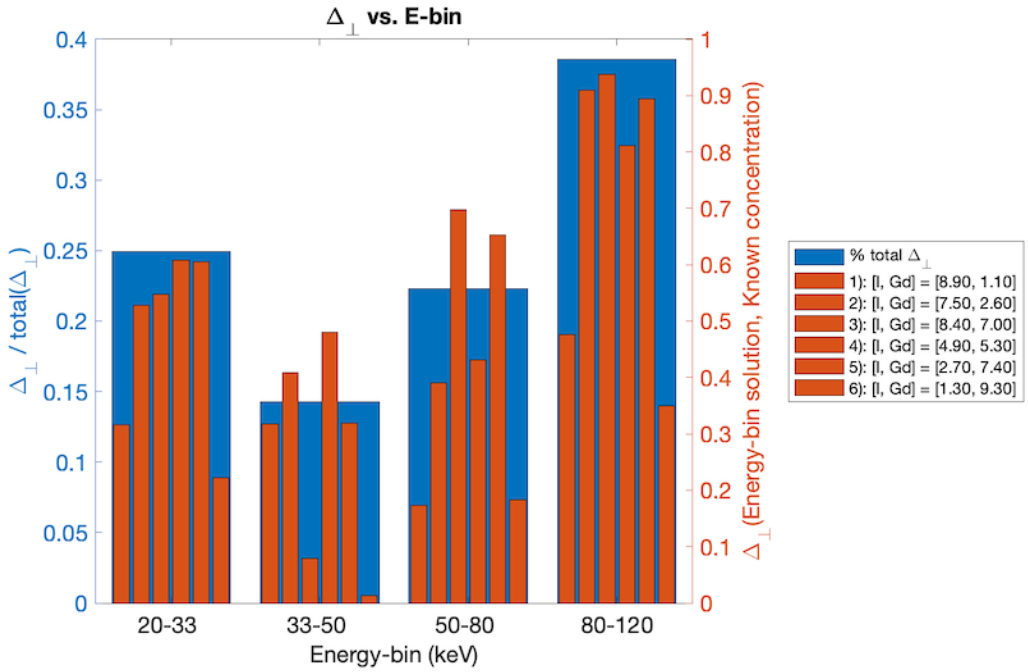


Figure 65: $\Delta_\perp$ summary plot for scan CL1

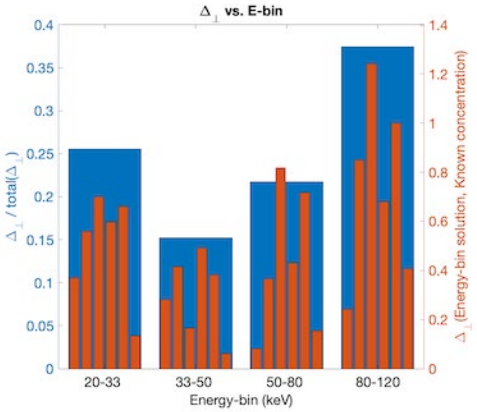


Figure 66: $\Delta_\perp$ summary plot for Scan CL2

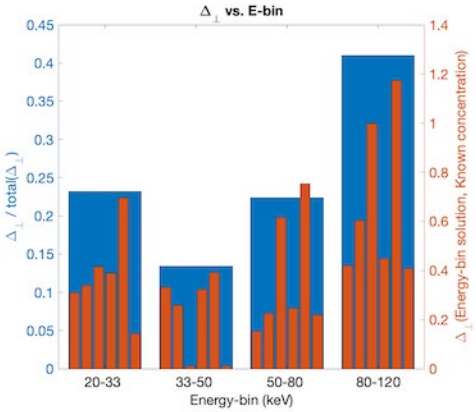


Figure 67: $\Delta_\perp$ summary plot for Scan CL3

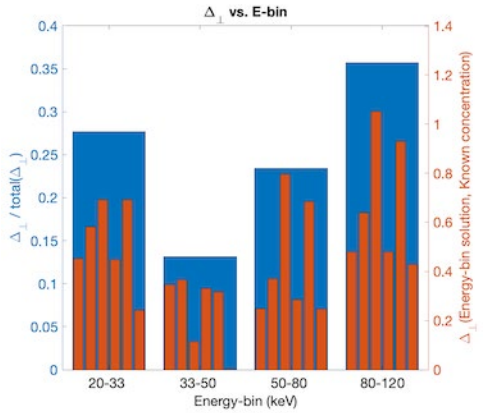


Figure 68: $\Delta_\perp$ summary plot for Scan CL3-2

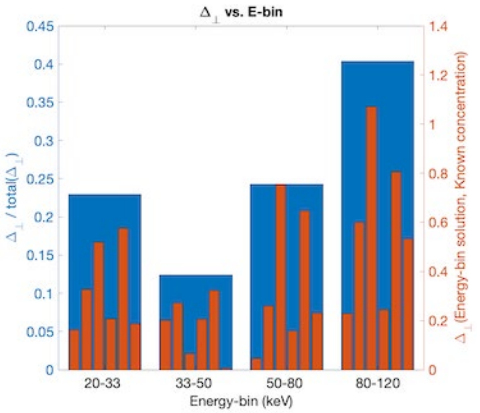


Figure 69: $\Delta_\perp$ summary plot for Scan CL3-3

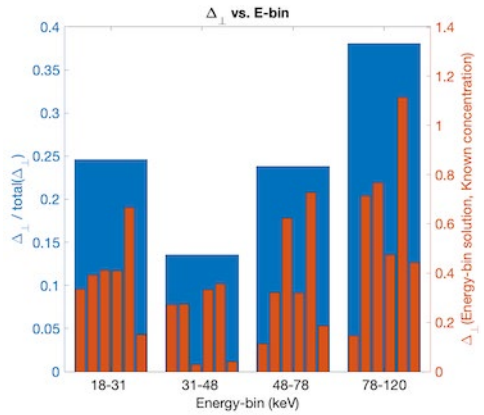


Figure 70: $\Delta_\perp$ summary plot for Scan CL4

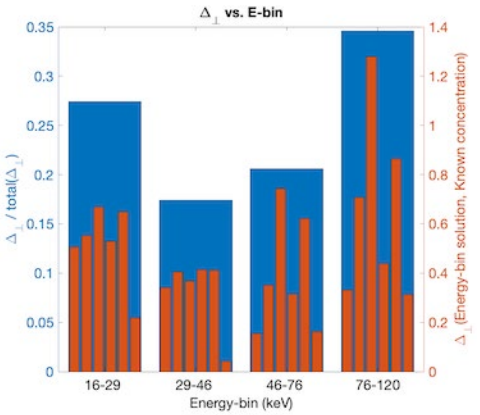


Figure 71: $\Delta_\perp$ summary plot for Scan CL5

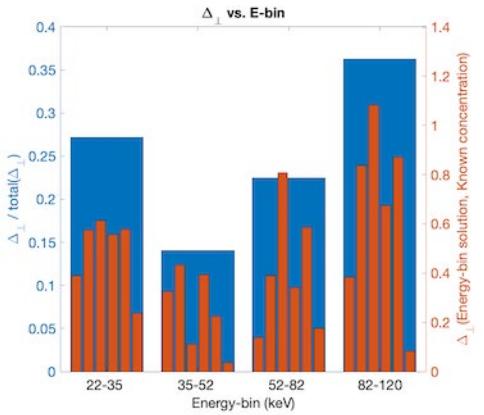


Figure 72: $\Delta_\perp$ summary plot for Scan CL6

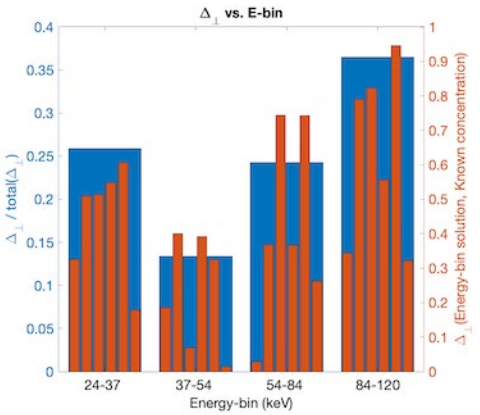


Figure 73: $\Delta_\perp$ summary plot for Scan CL7

4.5.1. Preliminary Analysis

4.5.1.1. Normality

The Anderson-Darling test was used to test normality of $\Delta_{\perp}$(E-bin i) with respect to Scan, Sample and the intersection of the Repeat Protocol × Sample. Results for each are given in Table 39, Table 40, and Table 41, respectively.

Results are summarized:

- $\Delta_{\perp}$ vs Scan: statistically significant p-values were found for scan CL3, all energy bins.
- $\Delta_{\perp}$ vs Sample: statistically significant p-values were found for Sample 2, Energy-bins 1, 2, and 3.
- $\Delta_{\perp}$ vs Repeat Protocol × Sample: no statistically significant p-values were found.

We decided to perform a two-way ANOVA even with these results because most groups had non-significant p-values, and an ANOVA is robust for violations of normality [75].

Table 39: Anderson Darling test results for $\Delta_{\perp}$(E-bin) vs. Scan

NORMALITY TEST DELTA(EBIN) VS SCAN FOR ALL SAMPLES				
SCAN	E1	E2	E3	E4
CL1	0.47205	0.47205	0.47205	0.47205
CL2	0.71999	0.71999	0.71999	0.71999
CL3	**0.041282**	**0.041282**	**0.041282**	**0.041282**
CL3-2	0.077001	0.077001	0.077001	0.077001
CL3-3	0.22325	0.22325	0.22325	0.22325
CL4	0.40202	0.40202	0.40202	0.40202
CL5	0.37969	0.37969	0.37969	0.37969
CL6	0.71278	0.71278	0.71278	0.71278
CL7	0.36042	0.36042	0.36042	0.36042

Table 40: Anderson Darling test results for $\Delta_\perp$(E-bin) vs. Sample

NORMALITY TEST $\Delta_\perp$(EBIN) VS SAMPLE FOR ALL SCANS

SAMPLE	E1	E2	E3	E4
1	0.46797	0.087263	0.84832	0.80356
2	**0.032156**	**0.019463**	**0.041723**	0.59613
3	0.47623	0.026594	0.26561	0.83256
4	0.25664	0.5976	0.79338	0.5751
5	0.39688	0.31314	0.72031	0.35514
6	0.42602	0.24171	0.53513	0.14313

Table 41: Anderson Darling test results for $\Delta_\perp$(E-bin) vs. Protocol 1 × Sample

NORMALITY TEST $\Delta_\perp$(EBIN) VS SAMPLE FOR REPEAT SCANS (1-5)

SAMPLE	E1	E2	E3	E4
1	0.67254	0.29199	0.89617	0.12405
2	0.10489	0.21956	0.14083	0.10247
3	0.4912	0.96108	0.74099	0.5382
4	0.58108	0.41852	0.44123	0.86946
5	0.40917	0.059735	0.66368	0.74961
6	0.60455	0.012703	0.68394	0.26544

4.5.1.2. Homogeneity of Variance

Bartlett's test was used to verify the assumption of homogeneity of variance of $\Delta_\perp$(E-bin i) with respect to Scans, Sample and the intersection of the Repeat Protocol × Sample. Results for each are given in Table 42, Table 43, and Table 44, respectively. Two statistically significant results were found for $\Delta_\perp$(E1) vs Sample 1 ($p = 0.0167$), and $\Delta_\perp$(E2) vs Sample 2 ($p = 0.004$). We believe that these significant p-values may be due to low sample size and do not indicate a pattern of violating homogeneity of variance, so a two-way ANOVA was used to test for homogeneity of $\Delta_\perp$ vs. Protocol, Sample and the Protocol × Sample interaction.

Table 42: Bartlett's test results for $\Delta_\perp$(E-bin) vs. Scan

HOMOGENEITY OF VARIANCE FOR $\Delta_\perp$(E) VS SCAN DISTRIBUTIONS

	E1	E2	E3	E4
BARTLETT'S STATISTIC	0.86292	1.01825	0.62837	1.78316
DEGREES OF FREEDOM	8	8	8	8
P-VALUE	0.99897	0.99813	0.99968	0.98695

Table 43: Bartlett's test results for $\Delta_\perp$(E-bin) vs. Sample

HOMOGENEITY OF VARIANCE FOR $\Delta_\perp$(E) VS SAMPLE DISTRIBUTIONS

	E1	E2	E3	E4
BARTLETT'S STATISTIC	13.8317	17.2702	5.23653	2.79752
DEGREES OF FREEDOM	5	5	5	5
P-VALUE	**0.0167**	**0.004**	0.3877	0.73117

Table 44: Bartlett's test results for $\Delta_\perp$(E-bin) vs. Protocol 1 × Sample

HOMOGENEITY OF VARIANCE FOR $\Delta_\perp$(E) VS SAMPLE DISTRIBUTIONS (REPEAT SCANS)

	E1	E2	E3	E4
BARTLETT'S STATISTIC	7.43842	10.0763	6.0392	4.92923
DEGREES OF FREEDOM	5	5	5	5
P-VALUE	0.19002	0.0731	0.30242	0.42458

4.5.2. ANOVA Results

Two-way ANOVA results for the protocol and sample effect on $\Delta_\perp$ are reported for Energy-bins 1 - 4 in Table 45 - Table 48. There was no statistically significant Protocol × Sample interactions found for $\Delta_\perp$ of any energy-bin, which indicates that the effect of sample composition is the same for each scan. There were also no statistically significant results for $\Delta_\perp$ vs Protocol. ANOVA results were statistically significant for $\Delta_\perp$ vs Sample of every energy-bin, with p-values of 9.28E-06, 8.06E-09, 6.12E-13, and 4.06E-09 for Bins 1-4, respectively. Relevant Tukey's HSD post hoc multiple comparison tests for these significant results are presented in Section 4.5.2.1.

Table 45: 2-way ANOVA results for $\Delta_{\perp}$(E-bin 1)

ENERGY-BIN 1

SOURCE	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
PROTOCOL	0.06124	4	0.01531	1.25	0.3174
SAMPLE	0.70806	5	0.14161	11.55	**9.28E-06**
PROTOCOL*SAMPLES	0.06503	20	0.00325	0.27	0.9981
ERROR	0.29436	24	0.01226		
TOTAL	1.56404	53			

Table 46: 2-way ANOVA results for $\Delta_{\perp}$(E-bin 2)

ENERGY-BIN 2

SOURCE	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
PROTOCOL	0.04843	4	0.01211	2.52	0.0673
SAMPLE	0.6042	5	0.12084	25.19	**8.06E-09**
PROTOCOL*SAMPLES	0.08584	20	0.00429	0.89	0.596
ERROR	0.11512	24	0.0048		
TOTAL	1.20519	53			

Table 47: 2-way ANOVA results for $\Delta_{\perp}$(E-bin 3)

ENERGY-BIN 3

SOURCE	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
PROTOCOL	0.00502	4	0.00125	0.21	0.9311
SAMPLE	1.86559	5	0.37312	62.07	**6.12E-13**
PROTOCOL*SAMPLES	0.05826	20	0.00291	0.48	0.9479
ERROR	0.14427	24	0.00601		
TOTAL	3.03641	53			

Table 48: 2-way ANOVA results for $\Delta_{\perp}$(E-bin 4)

ENERGY-BIN 4

SOURCE	Sum Sq.	d.f.	Mean Sq.	F	Prob>F
PROTOCOL	0.03074	4	0.00768	0.38	0.8224
SAMPLE	2.74646	5	0.54929	26.99	**4.06E-09**
PROTOCOL*SAMPLES	0.39141	20	0.01957	0.96	0.5308
ERROR	0.48843	24	0.02035		
TOTAL	4.82112	53			

4.5.2.1. Samples

Statistically significant Tukey's HSD results are reported below for $\Delta_{\perp}$ vs Sample of all energy-bins.

- Energy-bin 1: Sample 6 vs. Samples 2, 3, 4, and 5 (Sample 6 has smallest $\Delta_{\perp}$), and Samples 1 vs. 5.
- Energy-bin 2: Sample 6 vs. Samples 1, 2, 4, and 5, as well as Sample 3 vs. Samples 1, 2, 4, and 5, (Sample 6 and 3 have the smallest $\Delta_{\perp}$)
- Energy-bin 3: Statistically significantly differences were observed in 11 of 15 pairs. From the multiple comparison graph, Figure 76, Samples 1 and 6 have the smallest $\Delta_{\perp}$, and Sample 3 and 5 have the largest $\Delta_{\perp}$.
- Energy-bin 4: Statistically significant differences were observed in 8 of 15 pairs. From the multiple comparison graph, Figure 77, Samples 1 and 6 have the smallest $\Delta_{\perp}$, and Sample 3 and 5 have the largest $\Delta_{\perp}$.

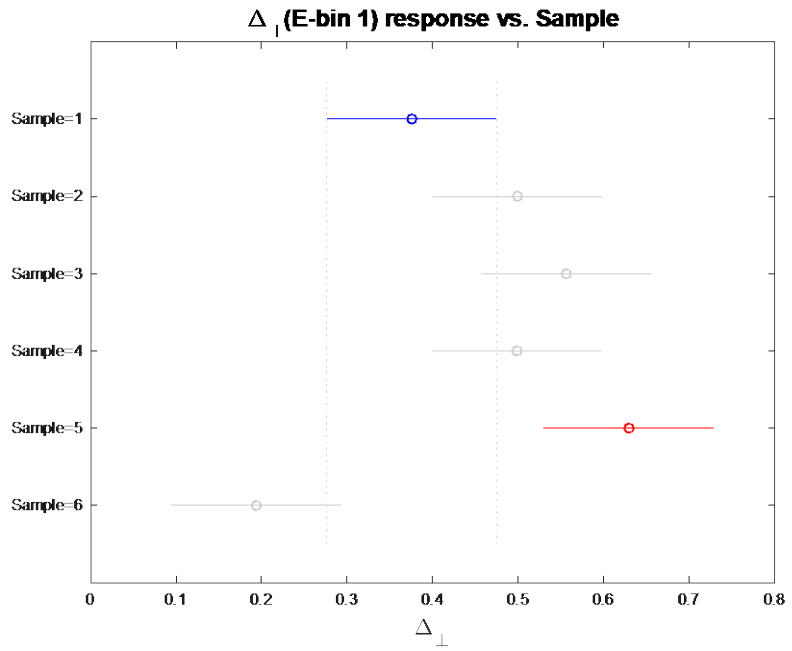


Figure 74: Multiple comparison for $\Delta_{\perp}$(E-bin 1) vs Sample.

Table 49: Tukey's HSD results for $\Delta_{\perp}$(E1) vs Sample

GROUP		P-VALUE
1	2	0.41
1	3	0.09
1	4	0.42
1	5	**0.01**
1	6	0.09
2	3	0.95
2	4	1.00
2	5	0.36
2	6	**9.73E-04**
3	4	0.94
3	5	0.86
3	6	**1.08E-04**
4	5	0.35
4	6	**9.98E-04**
5	6	**7.02E-06**

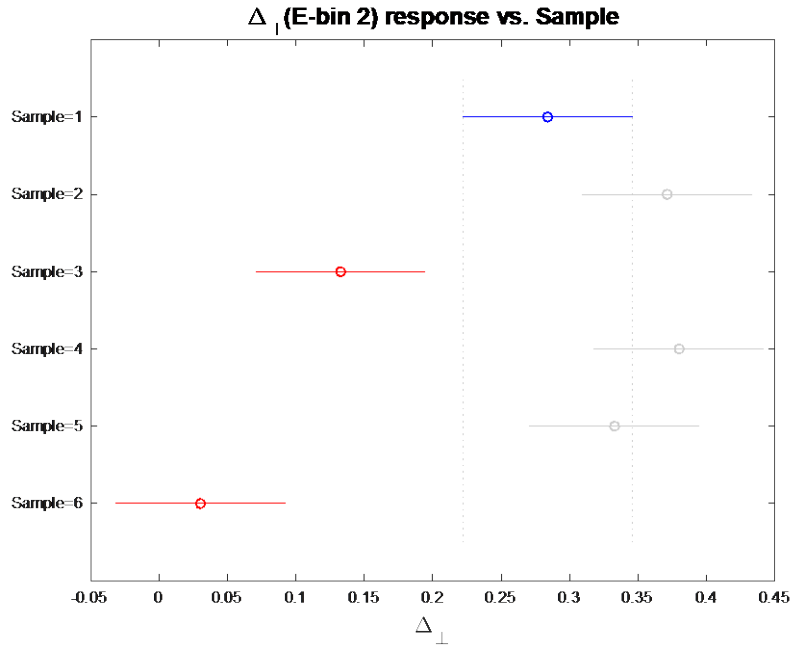


Figure 75: Multiple comparison for $\Delta_{\perp}$(E-bin 2) vs Sample.

Table 50: Tukey's HSD results for $\Delta_{\perp}$(E2) vs Sample

GROUP		P-VALUE
1	2	0.28
1	3	**0.01**
1	4	0.20
1	5	0.82
1	6	**2.12E-05**
2	3	**5.25E-05**
2	4	1.00
2	5	0.93
2	6	**1.71E-07**
3	4	**3.10E-05**
3	5	**5.53E-04**
3	6	0.15
4	5	0.84
4	6	**1.16E-07**
5	6	**1.26E-06**

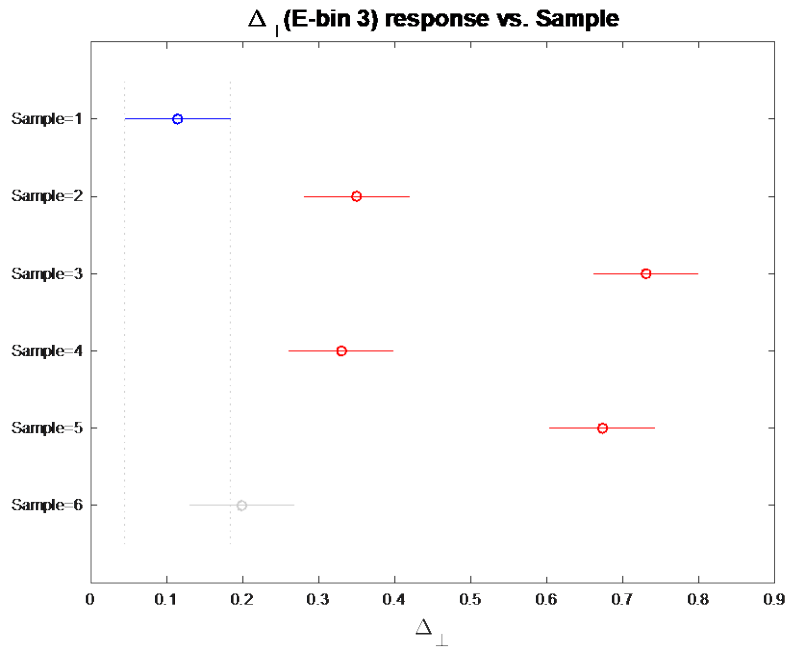


Figure 76: Multiple comparison for $\Delta_{\perp}$(E-bin3) vs Sample.

Table 51: Tukey's HSD results for $\Delta_{\perp}$(E3) vs sample

GROUP		P-VALUE
1	2	**2.97E-04**
1	3	**2.07E-08**
1	4	**8.86E-04**
1	5	**2.08E-08**
1	6	0.44
2	3	**1.79E-07**
2	4	1.00
2	5	**2.69E-06**
2	6	**0.03**
3	4	**8.28E-08**
3	5	0.80
3	6	**2.09E-08**
4	5	**9.97E-07**
4	6	0.07
5	6	**2.31E-08**

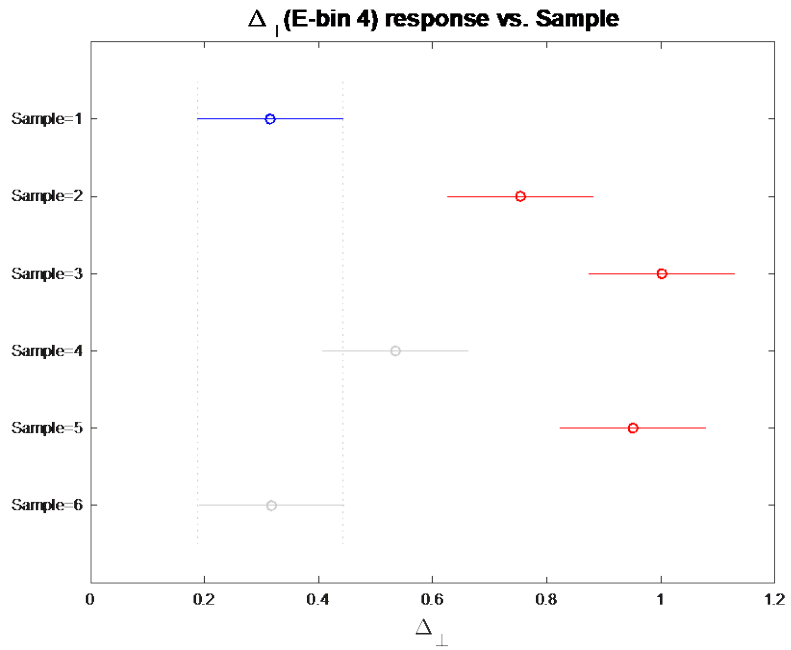


Figure 77: Multiple comparison for $\Delta_{\perp}$(E-bin 4) vs Sample.

Table 52: Tukey's HSD results for $\Delta_{\perp}$(E4) vs Sample

GROUP		P-VALUE
1	2	**2.48E-04**
1	3	**2.46E-07**
1	4	0.12
1	5	**8.87E-07**
1	6	1.00
2	3	0.06
2	4	0.12
2	5	0.20
2	6	**2.68E-04**
3	4	**1.08E-04**
3	5	0.99
3	6	**2.61E-07**
4	5	**4.88E-04**
4	6	0.13
5	6	**9.49E-07**

4.5.3. Repeat Test of GUI

Table 53 reports the mean, standard deviation, and coefficient of variation of $\Delta_\perp$(E-bin) of each sample for the GUI Repeat test (20 runs of the GUI for Scan CL1). The same data is provided for $\Delta_\perp$ of all scans in Table 54. For insight into the variation and spread of this data, the minimum, median, and maximum of mean($\Delta_\perp$), SD($\Delta_\perp$), and CoV($\Delta_\perp$) from the GUI repeat study are summarized in Table 55, and the same data is provided for all scans in Table 56.

Table 53: GUI repeat test: $\Delta_\perp$ for each sample and energy bin

$\Delta_\perp$(E-BIN) VS. SAMPLE, REPEAT GUI STUDY

	(1) 8.9 I + 1.1 GD			**(2) 7.5 I + 2.6 GD**			**(3) 8.4 I + 7 GD**		
	mean	SD	CoV	mean	SD	CoV	mean	SD	CoV
E1	0.33322	0.02486	7%	0.52206	0.02844	5%	0.52063	0.04685	9%
E2	0.361	0.02026	6%	0.40974	0.01742	4%	0.07533	0.02586	34%
E3	0.16613	0.02375	14%	0.38104	0.02819	7%	0.67229	0.03976	6%
E4	0.43679	0.07135	16%	0.90153	0.06772	8%	0.90923	0.06234	7%
	(4) 4.9 I + 5.3 GD			**(5) 2.7 I + 7.4 GD**			**(6) 1.3 I + 9.3 GD**		
	mean	SD	CoV	mean	SD	CoV	mean	SD	CoV
E1	0.58081	0.0462	8%	0.60438	0.03127	5%	0.18226	0.04093	22%
E2	0.46926	0.0184	4%	0.32303	0.01796	6%	0.02351	0.01917	82%
E3	0.41802	0.01841	4%	0.65406	0.0196	3%	0.14244	0.04272	30%
E4	0.79725	0.06289	8%	0.90029	0.05605	6%	0.28686	0.06697	23%

Table 54: All protocol scans: $\Delta_\perp$ for each sample and energy bin.

$\Delta_\perp$(E-BIN) VS. SAMPLE, ALL SCANS

	(1) 8.9 I + 1.1 GD			(2) 7.5 I + 2.6 GD			(3) 8.4 I + 7 GD		
	mean	SD	CoV	mean	SD	CoV	mean	SD	CoV
E1	0.35196	0.09703	28%	0.48487	0.1029	21%	0.56451	0.11098	20%
E2	0.28917	0.05936	21%	0.35926	0.07047	20%	0.11255	0.10658	95%
E3	0.12598	0.06781	54%	0.33765	0.05848	17%	0.73309	0.07426	10%
E4	0.33893	0.11485	34%	0.73905	0.11363	15%	1.02778	0.17081	17%
	(4) 4.9 I + 5.3 GD			**(5) 2.7 I + 7.4 GD**			**(6) 1.3 I + 9.3 GD**		
	mean	SD	CoV	mean	SD	CoV	mean	SD	CoV
E1	0.47721	0.12807	27%	0.63687	0.0463	7%	0.19029	0.04137	22%
E2	0.37395	0.08848	24%	0.339	0.0554	16%	0.02516	0.02069	82%
E3	0.32124	0.08682	27%	0.68134	0.0582	9%	0.20242	0.03855	19%
E4	0.53409	0.16767	31%	0.9555	0.12119	13%	0.36549	0.12546	34%

Table 55: Minimum, median, and maximum values of mean($\Delta_\perp$) SD($\Delta_\perp$) and CoV($\Delta_\perp$) of the repeat GUI study.

REPEAT GUI STUDY

	mean($\Delta_\perp$)	SD($\Delta_\perp$)	CoV($\Delta_\perp$)
MIN	0.02351	0.01742	3%
MEDIAN	0.427405	0.029855	7%
MAX	0.90923	0.07135	82%

Table 56: Minimum, median, and maximum values of mean($\Delta_\perp$) SD($\Delta_\perp$) and CoV($\Delta_\perp$) of all 9 scans (CL1- CL7).

ALL SCANS

	mean($\Delta_\perp$)	SD($\Delta_\perp$)	CoV($\Delta_\perp$)
MIN	0.02516	0.02069	7%
MEDIAN	0.362375	0.08765	21%
MAX	1.02778	0.17081	95%

5. Discussion

5.1. Δ, known-calculated distance

A better protocol was expected to result in more accurate material decomposition results, therefore Δ was expected to decrease for a more optimal protocol. However, results indicated that the sample composition, and not the scan protocol, affected Δ for the protocols and samples tested. This ANOVA result is illustrated by the boxplots for Δ vs Scan (Figure 54) and Δ vs Sample (Figure 55). Δ distributions for different Samples are distinct, while Δ distributions across all scans are similar. Bateman et al. states that a popular choice of protocol is to place thresholds just below the K-edge energies for the contrast materials in the scan [76]. This Δ vs Protocol result implies that this choice may be no different than placing the thresholds exactly on the K-edge energies.

5.1.1. Samples

The ANOVA and multiple comparison results for Δ vs Sample indicate that there are statistically significant differences in Δ for different sample compositions. This effect is also visibly apparent from the graphs of Δ vs iodine and gadolinium concentration, shown in Figure 45 - Figure 53. The figures and the data show that Δ is the largest for the highest combined iodine-gadolinium concentration that was tested, 8.4 mg/ml I + 7.0 mg/ml Gd. The next largest Δ value for all scans came from 2.7 mg/ml I + 7.4 mg/ml Gd. The smallest Δ value, on average, came from 8.9 mg/ml I + 1.1 mg/ml Gd, although this result was not statistically significantly different from 7.5 I + 2.6 Gd or 4.9 I + 5.3 Gd. This may mean that higher simultaneous concentration of I and Gd result in larger Δ values, and that Δ is preferentially affected by gadolinium concentrations for the protocols tested.

The trend of higher Δ value associated with higher simultaneous iodine and gadolinium concentration might be caused by two factors. (1) Small errors in calibration are magnified at higher material concentrations. (2) At high contrast-material concentrations, the displacement of water volume by contrast volume may play a role in increasing inaccuracy of the quantitative material decomposition.

(1) Small errors in slope result in larger errors in concentration at higher material concentrations, which would manifest as a larger Δ value. Calibrating to a higher maximum concentration could mitigate this effect by reducing uncertainty in the slope of the calibration curve. A study by Curtis and Roeder found that calibrating to a higher maximum concentration improved the quantitative material decomposition accuracy for gadolinium, reasoning that greater signal to noise ratio at higher concentrations resulted in a better calibration, which improved quantitative accuracy [53].

(2) As the iodine and gadolinium concentrations increase, the volume fraction of water in the voxel would necessarily decrease, therefore lowering the attenuation contribution from water. However, we assumed that the attenuation contribution from water was constant. This assumption is probably fine for small total contrast-material concentrations, especially because the calibration matrix was also created under the same assumption, but it is uncertain at which concentration this would have a significant detrimental effect. Additionally, it is possible that this error is enhanced by mixing the contrast materials together, resulting in higher total volume of contrast. In their study, Curtis and Roeder accounted for volume fraction, but they were calibrating for a two-material system of water and gadolinium [53]. The MARS MD algorithms consider volume fraction when calculating lipid plus water concentrations in soft tissue, and only account for one high-Z material (and water)

in each voxel [76]. Including a second contrast material and accounting for volume fraction would be more accurate, but would increase complexity.

5.1.2. GUI Repeatability

Although the variation in Δ was smaller on average for repeat runs of the GUI compared to scan-to-scan variation, in order to distinguish a change in Δ due to protocol from a change in Δ due to the Monte Carlo simulation in the GUI, a higher power study (more runs of each scan) or a larger change in Δ (more extreme protocol difference) is needed. This study covered limited protocol differences, and greater differences in protocols may result in a greater change in Δ. In conversation with Mayo Clinic medical physicists, they agreed with this assessment and shared their experience that energy-thresholds had to be moved by more than 10 keV before differences in protocol were apparent (Dr. Andrae Farraro and Dr. Christopher Favazza, personal communication, March 5, 2020).

It is also worth noting that Δ is small for all samples (Table 13 and Table 20), reflecting the fact that all protocols reproduced the concentration of I and Gd in the samples fairly well, and Δ is the residual sum squared. Table 9 and Table 10 report the known concentration, calculated concentration vs Scan, mean and SD of the calculated concentration for iodine and gadolinium, and a summary of average error for each sample is provided in Table 11. Average error for iodine ranged from 0.16 mg/ml (Sample 5) – 0.62 mg/ml (Sample 3), and the average error for gadolinium ranged from 0.05 mg/ml (Sample 1) – 1.48 mg/ml (Sample 3).

5.2. Confidence Interval Areas

An optimal protocol would minimize noise and maximize noncollinearity of the energy-bin-lines, resulting in smaller and rounder confidence ellipses. Therefore, it was expected that better protocols would have smaller CI Areas. The idea of using confidence region size to compare protocols was proposed by Nik et al. and tested with a simplified model in which only noise due to photon counting statistics in the sample image ($\sigma = \sqrt{N}$) was accounted for, and two energy-bins were considered [61]. This research differed from Nik et al. by using empirical data and a Monte Carlo method to draw confidence regions which included noise in the calibration matrix and the sample image and accounting for four energy-bins.

Of the protocols and samples tested, our results indicated that grouping CI Area by both sample composition and protocol resulted in statistically significant differences in confidence interval areas.

In retrospect, it was unnecessary to analyze multiple CI Areas because statistical results for both CI Areas are identical due to the calculation method: a multivariate gaussian is fitted to the distribution of simulated concentration solutions, then the covariance matrix of this gaussian is used to draw confidence ellipses. In other words, the 68% and 95% confidence ellipses are re-scalings of the same ellipse, scaled to represent different standard deviations of the data. Because of this, the analyses of the ellipse populations for groupings of Scan and Sample are the same even though the size of the ellipses are different.

Overall, CI Areas were very large for all samples and scans. This indicates that material images produced from these data sets would be extremely noisy, and the pixel-to-pixel material

concentration would be unreliable. This result is bad for diagnostic material measurement, and the source of this issue ought to be investigated.

A possible source may be Poisson noise due to low signal. Low SNR from low signal in each energy-bin would cause noisy energy-bin lines, and this in turn could cause large CI Areas. In this case, increasing the number of photons by increasing the mAs would decrease the CI Areas.

Another possible origin of the noisy result is detector inhomogeneity, in which different detector pixels to respond differently to the same signal. This is unlikely because a flatfield calibration is run before every scan to equalize pixel response. However, this issue could be ruled out by analyzing the uniformity of a flatfield image, and by examining the detector-pixel response collected in the flatfield calibration.

Beam hardening or other artifacts could also be the source of the large CI Areas. One way this could be investigated is by imaging a homogeneous phantom, then testing the images for beam hardening artifacts and image uniformity. Another interesting experiment would be to colorize each pixel in the sample areas by their concentration to see if there is a pattern of concentration nonuniformity indicative of other artifacts.

5.2.1. Protocol

The two protocols with statistically significantly larger CI Areas were CL4 (18, 31, 48, 78 keV), and CL5 (16, 29, 46, 76 keV), with CL5 having the largest CI Areas. This implies that lowering the energy-thresholds by a few keV off of the K-edges increases the uncertainty (noise) in the

calculated material concentrations, while raising the energy-thresholds by a few keV does not impact the spread of concentration values.

This effect may be caused by changing the number of photons in each energy-bin. Shown in Figure 44, Energy-bin 3 (~50 keV to ~ 80 keV) contained the lowest number of photons out of all energy-bins, and CL4 and CL5 has a visibly decreased number of photons in this bin, CL5 having the lowest number of photons. A lower number of photons would cause increased noise, and increased noise in one energy-bin may have contributed to the increased noise overall, i.e. larger CI Areas.

In this study we chose to move all energy-bins simultaneously to keep the energy-bins the same size (except, necessarily, Bin-4). The theoretical impact of noise in E-bin 3 might be mitigated by widening the energy-bin (raising the upper threshold while lowering the on-K-edge threshold). It would be interesting to study a larger variety of protocols to see if there was a connection between number of photons per energy-bin and confidence region size.

Based on these results, the -2 and -4 keV protocols exhibit suboptimal precision in comparison to the other protocols, and would therefore not be recommended. However, it appears that there is no statistically significant difference between the other 3 protocols, so any could be chosen as an optimal iodine-gadolinium protocol.

5.2.2. Sample

The sample with the highest combined gadolinium-iodine concentration (8.4 mg/ml I + 7.0 mg/ml Gd) had the largest CI Areas, on average. The rest of the samples had CI Area distributions that

were much more similar, although Sample 1 (8.9 mg/ml I + 1.1 mg/ml Gd) has a slightly larger average CI Area and Sample 5 (2.7 mg/ml I + 7.4 mg/ml Gd) had the smallest average CI Area.

As discussed in Section 5.1.1, larger CI Areas for higher combined I plus Gd concentrations could be due to increased calibration error at higher concentrations, or error due to not accounting for lower volume-fraction of water. Additionally, material decomposition for balanced concentrations of contrast agents could be less precise because of less distinct K-edges due to attenuation signal averaging; the mixed attenuation signal of I and Gd signal is perhaps less clear than a single-material signal.

5.2.3. GUI Repeatability

CI Areas were fairly repeatable for repeated runs of the GUI. The average CoV was 9% for the 20 repeated analyses of CL1, with a range of 6% - 12% for different samples. CoVs for different scans were 23% on average, with a range of 17% - 32% for different samples. Thus, CI Area may be a promising metric for differentiating between protocols for precise material decomposition.

5.3. $\Delta_{\perp}$, known-line distance

In an ideal linear system, all energy-bin-lines would intersect at the known concentration point, so a smaller $\Delta_{\perp}$(E-bin) was expected to indicate a better protocol. Inaccuracy in each energy-bin, represented by $\Delta_{\perp}$, was thought to arise mainly from noise in the system. Therefore, we sought to analyze the effect of binning by comparing $\Delta_{\perp}$ from each energy-bin for each protocol. However, results indicated that differences in energy-binning did not affect $\Delta_{\perp}$(E-bin). As seen with Δ, differences in sample composition, and not the scan protocol, affected $\Delta_{\perp}$ for the protocols and samples tested.

5.3.1. Samples

All energy-bins had statistically significant differences in $\Delta_\perp$ between samples. Sample 6 (1.3 I + 9.3 Gd) stood out in the multiple comparison figures (Figure 74 - Figure 77) as being consistently close to 0 for all scans. Sample 1 (8.9 mg/ml I + 1.1 mg/ml Gd) had the smallest average $\Delta_\perp$ for Energy-bins 3 and 4. This showed an apparent trend that the smallest $\overline{\Delta_\perp}$ values came from samples that were primarily one material – Sample 6 had the smallest $\overline{\Delta_\perp}(E1)$ and $\overline{\Delta_\perp}(E2)$, and Sample 1 (8.9 I + 1.1 Gd) had the smallest $\overline{\Delta_\perp}(E3)$ and $\overline{\Delta_\perp}(E4)$.

This trend shows that energy-bin-lines may be more accurate when the signal from only one material is strong. This finding is similar to the result in Section 5.2.2 (largest confidence intervals associated with high, balanced contrast concentrations), and gives more evidence that a more definite K-edge signal results in more precise linear system, while precision is lost for balanced concentrations due to attenuation signal averaging.

This may also be a factor of having a limited number of energy-bins. An ideal K-edge imaging system would have distinct pre-K-edge and post-K-edge energy-bins. In our case, the pre-K-edge bin for gadolinium is the post-K-edge bin for iodine. Based on these results, a better protocol may result from pairing iodine with a material with a higher K-edge than gadolinium, and using a 5-bin system with distinct pre- and post- K-edge bins for each material.

A recent publication by Ostadhossein et al. reported successful material decomposition of Hafnium (Hf), I, lipid, water, and air with a protocol of 30 – 45 – 65 –80 – 118 keV. This protocol sacrificed the iodine pre-K-edge bin in order to gain distinct bins for post-K-edge I, pre-K-edge Hf, and post-

K-edge Hf [77]. Differentiating multiple high-Z materials represents a significant improvement in material decomposition algorithms and is promising for multi-contrast imaging in the future. Additionally, Hf nanoparticles are promising for safe human use and disease targeting, and have a K-edge energy of 65.35 keV, which is low enough for CT imaging and high enough to allow for ideal K-edge energy-binning when imaged with iodine.

Interestingly, the smallest $\Delta_{\perp}$ values for every sample came from Energy-bin 2, and five out of six of these came from Sample 6 (1.3 I + 9.3 Gd). Energy-bin 2 also consistently had the highest photon count out of all bins. Many researchers have connected increased SNR to better results, and this may be another indicator that increased SNR correlates with a more accurate material signal. However, although Sample 6 had the smallest average $\Delta_{\perp}$ value and the smallest average CI Area, this did not correlate with the smallest $\overline{\Delta}$ (associated with Sample 1), or the most accurate concentration solutions (Iodine; Sample 5, Gadolinium; Sample 1). This indicates that a single accurate energy-bin does not result in an overall more accurate protocol, instead the system as a whole should be considered.

Variable energy-binning with PCCT creates the opportunity for use of new contrast materials or simultaneous use of multiple contrast agents. Novel contrast agents for PCCT are a current topic of research because contrast agents besides I, Ba, and Ga need to be studied and approved for future clinical use. In the future, it may even be possible to have a contrast material library from which to choose a contrast agent with a favorable K-edge based on the PCCT protocol. Also, combining a contrast chemical with a targeted agent would open up new possibilities in functional CT imaging similar to nuclear medicine and MRI. One issue that must be overcome for any PCCT

contrast imaging application is sensitivity. CT in general is much less sensitive than MRI or Nuclear Medicine, but some sensitivity improvements could be made with more advanced image processing algorithms, and even with lower sensitivity the new contrast imaging possibilities in PCCT may find an important role in clinical imaging.

5.3.2. GUI Repeatability

We compared the standard deviation and coefficient of variation for each sample and energy-bin of the 20 GUI repeat runs to the same data calculated for the 9 scans.

23 out of 24 values of SD($\Delta_\perp$) of the GUI repeatability test were smaller than SD($\Delta_\perp$) of all scans, and 22 out of 24 values of CoV($\Delta_\perp$) of the GUI repeatability test were smaller than CoV($\Delta_\perp$) of all scans. The minimum, median, and maximum SD($\Delta_\perp$) and CoV($\Delta_\perp$) for the repeatability test were smaller than those for all scans, indicating that the GUI results for $\Delta_\perp$ are repeatable and should be able to be distinguished between different protocols, given different enough protocols and high enough sample size.

5.4. Future Works

This GUI provides insight into the behavior of the linear system, and describes a method for assessing the accuracy and precision of a photon-counting CT system with empirical data. This method is promising for future studies and future clinical use. The following section describes many ways that this research could be continued. Sub-sections are organized by general topic: expand this study (Section 5.4.1), improve analysis GUI (Section 5.4.2), extend analysis (Section 5.4.3), new study ideas (Section 5.4.4), and clinical implementation (Section 5.4.5).

5.4.1. Expand This Study

Expanding this study by testing a greater variety of protocols and a larger sample size (more repeated scans for each protocol) would likely result in finding differences in Δ and $\Delta_\perp$ between protocols.

Although the ANOVA results for Δ and $\Delta_\perp$ failed to reject the null hypothesis for homogeneity of means across protocols, this does not mean that Δ and $\Delta_\perp$ are not affected by protocol differences. Our study encompassed a limited range of protocols and limited sample size. Testing a wider variety of protocols would give a better understanding of the response of these metrics for change in protocol. It would be interesting to study how much an energy threshold has to be moved in order to see a statistically significant difference in Δ and $\Delta_\perp$. Both more extreme protocol differences and a higher power study would help to investigate this effect.

The GUI repeatability tests showed that the spread in data due to repeated runs of the GUI was smaller on average than the spread in data across different protocols. This indicates that a higher power test (more repeated scans) would improve the model's sensitivity to changes in $\overline{\Delta}$, and $\overline{\Delta_\perp}$. Under the direction of a statistician, we repeated one protocol five times in order to get an estimate for the expected spread in data. Ideally, we would have repeated scans for every protocol, but devoting valuable scan time and the time of our colleagues to repeat studies is not practical, especially when material decomposition results are good even without extreme protocol optimization (Δ is small for all scans). If one wanted to improve the power of this study, I would recommend seeking the advice of a statistician for additional guidance on balancing statistical power with minimal necessary scans.

5.4.2. Improve Analysis GUI

The analysis GUI has the potential to continuously be improved and added to. Three interesting ideas for improvement include: increased automation, generalizing the system to include any number of materials, and adding previously overlooked analysis details.

5.4.2.1. Increase Automation

Currently, ROI selection is not automated which makes the GUI flexible for working with different phantom placement, numbers of vials, and vial sizes. However, when the same phantom is used in the same position within the scanner, analysis time could be saved by saving an ROI file to re-use for repeat analysis of the same phantom.

Another possible place to save time is loading the images. This currently represents the largest time spent idle at the computer (approximately 1 minute). Instead, if the user was performing the same analysis on multiple scans, the image loading could be consolidated. The user would select all the files to be analyzed, then load them all at once. With the addition of a ROI file, the user would have minimal computer interaction: once to load the images, and once to select ROIs and enter material contents. The same analysis could then be performed automatically on all files.

5.4.2.2. Generalize Analysis to Higher Dimensions

The GUI currently works for water plus one - two additional materials, allowing the energy-bins to be pictured as lines in concentration space. Visualization of the linear system was one of the goals of this analysis, but it becomes less helpful in 3D (three materials with water subtracted, or water plus two materials), and visualization is impossible with more materials. However, the report

metrics (Δ, $\Delta_{\perp}$, and CI Area) can still be calculated. Therefore, this analysis could be generalized to higher dimensions by leaving out the linear system visualization.

5.4.2.3. Add Analysis Details

Two details that could be incorporated in the analysis are: material concentration uncertainty, and volume fraction. We did not incorporate uncertainty in ICP-MS measurements in the Monte Carlo solution for the calibration matrix and corresponding calibration error matrix, but this detail could be added. As instructed by Dr. Ralle, Director of the USR Elemental Analysis Core at OHSU, we reported the element concentration to one significant figure, and in the future, we could use the CoV from the ICP-MS report to include concentration error in the Monte Carlo simulation. We also assumed that the volume of contrast materials was not large enough to impact the attenuation of water in each voxel. This assumption could be corrected to incorporate volume fraction in the calibration.

5.4.2.4. Increase the Number of Energy-bins

We used 4 energy-bins in this research, but in the future photon-counting detectors with many more energy-bins will be available. PCCT with near-continuous energy bins would significantly improve the energy-resolution, but signal in each bin would be an issue, as would the amount of data to be analyzed and stored. Assuming these issues could be overcome, an increased number of bins would allow an increased number of materials to be identified. Also, retrospective binning could be applied, which would negate the need for energy-threshold optimization. This would also increase the complexity of image processing, but results would probably be more accurate and precise for material concentration maps.

5.4.3. Extend Analysis

It would be possible to extend the analysis to include additional factors, such as calibrating to higher concentration, analyzing the number of photons per bin, and adding an eccentricity factor to the confidence ellipse analysis.

5.4.3.1. Calibrating to a Higher Concentration

Sections 5.1.1 and 5.2.2, suggested that calibrating to a higher concentration would decrease calibration uncertainty and decrease Δ and CI Areas for higher-concentration samples. A study could be done to test the effect of max calibration-material concentration on accuracy and precision of the linear system. It might also be helpful to use this information to develop a "rule of thumb" for calibration, i.e., calibrate to 3× the max expected concentration in your scan. A general rule of thumb for calibration that is based on empirical data would be useful for clinical implementation of photon-counting CT.

5.4.3.2. Include Photons Per Bin

Signal-to-noise ratio of the energy images are directly dependent on the number of photons captured in each energy-bin. Many protocol optimization methods use SNR-based figures of merit because the number of photons in each energy bin is likely to be important for material quantification. For example, in Section 5.2.1, we discussed a possible connection between increased CI area and lower photons in an energy-bin. A module could be built into the GUI to investigate correlation between number of photons and other report metrics.

5.4.3.3. Include Eccentricity Factor

In addition to analyzing only one confidence ellipse, it may be interesting to add a measurement of eccentricity to describe the overall collinearity of energy-bin lines. Eccentricity close to zero would be expected for perpendicular energy-bin lines, and eccentricity would approach 1 as the lines become more colinear. A very long ellipse would indicate that a protocol is poor for distinguishing between two materials. The 68% confidence ellipse for Sample 2, pictured in the sample report (Section 4.2) shows that the ellipse is stretched enough to intersect with both axes, intersecting with the Gd = 0 axis at about ~10 mg/ml I, and intersecting with the I = 0 axis at ~12 mg/ml Gd. This represents a fairly noisy and unreliable material image. If it were possible to make the confidence ellipse more circular (lower eccentricity) the material images should be less noisy. This may be possible in the future with detectors with higher energy resolution, and very narrow energy-bins combined with image processing for low-count reconstruction.

5.4.4. New Study Ideas

This section discusses the possible application of the linear-system analysis GUI developed in this research to a few other research projects.

5.4.4.1. K-edge Imaging, Compare I + Gd with I + Hf

Section 4.5.2.1 highlighted a possible problem with K-edge imaging of iodine in combination with gadolinium. The $\Delta_{\perp}$ results suggested that a K-edge blending issue may occur when one bin contains both the post-K-edge signal of one material and the pre-K-edge signal of another material, effectively averaging out distinguishing markers for mixed materials. Imaging of iodine in combination with hafnium appears to be a promising alternative, and Ostadhossein et al. demonstrated successful differentiation of these elements with a four-bin protocol of 30 – 45 – 65

– 80 – 118 keV [77]. It would be interesting to compare our iodine-gadolinium results to results of iodine plus hafnium, and also to compare Dr. Ostadhossein's four-bin protocol with a protocol using the extended five-bin capability of the MARS photon-counting CT.

5.4.4.2. Phantom Thickness Study

Clinical CT protocols are optimized for different-sized patient populations, and this tool could be used to study the energy-image linear system of how photon-counting CT changes with subject diameter. In a paper titled, "Robustness of Optimal Energy Thresholds in Photon Counting CT," Zheng et al looked at the effect of changing different parameters, like phantom thickness, on SNR^2 and $SDNR^2$ in computer models of photon-counting detectors [5]. It would be interesting to compare their results for optimal protocol with results using this empirical-analysis. This information could be used to create guidelines for photon-counting protocol adjustment for different patient populations.

5.4.4.3. Calibration Drift Study

In addition to studying the empirical impact of calibrating to a higher max concentration, it would be interesting to gather calibration data across a long period of time and look for calibration drift. It would also be clinically useful to develop a numerical estimation for calibration error projected to linear-system error. Additionally, studying calibration drift over time could be used to develop clinical guidelines for calibration acceptance and a timeline for re-calibration. Because calibration is time-consuming, it is clinically useful to determine a sufficient re-calibration frequency (daily vs. weekly vs. longer) to balance time spent calibrating with accuracy and precision of material decomposition results.

5.4.4.4. Response Surfaces

In conversation with an OHSU staff biostatistician, Jack Wiedrick, we discussed the possibility of using response surface analysis to use these metrics as a way to truly optimize protocols (not just to compare protocols and point to the best). For example, Δ could be descrbed as a 5-dimensional response surface, $\Delta(e1, e2, e3, e4, e5)$, where $e1$ = distance from 0 keV to the 1st threshold, $e2$ = the distance between the 1st and 2nd thresholds, and so on. Describing energy-threshold placement this way makes it impossible to describe a situation where the n^{th} threshold is lower than the $n - m$ threshold. Then the 5D Δ surface could be sampled across the protocol hyperplane, and a global search for minimums would result in the optimal protocol. This method is not practical due to the number of scans needed to complete a global search, but it is an interesting idea. For the interested reader, the book recommended to me on this subject was, "Response Surfaces, Mixtures, and Ridge Analysis", by George E.P. Box, and Norman R. Draper [78].

5.4.5. Clinical Implementation

5.4.5.1. Quality Control Tool

In the future, I imagine that this analysis GUI could be a useful clinical tool for protocol quality control (QC). Currently, CT technologists perform daily, weekly, and monthly CT QC, and physicists perform yearly CT QC. This research presents an analysis GUI which assesses empirical accuracy of the linear system as a whole (Δ), accuracy of each energy-bin ($\Delta_{\perp}$) and the precision of PCCT material decomposition (CI Area). A tool like this (especially with increased automation) could be paired with a dedicated spectral phantom and used to check and track tolerance levels of a clinical PCCT scanner.

This tool may also be useful for adjusting protocols for different-sized patients or different materials. The relative insensitivity of a protocol to 4 keV adjustments in energy-thresholds is promising for future clinical implementation. This means that protocols should be robust to changes in X-ray spectrum within a scan, and it is known that the X-ray spectrum changes with position in a CT scanner and angle of projection through a patient. Protocols could be adjusted for different-sized patients (infant, child, adult, bariatric) by scanning phantoms of different sizes and using this tool to check the efficacy of the linear system while adjusting protocols. This same method could be used to develop and check protocols for different materials.

Protocol optimization is a moving target in PCCT. I do not think you could expect the same accuracy and precision for different combinations of contrast material, patient thickness, machine, and protocol. In a clinical setting, it would be useful to have a tool that could be used to assure the efficacy of a scan when one of these parameters is changed. It would also be beneficial to have a measurement of material decomposition precision when interpreting images of material concentration, although this depends heavily on the material decomposition algorithm used.

5.4.5.2. Machine Learning use in PCCT

Machine learning is popular in any field that requires analyzing large amounts of data. At a basic level, machine learning is a type of data analysis in which computer algorithms are trained on large data sets, and the results improve with an increased amount of training data. This has been applied to image processing and analysis, and applying ML to PCCT image analysis is an attractive option. Some natural PCCT applications include: image noise reduction, increased accuracy and precision of material identification, and material identification of similarly attenuating materials (e.g.

different types of calcium, indicative of malignant vs benign breast calcifications). The challenge for these and any machine learning task is acquiring a large enough training set, but it is inevitable that machine learning will be applied to PCCT.

6. Summary and Conclusions

The main goal of this research was to develop an empirical method for analyzing the linear system of PCCT material-basis decomposition for use in protocol optimization. We aimed to create a user-friendly method, and test it on iodine-gadolinium contrast mixtures for the end goal of determining the effect of energy-threshold placement on linear-system accuracy and precision, and reaching a conclusion on an optimized protocol. These goals were accomplished with the creation of a protocol optimization GUI which was used to compare five protocols. A short explanation of the work accomplished for this project and notable results are described below.

A GUI was made using MATLAB's App Designer which allowed images to be loaded, viewed, and processed, in one platform. The GUI leads the user through the analysis process, all computations are performed in the background, and a report is produced which summarizes the results.

Analysis metrics were designed to give a measurement of linear system accuracy and precision, and accuracy of each energy-bin. Accuracy of the linear system was measured with Δ, the residual sum of squares of material concentration, equivalently described as the distance between the known and calculated sample concentration.

$$\Delta = \sqrt{(I_{known} - I_{calc})^2 + (Gd_{known} - Gd_{calc})^2}$$

Accuracy of each energy-bin was measured with $\Delta_{\perp}$(E-bin i), the distance between the known concentration point and Energy-bin-line i. Precision of the linear system was measured with the 68% and 95% confidence interval areas, measured by fitting 500 simulated solutions of the linear system with a multivariate gaussian then calculating the areas of 1 SD and 2 SD ellipses. It was

expected that a better protocol would produce a more accurate result (smaller Δ and $\Delta_\perp$), and would be less noisy (smaller CI areas).

The linear system analysis GUI was used to test five near-K-edge protocols to gain insight on placing energy-thresholds on the K-edge (a common result in theoretical protocol optimization research) vs placing energy-thresholds a few keV off of the K-edge (a popular protocol choice in practice). A phantom containing iodine and gadolinium calibration vials and iodine plus gadolinium sample mixtures was scanned using five near-K-edge protocols:

1. on-K-edge protocol: [20 – 33 – 50 – 80 – 118 keV], repeated 5×
2. -2 keV protocol: [18 – 31 – 48 – 78 – 118 keV]
3. $-$ 4 keV protocol: [22 – 35 – 52 – 82 – 118 keV]
4. $+2$ keV protocol: [16 – 29 – 46 – 76 – 118 keV]
5. $+4$ keV protocol: [24 – 37 – 54 – 84 – 118 keV]

From this study, we found that the only statistically significant difference between protocols resulted from CI Areas of the -2 and -4 keV protocols, which were larger than those of all other protocols. The increased size of CI Areas of these protocols was likely due to a reduction in the number of photons in Energy-bin 3. Bin 3 already had the lowest photon count out of all energy-bins for all protocols, and the further reduced count may have increased noise for this energy-bin, resulting in increased noise for the overall linear system and appearing as increased confidence interval size.

Additionally, differences in sample concentration produced statistically significant differences for all metrics (Δ, $\Delta_\perp$, and CI Areas). Notably, the largest Δ measurement came from Sample 3 (8.4 $\frac{mg\ I}{ml}$ + 7.0 $\frac{mg\ Gd}{ml}$) and second-largest from Sample 5 (2.7 $\frac{mg\ I}{ml}$ + 7.4 Gd). CI Areas were largest for

Sample 3 (8.4 $\frac{mg\ I}{ml}$ + 7.0 $\frac{mg\ Gd}{ml}$). Finally, $\Delta_{\perp}$ for Bins 1 and 2 were smallest for Sample 6 (1.3 $\frac{mg\ I}{ml}$ + 9.3 $\frac{mg\ Gd}{ml}$), and $\Delta_{\perp}$ for Bins 3 and 4 were smallest for Sample 1 (8.9 $\frac{mg\ I}{ml}$ + 1.1 $\frac{mg\ Gd}{ml}$).

These results suggest:

(1) Small errors in calibration are magnified at higher material concentrations which cause larger inaccuracy for higher concentrations. This may possibly be improved by calibrating to a higher maximum concentration.

(2) At high contrast-material concentrations, the displacement of water volume by contrast volume may play a role in increasing inaccuracy of the quantitative material decomposition, again causing increased inaccuracy at higher concentrations. This effect may be mitigated by including volume fraction in the calibration and in the material decomposition.

(3) It may not be ideal for one energy-bin to include both the post-K-edge region of one material (i.e. iodine) and the pre-K-edge region of another material (i.e. gadolinium). This may make the K-edge signal less distinct, especially for higher mixed concentrations. Using contrast materials with more separation between K-edges would allow for distinct pre- and post-K-edge bins for all scanned contrast materials, and may result in a more accurate and precise linear system.

In conclusion, linear-system accuracy proved to be relatively insensitive to energy-threshold adjustment of $\leq$ 4keV, but system precision was slightly worse for protocols with the lowest minimum energy-bin photon count. Therefore, placing thresholds on-K-edge vs off-K-edge likely makes no difference to accuracy, and the best overall result would likely arise from placing them near the K-edge, with slight adjustments to capture as many photons as possible in the bin with the least photon intensity to reduce overall system noise. As photon-counting detector systems

improve, the benefit of placing thresholds directly at the K-edge energy may change, and protocols will need to be checked an optimized with any change to the PCCT imaging chain.

Final Thoughts:

We presented a PCCT protocol analysis method, wrapped in a user-friendly GUI, which showed that protocol results were relatively insensitive to threshold adjustment of $\leq$ 4keV for the materials and system tested. This apparent insensitivity is promising for clinical implementation, because protocols need to be robust for intra-scan changes in X-ray beam intensity as the beam is projected through different angles around a patient.

There are many protocol optimization methods, and this research presents a new way of directly analyzing the material-basis linear system with empirical data. The use of empirical data and analysis automation makes this a promising tool for future clinical use for PCCT QC, testing theoretical protocol optimization results, and adjusting clinical protocols for different patient sizes or contrast materials. PCCT imaging is well on its way to clinical implementation, and evidence-based protocol optimization is necessary to deliver the optimum benefit of this technology.

7. Citations

[1] P. Sahbaee, E. Abadi, W. P. Segars, D. Marin, R. C. Nelson, and E. Samei, "The Effect of Contrast Material on Radiation Dose at CT: Part II. A Systematic Evaluation across 58 Patient Models," *Radiology*, vol. 283, no. 3, pp. 749–757, Mar. 2017, doi: 10.1148/radiol.2017152852.

[2] C. H. McCollough, S. Leng, L. Yu, and J. G. Fletcher, "Dual- and Multi-Energy CT: Principles, Technical Approaches, and Clinical Applications," *Radiology*, vol. 276, no. 3, pp. 637–653, Sep. 2015, doi: 10.1148/radiol.2015142631.

[3] B. M. Yeh *et al.*, "Opportunities for new CT contrast agents to maximize the diagnostic potential of emerging spectral CT technologies," *Adv. Drug Deliv. Rev.*, vol. 113, pp. 201–222, Apr. 2017, doi: 10.1016/j.addr.2016.09.001.

[4] E. Roessl and C. Herrmann, "Cramér–Rao lower bound of basis image noise in multiple-energy x-ray imaging," *Phys. Med. Biol.*, vol. 54, no. 5, p. 1307, 2009, doi: 10.1088/0031-9155/54/5/014.

[5] Y. Zheng *et al.*, "Robustness of Optimal Energy Thresholds in Photon-counting Spectral CT," *ArXiv180200965 Phys.*, Feb. 2018, Accessed: Oct. 19, 2018. [Online]. Available: http://arxiv.org/abs/1802.00965.

[6] R. K. Panta *et al.*, "Element-specific spectral imaging of multiple contrast agents: a phantom study," *J. Instrum.*, vol. 13, no. 02, p. T02001, 2018, doi: 10.1088/1748-0221/13/02/T02001.

[7] A. Raja *et al.*, "Measuring Identification and Quantification Errors in Spectral CT Material Decomposition," *Appl. Sci.*, vol. 8, no. 3, p. 467, Mar. 2018, doi: 10.3390/app8030467.

[8] "The Nobel Prize in Physiology or Medicine 1979," *NobelPrize.org*. https://www.nobelprize.org/prizes/medicine/1979/summary/ (accessed Jan. 08, 2020).

[9] L. W. Goldman, "Principles of CT and CT Technology," *J. Nucl. Med. Technol.*, vol. 35, no. 3, pp. 115–128, Sep. 2007, doi: 10.2967/jnmt.107.042978.

[10] M. Mahesh, "The AAPM/RSNA Physics Tutorial for Residents," *RadioGraphics*, vol. 22, no. 4, pp. 949–962, Jul. 2002, doi: 10.1148/radiographics.22.4.g02jl14949.

[11] J. T. Bushberg, Ed., *The essential physics of medical imaging*, Kindle edition. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2012.

[12] S. Ehn *et al.*, "Basis material decomposition in spectral CT using a semi-empirical, polychromatic adaption of the Beer–Lambert model," *Phys. Med. Biol.*, vol. 62, no. 1, p. N1, 2017, doi: 10.1088/1361-6560/aa4e5c.

[13] J. F. Williamson, S. Li, S. Devic, B. R. Whiting, and F. A. Lerma, "On two-parameter models of photon cross sections: Application to dual-energy CT imaging," *Med. Phys.*, vol. 33, no. 11, pp. 4115–4129, 2006, doi: 10.1118/1.2349688.

[14] J. A. Seibert and J. M. Boone, "X-Ray Imaging Physics for Nuclear Medicine Technologists. Part 2: X-Ray Interactions and Image Formation," *J. Nucl. Med. Technol.*, vol. 33, no. 1, pp. 3–18, Mar. 2005.

[15] H. Lusic and M. W. Grinstaff, "X-Ray Computed Tomography Contrast Agents," *Chem. Rev.*, vol. 113, no. 3, Mar. 2013, doi: 10.1021/cr200358s.

[16] G. Poludniowski, G. Landry, F. DeBlois, P. M. Evans, and F. Verhaegen, "SpekCalc : a program to calculate photon spectra from tungsten anode x-ray tubes," *Phys. Med. Biol.*, vol. 54, no. 19, p. N433, 2009.

[17] “NIST X-Ray Form Factor, Atten. Scatt. Tables Form Page.” https://physics.nist.gov/PhysRefData/FFast/html/form.html (accessed Jun. 09, 2020).
[18] K. T. Bae, “Intravenous Contrast Medium Administration and Scan Timing at CT: Considerations and Approaches,” *Radiology*, vol. 256, no. 1, pp. 32–61, Jul. 2010, doi: 10.1148/radiol.10090908.
[19] M. K. Kalra, A. D. Sodickson, and W. W. Mayo-Smith, “CT Radiation: Key Concepts for Gentle and Wise Use,” *RadioGraphics*, vol. 35, no. 6, pp. 1706–1721, Oct. 2015, doi: 10.1148/rg.2015150118.
[20] S. P. Raman, M. Mahesh, R. V. Blasko, and E. K. Fishman, “CT Scan Parameters and Radiation Dose: Practical Advice for Radiologists,” *J. Am. Coll. Radiol.*, vol. 10, no. 11, pp. 840–846, Nov. 2013, doi: 10.1016/j.jacr.2013.05.032.
[21] C. H. McCollough, “The AAPM/RSNA physics tutorial for residents. X-ray production.,” *RadioGraphics*, vol. 17, no. 4, pp. 967–984, Jul. 1997, doi: 10.1148/radiographics.17.4.9225393.
[22] W. Huda and R. B. Abrahams, “Radiographic Techniques, Contrast, and Noise in X-Ray Imaging,” *Am. J. Roentgenol.*, vol. 204, no. 2, pp. W126–W131, Feb. 2015, doi: 10.2214/AJR.14.13116.
[23] C. H. McCollough, A. N. Primak, N. Braun, J. Kofler, L. Yu, and J. Christner, “Strategies for Reducing Radiation Dose in CT,” *Radiol. Clin. North Am.*, vol. 47, no. 1, pp. 27–40, Jan. 2009, doi: 10.1016/j.rcl.2008.10.006.
[24] G. N. Hounsfield, “Computerized transverse axial scanning (tomography). 1. Description of system,” *Br. J. Radiol.*, vol. 46, no. 552, pp. 1016–1022, Dec. 1973, doi: 10.1259/0007-1285-46-552-1016.
[25] R. E. Alvarez and A. Macovski, “Energy-selective reconstructions in X-ray computerized tomography,” *Phys. Med. Biol.*, vol. 21, no. 5, pp. 733–744, Sep. 1976.
[26] T. R. C. Johnson, “Dual-energy CT: general principles,” *AJR Am. J. Roentgenol.*, vol. 199, no. 5 Suppl, pp. S3-8, Nov. 2012, doi: 10.2214/AJR.12.9116.
[27] C. H. McCollough *et al.*, “Principles and Applications of Multi-energy CT Report of AAPM Task Group 291,” *Med. Phys.*, vol. n/a, no. n/a, doi: 10.1002/mp.14157.
[28] E. S. R. Bs. Ms. FCAMRT, *Computed Tomography: Physical Principles, Clinical Applications, and Quality Control*, 4 edition. St. Louis, Missouri: Saunders, 2015.
[29] “Always on spectral results eliminates the patient selection dilemma.” http://clinical.netforum.healthcare.philips.com/global/Explore/Abstracts/CT/Always-on-spectral-results-eliminates-the-patient-selection-dilemma (accessed Jul. 19, 2019).
[30] J. R. Grajo, M. Patino, A. Prochowski, and D. V. Sahani, “Dual energy CT in practice: Basic principles and applications,” *Appl. Radiol.*, vol. 45, no. 7, pp. 6–12, Jul. 2016.
[31] P. R. S. Mendonça, P. Lamb, and D. V. Sahani, “A Flexible Method for Multi-Material Decomposition of Dual-Energy CT Images,” *IEEE Trans. Med. Imaging*, vol. 33, no. 1, pp. 99–116, Jan. 2014, doi: 10.1109/TMI.2013.2281719.
[32] X. Liu, L. Yu, A. N. Primak, and C. H. McCollough, “Quantitative imaging of element composition and mass fraction using dual-energy CT: Three-material decomposition,” *Med. Phys.*, vol. 36, no. 5, pp. 1602–1609, May 2009, doi: 10.1118/1.3097632.
[33] L. A. Lehmann *et al.*, “Generalized image combinations in dual KVP digital radiography,” *Med. Phys.*, vol. 8, no. 5, pp. 659–667, 1981, doi: 10.1118/1.595025.

[34] W. A. Kalender, W. H. Perman, J. R. Vetter, and E. Klotz, “Evaluation of a prototype dual-energy computed tomographic apparatus. I. Phantom studies,” *Med. Phys.*, vol. 13, no. 3, pp. 334–339, 1986, doi: 10.1118/1.595958.
[35] B. J. Heismann, J. Leppert, and K. Stierstorfer, “Density and atomic number measurements with spectral x-ray attenuation method,” *J. Appl. Phys.*, vol. 94, no. 3, pp. 2073–2079, Aug. 2003, doi: 10.1063/1.1586963.
[36] J. W. Lambert *et al.*, “The effect of patient diameter on the dual-energy ratio of selected contrast-producing elements,” *J. Comput. Assist. Tomogr.*, vol. 41, no. 3, pp. 505–510, 2017, doi: 10.1097/RCT.0000000000000557.
[37] X. Wang, D. Meier, K. Taguchi, D. J. Wagenaar, B. E. Patt, and E. C. Frey, “Material separation in x-ray CT with energy resolved photon-counting detectors,” *Med. Phys.*, vol. 38, no. 3, pp. 1534–1546, 2011, doi: 10.1118/1.3553401.
[38] B. Krauss, K. L. Grant, B. T. Schmidt, and T. G. Flohr, “The importance of spectral separation: an assessment of dual-energy spectral separation for quantitative ability and dose efficiency,” *Invest. Radiol.*, vol. 50, no. 2, pp. 114–118, Feb. 2015, doi: 10.1097/RLI.0000000000000109.
[39] M. Patino *et al.*, “Material Separation Using Dual-Energy CT: Current and Emerging Applications,” *RadioGraphics*, vol. 36, no. 4, pp. 1087–1105, Jul. 2016, doi: 10.1148/rg.2016150220.
[40] E. Shefer *et al.*, “State of the Art of CT Detectors and Sources: A Literature Review,” *Curr. Radiol. Rep.*, vol. 1, no. 1, pp. 76–91, Mar. 2013, doi: 10.1007/s40134-012-0006-4.
[41] K. Taguchi and J. S. Iwanczyk, “Vision 20/20: Single photon counting x-ray detectors in medical imaging,” *Med. Phys.*, vol. 40, no. 10, p. n/a-n/a, Oct. 2013, doi: 10.1118/1.4820371.
[42] P. Delpierre, “A history of hybrid pixel detectors, from high energy physics to medical imaging,” *J. Instrum.*, vol. 9, no. 05, pp. C05059–C05059, May 2014, doi: 10.1088/1748-0221/9/05/C05059.
[43] M. Garcia-Sciveres and N. Wermes, “A review of advances in pixel detectors for experiments with high rate and radiation,” *Rep. Prog. Phys.*, vol. 81, no. 6, p. 066101, May 2018, doi: 10.1088/1361-6633/aab064.
[44] M. R. Uddin, “Spectral Image Quality in MARS Scanners,” p. 160.
[45] “New Zealand start-up scans 1st humans with spectral CT,” *AuntMinnie.com*. https://www.auntminnie.com/index.aspx?sec=sup&sub=adv&pag=dis&ItemID=121308 (accessed Jan. 20, 2020).
[46] S. M. Jorgensen, A. J. Vercnocke, D. S. Rundle, P. H. Butler, C. H. McCollough, and E. L. Ritman, “Evaluation of a photon counting Medipix3RX CZT spectral x-ray detector,” *Proc. SPIE-- Int. Soc. Opt. Eng.*, vol. 9969, Aug. 2016, doi: 10.1117/12.2236501.
[47] A. Pourmorteza *et al.*, “Abdominal Imaging with Contrast-enhanced Photon-counting CT: First Human Experience,” *Radiology*, vol. 279, no. 1, pp. 239–245, Feb. 2016, doi: 10.1148/radiol.2016152601.
[48] R. Gutjahr *et al.*, “Human Imaging With Photon Counting–Based Computed Tomography at Clinical Dose Levels: Contrast-to-Noise Ratio and Cadaver Studies,” *Invest. Radiol.*, vol. 51, no. 7, pp. 421–429, Jul. 2016, doi: 10.1097/RLI.0000000000000251.
[49] Z. Yu *et al.*, “Initial results from a prototype whole-body photon-counting computed tomography system,” *Proc. SPIE-- Int. Soc. Opt. Eng.*, vol. 9412, 2015, doi: 10.1117/12.2082739.

[50] D. P. Cormode *et al.*, "Multicolor spectral photon-counting computed tomography: in vivo dual contrast imaging with a high count rate scanner," *Sci. Rep.*, vol. 7, no. 1, p. 4784, Jul. 2017, doi: 10.1038/s41598-017-04659-9.
[51] J. P. Schlomka *et al.*, "Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography," *Phys. Med. Biol.*, vol. 53, no. 15, p. 4031, 2008, doi: 10.1088/0031-9155/53/15/002.
[52] H. Q. Le and S. Molloi, "Least squares parameter estimation methods for material decomposition with energy discriminating detectors," *Med. Phys.*, vol. 38, no. 1, pp. 245–255, Jan. 2011, doi: 10.1118/1.3525840.
[53] T. E. Curtis and R. K. Roeder, "Effects of calibration methods on quantitative material decomposition in photon-counting spectral computed tomography using a maximum a posteriori estimator," *Med. Phys.*, vol. 44, no. 10, pp. 5187–5197, Oct. 2017, doi: 10.1002/mp.12457.
[54] J. S. Butzer, A. Butler, P. Butler, P. J. Bones, N. Cook, and L. Tlustos, "Medipix imaging - Evaluation of datasets with PCA," presented at the 2008 23rd International Conference Image and Vision Computing New Zealand, IVCNZ, Dec. 2008, pp. 1–6, doi: 10.1109/IVCNZ.2008.4762080.
[55] Y. Lu *et al.*, "A learning-based material decomposition pipeline for multi-energy x-ray imaging," *Med. Phys.*, vol. 46, no. 2, pp. 689–703, 2019, doi: 10.1002/mp.13317.
[56] E. Roessl and R. Proksa, "K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors," *Phys. Med. Biol.*, vol. 52, no. 15, p. 4679, 2007, doi: 10.1088/0031-9155/52/15/020.
[57] P. He, B. Wei, P. Feng, M. Chen, and D. Mi, "Material Discrimination Based on K-edge Characteristics," *Comput. Math. Methods Med.*, vol. 2013, 2013, doi: 10.1155/2013/308520.
[58] A. Brambilla, A. Gorecki, A. Potop, C. Paulus, and L. Verger, "Basis material decomposition method for material discrimination with a new spectrometric X-ray imaging detector," *J. Instrum.*, vol. 12, no. 08, p. P08014, 2017, doi: 10.1088/1748-0221/12/08/P08014.
[59] E. Roessl and R. Proksa, "Optimal energy threshold arrangement in photon-counting spectral x-ray imaging," in *2006 IEEE Nuclear Science Symposium Conference Record*, San Diego, CA, USA, 2006, pp. 1950–1954, doi: 10.1109/NSSMIC.2006.354276.
[60] P. He, B. Wei, W. Cong, and G. Wang, "Optimization of K-edge imaging with spectral CT," *Med. Phys.*, vol. 39, no. 11, pp. 6572–6579, Nov. 2012, doi: 10.1118/1.4754587.
[61] S. J. Nik, J. Meyer, and R. Watts, "Optimal material discrimination using spectral x-ray imaging," *Phys. Med. Biol.*, vol. 56, no. 18, p. 5969, 2011, doi: 10.1088/0031-9155/56/18/012.
[62] "Omnipaque (Iohexol Injection): Uses, Dosage, Side Effects, Interactions, Warning," *RxList*. https://www.rxlist.com/omnipaque-drug.htm (accessed Mar. 21, 2020).
[63] "Omniscan (Gadodiamide): Uses, Dosage, Side Effects, Interactions, Warning," *RxList*. https://www.rxlist.com/omniscan-drug.htm (accessed Mar. 21, 2020).
[64] "Milli-Q® Type 1 Ultrapure Water Systems | MilliporeSigma." https://www.emdmillipore.com/US/en/products/water-purification/type-1/5cSb.qB.89EAAAFAJqxkiQpx,nav (accessed Mar. 21, 2020).

[65] “Publish and Share MATLAB Code - MATLAB & Simulink.” https://www.mathworks.com/help/matlab/matlab_prog/publishing-matlab-code.html (accessed Mar. 23, 2020).
[66] “Publishing Markup - MATLAB & Simulink.” https://www.mathworks.com/help/matlab/matlab_prog/marking-up-matlab-comments-for-publishing.html (accessed Mar. 23, 2020).
[67] “Solve systems of linear equations Ax = B for x - MATLAB mldivide \.” https://www.mathworks.com/help/matlab/ref/mldivide.html (accessed Mar. 23, 2020).
[68] “mldivide \, mrdivide / (MATLAB Functions).” http://matlab.izmiran.ru/help/techdoc/ref/mldivide.html (accessed Mar. 23, 2020).
[69] “When should I use x = A\b and when should I use x = inv(A)*b ? - MATLAB Answers - MATLAB Central.” https://www.mathworks.com/matlabcentral/answers/91458-when-should-i-use-x-a-b-and-when-should-i-use-x-inv-a-b (accessed Mar. 23, 2020).
[70] E. W. Weisstein, “Point-Plane Distance.” https://mathworld.wolfram.com/Point-PlaneDistance.html (accessed Mar. 23, 2020).
[71] E. W. Weisstein, “Hessian Normal Form.” https://mathworld.wolfram.com/HessianNormalForm.html (accessed Mar. 23, 2020).
[72] “Polynomial curve fitting - MATLAB polyfit.” https://www.mathworks.com/help/matlab/ref/polyfit.html (accessed Mar. 24, 2020).
[73] “How to plot a covariance error ellipse • Open Source is Everything.” https://www.xarg.org/2018/04/how-to-plot-a-covariance-error-ellipse/ (accessed Mar. 24, 2020).
[74] “One-way ANOVA - Its preference to multiple t-tests and the assumptions needed to run this test | Laerd Statistics.” https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide-2.php (accessed Apr. 02, 2020).
[75] “Two-way ANOVA in SPSS Statistics | Laerd Statistics Premium.” https://statistics.laerd.com/premium/spss/twa/two-way-anova-in-spss-9.php#violation-normality (accessed Apr. 16, 2020).
[76] C. J. Bateman *et al.*, “MARS-MD: rejection based image domain material decomposition,” *J. Instrum.*, vol. 13, no. 05, pp. P05020–P05020, May 2018, doi: 10.1088/1748-0221/13/05/P05020.
[77] F. Ostadhossein *et al.*, “Multi-‘Color’ Delineation of Bone Microdamages Using Ligand-Directed Sub-5 nm Hafnia Nanodots and Photon Counting CT Imaging,” *Adv. Funct. Mater.*, vol. 30, no. 4, p. 1904936, 2020, doi: 10.1002/adfm.201904936.
[78] G. E. P. Box and N. R. Draper, *Response Surfaces, Mixtures, and Ridge Analyses*. John Wiley & Sons, 2007.