



Radiographic Radiomics Phantom

October 7, 2026

Biomedical Engineering 200/300

Section 302

Client: Dr. Jennie Crosby, Department of Medical Physics, University of Wisconsin-Madison

Advisor: Dr. Yang Lu

Team Members:

Pierson Trapskin (Leader) | petrapskin@wisc.edu
Presley Stellflue (BSAC) | pstellflue@wisc.edu
Ava Hodgson (Communicator) | amhodgson@wisc.edu
Matthew Casas (Co-BWIG) | mecasas2@wisc.edu
Sonja Pionek (Co-BWIG) | spionek@wisc.edu
Allie Ma (BPAG) | azma@wisc.edu

Abstract

Radiomics extracts quantitative texture features from medical images to support cancer diagnosis, prognosis, and prediction of treatment response. Clinical use has been limited because feature values change with scanner model, acquisition and reconstruction settings, and patient size, sometimes by as much as the biological differences the features are meant to detect. Commercial CT calibration phantoms contain homogeneous inserts with no internal texture, and existing texture phantoms are built in a single size, so neither can separate the effect of body size from other sources of variation. This project is designing a durable radiomics phantom for Dr. Jennie Crosby of the University of Wisconsin-Madison Department of Medical Physics that presents interchangeable textures in both head-sized and abdomen-sized configurations. Three preliminary designs were compared in a weighted design matrix, and an elliptical design scored highest (79 of 100). The proposed phantom has a 400 mm $\times$ 300 mm elliptical cast acrylic shell around a removable $\varnothing$ 180 mm acrylic core, 160 mm deep, with keyed bores for 3D-printed polylactic acid (PLA) inserts whose internal infill patterns create controlled texture. No prototype has been built yet, so no test results are reported. The testing plan covers dimensional fit, material CT numbers, body uniformity, spatial resolution and noise, test-retest and print-to-print reproducibility, robustness to scan protocol, the effect of phantom size, and stability over time, using the intraclass correlation coefficient, concordance correlation coefficient, and coefficient of variation. If successful, the phantom will give medical physicists a reusable reference for identifying which radiomic features are stable across scanners, protocols, and body sizes.

Table of Contents

Abstract	2
Table of Contents	2
Introduction	3
Motivation and Societal Impact	3
Current Methods	4
Problem Statement	5
Background	5
Relevant Physiology and Biology	5
Research Required for Design and Fabrication	6
CT Acquisition and Reconstruction Parameters	6
Radiomic Feature Extraction	6
Phantom Materials and Fabrication	7
Evaluation Metrics	7
Client	8
Design Specifications	8
Preliminary Designs	8
Design 1: Varied Radius Phantom	8

	3
Design 2: Elliptical Phantom	9
Design 3: Rectangular Phantom	10
Preliminary Design Evaluation	11
Design Matrix	11
Design Matrix Summary	11
Proposed Final Design	12
Fabrication	12
Materials	12
Methods	13
Design	13
Acrylic Shell and Core	14
Texture Inserts	15
Assembly and Pre-Scan Check	15
Final Prototype	15
Testing and Results	16
Image Acquisition and Feature Extraction	16
Planned Tests	17
Test 1: Dimensions and Fit	17
Test 2: Material CT Numbers	17
Test 3: Body Uniformity and Seam Visibility	17
Test 4: Repeatability (Test-Retest)	18
Test 5: Print-to-Print Reproducibility	18
Test 6: Protocol Robustness	18
Test 7: Size Dependence	18
Test 8: Texture Range and Clinical Realism	19
Test 9: Stability Over Time	19
Test 10: Spatial Resolution and Noise	19
Sources of Error and Variance	20
Statistical Analysis	20
Results	20
Conclusions	20
References	21
Appendix	25
Appendix A: Product Design Specifications	25
Function	25
Client Requirements	25
Design Requirements	25
Appendix A References	29
Appendix B: Materials and Budget	29
Appendix C: Engineering Drawings	30
Appendix D: Print Settings	30

Appendix E: Feature Extraction Settings	30
Appendix F: Raw Data	30

Introduction

Motivation and Societal Impact

Medical images have traditionally been interpreted visually, yet each computed tomography (CT) image contains quantitative information that the eye does not fully use. Radiomics is the high-throughput extraction of quantitative features, including intensity statistics, shape descriptors, and texture measures, from medical images so that they can be mined to support diagnosis, prognosis, and prediction of treatment response [1]. Texture features are particularly interesting because they quantify spatial heterogeneity within tumors, an established prognostic factor, and early evidence suggests texture analysis can improve tumor characterization, staging, and response assessment [2].

The potential reach of CT radiomics is large. In 2023, about 61.5 million patients in the United States underwent an estimated 93 million CT examinations, including about 2.57 million examinations in children [3]. Each scan is a possible source of radiomic data at no added radiation dose. Because CT carries a radiation cost, with the same analysis projecting about 103 000 future cancers from 2023 CT use [3], extracting more clinically useful information from scans that are already being performed is an attractive way to increase the value of every examination.

Clinical translation has been slow because radiomic features are often not reproducible. A systematic review of 41 studies found that feature values are sensitive to acquisition settings, reconstruction algorithm, image preprocessing, and extraction software, and that texture features were generally less reproducible than first-order features [4]. In a multi-center study, the variability caused by using different CT scanners was, for some features, comparable to the variability between non-small cell lung cancer tumors in different patients [5]. In that situation a feature difference may reflect the scanner rather than the biology. Phantom work has shown that the share of reproducible features can fall to 43.1 % of 177 features when only the reconstruction kernel is changed [6], and patient data confirm that dose and reconstruction settings change feature values within the same person [7]. The Image Biomarker Standardization Initiative (IBSI) has standardized how 174 features are calculated [8], but software standardization cannot correct differences in the images themselves. A physical reference object, or phantom, is needed to measure those differences.

Patient size is a source of variation that has received less attention. Larger bodies attenuate and scatter more x-rays, which raises image noise and changes beam hardening, and they also lead to a larger reconstruction field of view and therefore larger pixels, a parameter that texture features depend on directly [9]. In a recent cone-beam CT phantom study, only 8 of 93 features from a standard clinical protocol were robust between phantoms representing small and large body habitus [10]. Radiomic models are applied to children and adults and to the head, chest, and abdomen, so a phantom that tests the same tissue surrogate at different body sizes addresses a real and growing need.

Current Methods

Phantoms used for CT radiomics fall into three groups: homogeneous tissue-equivalent calibration phantoms, purpose-built texture phantoms, and anthropomorphic phantoms derived from patient scans. Representative examples are compared in Table 1.

Table 1. Representative phantoms for CT calibration and radiomics, compared by size adaptability and insert type. Specifications are taken from manufacturer datasheets and published descriptions.

Device	Type	Size adaptability	Insert type	Main limitation for this project
CIRS Model 062M [11]	Electron density calibration	Removable 180 mm diameter head disk nested in a 330 mm $\times$ 270 mm body	Homogeneous tissue-equivalent epoxy plugs	No internal texture
Sun Nuclear Advanced Electron Density Phantom [12]	Electron density calibration	Removable 200 mm head section in a 400 mm $\times$ 300 mm body	16 homogeneous plugs, relative electron density 0.28 to 1.77	No internal texture
Credence Cartridge Radiomics phantom [5], [13]	Texture	Single configuration	10 cartridges: ABS honeycomb, wood, cork, rubber, acrylic, plaster	Natural materials are hard to reproduce; no size change
3D-printed texture and liver phantoms [14], [15]	Texture	Single configuration	Printed patterns of graded heterogeneity	No size change; liver-focused range
Open-source modular texture phantom [16]	Texture, modular	Effect of patient size not reported	8 interchangeable printed inserts	Size dependence untested
Patient-derived printed phantoms [17], [18], [19]	Anthropomorphic	Fixed to one patient	Voxel-level printed patient anatomy	Specialized printers or workflows
U.S. Patent 12 089 982 B2 [20]	Radiotherapy calibration	Ellipsoid base	Removable tissue-mimicking inserts	Aimed at dose calculation, not texture

Electron density phantoms. Commercial CT-to-electron density calibration phantoms contain cylindrical plugs of tissue-equivalent material. The CIRS Model 062M uses two nested disks, a 180 mm diameter head section that fits inside a 330 mm $\times$ 270 mm body section, with plugs mimicking lung, adipose, breast, muscle, liver, and bone [11]. The Sun Nuclear Advanced Electron Density Phantom similarly has a removable 200 mm head section within a 400 mm $\times$ 300 mm body and 16 insert positions with relative electron densities from 0.28 (lung) to 1.77 (cortical bone) [12]. These products already use a nested, size-swapping layout, but their inserts are homogeneous. Homogeneous electron density plugs have been repurposed for texture studies [21], but in a direct comparison only 15.7 % to 19.0 % of features measured on a homogeneous commercial phantom fell within the range seen in patient lung tumors, compared with 55.6 % to 90.2 % for heterogeneous inserts [22].

Texture phantoms. The Credence Cartridge Radiomics (CCR) phantom, built for the multi-center study described above, holds 10 cartridges containing 3D-printed ABS with 20 % to 50 % honeycomb fill, sycamore wood, cork, shredded rubber, acrylic, and 3D-printed plaster, and spans mean CT numbers from -864 HU to 652 HU [5], [13]. Natural materials such as wood and cork give realistic texture but cannot be duplicated exactly from one phantom to the next.

Later groups 3D-printed patterns of progressively increasing heterogeneity [14] and liver-mimicking inserts with CT numbers between -30 HU and 120 HU [15]. Most recently, an open-source modular phantom with an epoxy-filled cylindrical body and eight cavities for interchangeable 3D-printed inserts showed body uniformity within 5 HU and an intraclass correlation coefficient (ICC) above 0.999 between independently printed insert sets [16]. Its published evaluation addressed manufacturing reproducibility and dose and kernel changes, not patient size.

Anthropomorphic phantoms. Patient-derived phantoms reproduce real anatomy using multi-material voxel printing [17], synthetic lesion textures printed into lung nodules [23], iodinated ink printed on paper [18], or fused deposition printing that varies filament density voxel by voxel [19]. These phantoms are highly realistic, but they generally require specialized printers or workflows, and each represents a single patient.

Patents. U.S. Patent 12 089 982 B2, assigned to the University of Cincinnati, claims an ellipsoid phantom base with cylindrical voids that accept removable tissue-mimicking inserts for radiotherapy calibration [20]. It supports the general approach of interchangeable inserts but is designed for dose calculation rather than radiomic texture.

Problem Statement

This project aims to design, fabricate, and test a long-lasting radiomics phantom that enables medical physicists to benchmark and standardize quantitative radiomic texture features across clinical CT imaging systems. The phantom must incorporate a variety of non-homogeneous, tissue-like textures, accommodate clinical scanning protocols for both small and large anatomical regions, and maintain mechanical and imaging stability over a multi-year service life. By systematically evaluating the feature reproducibility, spatial uniformity, and Hounsfield unit (HU) consistency across varying scanner models and imaging parameters, the system will establish reliable feature calibration metrics. Ultimately, this phantom will provide Dr. Jennie Crosby and the Department of Medical Physics with a reliable research platform that ensures radiomic feature stability and enhances the clinical reliability of radiomics-based medical decisions.

Background

Relevant Physiology and Biology

Tumor heterogeneity. Solid tumors are spatially heterogeneous. Regions of high cellularity, necrosis, hemorrhage, fibrosis, and variable blood supply sit side by side, and this heterogeneity is associated with aggressive behavior and resistance to treatment [2]. On CT, these regions differ slightly in attenuation and produce a pattern of gray levels that texture features quantify. Clinical CT pixels are typically 0.39 mm to 0.98 mm across [9], so a phantom insert must contain structure at roughly this scale and larger to generate measurable texture.

How texture appears on CT. CT reports attenuation in Hounsfield units, $HU = 1000 \times (\mu - \mu_{\text{water}}) / \mu_{\text{water}}$, where μ is the linear attenuation coefficient; water is defined as 0 HU and air is approximately -1000 HU [24]. Texture in a CT image is the spatial pattern of HU values from voxel to voxel, so it depends on two things: the attenuation contrast between neighboring structures and the size and arrangement of those structures relative to the voxel. This phantom is designed to reproduce physical texture only. It is not intended to match the electron density of any specific tissue, so the inserts do not need to be tissue-equivalent in composition. Instead, each insert creates texture through its internal geometry, such as pores, lattices, or randomly placed voids within a solid material, in the same way that the honeycomb, cork, and rubber cartridges of the CCR phantom produce distinct feature values [5]. Keeping the mean CT number of the inserts within a clinically realistic range is still useful, because gray-level discretization and texture calculations behave differently at very high or very low contrast [9]. For reference, non-small cell lung tumors are measured from -186 HU to 35 HU in mean CT number [5], and normal liver falls within about -30 HU to 120 HU [15].

Body size. Body dimensions vary widely across the patient population. AAPM Report 204 tabulates size-dependent dose conversion factors for effective diameters from about 8 cm in small children to 45 cm in large adults [25], and standard CT dosimetry uses separate 16 cm head and 32 cm body cylinders to represent this range [24]. As diameter increases, fewer photons reach the detector, noise increases, and scanners respond by changing tube current and field of view, all of which alter radiomic feature values [9], [10]. A removable core that allows the same inserts to sit inside a small or a large outer shell reproduces this anatomical difference directly.

Research Required for Design and Fabrication

CT Acquisition and Reconstruction Parameters

The team must confirm and understand which scan settings the phantom will be tested against. Reconstruction kernels, slice thickness, pitch, field of view, tube voltage, and dose have all been shown to change feature values [6], [7], and pixel size and gray-level discretization have intrinsic effects on texture features [9]. These parameters define the testing matrix and inform how large the outer shells must be to produce clinically realistic noise and field of view.

Radiomic Feature Extraction

Radiomic features are grouped into first-order intensity statistics, shape features, and higher-order texture features calculated from matrices such as the gray-level co-occurrence matrix (GLCM) and neighborhood gray-tone difference matrix (NGTDM) [1]. Results depend on the extraction software and preprocessing [4], so testing should use an IBSI-compliant tool and report its settings [8]. Open-source packages such as PyRadiomics have been used in comparable phantom studies [22].

Phantom Materials and Fabrication

Fused deposition 3D printing is a low-cost way to build texture inserts because internal structure can be specified exactly and reprinted. Texture can be controlled through infill percentage and pattern, as in the 20 % to 50 % honeycomb cartridges of the CCR phantom [5], [13], or through mathematically defined lattices and random shape distributions, which produced distinct, non-redundant texture stimuli with ICC above 0.999 between independent prints [16]. Infill also shifts the mean CT number: 13 commercial filaments printed at 50 % to 100 % infill ranged from -732 HU to $10\,047$ HU [26], so each texture pattern must be characterized on the scanner rather than assumed. When a uniform region is needed, such as the phantom body around the inserts, a gyroid infill pattern gives the least CT number variation [27], and the open-source modular phantom achieves body uniformity within 5 HU using an epoxy-filled body [16]. A ten-center audit of printed polylactic acid (PLA) reported a mean dimensional error of (-0.2 ± 0.3) mm [28], which supports printing fine texture features and alignment keys with sub-millimeter accuracy. Because feature reproducibility depends on the phantom material itself [29], material choice, infill, and print settings should be recorded and held constant for every insert.

For the same reason, the phantom body will be machined from cast acrylic (polymethyl methacrylate, PMMA) rather than cast from epoxy. Acrylic is the established material for CT dosimetry phantoms, which use nested 160 mm head and 320 mm body cylinders of PMMA with a density of 1.19 g/cm³ [30], and it has a consistent, well-characterized CT number of approximately $+118$ HU to $+124$ HU at 120 kV [31]. A cast epoxy body can achieve comparable uniformity [16], but the casting process introduces sources of error that a purchased acrylic blank avoids. Air bubbles trapped during mixing appear as low-attenuation spots, incomplete mixing or batch-to-batch differences can shift the CT number, and the heat released while thick pours cure can cause shrinkage, distortion, and cracking [32]. Deep-pour casting resins are also limited to roughly 2.5 cm to 10 cm (1 in to 4 in) per pour [32], [33], so a 160 mm deep body would require several layers, each forming an internal interface that could appear in the image. Machining acrylic instead produces tight, repeatable tolerances between the removable core and the outer shell, which minimizes air gaps that would appear as dark rings around the core. Its transparency also allows the team to confirm visually that each insert is fully seated and free of trapped air before scanning.

Evaluation Metrics

Phantom performance is judged by repeatability (the same scan repeated) and reproducibility (different scanners or settings). Common statistics are the ICC, the concordance correlation coefficient (CCC), and the coefficient of variation (CV) [6]; for example, features have been called repeatable when $CV \leq 0.10$ [22]. Phantom feature values should also be compared against the range found in patient images to confirm clinical relevance [5], [22].

Client

The client, Dr. Jennie Crosby, is in the Department of Medical Physics at the University of Wisconsin-Madison. Her work focuses on adaptive and MR-guided radiotherapy. She has also completed research in patient safety and quality assurance, including analysis of trends and workflow inefficiencies in radiology.

Design Specifications

Design specifications were determined by a combination of requests from the client, literature reviews, and formal standards (see Appendix A). This design should focus on the incorporation of non-homogeneous textures to replicate characteristics seen in tissue imaging and for easy identification in images. The phantom should have an outer shell and a removable inner core with size requirements specified by the client (see Appendix A). The outer shell should represent the size of a human abdomen and the removable inner core should represent the size of a human head. Removing the inner core and being able to replace the material allows for more types of images to be collected from different tissue-like patterns, which is necessary when calibrating and testing the CT scanner. Regarding fixed performance requirements, the difference in HU between the center of the phantom and the outer edges should be less than or equal to 5 HU [34]. After repeated use and scanning of the phantom, the recorded values and standard deviation should remain within $\pm 10\%$ of the baseline noise. For spatial resolution, the phantom should be evaluated using a line-pair or high-contrast pattern, with the highest resolvable spatial frequency reported in line pairs per centimeter (lp/cm). American College of Radiology recommendations include a minimum spatial resolution of 5 lp/cm for adult abdomen imaging and 6 lp/cm for high-resolution chest imaging [34]. Dimensional accuracy and repeatability are also important to ensure consistent imaging results between scans. The critical dimensions of the phantom, including the outer shell and removable inner core, should remain within ± 1 mm of the specified dimensions. The removable core should also remain securely positioned during scanning to prevent changes in material location that could affect the resulting CT images. The phantom should maintain its dimensions, material properties, and imaging characteristics over its intended five-year service life. Repeated scans should be used to evaluate whether the phantom maintains consistent HU values, image noise, and radiomic features over time [34], [35].

Preliminary Designs

Design 1: Varied Radius Phantom

The varied radius phantom provides the most anatomically realistic design (Figure 1). The removable core still remains in the center of the body, with relative size requirements still applied. The removable inserts are included on the face of the phantom and would be placed in a circular pattern around the head. The differing radius throughout the body of the phantom would also serve as a visible depth marker, signifying the physical location of the scan occurring inside the phantom. This would aid in eliminating possible errors in imaging occurring from

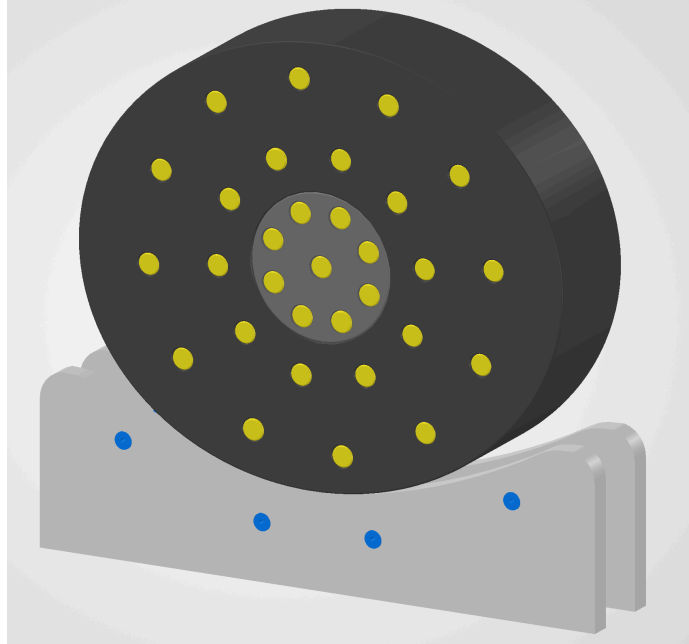


Figure 2. *Elliptical phantom.*

Design 3: Rectangular Phantom

The modular box phantom focuses on simple fabrication and low production cost by utilizing a rectangular block geometry with flat planar faces (Figure 3). The central core sits inside the block, surrounded by a straight grid of interchangeable inserts. The flat bottom allows the phantom to rest securely on the CT couch without an extra support cradle. However, the box shape fails to replicate realistic human beam attenuation, and the sharp corners restrict insert placement to central regions, limiting peripheral testing options.

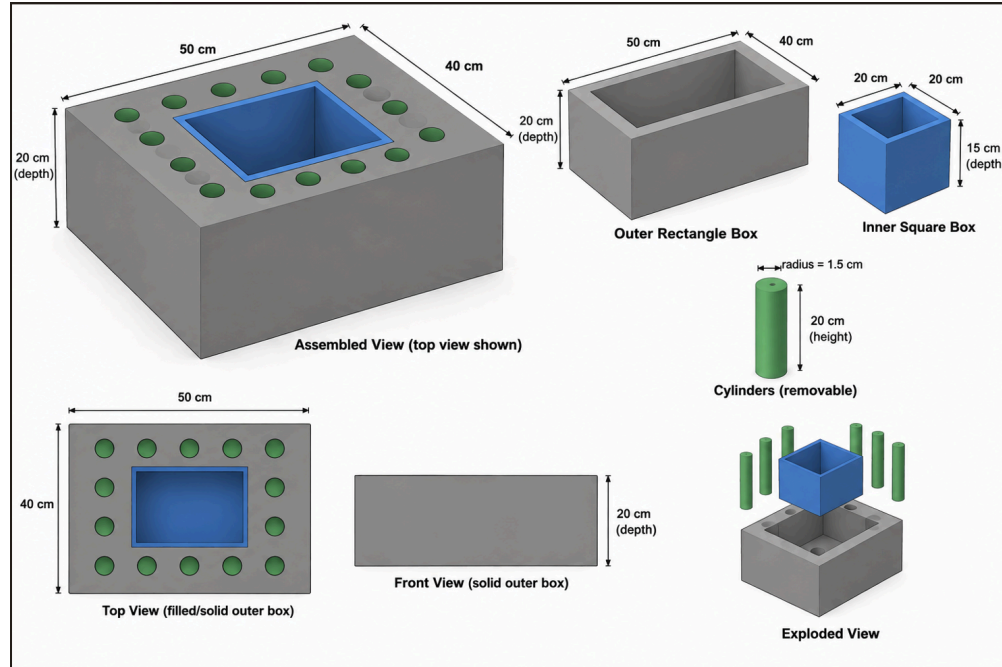


Figure 3. *Rectangular phantom.*

Preliminary Design Evaluation

Design Matrix

The most critical design specifications were considered when creating the design matrix (Table 2). Criteria are weighted and ordered based on importance. The winning design was the elliptical design (Design 2).

Table 2. *Design matrix comparing the three preliminary phantom designs. Each design is scored out of 5 for each criterion, and the score is scaled by the criterion weight.*

Criteria	Weight	Design 1 (Varied radius)	Design 2 (Elliptical)	Design 3 (Rectangular)
Effectiveness	25	4/5 → 20	5/5 → 25	3/5 → 15
Ease of use	20	4/5 → 16	4/5 → 16	3/5 → 12
Cost	15	4/5 → 12	3/5 → 9	3/5 → 9
Ease of production	15	2/5 → 6	3/5 → 9	4/5 → 12
Modularity	15	3/5 → 9	4/5 → 12	3/5 → 9
Safety	10	2/5 → 4	4/5 → 8	3/5 → 6
Total	100	67	79	63

Design Matrix Summary

The design matrix was used as an objective and quantitative method for determining the best phantom design. The criteria chosen for this evaluation included: effectiveness, ease of use, cost, ease of production, modularity, and safety. Effectiveness was ranked highest with a weight

of 25 % and was defined as the ability of the phantom to accurately replicate realistic anatomical geometry and support comprehensive evaluation across varying radial distances. Because robust, versatile radiomic feature evaluation is the primary technical objective, a design providing dense sampling positions was deemed paramount. Design 2 was ranked highest in this category because its high-density grid of multiple insert channels allows for thorough radiomic evaluation across varying radial distances, outperforming the constrained peripheral layouts of Design 1 and Design 3. Ease of use was assigned a weight of 20 % to account for the operational workflow efficiency during CT image acquisition. Design 1 and Design 2 tied for the highest rank in this category. Design 1 excelled due to its compact, manageable form factor, while Design 2 achieved an equal score through its contoured base cradle, which ensures stable mounting and simple alignment on the scanner couch. In contrast, Design 3 scored lower due to its bulky box-shaped design. Cost, ease of production, and modularity were each weighted at 15 %. Cost favored Design 1 due to its lower material usage and reduced overall machining expenses compared to larger alternatives. Ease of production favored Design 3 due to its straight planar faces and simple orthogonal geometry, which make it significantly easier to machine and assemble from standard stock materials than the curves of Design 1 or Design 2. For modularity, Design 2 ranked highest due to its extensive layout of inner and outer ring insert slots, providing maximum flexibility to swap and test various material samples compared to the surface-restricted designs. The final category, safety, weighted at 10 %, favored Design 2 due to its flat cradle support base that provides reliable stability and prevents unintended rolling during scanner operation.

Proposed Final Design

The design that will be focused on moving forward is Design 2 (elliptical). This design scored highest, or joint highest, in four of the six evaluation criteria, earning the highest overall score of 79 out of 100. Design 2 utilizes a high-density grid of multiple insert channels to enable comprehensive, high-resolution radiomic feature evaluation across varying radial distances. This system is reliable, versatile, and provides safe and stable operation during CT image acquisition through the integration of a contoured base cradle. The design consists of an elliptical phantom body featuring a layout of inner and outer ring insert slots that accommodate interchangeable texture inserts. The quantitative CT metric values and radiomic data derived from these modular inserts will allow for precise scanner calibration and robust feature evaluation, producing a versatile phantom platform for optimized image acquisition protocols.

Fabrication

Materials

The phantom has three components: an outer acrylic shell, a removable acrylic core, and interchangeable 3D-printed texture inserts (Figure 4). Table 3 summarizes each material and the

reason it was selected. The complete materials list, with part numbers, vendors, quantities, and costs, is in Appendix B.

Table 3. *Phantom components, materials, and rationale for selection. The shell and core share one material so that the small and large configurations differ only in size.*

Component	Material	Rationale
Outer shell, elliptical, 400 mm × 300 mm × 160 mm	Cast acrylic (PMMA), 1.19 g/cm ³	Elliptical outline matches a commercial body phantom with a removable head section [12] and better approximates the wider-than-deep human torso; PMMA is the established material for CT dosimetry phantoms [30]; uniform CT number of about +118 HU to +124 HU at 120 kV [31]; machinable to tight tolerances; transparent, so insert seating can be checked visually
Removable core, Ø180 mm × 160 mm	Cast acrylic (PMMA)	Head-size core matching the removable head section of a commercial calibration phantom [11]; same material as the shell, so the only variable between configurations is phantom size; machined for a precise sliding fit
Slab bonding (only if full-thickness stock is unavailable)	Weld-On 40 two-part reactive acrylic cement	Two-component reactive acrylic cement [36] that cures solid and fills small gaps, avoiding the trapped air and dry spots that water-thin solvent cements tend to leave on large faces
Texture inserts, Ø30 mm × 160 mm	PLA filament, one brand and lot [brand and lot undecided]	Low cost; dimensionally accurate to (-0.2 ± 0.3) mm [28]; internal texture set by infill percentage and pattern [26]
Low-contrast inserts (optional)	PLA with ABS, dual-material print	ABS measures about -20 HU to -41 HU [31], giving texture with lower contrast than plastic against air

Methods

Design

All components will be modeled in SolidWorks. The proposed layout (Figure 4) uses an elliptical shell 400 mm wide and 300 mm deep around a circular Ø180 mm core. The core holds five insert bores (one central and four at a radius of 55 mm), and the shell holds eight bores on an elliptical ring with semi-axes of 160 mm and 118 mm, so that each insert can be scanned at the center, mid-radius, and periphery. These positions leave at least 13 mm of acrylic between any bore and the core boundary or outer surface. The two configurations have effective diameters of about 180 mm (core alone) and 346 mm (assembled), calculated as the square root of the product of the two axes [25], which spans head-size to adult-torso-size anatomy. Each bore has a key flat that fixes insert orientation, because printed texture is directional and rotating an insert could change texture features. Engraved crosshair marks on the shell, similar to those on commercial CT dosimetry phantoms [30], allow repeatable laser alignment. Engineering drawings are in Appendix C.

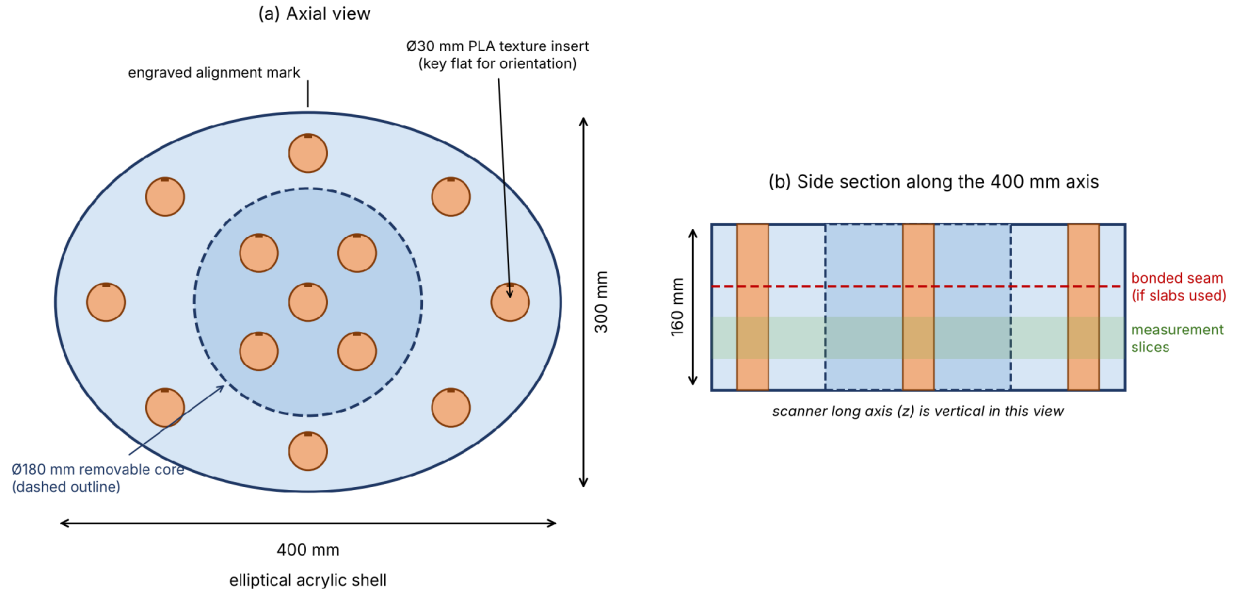


Figure 4. Proposed phantom layout. (a) Axial view showing the 400 mm × 300 mm elliptical acrylic shell, the removable Ø180 mm acrylic core (dashed), and 13 bores for Ø30 mm PLA texture inserts, each with a key flat for orientation. (b) Side section along the 400 mm axis showing the 160 mm depth along the scanner axis, the location of a bonded seam if stacked slabs are used, and the planned measurement slices placed away from the seam. Insert number, diameter, and positions are proposed values.

Acrylic Shell and Core

1. *Stock preparation.* Purchase cast acrylic blanks of at least 165 mm thickness: 410 mm × 310 mm for the shell and 190 mm × 190 mm for the core, which must be separate stock because a plug cut from the shell center is undersized by the cutter width. If full-thickness stock is unavailable, face-mill two slabs (for example, 102 mm and 64 mm) flat, clean them with mild soap and water, bond them with Weld-On 40 under even clamping pressure, and cure according to the manufacturer's instructions [36]. The seam is placed about 100 mm from one end so that it falls outside the planned measurement slices.
2. *Shell profiling and boring.* An ellipse cannot be produced on a lathe, so the outer 400 mm × 300 mm profile will be cut on a CNC mill or CNC router with a finishing pass to leave a smooth surface. The Ø180 mm center bore will be cut in the same setup, using circular interpolation or a boring head, so that the bore is centered on the ellipse.
3. *Core turning.* Measure the finished bore, then turn the core on a lathe to a sliding fit with a target diametral clearance of 0.1 mm. Face both parts to a final depth of 160 mm.
4. *Insert bores and alignment features.* Drill and ream the 13 insert bores after any bonding, so each bore is continuous through the seam. Mill the key flats, engrave the crosshair alignment marks, and add a matching index mark on the core and shell so the core is always inserted in the same rotation.

5. *Annealing.* Wait at least 6 h after cementing, then anneal at 80 °C to relieve machining and bonding stress, which reduces crazing and can increase bond strength by more than 50 % [37]. The manufacturer recommends 1 h of heating per millimeter of thickness and slow cooling until the oven is below 60 °C [37]. For thick sections this implies a multi-day cycle, so the schedule and oven size will be confirmed before this step. This step is optional but will increase the longevity of the product.
6. *Finishing.* Clean with mild soap and water only, because alcohol-based cleaners can craze acrylic. Record the final dimensions of the shell, core, and each bore, measured three times with calipers and reported as mean $\pm$ standard deviation.

Texture Inserts

1. *Pattern design.* Create 20 texture patterns ranging from fine to coarse: honeycomb infill at 20 %, 30 %, 40 %, and 50 %, following the CCR phantom [5], [13]; gyroid and rectilinear infill; and CAD-defined lattices and random void distributions [16].
2. *Printing.* Print all inserts on Bambu Lab X1 Carbon printers from a single filament lot with fixed settings (nozzle diameter, layer height, temperature, speed, and build orientation), recorded in Appendix D. Each insert has solid perimeter walls and solid end caps of at least 1 mm to seal the internal air spaces against moisture uptake.
3. *Fit adjustment.* Print test cylinders, measure each diameter at three locations, and adjust the CAD diameter until inserts slide into the bores without a visible gap, because any air gap appears as a dark ring around the insert. Because printed parts run slightly undersized [28], final insert dimensions are set from measured prints rather than nominal CAD values.
4. *Duplicates and labeling.* Print at least two copies of each pattern to allow print-to-print reproducibility testing [16]. Label each insert on its end cap, record its mass and dimensions, and store inserts sealed with desiccant.
5. *Optional low-contrast set.* Print selected patterns as PLA and ABS dual-material inserts on a dual-extrusion printer, without internal air.

Assembly and Pre-Scan Check

Before each scan, the core is inserted using the index mark and each insert is placed with its key flat aligned. Because the acrylic is transparent, the team will visually confirm that every insert is fully seated and that no air is trapped at the core-to-shell interface. A short scout scan will confirm that no seam or gap artifacts appear in the measurement slices. Small acrylic and PLA test coupons, including a bonded acrylic seam, will be scanned before full fabrication to confirm CT numbers and seam visibility on the client's scanner.

Final Prototype

The prototype has not yet been fabricated. When complete, this section will include high-quality photographs of the assembled phantom from multiple angles, the core removed

from the shell, and close-ups of representative texture inserts. Each image will have labels, dimensions, a scale reference, and a descriptive caption.

Testing and Results

The phantom has not yet been fabricated, so this section describes the planned tests, the design specification each test evaluates, and the expected outcomes. Testing proceeds in four stages: (1) bench measurements of the fabricated parts, (2) CT characterization of the phantom materials and image quality, (3) radiomic feature reliability, including the effect of phantom size, and (4) stability over time. Table 4 maps each test to its design specification, metric, and acceptance threshold.

Table 4. *Planned tests, the design specification each evaluates, primary metric, and proposed acceptance threshold. Highlighted thresholds will be confirmed with the client.*

Test	Design specification	Primary metric	Acceptance threshold
1. Dimensions and fit	Interchangeable, keyed inserts; precise core fit; ± 1 mm critical dimensions	Measured diameters and clearances (mean $\pm$ SD, $n = 3$)	Critical dimensions within ± 1 mm; inserts and core seat fully with no visible gap; core clearance ≤ 0.1 mm
2. Material CT numbers	Known, clinically realistic CT numbers	Mean HU $\pm$ SD in each material	Acrylic within $+118$ HU to $+124$ HU at 120 kV [31]; each insert characterized
3. Body uniformity and seam	Uniform background	Center-to-periphery HU difference; seam visibility	Difference ≤ 5 HU [34]; no seam in measurement slices
4. Repeatability (test-retest)	Reproducible feature values	ICC, CCC, CV across repeated scans	ICC > 0.75 [38]; CV $\leq 10\%$ [22]
5. Print-to-print reproducibility	Reproducible insert manufacture	ICC between duplicate inserts	ICC > 0.75
6. Protocol robustness	Benchmarks features across scan settings	CV and ICC across protocol changes; fraction of robust features	Reported per feature; no fixed pass value
7. Size dependence	Small and large configurations	Paired feature differences, core alone vs. assembled	Reported per feature; demonstrates measurable size effect
8. Texture range and realism	Range of heterogeneous textures	PCA separation of patterns; fraction of features within patient range	Patterns separate in PCA [16]; features within patient 10th to 90th percentile [22]
9. Stability over time	Five-year service life	Change in HU, insert mass, and features from baseline	HU drift ≤ 5 HU; mass change $\leq 1\%$
10. Spatial resolution and noise	Spatial resolution in lp/cm; noise within $\pm 10\%$ of baseline	Highest resolvable lp/cm; SD of HU in a uniform region	≥ 5 lp/cm (adult abdomen protocol) and ≥ 6 lp/cm (high-resolution chest protocol) [34]; noise within $\pm 10\%$ of baseline

Image Acquisition and Feature Extraction

All scans will be acquired on the various scanner models at the UW-Madison Department of Medical Physics. The reference protocol is 120 kV, fixed tube current of X mAs, 1 mm slice thickness, standard soft-tissue kernel, 250 mm display field of view, chosen to match a routine clinical abdominal protocol. Before each session the scanner's daily quality assurance will be confirmed, and the phantom will be aligned to the gantry lasers using its engraved marks.

Features will be extracted with PyRadiomics [39], an open-source package, using fixed settings recorded in Appendix E: resampling to 1 mm isotropic voxels, a fixed gray-level bin width of 25 HU, and identical regions of interest (ROIs) in every scan. Each ROI is a cylinder of 18 mm diameter, centered in its insert and inset from the insert wall to avoid partial-volume edge effects, spanning 10 mm within the measurement slices away from any seam. ROIs are placed from a saved template registered to the phantom alignment marks, so placement does not depend on the operator. First-order and texture features (GLCM, GLRLM, GLSZM, GLDM, and NGTDM classes) will be analyzed; shape features are excluded because the ROI geometry is fixed. Settings will follow IBSI definitions where available [8].

Planned Tests

Test 1: Dimensions and Fit

Purpose. Verify that machining and printing meet the fit requirements before any scanning, since gaps around inserts or the core create dark ring artifacts. *Method.* Measure the shell bore, core diameter, each insert bore, and each insert with digital calipers (resolution 0.01 mm), three times at different positions along the length, and report mean $\pm$ SD. Record the mass of each insert. Check that every insert seats fully with its key aligned. *Expected outcome.* Printed inserts are expected to run slightly undersized [28], so first prints may need diameter adjustment; final parts should meet the clearance target.

Test 2: Material CT Numbers

Purpose. Establish the CT number of each material, which sets the contrast available for texture. *Method.* Before full fabrication, scan small coupons of acrylic, a bonded acrylic seam, and PLA at each planned infill, three times each with the reference protocol and at 100 kV and 140 kV. Measure the mean HU and SD in a central ROI. *Expected outcome.* Acrylic should measure about +118 HU to +124 HU at 120 kV [31]. PLA brands vary widely, from about +16 HU to +124 HU when solid [31], and mean HU will fall as infill decreases [26], so each pattern's value will be recorded as its reference.

Test 3: Body Uniformity and Seam Visibility

Purpose. Confirm that the acrylic background is uniform so that differences between inserts are not caused by the body. *Method.* With all bores filled by solid acrylic plugs, scan the assembled phantom three times. Place ROIs at the center and at four peripheral positions in the shell and core and compare mean HU. Review sagittal and coronal reconstructions for visibility of any bonded seam or core-to-shell gap. *Expected outcome.* A slight center-to-edge HU difference is expected from beam hardening, larger in the assembled 400 mm $\times$ 300 mm configuration; the target is within 5 HU, the American College of Radiology uniformity limit [34], comparable to the open-source epoxy phantom [16].

Test 4: Repeatability (Test-Retest)

Purpose. Measure how much feature values change when nothing but positioning changes, which sets the noise floor for every other test. *Method.* Scan the assembled phantom five times with the reference protocol. Between scans, remove and reinsert the core and all inserts and realign the phantom. *Analysis.* For each insert and feature, calculate the coefficient of variation (CV), the intraclass correlation coefficient (ICC, two-way, absolute agreement, reported with 95 % confidence intervals) [38], and Lin's concordance correlation coefficient (CCC) [40]; Bland-Altman plots [41] will show agreement for selected features. Features with $CV \leq 10\%$ [22] and ICC above 0.75, which indicates good reliability [38], are classed as repeatable. *Expected outcome.* Most first-order features should be repeatable; some texture features are expected to be less stable, consistent with published reviews [4].

Test 5: Print-to-Print Reproducibility

Purpose. Show that a reprinted insert gives the same features as the original, so worn or damaged inserts can be replaced. *Method.* Scan the two printed copies of each pattern in the same bore position, three times each. *Analysis.* ICC between copies for each feature. *Expected outcome.* High agreement, given that comparable printed inserts reached ICC above 0.999 between independent prints [16].

Test 6: Protocol Robustness

Purpose. Characterize how each feature responds to common changes in scan settings, which is the core benchmarking function of the phantom. *Method.* Change one parameter at a time from the reference protocol, three scans per setting (Table 5). *Analysis.* For each feature, CV and ICC across settings within each parameter, and the percentage of features meeting the repeatability thresholds from Test 4. *Expected outcome.* Reconstruction kernel and slice thickness are expected to cause the largest changes, consistent with prior phantom work [6], [9].

Table 5. Protocol robustness matrix. One parameter is changed at a time while the others stay at the reference value (**bold**). Values will be adjusted to the protocols available on the client's scanner.

Parameter	Settings tested	Reason
Tube voltage	100, 120, 140 kV	Changes contrast and noise
Dose (tube current)	50 %, 100 %, 200 % of reference	Changes noise level [7]
Reconstruction kernel	Smooth, standard, sharp	Largest reported effect on texture [6]
Slice thickness	1, 2.5, 5 mm	Changes voxel size [9]
Display field of view	250, 350, 500 mm	Changes pixel size [9]
Reconstruction algorithm	Filtered back projection, iterative levels, e.g. low, medium, high	Changes noise texture

Test 7: Size Dependence

Purpose. Measure the effect of phantom size on features, the question the removable core was designed to answer. *Method.* With the same five inserts in the core, scan the core alone

(effective diameter about 180 mm) and the assembled phantom (about 346 mm), three times each. Repeat with automatic exposure control on and off, so that changes from patient size can be separated from changes in tube current. *Analysis.* For each feature, the percent difference between configurations, and a linear mixed-effects model with configuration as a fixed effect and insert as a random effect, using all repeated scans ($5 \text{ inserts} \times 3 \text{ scans} \times 2 \text{ configurations} = 30$ observations per feature). A simple paired nonparametric test across only five inserts is not used, because with $n = 5$ the smallest possible two-sided p-value from a Wilcoxon signed-rank test is 0.0625. Because hundreds of features are tested, p-values will be corrected for multiple comparisons with the Benjamini-Hochberg false discovery rate method at 0.05 [42]. CCC between configurations will identify size-robust features, as in a comparable small-versus-large phantom study [10]. *Expected outcome.* The larger configuration is expected to raise noise and change many texture features; only a fraction of features are expected to be robust to size [10].

Test 8: Texture Range and Clinical Realism

Purpose. Confirm that the insert patterns produce distinct textures and that the feature values fall in a clinically relevant range. *Method and analysis.* Principal component analysis (PCA) of features from all patterns will test whether patterns are separable and non-redundant [16]. Phantom feature values will be compared with values extracted from de-identified clinical CT scans provided by the client, or a public dataset such as The Cancer Imaging Archive, counting the fraction of features that fall between the 10th and 90th percentiles of patient values [22]. *Expected outcome.* Air-containing PLA inserts have much higher contrast than tumors, so some features may fall outside patient ranges; the optional PLA and ABS low-contrast inserts are expected to match patient ranges more closely.

Test 9: Stability Over Time

Purpose. Provide early evidence for the five-year service life requirement. *Method.* Repeat Tests 1, 2, and 4 on a reduced set of inserts at baseline and at 1 and 3 months, storing the phantom under its intended conditions. In parallel, place spare inserts in a humid environment [e.g., 40 °C and high humidity for 2 weeks] and record mass and HU change as an accelerated check on PLA moisture uptake. *Expected outcome.* Sealed inserts in dry storage should show no measurable change; this test can show short-term stability within the semester but cannot prove a multi-year life, which will be noted as future work.

Test 10: Spatial Resolution and Noise

Purpose. Verify the image quality requirements in the design specifications. *Method.* Scan a high-contrast line-pair pattern [a line-pair insert in one bore, or the client's ACR CT accreditation phantom in the same session] with the adult abdomen and high-resolution chest protocols, three times each, and record the highest resolvable spatial frequency. Measure noise as the standard deviation of HU in a uniform acrylic ROI at the center of the phantom in every session. *Analysis.* Compare resolution with the American College of Radiology minimums of 5

lp/cm for adult abdomen and 6 lp/cm for high-resolution chest imaging [34], and track noise against the baseline session, following the baseline-tolerance approach used for routine ACR phantom quality assurance [34]. *Expected outcome.* Clinical CT scanners routinely meet the ACR resolution minimums [34], so this test mainly confirms that the phantom and protocol are set up correctly and that noise stays within $\pm 10\%$ of baseline.

Sources of Error and Variance

Several sources of error are built into the test design so that they can be measured rather than assumed. Positioning and ROI placement error are measured directly by the repeated reinsertion in Test 4 and controlled with the alignment marks and a saved ROI template. Partial-volume effects at insert edges are limited by the inset ROI. Scanner drift is checked with daily quality assurance and by rescanning the reference protocol at the end of each session. Beam hardening differs between the two configurations and is quantified in Test 3. Manufacturing sources, including air gaps, the bonded seam, and filament lot differences, are addressed by Tests 1, 3, and 5. Every measurement will be repeated at least three times and reported as mean $\pm$ SD.

Statistical Analysis

Analyses will be performed in Python. Results will be reported as mean $\pm$ SD with three significant figures. ICC form and confidence intervals will be reported following published guidance [38]. Statistical significance is set at $\alpha = 0.05$ with false discovery rate correction for multiple features [42]. Raw feature tables will be placed in Appendix F, and the body of the report will present results as figures, such as heat maps of feature CV by protocol and scatter plots of features in the small versus large configuration.

Results

Results will be added after fabrication and scanning are complete.

Conclusions

The radiographic radiomics phantom for Dr. Jennie Crosby is being designed based on current client information and research from an extensive literature search. The design will be capable of evaluating radiomic features under different imaging conditions. Current phantoms on the market can evaluate image quality, but they do not provide the multiple non-homogeneous textures throughout the phantom that radiomics analysis needs. The proposed design incorporates an outer abdomen-sized shell and an inner head-sized removable core. Research also showed that CT image characteristics, including Hounsfield units, image resolution, ROI size, and scanning parameters, can affect radiomic feature reproducibility [35]. Therefore, controlling these variables and designing a phantom with consistent material placement and geometry will be important for obtaining reliable results. Moving forward, material and fabrication decisions will be finalized within budget and time constraints, and testing will begin once the prototype is

completed, following the plan in the Testing and Results section. Testing variables include Hounsfield units, spatial resolution, image noise, and radiomic feature repeatability. Based on the design's performance against these requirements, the phantom will be handed over to the client, Dr. Crosby, or refined further.

References

- [1] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: Images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, pp. 563-577, 2016, doi: 10.1148/radiol.2015151169.
- [2] F. Davnall et al., "Assessment of tumor heterogeneity: An emerging imaging tool for clinical practice?" *Insights Imaging*, vol. 3, no. 6, pp. 573-589, 2012, doi: 10.1007/s13244-012-0196-6.
- [3] R. Smith-Bindman et al., "Projected lifetime cancer risks from current computed tomography imaging," *JAMA Intern. Med.*, vol. 185, no. 6, pp. 710-719, 2025, doi: 10.1001/jamainternmed.2025.0505.
- [4] A. Traverso, L. Wee, A. Dekker, and R. Gillies, "Repeatability and reproducibility of radiomic features: A systematic review," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 102, no. 4, pp. 1143-1158, 2018, doi: 10.1016/j.ijrobp.2018.05.053.
- [5] D. Mackin et al., "Measuring computed tomography scanner variability of radiomics features," *Invest. Radiol.*, vol. 50, no. 11, pp. 757-765, 2015, doi: 10.1097/RLI.0000000000000180.
- [6] R. Berenguer et al., "Radiomics of CT features may be nonreproducible and redundant: Influence of CT acquisition parameters," *Radiology*, vol. 288, no. 2, pp. 407-415, 2018, doi: 10.1148/radiol.2018172361.
- [7] M. Meyer et al., "Reproducibility of CT radiomic features within the same patient: Influence of radiation dose and CT reconstruction settings," *Radiology*, vol. 293, no. 3, pp. 583-591, 2019, doi: 10.1148/radiol.2019190928.
- [8] A. Zwanenburg et al., "The Image Biomarker Standardization Initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping," *Radiology*, vol. 295, no. 2, pp. 328-338, 2020, doi: 10.1148/radiol.2020191145.
- [9] M. Shafiq-ul-Hassan et al., "Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels," *Med. Phys.*, vol. 44, no. 3, pp. 1050-1062, 2017, doi: 10.1002/mp.12123.
- [10] F. Bayat and C. Altunbas, "Improving the robustness of radiomic features to patient size variations in CBCT imaging for radiotherapy," *J. X-Ray Sci. Technol.*, vol. 34, no. 5, pp. 622-635, 2026, doi: 10.1177/08953996261459392.
- [11] Sun Nuclear Corp., "CIRS Model 062M electron density phantom," product datasheet. [Online]. Available: https://www.sunnuclear.com/uploads/documents/datasheets/Diagnostic/ElectronDensityPhantom_090122.pdf. Accessed: Oct. 6, 2026.
- [12] Sun Nuclear Corp., "Advanced electron density phantom," product datasheet. [Online]. Available: https://www.sunnuclear.com/uploads/documents/datasheets/Diagnostic/Advanced-Electron-Density-Phantom_052422.pdf. Accessed: Oct. 6, 2026.

- [13] D. Mackin, "Data from Lung_Phantom," The Cancer Imaging Archive, 2017. [Online]. Available: <https://wiki.cancerimagingarchive.net/x/eoCBAQ>. Accessed: Oct. 6, 2026.
- [14] B. A. Varghese et al., "Identification of robust and reproducible CT-texture metrics using a customized 3D-printed texture phantom," *J. Appl. Clin. Med. Phys.*, vol. 22, no. 2, pp. 98-107, 2021, doi: 10.1002/acm2.13162.
- [15] B. A. Varghese et al., "Technical and clinical considerations of a physical liver phantom for CT radiomics analysis," *J. Appl. Clin. Med. Phys.*, vol. 25, no. 4, p. e14309, 2024, doi: 10.1002/acm2.14309.
- [16] M. A. Daly, M. F. McNitt-Gray, and J. M. Hoffman, "An open-source CT texture phantom for radiomic imaging studies: Design, manufacture, and initial evaluation," *Med. Phys.*, vol. 53, no. 8, p. e70613, 2026, doi: 10.1002/mp.70613.
- [17] U. Mahmood et al., "Quality control of radiomic features using 3D-printed CT phantoms," *J. Med. Imaging*, vol. 8, no. 3, p. 033505, 2021, doi: 10.1117/1.JMI.8.3.033505.
- [18] M. Bach et al., "3D-printed iodine-ink CT phantom for radiomics feature extraction: Advantages and challenges," *Med. Phys.*, vol. 50, no. 9, pp. 5682-5697, 2023, doi: 10.1002/mp.16373.
- [19] J. Y. Im et al., "PixelPrint 4D: A 3D printing method of fabricating patient-specific deformable CT phantoms for respiratory motion applications," *Invest. Radiol.*, vol. 60, no. 10, pp. 636-646, 2025, doi: 10.1097/RLI.0000000000001182.
- [20] J. S. Stringer III, H. Spitz, P. Sandwall, M. Lamba, and S. Glover, "Calibration phantom for radiotherapy," U.S. Patent 12 089 982 B2, Sep. 17, 2024.
- [21] V. Nardone et al., "Delta-radiomics increases multicentre reproducibility: A phantom study," *Med. Oncol.*, vol. 37, no. 5, p. 38, 2020, doi: 10.1007/s12032-020-01359-9.
- [22] L. Rinaldi et al., "HeLLePhant: A phantom mimicking non-small cell lung cancer for texture analysis in CT images," *Phys. Med.*, vol. 97, pp. 13-24, 2022, doi: 10.1016/j.ejmp.2022.03.010.
- [23] E. Samei, J. Hoyer, Y. Zheng, J. B. Solomon, and D. Marin, "Design and fabrication of heterogeneous lung nodule phantoms for assessing the accuracy and variability of measured texture radiomics features in CT," *J. Med. Imaging*, vol. 6, no. 2, p. 021606, 2019, doi: 10.1117/1.JMI.6.2.021606.
- [24] J. T. Bushberg, J. A. Seibert, E. M. Leidholdt Jr., and J. M. Boone, *The Essential Physics of Medical Imaging*, 4th ed. Philadelphia, PA, USA: Wolters Kluwer, 2021.
- [25] J. M. Boone et al., "Size-specific dose estimates (SSDE) in pediatric and adult body CT examinations," American Association of Physicists in Medicine, College Park, MD, USA, AAPM Rep. 204, 2011.
- [26] M. Kozee et al., "Methodology for computed tomography characterization of commercially available 3D printing materials for use in radiology/radiation oncology," *J. Appl. Clin. Med. Phys.*, vol. 24, no. 6, p. e13999, 2023, doi: 10.1002/acm2.13999.
- [27] J. P. O. Bustillo et al., "Characterization and evaluation methods of fused deposition modeling and stereolithography additive manufacturing for clinical linear accelerator photon and electron radiotherapy applications," *Phys. Med.*, vol. 130, p. 104904, 2025, doi: 10.1016/j.ejmp.2025.104904.

- [28] M. R. Jones et al., "A multicentre audit of the geometric accuracy and water equivalence of 3D printed thermoplastic objects for use in radiotherapy," *Phys. Imaging Radiat. Oncol.*, vol. 39, p. 101006, 2026, doi: 10.1016/j.phro.2026.101006.
- [29] Y. Li et al., "The impact of phantom design and material-dependence on repeatability and reproducibility of CT-based radiomics features," *Med. Phys.*, vol. 49, no. 3, pp. 1648-1659, 2022, doi: 10.1002/mp.15491.
- [30] RaySafe, "Nested adult head and body CTDI phantom," product leaflet 5200117-1.0. [Online]. Available: https://www.raysafe.com/sites/default/files/2021-04/5200117-1.0_RaySafe_Nested_Adult_Head_and_Body_CTDI_leaflet_EN.pdf. Accessed: Oct. 6, 2026.
- [31] F. Okkalidis et al., "Characterization of commercial and custom-made printing filament materials for computed tomography imaging of radiological phantoms," *Technologies*, vol. 12, no. 8, p. 139, 2024, doi: 10.3390/technologies12080139.
- [32] MAS Epoxies, "Deep pour epoxy resin," product page. [Online]. Available: <https://masepoxies.com/products/deep-pour-epoxy-resin>. Accessed: Oct. 6, 2026.
- [33] Let's Resin, "Deep pour epoxy resin," product page. [Online]. Available: <https://letsresin.com/products/deep-pour-epoxy-resin>. Accessed: Oct. 6, 2026.
- [34] M. A. Hobson, E. T. Soisson, S. D. Davis, and W. Parker, "Using the ACR CT accreditation phantom for routine image quality assurance on both CT and CBCT imaging systems in a radiotherapy environment," *J. Appl. Clin. Med. Phys.*, vol. 15, no. 4, p. 4835, 2014, doi: 10.1120/jacmp.v15i4.4835.
- [35] Y. Soleymani, Z. Valibeiglou, M. Fazel Ghaziani, A. Jahanshahi, and D. Khezerloo, "Radiomics reproducibility in computed tomography through changes of ROI size, resolution, and Hounsfield unit: A phantom study," *Radiography*, vol. 30, no. 6, pp. 1629-1636, 2024, doi: 10.1016/j.radi.2024.10.003.
- [36] ePlastics, "IPS Weld-On 40 pint kit," product page. [Online]. Available: <https://www.eplastics.com/ips40-pt>. Accessed: Oct. 6, 2026.
- [37] ACRYLITE, "Annealing ACRYLITE acrylic sheet," fabrication manual. [Online]. Available: <https://www.acrylite.co/resources/fabrication-manuals/annealing-acrylite-acrylic-sheet>. Accessed: Oct. 6, 2026.
- [38] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *J. Chiropr. Med.*, vol. 15, no. 2, pp. 155-163, 2016, doi: 10.1016/j.jcm.2016.02.012.
- [39] J. J. M. van Griethuysen et al., "Computational radiomics system to decode the radiographic phenotype," *Cancer Res.*, vol. 77, no. 21, pp. e104-e107, 2017, doi: 10.1158/0008-5472.CAN-17-0339.
- [40] L. I.-K. Lin, "A concordance correlation coefficient to evaluate reproducibility," *Biometrics*, vol. 45, no. 1, pp. 255-268, 1989.
- [41] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 1, no. 8476, pp. 307-310, 1986.

- [42] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. R. Stat. Soc. Ser. B*, vol. 57, no. 1, pp. 289-300, 1995.

Appendix

Appendix A: Product Design Specifications

Function

Radiomics, the extraction and analysis of high-dimensional quantitative texture features from medical images, has shown promise in oncology for prognosis, treatment response assessment, and outcome prediction. However, radiomic features are highly sensitive to image acquisition specifications and hardware-specific characteristics. In simpler terms, models developed on one imaging platform cannot be reliably applied to another without proper validation and benchmarking. This benchmarking requires a phantom containing a variety of heterogeneous, tissue-like textures; however, commercially available phantoms lack a sufficient range of materials and are not designed for this purpose.

Dr. Jennie Crosby requires a phantom that incorporates multiple clinically relevant texture patterns, is compatible with imaging protocols for both small anatomy (e.g., head) and large anatomy (e.g., abdomen), and is constructed from materials that can maintain mechanical stability over a multi-year period to support longitudinal radiomic feature evaluation across radiographic imaging systems.

Client Requirements

Dr. Jennie Crosby outlined several key specifications for the development of the Radiographic Radiomics Phantom:

1. Incorporate a variety of different textures to represent the variability and non-homogeneous characteristics of real human tissue.
2. Size similar to that of a human abdomen, and the inner core should resemble the size of a human head.
3. Withstand research usage for multiple years and ideally provide consistent results through repeated usage.
4. The approximate size of the outer abdomen shell should be a diameter of 38 cm and a height of 20 cm, and the head should have a diameter of approximately 16 cm and a height of 25 cm.

Design Requirements

1. Physical and Operational Characteristics

a. Performance Requirements

- Each phantom part should replicate the body part's mechanical properties with tissue texture. The non-homogeneous nature of the phantom material should be visible and representable in the imaging done by clinical CT imaging systems.

- The phantom should have a long lifespan, meaning that the material shall not have noticeable degradation over the intended period of usage time (around five years; see section d, Life in Service). The material for emulating the texture of a phantom body part should reliably maintain its texture for at least "a couple of years."

b. Safety

- The phantom must contain materials that are safe to use under normal clinical conditions and must not contain any radioactive materials that could cause harm to the user or be left as residue on the machine [1].
- The user should be able to easily and manually transport the phantom through a typical indoor, sterile laboratory environment.
- All sharp edges and exposed hazardous materials should be either removed or concealed before handing over to the client. All components must be identified clearly without instruction and easily assembled, with a removable core that remains stable and prevents movement during scanning.

c. Accuracy and Reliability

- Fabrication and manufacturing reliability: The phantom must maintain consistent dimensions and placement of materials between scans and produce repeatable CT texture characteristics. Manufacturing inconsistencies should be minimized, and controlled material compositions should be used to improve reproducibility.
- Testing reliability and accuracy: The phantom should be tested repeatedly in CT scans to determine consistency and accuracy. The assembly of the outer shell should also contain no gaps or voids in the assembly of material, as this can cause errors in images and false data.

d. Life in Service

- At least 5 years. The phantom must maintain its geometry, material composition, and imaging performance during repeated use over an extended period of time. Existing CT phantoms are intended for long-term clinical and research use. Competing designs have acquired consistent and accurate scans after a 7-month period [2].

e. Shelf Life

- The phantom must maintain its physical and imaging properties during storage in recommended conditions. In reference to section d, the phantom should last for approximately 5 years in service. To achieve this, materials that do not significantly degrade, wear out, or change during normal storage should be used to create all parts of the phantom.

f. Operating Environment

- The phantom must be suitable for an indoor laboratory environment and hospital imaging. Repeated handling of the phantom must be taken into account, as well as the ability to easily place and remove the phantom from the imaging equipment.

- Dr. Crosby will be the intended user of this phantom and will be storing it in her lab. Possible transportation of the phantom to other labs is anticipated.

g. Ergonomics

- The phantom should be designed for safe and easy handling, staining, and removal of the core by the user when possible. The removable pieces of the phantom should have an easily accessible grasping and positioning feature that prevents any harm to the device itself.
- The shape of the phantom also must be considered for ergonomics, since an abnormal shape or cut of any section may require repositioning to acquire a proper image. An elliptical cylinder will be the main form of the phantom, as it should decrease any physical labor associated with moving and positioning the phantom on the CT machine.

h. Size

- Multiple different scales of human anatomy should be represented. The phantom should be able to approximately represent a human head as the core and a human torso as the width and outer diameter dimensions. Dimensions shall be selected based on anthropometric and medical image references [3].
- The client requested the following dimensions for the phantom: the head should have a diameter of 16 cm by 25 cm height, the abdomen should have a diameter of 38 cm by 20 cm height, and the depth should be 30cm.

i. Weight

- There are no specified weight requirements for the device, other than that it must be movable by the client. This limit should not be exceeded as long as standard materials are used and the current sizing guidelines are followed. The density of the material should be roughly that of water (1 g/mL) [1].

j. Materials

- Materials must be cost-effective, corresponding to the size of the design and the budget. However, the material must maintain all shaping and mechanical properties even after handling for its entire service life.
- The main material options for the base of the phantom are acrylic and PLA 3D-printed polymer. The 3D printing approach is a cost-effective and time-efficient solution; however, it is not as standard and reputable as acrylic [2]. The final design will likely be composed of acrylic.
- The tissue-replicating material should have an electron density similar to water to help mimic tissue under CT scanning.

k. Aesthetics, Appearance, and Finish

- The finish of the phantom should be clean and smooth. There must be no gaps or voids in the assembly of any modules to prevent errors in scanning.
- The removable core of the phantom should be uniform in size and free of harsh edges.

- The entire system should be as simple and organized as possible for ease of use, without sacrificing functionality. The entire phantom should look cohesive and should carry a professional aesthetic.

2. Product Characteristics

a. Quantity

- Only one phantom will be fabricated and provided to the client upon completion. The one phantom will have multiple uses and can be used to mimic a range of different human anatomy depending on the client's imaging goals.

b. Target Product Cost

- With consideration of the various components required for the phantom already available at the university, and an evaluation of prices for additional material, a target cost for the project is approximately \$400. The main cost factor of this design is the acrylic base of the phantom.
- There are other materials with similar properties and behaviors that are sold at a lower price, and the final material selection may significantly change the estimated cost. This figure includes the \$50 stipend given for the project by the department as well as further monetary support that will be provided by the client if and when required.
- Project specifications can be adjusted to lower cost to fit the budget if necessary.

3. Miscellaneous

a. Standards and Specifications

The texture phantom design will need to be usable across various CT systems. To ensure that medical and research decisions based on these CT images are sound, the phantom design will need to comply with the following standards:

- ASTM E1935-26, which details the procedures and requirements for calibrating and measuring material densities from CT images [4].
- IEC 61223-3-5:2019, which details acceptance and constancy tests of the imaging performance of CT X-ray equipment [5].

b. Customer

- The user of this device will be Dr. Jennie Crosby in the Department of Medical Physics at the University of Wisconsin-Madison. This device is for research and testing purposes with CT imaging technology.

c. Patient-Related Concerns

- Degradation reduces the quality of medical decisions. Therefore, it is important that the phantom is able to last for a long time with no degradation while also being able to mimic various body parts. Any changes in the phantom should be easily identifiable for clients.
- Modules should be safe for handling. No toxic materials. The weight of each module should be easy to handle.

d. Competition

- This radiomic phantom should last longer and provide more consistent results in CT imaging than biological phantoms [6].
- CT phantoms on the market tend to hold a five-year warranty. Such phantoms are typically constructed from acrylic or lower-density common plastics such as low-density polyethylene [7].

Appendix A References

- [1] N. Nonoura, M. Yamada, N. Nagasawa, T. Yamashita, H. Hisazumi, and H. Yamazaki, "Influence of body size on the accuracy of lumbar bone mineral density by dual-energy CT: A phantom study," *Eur. J. Radiol.*, vol. 200, p. 112878, 2026, doi: 10.1016/j.ejrad.2026.112878.
- [2] A. P. Hariyanto, E. Endarko, M. D. Mardat, L. Almira, and K. H. Ng, "Development and dosimetric evaluation of a low-cost 3D-printed PLA head CTDI phantom and CT parameter comparison with commercial phantom for CT quality assurance," *Radiography*, vol. 32, no. 7, p. 103543, 2026, doi: 10.1016/j.radi.2026.103543.
- [3] S. J. Lokitz, J. M. Wilson, and T. G. Turkington, "Development of a phantom to simulate a wide range of body sizes," in *Proc. IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 5, 2007, pp. 3713-3717, doi: 10.1109/NSSMIC.2007.4436928.
- [4] Standard Practice for Calibrating and Measuring CT Density, ASTM E1935-26, Jul. 16, 2026. [Online]. Available: <https://compass.astm.org/content-access?contentCode=ASTM%7CE1935-26%7Cen-US>
- [5] Evaluation and Routine Testing in Medical Imaging Departments, Part 3-5: Acceptance Tests and Constancy Tests, Imaging Performance of Computed Tomography X-Ray Equipment, IEC 61223-3-5:2019, 2019. [Online]. Available: <https://webstore.iec.ch/publication/59789>
- [6] J. Cammin, G. S. K. Fung, E. K. Fishman, J. H. Siewerdsen, J. W. Stayman, and K. Taguchi, "A biological phantom for evaluation of CT image reconstruction algorithms," *Proc. SPIE*, vol. 9033, p. 903307, 2014, doi: 10.1117/12.2043714.
- [7] Sun Nuclear, "CT ACR 464 phantom," product page. [Online]. Available: <https://www.sunnuclear.com/products/ct-acr-464-phantom>. Accessed: Sep. 19, 2026.

Appendix B: Materials and Budget

Costs marked as estimates will be replaced with supplier quotes. Free and borrowed items are listed with a cost of \$0. All total costs are estimated.

Item	Description	Vendor	Part #	Qty	Cost each [\$]	Total [\$]
Cast acrylic blank, shell	Clear cast PMMA, 410 × 310 × 165 mm (or bonded slabs)	[TBD]	[TBD]	1	\$440-1,130 est.	600
Cast acrylic blank, core	Clear cast PMMA, 190 × 190 × 165 mm	[TBD]	[TBD]	1	130-320 est.	200

Item	Description	Vendor	Part #	Qty	Cost each [\$]	Total [\$]
Weld-On 40 cement	Two-part reactive acrylic cement, 1 pt kit (only if slabs are bonded)	ePlastics	IPS40-PT	1	46.71	46.71
PLA filament	1.75 mm, 1 kg, single lot	[TBD]	[TBD]	[3]	70	210
ABS filament (optional)	1.75 mm, 1 kg	[TBD]	[TBD]	[1]	50	50
Machine shop and annealing	CNC mill or router, lathe, and oven time	ECB Lab or Wendt Lab	N/A	1	\$0	\$0
Miscellaneous	-	[TBD]	N/A	[TBD]	\$300	\$300
					Total	1406.71

Appendix C: Engineering Drawings

[SolidWorks drawings of the shell, core, and inserts to be added.]

Appendix D: Print Settings

[Filament brand and lot, nozzle diameter, layer height, temperatures, speed, infill pattern and percentage, and build orientation for each insert.]

Appendix E: Feature Extraction Settings

[PyRadiomics version and parameter file: resampling, bin width, ROI definition, and enabled feature classes.]

Appendix F: Raw Data

[Raw feature tables to be added after testing.]