Check for updates

Received: 7 April 2026 | Revised: 29 June 2026 | Accepted: 6 July 2026

DOI: 10.1002/mp.70613

RESEARCH ARTICLE

MEDICAL PHYSICS

An open-source CT texture phantom for radiomic imaging studies: Design, manufacture, and initial evaluation

Morgan A. Daly | **Michael F. McNitt-Gray** | **John M. Hoffman**

Department of Radiological Sciences, University of California, Los Angeles, California, USA

Correspondence
John M. Hoffman, Department of Radiological Sciences, University of California, Los Angeles, CA 90024, USA.
Email: jmhoffman@mednet.ucla.edu

Funding information
National Institutes of Health, Grant/Award Number: T32EB002101-46

Abstract

Background: The lack of standardized, reproducible, and publicly available reference objects limits the interpretability and reproducibility of radiomic imaging studies across institutions. Current handcrafted phantoms lack the precision, reproducibility, and accessibility required for large-scale, multi-center validation studies.

Purpose: This work presents the design, manufacture, and initial evaluation of an open-source modular CT texture phantom intended as a reproducible, accessible platform for quantitative imaging research.

Methods: A modular phantom system was designed comprising an epoxy-filled cylindrical body and eight cavities for interchangeable 3D-printed texture inserts. The inserts utilized mathematically defined geometric primitives (e.g., periodic rod lattices, and random shape distributions) to provide standardized and reproducible printable objects. The design files for the phantom body and inserts were released in the public domain. Initial evaluation of the phantom body and inserts was done by manufacturing three copies of the phantom body and four sets of the inserts and then performing four assessments: (a) phantom body uniformity quantified through intra- and inter-phantom HU evaluations; (b) texture insert reproducibility based on an evaluation of the intraclass correlation coefficient (ICC) and coefficient of variation (CV) of radiomic feature values across insert replicates; (c) evaluating that the set of inserts provide non-redundant texture stimuli by performing a principal component analysis (PCA) on radiomic features extracted from each insert; and (d) evaluating the sensitivity of inserts to CT imaging protocol variation across two dose levels and three reconstruction kernels.

Results: The manufacturing pipeline demonstrated high intra- and inter-phantom uniformity across all four phantom bodies (<5 HU). The texture inserts demonstrated exceptional reproducibility of radiomic features across independent manufacturing runs with a ICC > 0.999 both across insert set prints and between slices. The PCA demonstrated that the set of inserts provided distinct texture stimuli. The protocol variation investigation demonstrated that the inserts showed anticipated variation in texture-based features across CT dose levels and reconstruction kernels.

Conclusions: We present an open-source phantom design and manufacturing approach capable of producing high-fidelity, reproducible reference objects for radiomic imaging studies using consumer-grade equipment and low-cost materials. The public domain release of design files removes licensing barriers and enables any research group to replicate, extend, and share phantom

All rights reserved, including rights for text and data mining and training of artificial intelligence technologies or similar technologies.
© 2026 American Association of Physicists in Medicine.

designs, supporting standardized phantom-based investigation of radiomic feature behavior across institutions.

KEYWORDS
3D printing, CT phantoms, quantitative imaging, radiomics

1 | INTRODUCTION

Quantitative imaging in computed tomography (CT) has produced promising results in medical imaging research, with applications ranging from tumor characterization to treatment response assessment. The field of radiomics, which involves the high-throughput extraction of quantitative features from regions of interest in medical images, has demonstrated potential to improve diagnostic accuracy and clinical decision making.[1–3] Texture-based radiomic features in particular, which capture spatial patterns of intensity variation within a region of interest, have shown utility across a range of clinical applications including the CT image-based diagnoses and prognoses in lung nodules,[4–6] interstitial lung disease,[7,8] and bone.[9–11] By converting medical images into mineable data, radiomics may enable the discovery of imaging biomarkers that reflect underlying pathophysiology not apparent to visual inspection alone.

However, the clinical translation of radiomic biomarkers has been impeded by challenges in reproducibility and standardization. Multiple studies have documented substantial variability in radiomic features due to differences in CT acquisition parameters, reconstruction algorithms, and scanner characteristics.[12–15] This variability undermines the generalizability of radiomic models and limits their utility in multi-institutional studies and general clinical practice. The need for improved standardization approaches in efforts to develop imaging biomarkers has been highlighted by the Quantitative Imaging Biomarkers Alliance (QIBA) and the Image Biomarker Standardisation Initiative (IBSI).[16–18]

One barrier to the advancement of quantitative imaging is the lack of accessible standardized phantoms for validating radiomic measurements. Conventional CT phantoms, such as the American College of Radiology (ACR) CT accreditation phantom, serve important roles in routine quality assurance; however, they typically consist of uniform materials that are not suitable for evaluating texture-based radiomic features.[19–21] This limitation has led researchers to develop custom "handcrafted" phantoms consisting of cartridges filled with different materials to produce textures, such as the Credence Cartridge Radiomics (CCR) Phantom.[22–24] Because these handcrafted phantoms rely on non-standardized, often organic, granular filler material, they are structurally irreproducible. This makes it challenging to directly compare results across different CT scanners, technical parameter settings, or institutions, as differences could be due to variation in the physical texture objects rather than strictly related to the variable of interest. Consequently, independent validation or extension of experimental results based on these handcrafted phantoms becomes reliant on transferring the original phantom between institutions, requiring logistical coordination which may not always be feasible.

Recent advances in three-dimensional (3D) printing technology have made it a promising solution for producing the custom phantoms used for studying the behavior of radiomic features in CT.[25] This approach enables the physical realization of computer-designed textures and shapes– both geometric and anatomically derived from clinical image data– providing a pathway to a substantially more reproducible manufacturing approach than traditional methods, such as filled cartridges. Physicists can utilize computer-aided design software and programming tools such as Python to generate the 3D models for printing, which allows for custom, reproducible task-specific textures that can then be shared digitally with colleagues. While 3D printed infill has already been included as a subset of cartridges in some CT radiomics cartridge phantoms,[22,26,27] this provides a limited set of possible textures. Other 3D-printed phantom approaches have focused on directly printing entire phantom bodies or employed sophisticated approaches to print anatomically realistic phantoms for specific clinical applications.[28–30] These approaches often require professional-grade equipment, materials, and expertise that may not be readily available to many research groups. More importantly, these approaches were not developed with widespread reproducibility as a goal, and the designs have not been made publicly available.

The lack of a shared, reproducible phantom represents a substantial gap in the radiomics workflow. Without common reference objects, it becomes challenging to determine whether observed variations in radiomic measurements arise from biological differences, technical factors, or phantom inconsistencies. To address this gap, we first present the design and manufacturing approach for an open-source, modular CT phantom body and set of eight texture inserts. We then provide an initial evaluation of the reproducibility of four manufactured phantom bodies based on measurements of uniformity, as well as the reproducibility across three replicate sets of inserts based on measurements of radiomic features. Finally, we characterize the phantom's utility as a platform for radiomic imaging studies by evaluating that the set of inserts provide non-redundant

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

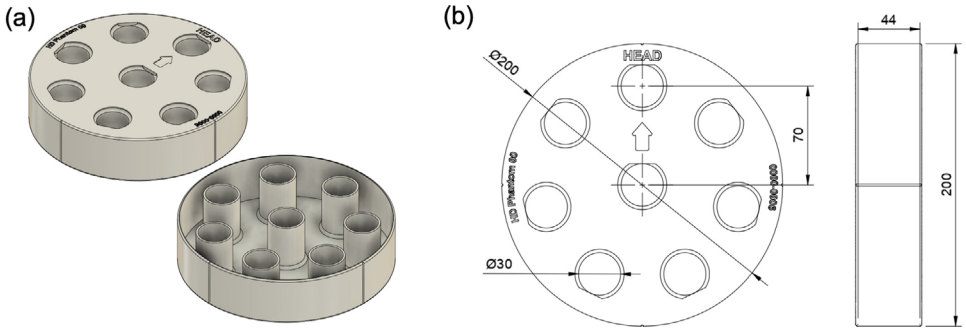

FIGURE 1 Schematic of phantom body Renderings (a) and schematics (b) of the 3D printed phantom body shell design. The shell is printed with a depth of 44 mm to allow for shrinkage of the epoxy resin to a target depth of 40 mm. Both show the alignment features, including the keyed design for the insert cavities, grooves for laser leveling and alignment, and labels for positioning.

texture stimuli based on measured radiomic feature values, and evaluating the sensitivity of the inserts to CT imaging protocol variation.

2 | MATERIALS AND METHODS

The primary focus of this work is the scientific evaluation and validation of the proposed phantom system. Accordingly, Section 2.1 presents the physical properties of the platform, and Section 2.2 describes the evaluations performed to assess the reproducibility of the phantom manufacturing approach, as well as the phantom's utility as a platform for radiomic imaging studies. Additional manufacturing instructions and methods are provided in Appendices A, B. The digital phantom design components (STL files) have been made publicly available at https://doi.org/10.5281/zenodo.21252065.[31]

2.1 | Phantom design and physical realization

2.1.1 | Phantom body specification and manufacturing

The phantom body consists of a hollow, cylindrical 3D-printed shell filled with a liquid casting resin (200 mm diameter × 40 mm depth, shown in Figure 1). It features eight cylindrical cavities (30 mm diameter) designed to accept the modular texture inserts: seven arranged peripherally on a 70 mm radius circle, and one located centrally. The geometric configuration was optimized to minimize beam hardening and photon starvation artifacts by ensuring that no more than two (potentially high-density) inserts align along any single ray path when the phantom is positioned at the scanner isocenter. The phantom body dimensions were scaled to match standard commercial modules (e.g., Gammex 464, Sun Nuclear, Melbourne, FL), ensuring cross-compatibility with existing clinical positioning equipment and scatter-inducing body rings. A photo of a completed phantom body is shown in Figure 2.

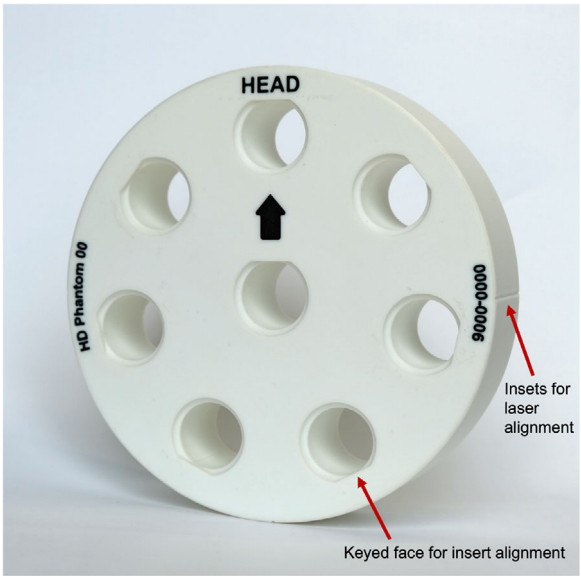

FIGURE 2 Manufactured phantom body. Shows the alignment features, including the keyed design for the insert cavities, grooves for laser leveling and alignment, and labels for positioning.

The specific properties of phantoms produced using this method will depend on the materials used for construction. This work employs polylactic acid (PLA) for printing, which has been characterized as having CT attenuation values broadly similar to soft tissue, with reported HU values ranging from 115 to 200 HU when printed with 100% infill, varying linearly with density.[30,32,33] Prior work has found that the attenuation properties of PLA are stable over time and do not require special storage conditions, suggesting the inserts are suitable for longitudinal use.[32] The phantom body is filled using a clear, two-part, slow-curing liquid epoxy resin. A characterization of clear epoxy resin for phantom applications reported measurements of density as 1.17 g/cm^3 and an average CT number of 116±6.7 HU, although this was a UV-cured 3D printed resin rather than a two-part liquid resin.[34] The resulting water equivalent diameter for the phantom was 188 mm. The materials utilized in this work are not expected to be water equivalent.

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

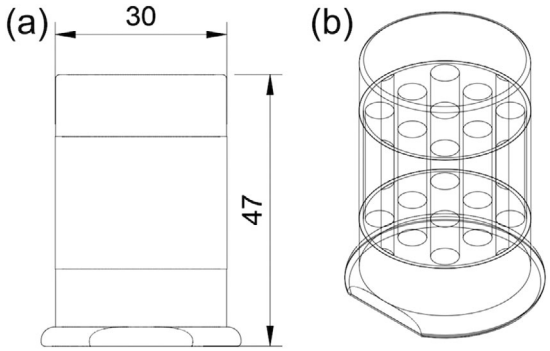

FIGURE 3 Schematics of the 3D printed modular inserts. The measurements of the insert exterior are shown in (a), with the corresponding wireframe showing the interior of an example texture design in (b). Both show the keyed lid design that ensures reproducible positioning of the inserts within the phantom body.

This hollow-shell-and-cast approach reduces total print time, mitigates thermal warping risk associated with large-volume fused deposition modeling (FDM) prints, and yields a more uniform attenuating background than 3D-printed infill; full material specifications and casting procedures are provided in Appendix A.

Several phantom and insert setup constraints were integrated directly into the physical design to maximize the reproducibility of the phantom alignment and positioning across imaging sessions. External insets (or "grooves") for horizontal and vertical alignment with the CT scanner's external laser system were included to facilitate consistent phantom placement at isocenter. The phantom surface features a directional arrow and the text "HEAD" to enforce consistent superior-inferior orientation and scan direction. Importantly, the cavities for the modular inserts feature a flat, mechanically keyed face that permits only a single, predetermined orientation for each test object. This eliminates rotational variability during physical teardown and re-assembly, isolating sources of texture variance to desired experimental variables.

2.1.2 | Modular insert design and 3D printing

The modular inserts' exterior dimensions were designed to match the phantom body, including a "lid" (a 2 mm protrusion at the base of the insert, shown in Figure 3), to ensure consistent insertion depth, and a keyed face in the lid to ensure reproducible angular orientation across experimental setups. Test objects of interest (i.e., 3D-printable shapes and textures) were then modeled within the cylindrical insert body. A schematic is shown in Figure 3.

The eight texture inserts were designed as binary density distributions (utilizing 3D printed material and air), shown in Figure 4. All textures were extruded uniformly in the z-direction to eliminate partial-volume effects along the longitudinal length of the insert, ensuring that minor translational offsets in phantom setup or slice selection during ROI placement do not introduce slice-to-slice morphological variations during the reproducibility assessment.

Six inserts (Figure 4, Textures 2–7) were designed as mathematically defined geometric primitives: periodic rectangular and circular grid lattices of varying pitch (2.5–8.0 mm) and feature size (0.75–7.0 mm). Designs were intended to provide some range in texture appearance, such as scale, ratio of 3D printed material to air, and rounded or squared shapes. These regular structures were selected because their parameterization provides both a known digital ground truth and flexibility for future experiments. By defining the geometries through discrete variables (e.g., rod diameter, lattice spacing), the scale and spatial frequency of the textures can be easily adjusted as an additional experimental variable.

One insert was designed with a stochastic distribution of cylindrical elements with diameters ranging from 0.5 mm to 2.0 mm (Figure 4, Texture 1). Unlike handcrafted approaches to producing stochastic textures (e.g., granular fill materials), this insert retains a deterministic digital model, providing a reproducible and shareable ground truth despite its randomized appearance. This serves as a proof-of-concept for the platform's extensibility toward more clinically realistic texture models. Finally, a solid 100% infill insert was included as a reference with no underlying structural heterogeneity (Figure 4, Texture 8).

All inserts were 3D printed using PLA, with air as the texture background. Definitions and print settings for all inserts are provided in Appendix B.

2.2 | Evaluation of manufacturing and metrology

To establish the viability of this design and manufacturing approach for radiomics applications, four phantom bodies and three modular insert sets were produced and evaluated. Manufacturing consistency was established via phantom body uniformity measurements (Section 2.2.1) and radiomic reproducibility evaluation (Section 2.2.2). Utility of the platform for radiomics investigations is established via principal component analysis (Section 2.2.3 to ensure different radiomic signatures are expressed by our inserts, as well as ensuring features show response to realistic CT protocol variation (Section 2.2.4).

All images were acquired in helical mode with a Siemens Definition AS 64 (Siemens Healthineers, Forchheim, Germany) using a tube voltage of 120 kV, fixed tube current, and a pitch of 1.0. All images were reconstructed at a slice thickness of 1.0 mm, a reconstruction diameter of 219 mm, and a matrix size of 512 $\times$ 512. The tube current and reconstruction kernel used are specified in Table 1. Doses were selected to be

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

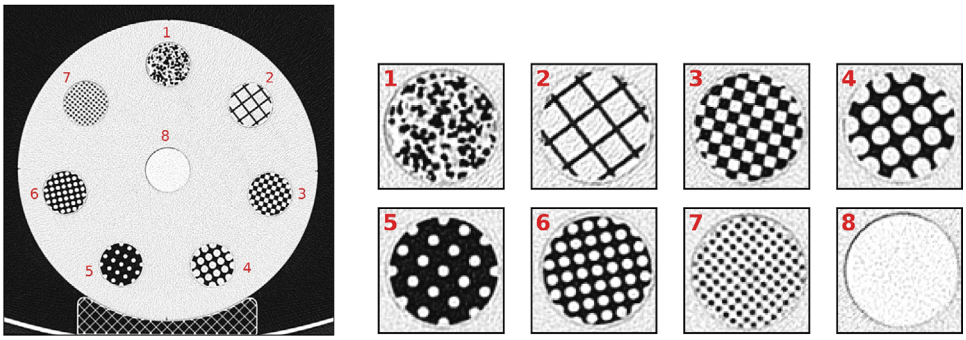

FIGURE 4 Images of the eight texture patterns that were designed and manufactured. The full CT image is shown left, with the inserts labeled. An enlarged view of the inserts is shown at right. Images are shown with window width 1300 HU, window level -450 HU.

TABLE 1 Imaging protocols (tube current and reconstruction kernel) used for each experiment.

Experiment	Tube current [mA]	Reconstruction kernel
Phantom body spatial uniformity (Section 2.2.1)	360	Br46f
Texture insert reproducibility (Section 2.2.2)	360	Br62f
Texture insert separability (Section 2.2.3)	360	Br62f
Sensitivity to imaging protocol (Section 2.2.4)	60, 360	Br32f, Br46f, Br62f

consistent with our institution's routine chest protocol (≈10–15 mGy) and a low-dose lung screening protocol (≈2 mGy).

2.2.1 | Phantom body spatial uniformity

The spatial uniformity of the epoxy phantom bodies was measured to assess whether the casting process produced a homogeneous, reproducible attenuating background. The four phantom bodies were scanned using a tube current of 360 mA ($CTDI_{vol}$ of 13.68 mGy) and images were reconstructed using a Siemens Br46f kernel. Insert holes were left unfilled; considering the modest phantom size (20 cm) and robust scatter rejection mechanisms on modern diagnostic CT scanners (anti-scatter grids, software correction), scatter-related effects on uniformity measurements were assumed to be negligible. Uniformity within each phantom body (intra-phantom uniformity) was evaluated through HU measured from different regions of interest (ROIs). Two sets of seven ROIs were placed at 40 mm and 80 mm radial distances to capture potential differences between the interior and periphery positions within the phantom. Three slices were measured within each phantom to capture potential axial heterogeneity in the *z*-direction. The measurements were then also used to assess the uniformity across four manufactured phantom bodies (inter-phantom uniformity).

The acceptance criteria was based on the ACR CT accreditation guidelines for uniformity, which uses a maximum absolute deviation of ±5 HU (and not exceeding ±7 HU) between ROIs in a homogeneous phantom.[35] Though ACR uniformity only assesses differences between center and each periphery position in an assumed uniform object, we apply this ±5 HU tolerance to the epoxy bodies to serve as a benchmark for phantom body homogeneity and agreement.

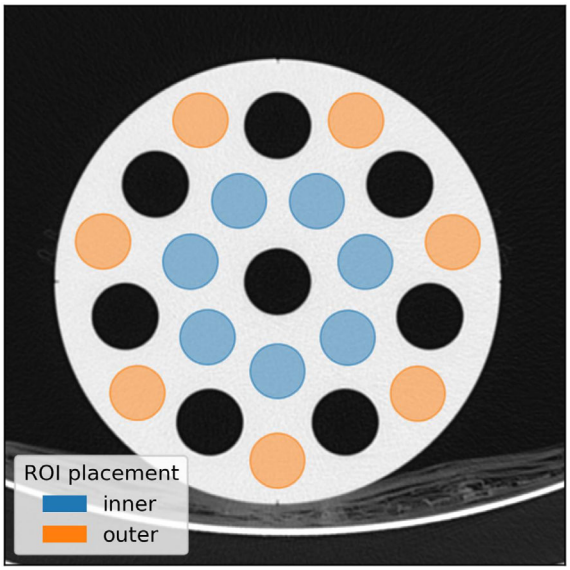

FIGURE 5 Placement of uniformity ROIs. Shown with window width 1300 HU, window level -450 HU.

Consistent ROI placement across all scans was ensured through the use of a 2D programmatic mask based on the digital ground-truth geometry, as shown in Figure 5; each mask was manually aligned with the target slice. On each slice, fourteen circular ROIs (24 mm diameter) were extracted: seven positioned at a

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

radial distance of 40 mm ("inner") and seven at 80 mm ("outer") from the phantom center. To capture potential axial heterogeneity, measurements were extracted from three image slices per phantom: the central slice, and slices offset by +5 mm and -5 mm. This sampling strategy was selected to provide the minimum necessary data to calculate standard deviation and isolate z-axis variance while maintaining an efficient manual segmentation workflow.

The mean HU was calculated from each ROI. The overall mean HU for each phantom was reported ($\mu_{overall}$, averaged over all ROIs within a phantom), as well as the inner mean (μ_{inner}, averaged over the ROIs at the inner position) and outer mean (μ_{outer}, averaged over the ROIs at the outer position) to evaluate radial gradients. The intra-phantom uniformity (Δ_{max}) was defined as the maximum absolute difference in mean CT number between any two ROIs within a given phantom. To assess the reproducibility of the liquid-casting pipeline across independent builds, inter-phantom uniformity was evaluated by calculating the maximum absolute difference in $\mu_{overall}$ between any pair of the four manufactured bodies.

Additionally, to evaluate both inter-phantom and intra-phantom equivalence, a pairwise two one-sided tests (TOST) procedure was applied.[36] The TOST procedure was performed comparing the inner and outer ROIs within each individual phantom body to verify the absence of radial manufacturing gradients or severe beam hardening artifacts, as well as across all included phantom bodies to evaluate inter-phantom reproducibility. The TOST procedure tests whether the observed difference between groups is small enough to be considered clinically irrelevant— here, whether phantom-to-phantom and inner-to-outer differences fall within the ±5.0 HU criteria. To account for potentially unequal variance between phantom bodies and between radial regions, the procedure used Welch's equations for standard error and degrees of freedom.

2.2.2 | Texture insert reproducibility

To evaluate the reproducibility of the 3D-printing manufacturing approach, radiomic feature extraction was performed and compared across three independently manufactured sets of phantom inserts. While assessing radiomic stability serves as an indirect surrogate for evaluating the microscopic dimensional accuracy of the printed geometries, it represents a critical metric for validation; the primary utility of this phantom is to standardize quantitative imaging pipelines. Consequently, agreement of the extracted radiomic features across the replicate inserts was established as a benchmark for manufacturing reproducibility.

Three insert sets were independently printed and imaged with 360 mA ($CTDI_{vol}$ of 13.68 mGy) and a Br62f reconstruction kernel. This high-frequency kernel was selected to maximize spatial resolution and preserve small details and edges within the textures, more accurately reflecting the true physical dimensionality of the printed objects. Between acquisitions the phantom was removed from the CT couch, all inserts were extracted and replaced with the next set, and the assembly was re-positioned using the scanner's external laser system. This approach ensures that measurements reflected the combined, realistic sources of variability: 3D printing, insert seating, and phantom positioning and alignment.

For image analysis, a rigid mask of 25.7 mm diameter circular ROIs (slightly smaller than the 30.0 mm diameter of the insert itself) derived from the digital ground truth geometry was manually placed over the eight texture inserts. Annotations were centered on the mid-plane of the insert volume and applied to slices at -3 mm and +3 mm from the center, yielding three non-adjacent slices per insert.

From these ROIs, all available 2D radiomic texture features were extracted using *PyRadiomics*.[37] The extraction yielded a total of 91 features per ROI: 18 first-order features, and 73 second order features including gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRLM), gray level size zone matrix (GLSZM), neighboring gray tone difference matrix (NGTDM), and gray level dependence matrix (GLDM) features.

Reproducibility across the three independent print-and-assembly cycles was quantified using the intra-class correlation coefficient (ICC), providing a measure of agreement where ICC >0.90 is considered excellent.[38,39]

Additionally, the reproducibility was quantified using the coefficient of variation ($CV = \frac{\sigma}{\mu}$) for each texture object. To evaluate the physical fidelity of the 3D printing process, or the "inter-print" reproducibility, the CV for each radiomic feature was calculated across the three physical insert set replicates (e.g., Print 1, Print 2, and Print 3); the measurement was repeated at three slice locations, then averaged. To evaluate inter-slice (i.e., *z*-axis) uniformity, the CV was calculated across three distinct slice locations; the measurement was repeated for each of the three insert sets, then averaged. The features were subsequently stratified by radiomic feature class (first-order vs. second-order). The percentage of features meeting established CV reproducibility thresholds ($\leqslant$ 10% and $\leqslant$ 15%)[40] was reported to provide an initial evaluation of platform manufacturing variability.

2.2.3 | Evaluation of texture insert set for non-redundancy

As the texture inserts were designed with relatively simple two-material geometric patterns, it was necessary to confirm that geometric variations successfully

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

translated into distinct radiomic signatures, rather than exhibiting high feature redundancy that would limit their utility.

To characterize the radiomic feature space coverage of the eight-insert set and confirm that the inserts provide non-redundant stimuli, a principal component analysis (PCA) was performed on the radiomic features extracted from each texture insert. Radiomic feature extraction was performed as described in Section 2.2.2 from images acquired with 360 mA ($CTDI_{vol}$ of 13.68 mGy) and a Br62f reconstruction kernel. Features were standardized to zero mean and unit variance prior to decomposition, due to the sensitivity of PCA to variable scaling. The PCA was performed using the *scikit-learn* library.[41] The first two components were kept and used to visualize the relative spatial distribution of the eight textures in the rescaled PCA space. A well-dispersed, non-overlapping separation of the inserts across the principal axes of variance demonstrates that this simplified fabrication method successfully yields a diverse spectrum of radiomic stimuli.

2.2.4 | Sensitivity to imaging protocol variation

Finally, we evaluated the sensitivity of the texture inserts to clinically relevant variations in acquisition and reconstruction protocol, assessing the platform's ability to detect and express the protocol-induced radiomic variation it is intended to measure.

To evaluate protocol sensitivity, a single set of the 3D-printed inserts was imaged across six distinct clinical conditions, capturing common ranges of image noise and spatial resolution. Scans were acquired at 60 mA and 360 mA (yielding a $CTDI_{vol}$ of 2.28 mGy and 13.68 mGy, respectively) and reconstructed using three kernels spanning smooth to edge-enhancing frequencies (Br32f, Br46f, and Br62f). Radiomic feature extraction was performed identically to the reproducibility assessment described in Section 2.2.2 for each of these imaging conditions. Using Br46f and 360 mA as a reference condition (roughly equivalent to our institutional routine chest protocol), the radiomic feature reproducibility was quantified using concordance correlation coefficient (CCC) to assess pairwise agreement for each tube current and reconstruction kernel setting.[42] Consistent with established guidelines, a CCC $\geqslant$ 0.90 was used as the threshold for reproducibility.[43]

3 | RESULTS

3.1 | Phantom body spatial uniformity

Across all 168 ROIs measured within the four manufactured phantom bodies, the overall dataset mean attenuation was 104.9 ± 1.5 HU. A summary of the measurements is provided in Table 2. All four phantom bodies met the target ±5 HU criteria for intra-phantom uniformity, with Δ_{max} ranging from 2.7 to 4.4 HU. Additionally, all four phantom bodies met the inter-phantom reproducibility target of all ROIs falling within ±5 HU of the overall mean, as shown in Figure 6. Furthermore, no visual beam hardening or streak artifacts were observed across the epoxy volumes.

The pairwise TOST procedure confirmed both intra- and inter-phantom uniformity within the ±5.0 HU criteria. For the worst-case inter-phantom comparison (Phantoms 2 and 4), the maximum mean divergence was -3.11 HU (90% CI: [-3.45, -2.76] HU), well within the equivalence criteria ($p < 0.001$). Intra-phantom spatial uniformity was similarly strong: for the worst-case phantom (Phantom 4), the inner-to-outer mean difference was -0.90 HU (90% CI: [-1.34, -0.45] HU), also falling entirely within the ±5.0 HU criteria ($p < 0.001$). Thus, the four manufactured phantom bodies demonstrated excellent reproducibility and consistently high internal spatial uniformity.

TABLE 2 Mean value ± standard deviation of the HU number measurements for each phantom body.

Phantom body	μ_{inner} [HU]	μ_{outer} [HU]	$\mu_{overall}$ [HU]	Δ_{max} [HU]
1	103.5 ± 0.5	104.6 ± 0.6	104.1 ± 0.8	3.0
2	102.7 ± 0.6	103.7 ± 0.9	103.2 ± 0.9	4.4
3	105.3 ± 0.5	106.4 ± 0.7	105.8 ± 0.8	2.7
4	105.9 ± 0.7	106.8 ± 1.0	106.3 ± 1.0	3.7
Overall mean	104.4 ± 1.4	105.4 ± 1.5	104.9 ± 1.5	3.5 ± 0.7

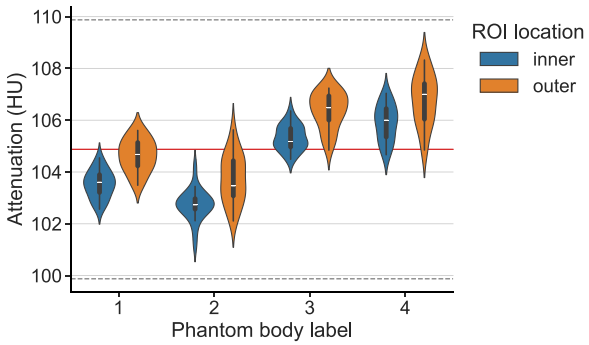

FIGURE 6 Uniformity measurements (mean values of ROIs in HU) for the four tested phantom bodies. Red, solid line illustrates the overall mean value (104.9 HU) and the gray, dashed lines illustrate ±5 HU from the mean.

3.2 | Texture insert reproducibility

Visual inspection of CT images confirmed consistent insert alignment and orientation across all three

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

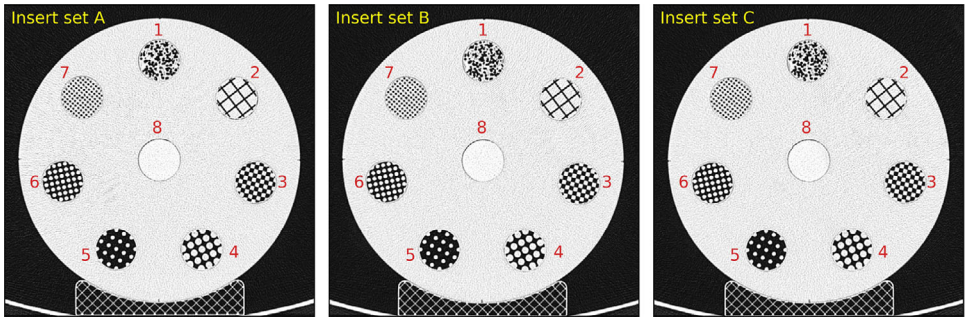

FIGURE 7 The three sets of 3D printed modular inserts, imaged at 360 mA and reconstructed with the Br62f kernel. Visual agreement on CT is excellent, including the alignment and orientation of the texture inserts relative to the phantom body across the three experimental setups and acquisitions. Images shown with window width 1300 HU and level -450 HU.

TABLE 3 Texture insert reproducibility results.

Texture	Group	Inter-print CV		Inter-slice CV	
		< 15%	< 10%	< 15%	< 10%
1	1st Order	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)
	2nd Order	100.0 (73/73)	95.9 (70/73)	100.0 (73/73)	98.6 (72/73)
2	1st Order	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)
	2nd Order	100.0 (73/73)	97.3 (71/73)	100.0 (73/73)	98.6 (72/73)
3	1st Order	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)
	2nd Order	100.0 (73/73)	100.0 (73/73)	100.0 (73/73)	100.0 (73/73)
4	1st Order	100.0 (18/18)	94.4 (17/18)	100.0 (18/18)	94.4 (17/18)
	2nd Order	100.0 (73/73)	95.9 (70/73)	98.6 (72/73)	95.9 (70/73)
5	1st Order	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)
	2nd Order	100.0 (73/73)	95.9 (70/73)	95.9 (70/73)	94.5 (69/73)
6	1st Order	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)
	2nd Order	100.0 (73/73)	100.0 (73/73)	98.6 (72/73)	93.2 (68/73)
7	1st Order	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)	100.0 (18/18)
	2nd Order	94.5 (69/73)	89.0 (65/73)	100.0 (73/73)	94.5 (69/73)
8[a]	1st Order	94.4 (17/18)	94.4 (17/18)	94.4 (17/18)	94.4 (17/18)
	2nd Order	71.2 (52/73)	63.0 (46/73)	67.1 (49/73)	60.3 (44/73)

Note: Inter-print CVs were averaged across all slice depths, and inter-slice CVs were averaged across all insert set replicates. Results are reported as % of features meeting CV thresholds (count/total).
aSolid insert, not a heterogeneous texture pattern

independent manufacturing sets (shown in Figure 7), with no macroscopic defects or differences between the sets observed.

The extracted radiomic features demonstrated high agreement across the independent print-and-assembly cycles. The inter-print ICC across all eight textures exceeded 0.998 in all cases (mean ICC > 0.999), demonstrating very strong agreement well above the established threshold for excellent reproducibility (ICC > 0.90). Inter-slice ICC within individual inserts performed similarly, yielding a mean ICC > 0.999.

This stability is further reflected in the CV analysis, summarized in Table 3. For the first order features, all the heterogeneous texture inserts except for Texture 4 had 100.0% (18/18) of the first order features meet the strict CV < 0.10 reproducibility threshold across both inter-print and inter-slice evaluations. For Texture 4, 100.0% of the first order features met the CV < 0.15 threshold and 94.4% (17/18) of the first order features met the CV < 0.10 threshold in both inter-print and inter-slice evaluations. For Texture 8, the homogeneous solid insert, 94.4% of the first order features met the CV < 0.10 threshold. The radiomic feature that did not meet the reproducibility criteria in each of these cases was first order skewness, which relies on the third central moment of the intensity distribution and can be highly sensitive to minor fluctuations in image noise.

For the second order texture features, Textures 1–6 had 100.0% (73/73) of the second order features meet the inter-print CV ⩽ 0.15 criteria, and between 95.9%

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

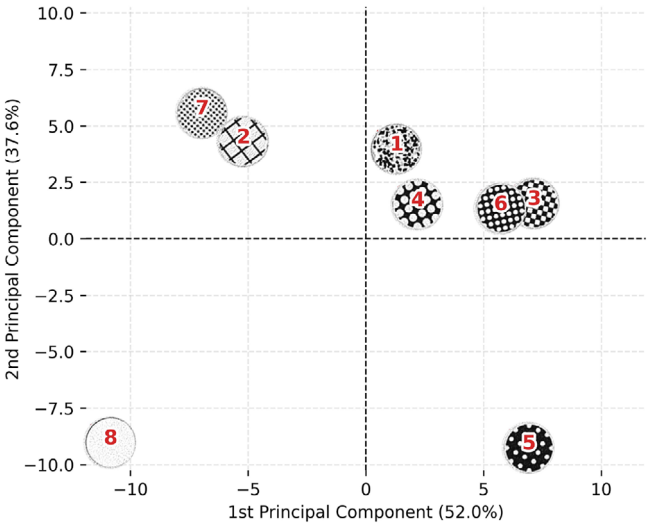

FIGURE 8 Principal component analysis of radiomic feature values measured on the eight texture inserts. Images of the textures themselves with labels overlaid are used as markers.

and 100.0% of the second order features meet the stricter CV $\leqslant$ 0.10 criteria. Texture 7 exhibited slightly lower inter-print reproducibility, with 94.5% of the second order features meeting the CV $<$ 0.15 threshold, and 89.0% of the second order features meeting the CV $<$ 0.10 threshold. The inter-slice reproducibility for the second order features was also high, with 93.2% to 100.0% of the second order features meeting the CV $<$ 0.10 threshold for all of the heterogeneous texture inserts. While Texture 8 displayed lower reproducibility for second order features, these results were not considered meaningful measures of the reproducibility of the physical objects, as there was not a heterogeneous texture pattern to measure and these variations likely reflect nuance in how radiomic textures are calculated in uniform objects.

Distributions of the inter-print and inter-slice CV, as well as the complete report of the features that did not meet the CV criteria, are provided in Appendix D. Overall, the inter-print reproducibility was comparable to inter-slice reproducibility, indicating that very low physical print-to-print manufacturing variance is similar to (or less than) the variation across printing layers.

3.3 | Evaluation of texture insert set for non-redundancy

The PCA of the radiomic features showed that the first two principal components captured 89.6% of total radiomic variance (PC1: 52.0%; PC2: 37.6%), indicating the set of texture inserts selected is well-described by two dominant axes of variation. The distribution of inserts within this reduced space is shown in Figure 8. The visualization shows that there is a spread in the distribution of the texture inserts within the feature space, confirming that the insert set provides a range of radiomic stimuli rather than redundant first-order variations. Most of the textures are well resolved along PC1, with PC2 separating Textures 5 and 8 from the other textures distinctly. These results show that Textures 5 and 8 are the most unique in the set, as they are each alone within their respective quadrants in the feature space with the greatest separation from any other texture insert. Conversely, certain inserts exhibit closer proximity in the feature space; for example, Textures 3 and 6 cluster together, reflecting their similar underlying physical geometry in texture pitch and element size. Rather than indicating a problematic redundancy across the entire insert set, this proximity is beneficial as it establishes a realistic spectrum of both highly distinct and distinct-but-similar radiomic stimuli. Overall, the visualization confirms that the insert set provides diverse radiomic stimuli rather than redundant texture designs.

3.4 | Sensitivity to imaging protocol variation

Visual assessment of the acquisitions (Figure 9) confirms that the selected protocols produced a clinically realistic spectrum of image noise and spatial resolution, with fine physical textures visibly blurring under smooth kernels and sharpening under edge-enhancing kernels.

The designed textures proved highly sensitive to variation in imaging protocol parameters, particularly reconstruction algorithms. Transitioning from the standard Br46f to the smooth Br32f kernel (with the reference tube current, 360 mA, held constant) yielded only 8.8% reproducible features, and transitioning to the sharp Br62f kernel yielded only 7.7%. When tube current was reduced from 360 to 60 mA (with the reference kernel, Br46f, held constant), 58.2% of the extracted radiomic features remained reproducible. Distributions of the CCC values are provided in Appendix E.

Together, Figure 10 and Figure 11 illustrate the measured radiomic feature values across all eight textures and imaging conditions. In Figure 10, raw feature values are shown for a representative subset of features, highlighting specific directional sensitivities to protocol variations. For example, first-order mean attenuation remains exceptionally stable across both dose and kernel changes, aligning with the physical reality that mean HU is driven by the textures' material ratios (PLA vs. air) rather than reconstruction algorithms or dose. Conversely, first-order variance increases significantly under sharper kernels, consistent with the expected increase in image noise. Figure 11 extends this view to all extracted features, with feature values min-max normalized to [0, 1] to enable visual comparison across features of differing scales. In this heatmap, macro-level

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

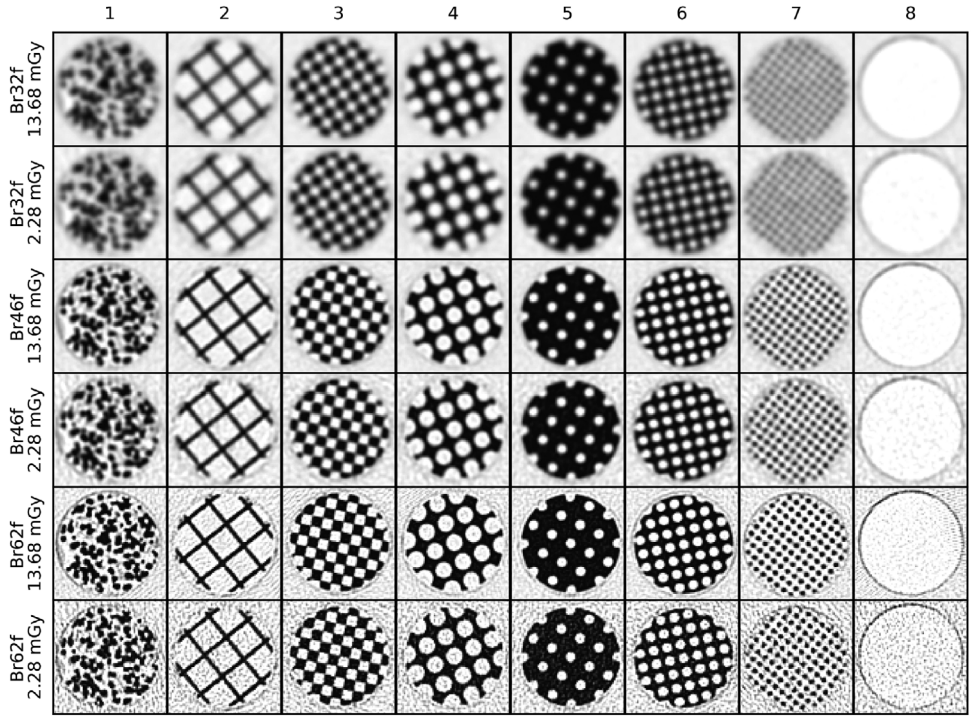

FIGURE 9 Images of the eight texture objects (labeled above) under each of the six imaging conditions (labeled left). Rows are ordered with the highest noise condition at the bottom to the lowest noise condition at the top. Images shown with window width 1300 HU and level -450 HU.

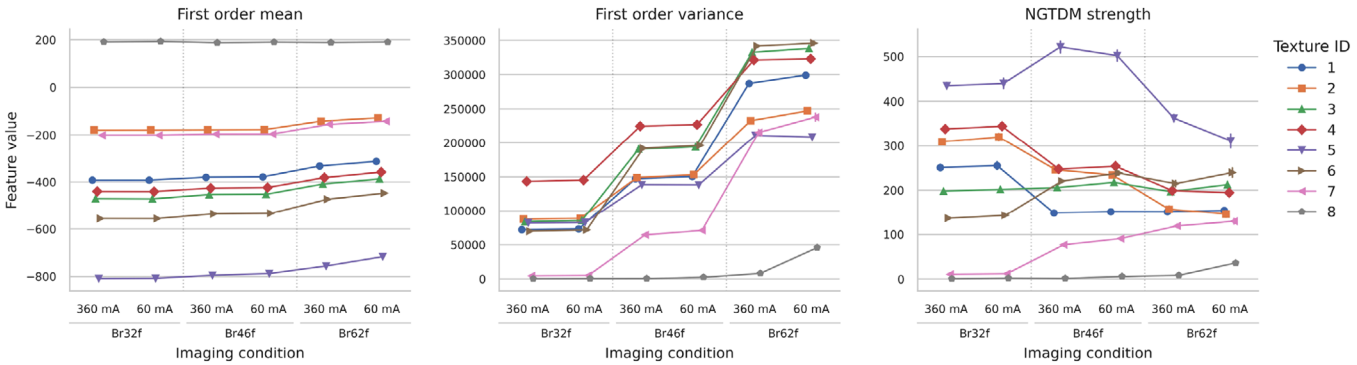

FIGURE 10 Radiomic features values for three illustrative features (first order mean, first order variance, NGTDM strength) for each texture insert across the six imaging conditions. Error bars indicate standard deviation where large enough to be shown.

color differences across texture groupings (columns, labeled above) reflect the degree to which features differ between texture objects, while color variation within a texture grouping reflects sensitivity to protocol. Features that appear uniform in color within a texture grouping are robust to protocol changes, whereas those exhibiting strong within-group variation are more sensitive to reconstruction kernel and dose. While investigations into the reproducibility of individual radiomic features was beyond the scope of this work, these figures illustrate the phantom's capacity to detect and measure this type of protocol-induced variation.

4 | DISCUSSION

This study demonstrates the feasibility of producing high-quality, reproducible phantoms for quantitative imaging using an accessible hybrid manufacturing approach. Results indicate that combining 3D-printed molds with epoxy resin casting yields phantom bodies with high uniformity and reproducibility. The modular insert design enables rapid development and sharing of texture patterns, addressing a critical need in the radiomics community for standardized yet customizable validation tools. Importantly, the phantom

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

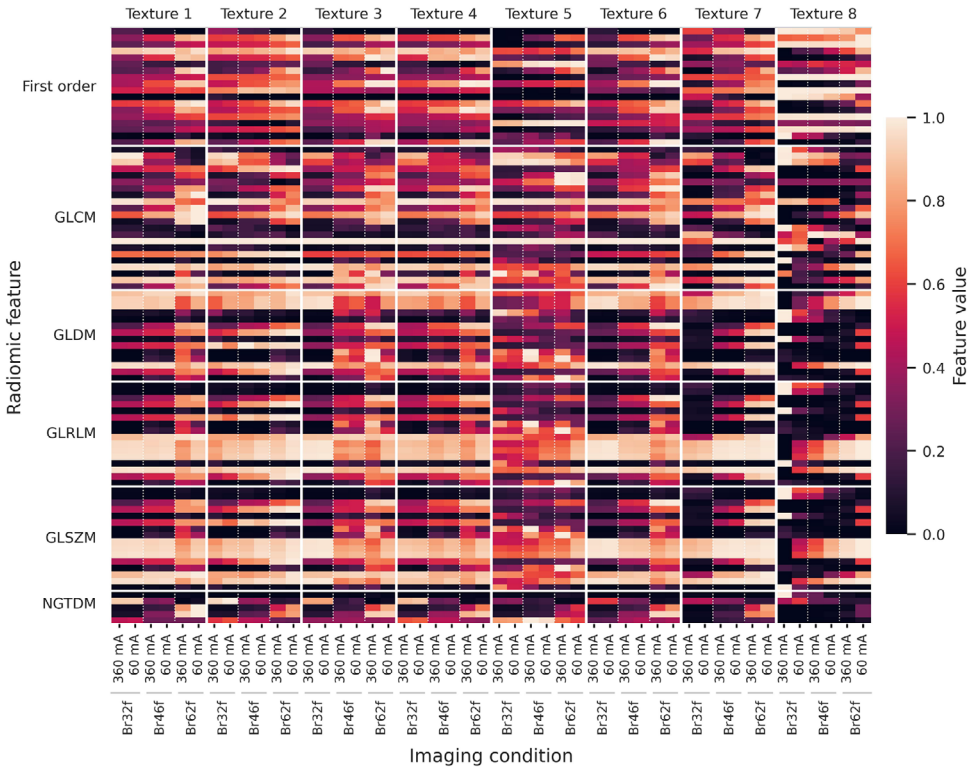

FIGURE 11 Radiomic feature values of all features (grouped by feature class on the *y*-axis), shown for each texture object and imaging condition (*x*-axis), using the same imaging condition order as in Figure 10. Feature values were min-max normalized to [0, 1]. Color variation between texture groupings reflects differences between the texture objects, while within-group variation reflects sensitivity to imaging protocol.

system design and manufacturing provides a pathway to those tools with dramatically reduced costs and improved accessibility compared to commercial alternatives (approximately $50 per phantom system; a cost breakdown is provided in Appendix C).

The epoxy casting approach produced four phantom bodies with spatial uniformity meeting our acceptance criteria and high inter-phantom consistency (all ROIs within ±5 HU, and inter-phantom differences statistically indistinguishable within ±5 HU with $p < 0.001$). A small but consistent radial gradient was observed, with outer ROIs approximately 1 HU higher than inner ROIs on average; this is likely attributable to minor beam hardening at the phantom periphery and falls well within acceptable limits. While this manufacturing run represents a best-case single-group scenario, it establishes that an individual research group can reliably produce a small batch of self-consistent phantom bodies using accessible equipment and standardized procedures.

The 3D printed texture inserts demonstrated excellent radiomic reproducibility across three independently manufactured and assembled sets, with a mean inter-print ICC > 0.999 across 91 radiomic features. This result is comparable to the ICC reported by Li et al. in test-retest analyses on a single texture phantom,[26] suggesting that the print-to-print manufacturing variance achieved here is at the same level as the measurement repeatability floor for this class of phantom. In contrast, clinical test-retest studies in patients have reported much lower radiomic feature reproducibility, underscoring the value of a controlled phantom reference for isolating protocol-induced variation from biological and positioning noise.[44]

We note that Texture 7, which has the smallest and most densely packed grid elements of any structured insert, exhibited slightly lower inter-print reproducibility (94.5% at CV < 0.15). This suggests the 0.75 mm feature diameter of this insert approaches the practical resolution limits of the FDM process, where small dimensional deviations between print cycles have proportionally greater impact on the resulting radiomic signature. This is a useful boundary condition for future insert design: features substantially below 1.0 mm in any dimension should be treated with caution when using standard 0.4 mm nozzle FDM printing, and the 0.2 mm nozzle used for Texture 7 represents the lower practical limit with consumer-grade equipment. Future work should further characterize these limits for sub-millimeter texture designs.

The PCA of the radiomics features provides strong evidence that the insert set provides genuinely non-redundant radiomic stimuli, with each insert contributing unique information to the feature space. Textures 5 and 8 were particularly distinct, which could be attributed

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

to each representing an extreme in terms of material compositions within the set, with Texture 5 having the highest ratio of air, and Texture 8 (the solid, uniform insert) having the highest ratio of PLA.

Finally, the phantom platform demonstrated the protocol sensitivity required of a quantitative imaging tool. While both tube current and reconstruction kernel had dramatic impacts on radiomic reproducibility, the effect of reconstruction kernel was greater: changing from the reference Br46f kernel to either the smooth Br32f or sharp Br62f kernel left only 8.8% or 7.7% of features reproducible, respectively, while reducing tube current from 360 mA to 60 mA left 58.2% of features reproducible. This sensitivity to changes in imaging protocol, as well as this asymmetry in particular, is consistent with the broader radiomics literature, which identified reconstruction kernel as a dominant source of protocol-induced radiomic variation.[12–14] The fact that our synthetic textures exhibit this well-established pattern validates their suitability as reference stimuli: a phantom that did not respond to kernel changes in the expected direction would be of limited utility for protocol harmonization studies. Furthermore, the responses to protocol observed in the individual radiomic features shown in Figure 10 were also consistent with physical intuition: first order mean was relatively unchanged by differences in the CT protocol, whereas first order variance increased substantially with kernel sharpness, and more modestly with decrease in tube current. The radiomic feature variation observed across the eight texture objects also supports the results from the texture insert set non-redundancy PCA evaluation.

A key distinguishing contribution of this platform relative to prior 3D-printed phantom work is its foundational commitment to open science. By dedicating the complete design files and manufacturing specification to the public domain under a CC0 license, we remove the licensing barriers that have historically restricted phantom-based harmonization efforts to well-resourced groups with access to proprietary hardware. Unlike prior approaches that employed 3D printing without public release of designs, this framework allows any research group with a consumer-grade FDM printer to replicate, validate, modify, and redistribute phantom designs without restriction. The modular architecture amplifies this benefit: new insert designs targeting specific clinical tasks or tissue types can be developed, validated, and shared independently of the phantom body, enabling community-driven expansion of the platform's insert library without requiring each group to develop a complete phantom system from scratch. This open approach aligns with broader efforts in the imaging community to improve reproducibility and accessibility of research tools. We aim to encourage community-driven use and development and ensure more universal access to essential validation tools.

The eight texture inserts evaluated in this work were selected to provide diverse and complementary radiomic signatures rather than to mimic specific pathologies. Consequently, the phantom is intended as a tool for studies involving radiomic feature reproducibility, protocol optimization methodologies, and scanner harmonization or quality assurance. The mean attenuation of the texture inserts evaluated ranged from -800 to 200 HU (shown in Figure 10), depending primarily on the texture's ratio of PLA to air. The materials used (PLA and clear epoxy resin) exhibited attenuation values within the range of 100 to 200 HU. We expect the inserts to be most relevant to quantitative imaging tasks involving heterogeneous tissue and, based on these attenuation characteristics, may be most applicable to contrast-enhanced soft tissues, lung, and abdominal organs. Applications in lung imaging may be of particular interest, as the lung is composed of heterogeneous patterns soft tissue and air that mirrors the PLA-and-air compositions of the texture inserts. However, the suitability of these textures for clinical applications should be evaluated in future work. Users are encouraged to design custom texture inserts tailored to specific clinical pathologies or modeling endpoints, which is supported by the phantom's modular nature and open-source design.

This initial evaluation was performed on a typical clinical CT scanner system, using the conditions described in Table 1. While the phantom body and modular insert architecture are expected to remain applicable across CT platforms, the specific texture inserts evaluated here represent only an initial validation set and can be refined or expanded as scanner resolution and quantitative imaging requirements evolve. Whether the measured reproducibility of the manufacturing approach presented in this work is impacted by substantially higher spatial resolutions available with systems such as photon-counting CT (PCCT) remains to be evaluated. However, because the phantom specification is fundamentally digital and does not mandate a particular fabrication method, if needed the platform can be adapted to match the ultra-high spatial resolutions achievable with next-generation PCCT systems through high-resolution additive manufacturing, micromachining, or other computer-aided manufacturing techniques.

Several limitations of this work should be noted. First, the reproducibility analysis intentionally represents a best-case manufacturing scenario: conducted by a single group using a single printer and single material supplier. This baseline was a necessary first step, but future multi-center studies will be needed to quantify the inter-institution variability introduced by different printers, filament batches, and casting environments. Some increase in variability is expected across sites, and it remains an open question whether stricter material standardization (e.g., specifying filament density tolerances or epoxy formulation) will be required for cross-site

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

harmonization. Second, this was a single time point investigation, and thus did not include evaluation of the longitudinal stability of the phantom materials. While prior work has found that the attenuation properties of PLA are stable over time and do not require special storage conditions,[32] a longitudinal stability study is warranted before this platform is used for studies spanning multiple years. Finally, the current insert set was intentionally designed as a simplified baseline: binary material distributions with no *z*-axis structural variation and no attempt to mimic specific biological tissues or clinical pathologies. This simplification was appropriate for the initial evaluation of the reproducibility of the manufacturing approach and validating the platform's basic metrological properties, but it means the inserts do not yet provide the clinically relevant or application-specific stimuli that would be needed for targeted radiomic studies (e.g., inserts designed to emulate the textural properties of lung nodules, liver lesions, or other tissues of clinical interest). Developing such inserts, including designs that incorporate *z*-axis structural complexity or stochastic heterogeneity representative of specific pathologies, represents the most important next step for the platform. The modular open-source architecture is specifically designed to support these types of extensions.

5 | CONCLUSION

The open-source, accessible phantom presented here addresses critical barriers to quantitative imaging research. By combining 3D printing technology with epoxy resin casting, we have created a solution that is simultaneously affordable, customizable, and reproducible. Serving as a rigorous middle ground between expensive commercial products and non-reproducible handmade alternatives, this platform successfully demonstrated the key properties we require for our radiomic texture experiments: uniformity, radiomic reproducibility, and radiomic sensitivity to clinical protocol variations. By releasing the phantom and insert design into the public domain (https://doi.org/10.5281/zenodo.21252065 [31]), this tool remains unencumbered for the entire research community, fostering collaborative development and long-term standardization efforts. The immediate practical consequence is that any research group with access to a consumer-grade FDM printer can replicate, validate, and extend this platform without restriction.

ACKNOWLEDGMENTS

This work was supported in part by the NIBIB of the National Institutes of Health under award number T32EB002101-46. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CONFLICT OF INTEREST STATEMENT

UCLA has a Master Research Agreement with Siemens Healthineers (Forchheim, Germany). Authors JMH and MMG have received research funding unrelated to this project under this agreement. The other authors have no relevant conflicts to disclose.

REFERENCES

1. Aerts HJWL, Velazquez ER, Leijenaar RTH, et al., Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006. doi:10.1038/ncomms5006
2. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278:563-577. doi:10.1148/radiol.2015151169
3. Fave X, Zhang L, Yang J, et al. Delta-radiomics features for the prediction of patient outcomes in non-small cell lung cancer. *Sci Rep*. 2017;7:588. doi:10.1038/s41598-017-00665-z
4. Fried DV, Tucker SL, Zhou S, et al. Prognostic value and reproducibility of pretreatment CT texture features in stage III non-small cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2014;90:834-842. doi:10.1016/j.ijrobp.2014.07.020
5. Wu G, Woodruff HC, Shen J, et al. Diagnosis of invasive lung adenocarcinoma based on chest CT radiomic features of part-solid pulmonary nodules: a multicenter study. *Radiology*. 2020;297:451-458. doi:10.1148/radiol.2020192431
6. Selvam M, Chandrasekharan A, Sadanandan A, Anand VK, Murali A, Krishnamurthi G. Radiomics as a non-invasive adjunct to Chest CT in distinguishing benign and malignant lung nodules. *Sci Rep*. 2023;13:19062. doi:10.1038/s41598-023-46391-7
7. Park HJ, Lee SM, Song JW, et al. Texture-based automated quantitative assessment of regional patterns on initial CT in patients with idiopathic pulmonary fibrosis: relationship to decline in forced vital capacity. *Am J Roentgenol*. 2016;207:976-983. doi:10.2214/AJR.16.16054
8. Budzikowski JD, Foy JJ, Rashid AA, Chung JH, Noth I, Armato SG. Radiomics-based assessment of idiopathic pulmonary fibrosis is associated with genetic mutations and patient survival. *J Med Imaging*. 2021;8:031903. doi:10.1117/1.JMI.8.3.031903
9. Nardone V, Tini P, Croci S, et al. 3D bone texture analysis as a potential predictor of radiation-induced insufficiency fractures. *Quant Imaging Med Surg*. 2018;8:14-24. doi:10.21037/qims.2018.02.01
10. Y.-C,. Chen, Y.-T. Li, P.-C. Kuo, et al. Automatic segmentation and radiomic texture analysis for osteoporosis screening using chest low-dose computed tomography. *Eur Radiol*. 2023;33:5097-5106. doi:10.1007/s00330-023-09421-6
11. Lee Y, Hong S, Lee M, et al. Radiomics identifies distinct cortical bone texture alterations in patients with CKD using HR-pQCT. *Bone Res*. 2026;14:36. doi:10.1038/s41413-026-00515-7
12. Lo P, Young S, Kim HJ, Brown MS, McNitt-Gray MF. Variability in CT lung-nodule quantification: effects of dose reduction and reconstruction methods on density and texture based features. *Med Phys*. 2016;43:4854-4865. doi:10.1118/1.4954845
13. Berenguer R, Pastor-Juan MdR, Canales-Vázquez J, et al. Radiomics of CT features may be nonreproducible and redundant: influence of CT acquisition parameters. *Radiology*. 2018;288:407-415. doi:10.1148/radiol.2018172361
14. Midya A, Chakraborty J, Gönen M, Do RKG, Simpson AL. Influence of CT acquisition and reconstruction parameters on radiomic feature reproducibility. *J Med Imaging*. 2018;5:011020. doi:10.1117/1.JMI.5.1.011020
15. Hoye J, Solomon J, Sauer TJ, Robins M, Samei E. Systematic analysis of bias and variability of morphologic features for lung lesions in computed tomography. *J Med Imaging*. 2019;6:013504. doi:10.1117/1.JMI.6.1.013504

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

16. Sullivan DC, Obuchowski NA, Kessler LG, et al. Metrology standards for quantitative imaging biomarkers. *Radiology*. 2015;277:813-825. doi:10.1148/radiol.2015142202
17. McNitt-Gray M, Napel S, Jaggi A. et al. Standardization in quantitative imaging: a multicenter comparison of radiomic features from different software packages on digital reference objects and patient data sets. *Tomography*. 2020;6:118-128. doi:10.18383/j.tom.2019.00031
18. Zwanenburg A, Vallières M, Abdalah MA, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. 2020;295:328-338. doi:10.1148/radiol.2020191145
19. McCollough CH, Bruesewitz MR, McNitt-Gray MF, et al. The phantom portion of the American College of Radiology (ACR) Computed Tomography (CT) accreditation program: practical tips, artifact examples, and pitfalls to avoid. *Med Phys*. 2004;31:2423-2442. doi:10.1118/1.1769632
20. Gulliksrud K, Stokke C, and Trægde Martinsen AC. How to measure CT image quality: variations in CT-numbers, uniformity and low contrast resolution for a CT quality assurance phantom. *Physica Med*. 2014;30:521-526. doi:10.1016/j.ejmp.2014.01.006
21. Kalendralis P, Traverso A, Shi Z, et al. Multicenter CT phantoms public dataset for radiomics reproducibility tests. *Med Phys*. 2019;46:1512-1518. doi:10.1002/mp.13385
22. Mackin D, Fave X, Zhang L, et al. Measuring computed tomography scanner variability of radiomics features. *Invest Radiol*. 2015;50:757. doi:10.1097/RLI.0000000000000180
23. Buch K, Li B, Qureshi MM, Kuno H, Anderson SW, Sakai O. Quantitative Assessment of variation in CT parameters on texture features: pilot study using a nonanatomic phantom. *Am J Neuroradiol*. 2017;38:981-985. doi:10.3174/ajnr.A5139
24. Varghese BA, Cen SY, Jensen K, et al. Technical and clinical considerations of a physical liver phantom for CT radiomics analysis. *J Appl Clin Med Phys*. 2024;25:e14309. doi:10.1002/acm2.14309
25. Ehler E, Craft D, Rong Y. 3D printing technology will eventually eliminate the need of purchasing commercial phantoms for clinical medical physics QA procedures. *J Appl Clin Med Phys*. 2018;19:8-12. doi:10.1002/acm2.12392
26. Li Y, Reyhan M, Zhang Y, et al. The impact of phantom design and material-dependence on repeatability and reproducibility of CT-based radiomics features. *Med Phys*. 2022;49:1648-1659. doi:10.1002/mp.15491
27. Varghese BA, Hwang D, Cen SY, et al. Identification of robust and reproducible CT-texture metrics using a customized 3D-printed texture phantom. *J Appl Clin Med Phys*. 2021;22:98-107. doi:10.1002/acm2.13162
28. Solomon J, Ba A, Bochud F, Samei E. Comparison of low-contrast detectability between two CT reconstruction algorithms using voxel-based 3D printed textured phantoms. *Med Phys*. 2016;43:6497-6506. doi:10.1118/1.4967478
29. Samei E, Hoye J, Zheng Y, Solomon JB, Marin D. Design and fabrication of heterogeneous lung nodule phantoms for assessing the accuracy and variability of measured texture radiomics features in CT. *J Med Imaging*. 2019;6:021606. doi:10.1117/1.JMI.6.2.021606
30. Mei K, Geagan M, Roshkovan L, et al. Three-dimensional printing of patient-specific lung phantoms for CT imaging: emulating lung tissue with accurate attenuation profiles and textures. *Med Phys*. 2022;49:825-835. doi:10.1002/mp.15407
31. Hoffman J, Daly M, HD phantom body and radiomics texture inserts for CT imaging, 2026. doi:10.5281/zenodo.21252065
32. Craft DF, Kry SF, Balter P, Salehpour M, Woodward W, Howell RM. Material matters: analysis of density uncertainty in 3D printing and its consequences for radiation oncology. *Med Phys*. 2018;45:1614-1621. doi:10.1002/mp.12839
33. Solc J, Vrba T, Burianova L. Tissue-equivalence of 3D-printed plastics for medical phantoms in radiology. *J Instrum*. 2018:13:P09018-P09018. doi:10.1088/1748-0221/13/09/P09018
34. Badiuk SR, Sasaki DK, Rickey DW. An anthropomorphic maxillofacial phantom using 3-dimensional printing, polyurethane rubber and epoxy resin for dental imaging and dosimetry. *Dentomaxillofacial Radiology*. 2022;51:20200323. doi:10.1259/dmfr.20200323
35. Dillon C, Iii WB, Clements J, et al. Computed tomography quality control manual. *Technical report*, American College of Radiology, 2017.
36. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinet Biopharm*. 1987;15:657-680. doi:10.1007/BF01068419
37. Griethuysen Van JJ, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77:e104-e107. doi:10.1158/0008-5472.CAN-17-0339
38. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155-163. doi:10.1016/j.jcm.2016.02.012
39. Bartko JJ. The intraclass correlation coefficient as a measure of reliability. *Psychol Rep*. 1966;19:3-11. doi:10.2466/pr0.1966.19.1.3
40. Kessler LG, Barnhart HX, Buckler AJ, et al. The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. *Stat Methods Med Res*. 2015;24:9-26. doi:10.1177/0962280214537333
41. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python. *J Mach Learn Res*. 2011;12:2825-2830.
42. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45:255-268.
43. McBride GB. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. *NIWA Client Report: HAM2005-062*. 2005;45:307-310.
44. Zhao B, Tan Y, Tsai W-Y, Qi J, Xie C, Lu L, Schwartz LH. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep*. 2016;6. doi:10.1038/srep23428

How to cite this article: Daly MA, McNitt-Gray MF, Hoffman JM. An open-source CT texture phantom for radiomic imaging studies: Design, manufacture, and initial evaluation. *Med Phys*. 2026;53:e70613.
https://doi.org/10.1002/mp.70613

APPENDIX A: MATERIALS AND MANUFACTURING

The phantom platform utilizes a hybrid manufacturing approach, combining fused deposition modeling (FDM)-based 3D printing for the hollow outer shell and modular inserts with a liquid two-part epoxy resin cast for the core. This method dramatically reduces print time and mitigates the severe thermal warping that frequently causes dense FDM prints to fail.

The printable 3D models (STL files) are publicly available at https://doi.org/10.5281/zenodo.21252065.[31] To ensure the interchangeable inserts slide smoothly into the phantom body without jamming, a 0.15 mm radial tolerance (0.3 mm total diametric clearance) is already incorporated into the provided design files.

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The permanent outer shells and inserts were printed using a Bambu Lab X1-Carbon printer (0.4 mm nozzle, 0.28 mm layer height, 100% infill) and Jade White polylactic acid (PLA) filament (Bambu Lab, Shenzhen, China). The phantom body mold took approximately four hours to print using these settings.

To maximize metrological reproducibility across institutions, we strongly recommend utilizing un-dyed or pure white PLA; filaments containing heavy dyes or specialized additives (e.g., "glow-in-the-dark" blends) introduce unpredictable variations in CT attenuation. Furthermore, alternative FDM materials such as PET-G were found to be chemically incompatible with the epoxy resin and should be avoided.

The core of the phantom was cast using a two-part, clear "Deep Pour" epoxy resin system (Teexpert, https://teexpert-resin.com) to ensure a low-exotherm cure that does not melt the PLA shell. The resin and hardener were measured strictly by weight in a 2:1 ratio to ensure stoichiometric accuracy, and vigorously mixed for 10 minutes using a power drill equipped with a mixing paddle. Following mixing, the resin was placed inside a 3-gallon vacuum chamber powered by a 5 CFM dual-stage vacuum pump (Kozyvacu, Kozyard, Fontana, CA) to extract introduced air. Because epoxy foams and expands rapidly under vacuum, we strongly recommend calculating the volume required for a single phantom body, mixing it in an oversized disposable container, and degassing one batch at a time to prevent boiling over into the pump. For the phantoms manufactured in this study, the degassing phase was terminated once visual off-gassing subsided; however, to improve batch-to-batch reproducibility, we recommend future users adopt a strict 30-minute vacuum duration per batch to standardize air removal rather than relying on subjective visual cues.

The degassed epoxy was slowly poured over wooden stir stick resting inside the mold to safely break the surface tension and minimize the re-introduction of air. The liquid epoxy should be filled flush with the top rim of the PLA shell; the mold design intentionally includes 4 mm of excess vertical depth to accommodate the approximate 10% volumetric shrinkage that occurs during the 48-hour room-temperature cure, ensuring a consistent final phantom depth of 40 mm. Minor bubbles should generally be left alone, as attempting to manually remove them (e.g., via toothpick) or utilizing a heat gun risks introducing more air or warping the PLA shell. The phantoms must be left undisturbed on a leveled surface during curing to prevent a slanted top surface that could introduce geometric misalignments.

For final image acquisition, the cylindrical phantom body can be stabilized using either the standard cradle provided with the ACR CT accreditation phantom or a custom 3D-printed stand, and aligned orthogonally to the scan plane utilizing the scanner's standard lasers and the physical alignment markings integrated directly into the phantom shell.

APPENDIX B: TEXTURE DEFINITIONS AND PRINT SETTINGS

All design work was done using Fusion 360 (version 2.0.17954; Autodesk, San Francisco, CA). This parametric computer-aided design (CAD) software was selected to ensure the precise, mathematically driven definition of the 3D structures, allowing for highly reproducible spatial dimensions prior to slicing and fabrication. Detailed design and print parameters are provided in Table B1. All model slicing and print configuration was done in Bambu Studio (version 2.1.0.59; Bambu Lab, Shenzhen, China). Most textures (2-6, 8) were printed using a 0.4 mm nozzle with a layer height of 0.2 mm. Two textures (1 and 7) utilized a 0.2 mm print nozzle to achieve smaller details. The printed layer heights were chosen because they were the printer-optimized default settings in our system and is below the minimum slice thickness of most CT scanners.

All inserts were manufactured using Jade White PLA (Bambu Lab, Shenzhen, China) at 100% infill, utilizing the default recommended infill pattern generated by the slicing software. Default print profiles were utilized for all inserts, with the exception of those requiring a 0.2 mm nozzle (Textures 1 and 7), for which the highest-quality print preset was selected.

TABLE B1 Detailed geometric definitions and printing specifications for the eight modular texture inserts.

Texture ID	Geometry	Dimensions & pitch	Nozzle/Layer (mm)
1	Random circles (jittered)	base $r = 0.3$ mm, jitter $\leqslant 0.8$ mm	0.2 / 0.08
2	Square rods grid	7.0 mm edge, 8.0 mm pitch	0.4 / 0.20
3	Checkerboard	3.0 mm squares	0.4 / 0.20
4	Circular rod grid	5.0 mm dia, 6.0 mm pitch	0.4 / 0.20
5	Circular rod grid	3.0 mm dia, 6.0 mm pitch	0.4 / 0.20
6	Circular rod grid	3.0 mm dia, 4.0 mm pitch	0.4 / 0.20
7	Inverted rods (air) grid	0.75 mm dia, 2.5 mm pitch	0.2 / 0.08
8	Solid block	N/A	0.4 / 0.20

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

APPENDIX C: ESTIMATED PRODUCTION COSTS

The itemized manufacturing costs for the hardware and materials utilized in this study are provided in Table C1. All listed values represent prices at the time of purchase, excluding taxes, tariffs, and shipping. In addition to the listed equipment, fabrication requires a standard power drill equipped with a mixing paddle for vigorous blending of the epoxy resin and hardener, as well as standard laboratory safety equipment (gloves, eye protection, and respiratory protection such as a ventilated fume hood).

Assuming a research group already possesses the necessary capital equipment (a 3D printer and vacuum degassing chamber), the marginal consumable cost to produce one complete phantom system (one epoxy body and one complete set of eight texture inserts) is highly economical. The material yields from this experiment were four phantom bodies and approximately four of the eight-texture inset sets (three sets described in the experiment, plus approximately one additional set in test prints of texture designs), produced from 3 gallons of resin and 3 spools of filament (with some filament remaining). With a total materials cost of $202.31 to produce four phantom systems, the estimated material cost was about $50 per complete system.

Even when factoring in the total capital equipment investment required to initiate this project from scratch, the amortized cost across the limited four-phantom manufacturing run was approximately $401 per phantom (based on the estimated total cost of $1,604.51). By comparison, standard proprietary quantitative imaging phantoms typically range from thousands to tens of thousands of dollars. Therefore, this open-source methodology provides an established research lab with a tool that is roughly two orders of magnitude cheaper (for marginal materials) than commercial alternatives, while lowering the initial barrier to entry for an unequipped lab by at least one order of magnitude. Furthermore, the lab now possesses the equipment required to both produce additional phantoms and extend the platform. Cost estimates exclude personnel time required for printing, casting, and post-processing.

TABLE C1 Itemized cost breakdown of the capital equipment and consumable materials required to manufacture the four phantom bodies and three complete insert sets evaluated in this study.

Item	Source	Quantity	Cost (USD)
Capital Equipment			
X1 Carbon 3D Printer	Bambu Lab	1	1199.00
Dual-stage Vacuum Pump (5 CFM)	Kozyvacu (Amazon)	1	119.99
3-Gallon Vacuum Chamber	Vevor (Amazon)	1	83.21
Total equipment cost			**1402.20**
Consumable Materials			
Deep Pour Epoxy Resin	Teexpert (Amazon)	3 gallons	142.34
Jade White PLA Filament	Bambu Lab	3 spools	59.97
Total materials cost			**202.31**
Total cost (4 phantom systems)			**1604.51**

Note: Costs reflect prices at the time of purchase and exclude local taxes, shipping, and tariffs.

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

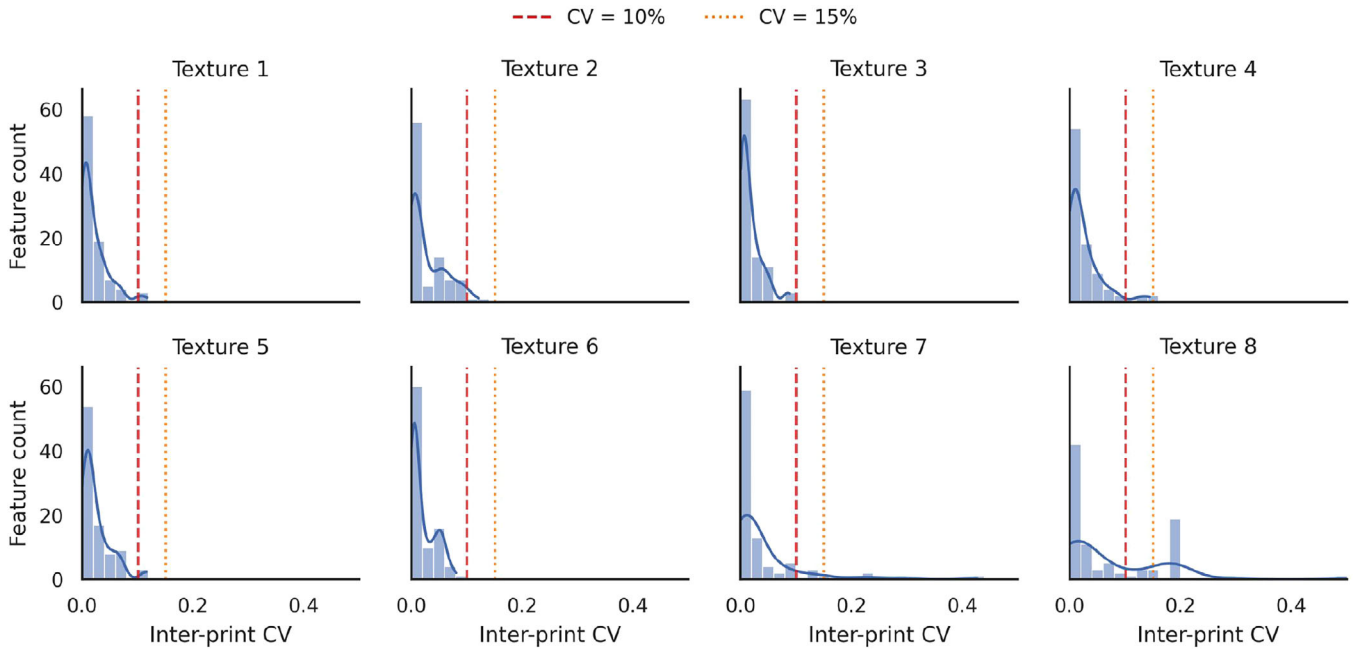

FIGURE D1 Distributions of the inter-print CV values measured for each texture insert. Bars correspond to the number of radiomic features at a given CV value. The two thresholds used in the evaluation are indicated with vertical lines.

APPENDIX D: ADDITIONAL RESULTS FOR TEXTURE INSERT REPRODUCIBILITY

See Figures D1, D2 and Table D1.

APPENDIX E: ADDITIONAL RESULTS FOR SENSITIVITY TO IMAGING PROTOCOL VARIATION

See Figure E1.

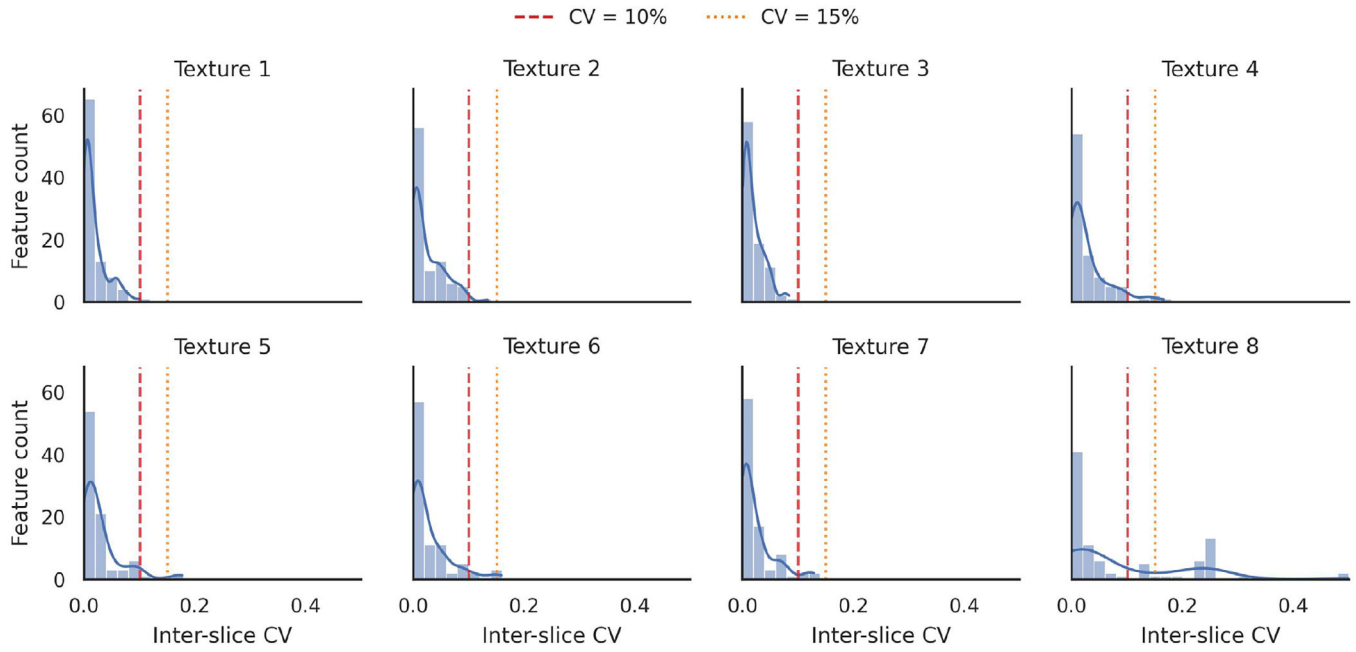

FIGURE D2 Distributions of the inter-slice CV values measured for each texture insert. Bars correspond to the number of radiomic features at a given CV value. The two thresholds used in the evaluation are indicated with vertical lines.

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE D1 Radiomic features with either inter-print or inter-slice CV above the stricter 0.1 threshold for Textures 1-7.

Texture	Radiomic feature	Inter-print CV	Inter-slice CV
1	GLCM joint energy	0.117	0.100
1	GLRLM run variance	0.104	—
1	GLSZM zone variance	0.104	—
2	GLCM inverse variance	—	0.134
2	GLCM joint energy	0.121	—
2	GLCM maximum probability	0.103	—
4	First order skewness	0.127	0.126
4	GLCM joint energy	0.122	0.166
4	GLRLM run variance	0.144	0.141
4	GLSZM zone variance	0.144	0.141
5	GLCM joint energy	0.115	0.158
5	GLDM dependence variance	—	0.107
5	GLRLM run variance	0.116	0.177
5	GLSZM zone variance	0.116	0.177
6	GLCM joint energy	—	0.159
6	GLRLM long run low gray level emphasis	—	0.115
6	GLRLM run variance	—	0.146
6	GLSZM large area low gray level emphasis	—	0.115
6	GLSZM zone variance	—	0.146
7	GLCM maximum probability	0.426	0.116
7	GLDM dependence variance	0.139	0.115
7	GLDM large dependence low gray level emphasis	—	0.078
7	GLDM low gray level emphasis	0.147	—
7	GLRLM long run low gray level emphasis	0.224	—
7	GLRLM run variance	0.139	0.129
7	GLSZM large area low gray level emphasis	0.224	—
7	GLSZM zone variance	0.139	0.129

Note: The CV is provided where >0.1. Texture 8 was excluded, as it was solid insert rather than a heterogeneous texture pattern, making the reproducibility of radiomic texture features a less meaningful evaluation of manufacturing consistency.

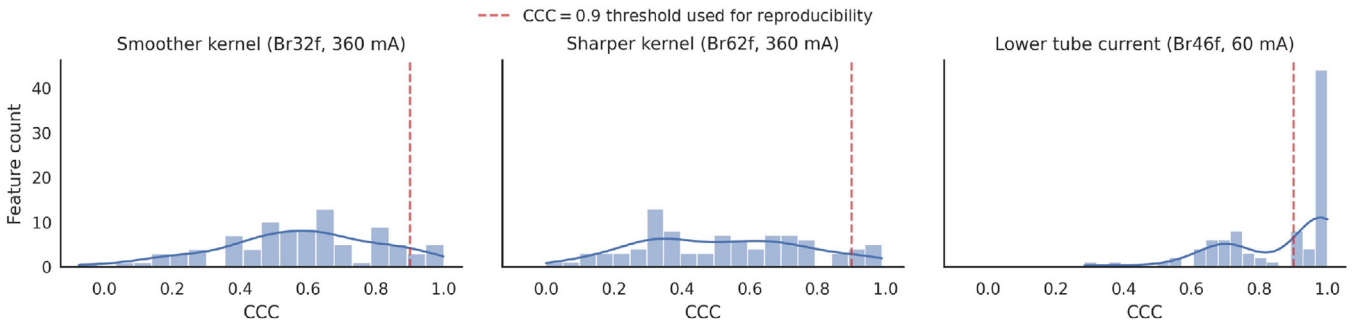

FIGURE E1 Distributions of CCC values measured between radiomic feature sets for each CT protocol comparison (relative to the Br46f, 360 mA reference protocol). Bars correspond to the number of radiomic features at a given CCC value. The 0.9 threshold used in the evaluation is indicated with a vertical line.

24734209, 2026, 8, Downloaded from https://aapm.onlinelibrary.wiley.com/doi/10.1002/mp.70613 by University Of California, Los Angeles, Wiley Online Library on [09/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License