Received: 17 December 2021 | Revised: 7 February 2023 | Accepted: 20 February 2023

DOI: 10.1002/mp.16373

RESEARCH ARTICLE

MEDICAL PHYSICS

3D-printed iodine-ink CT phantom for radiomics feature extraction - advantages and challenges

Michael Bach[1] | **Christoph Aberle**[1] | **Adrien Depeursinge**[2,3] | **Oscar Jimenez-del-Toro**[2] | **Roger Schaer**[2] | **Kyriakos Flouris**[4] | **Ender Konukoglu**[4] | **Henning Müller**[2,5] | **Bram Stieltjes**[1] | **Markus M. Obmann**[1]

[1]Clinic of Radiology and Nuclear Medicine, University Hospital Basel, University of Basel, Basel, Switzerland

[2]University of Applied Sciences Western Switzerland (HES-SO) Valais, Sierre, Switzerland

[3]Department of Nuclear Medicine and Molecular Imaging, Lausanne University Hospital, Lausanne, Switzerland

[4]Computer Vision Lab, ETH Zurich, Zurich, Switzerland

[5]Faculty of Medicine, University of Geneva (UNIGE), Geneva, Switzerland

Correspondence
Markus M. Obmann, Clinic of Radiology and Nuclear Medicine, University of Basel Hospital, Petersgraben 4, CH-4031 Basel, Switzerland.
Email: Markus.Obmann@usb.ch

Funding information
The Swiss Personalized Health Network, Grant/Award Number: QA4IQI

Abstract

Background: To test and validate novel CT techniques, such as texture analysis in radiomics, repeat measurements are required. Current anthropomorphic phantoms lack fine texture and true anatomic representation. 3D-printing of iodinated ink on paper is a promising phantom manufacturing technique. Previously acquired or artificially created CT data can be used to generate realistic phantoms. Purpose: To present the design process of an anthropomorphic 3D-printed iodine ink phantom, highlighting the different advantages and pitfalls in its use. To analyze the phantom's X-ray attenuation properties, and the influences of the printing process on the imaging characteristics, by comparing it to the original input dataset.

Methods: Two patient CT scans and artificially generated test patterns were combined in a single dataset for phantom printing and cropped to a size of $26 \times 19 \times 30$ cm^3. This DICOM dataset was printed on paper using iodinated ink. The phantom was CT-scanned and compared to the original image dataset used for printing the phantom. The water-equivalent diameter of the phantom was compared to that of a patient cohort ($N = 104$). Iodine concentrations in the phantom were measured using dual-energy CT. 86 radiomics features were extracted from 10 repeat phantom scans and the input dataset. Features were compared using a histogram analysis and a PCA individually and overall, respectively. The frequency content was compared using the normalized spectrum modulus.

Results: Low density structures are depicted incorrectly, while soft tissue structures show excellent visual accordance with the input dataset. Maximum deviations of around 30 HU between the original dataset and phantom HU values were observed. The phantom has X-ray attenuation properties comparable to a lightweight adult patient (~54 kg, BMI 19 kg/m^2). Iodine concentrations in the phantom varied between 0 and 50 mg/ml. PCA of radiomics features shows different tissue types separate in similar areas of PCA representation in the phantom scans as in the input dataset. Individual feature analysis revealed systematic shift of first order radiomics features compared to the original dataset, while some higher order radiomics features did not. The normalized frequency modulus $|f(\omega)|$ of the phantom data agrees well with the original data. However, all frequencies systematically occur more frequently in the phantom compared

to the maximum of the spectrum modulus than in the original data set, especially for mid-frequencies (e.g., for $\omega = 0.3942\ mm^{-1}$, $|f(\omega)|_{original} = 0.09 * |f_{max}|_{original}$ and $|f(\omega)|_{phantom} = 0.12 * |f_{max}|_{phantom}$).

Conclusions: 3D-iodine-ink-printing technology can be used to print anthropomorphic phantoms with a water-equivalent diameter of a lightweight adult patient. Challenges include small residual air enclosures and the fidelity of HU values. For soft tissue, there is a good agreement between the HU values of the phantom and input data set. Radiomics texture features of the phantom scans are similar to the input data set, but systematic shifts of radiomics features in first order features, due to differences in HU values, need to be considered. The paper substrate influences the spatial frequency distribution of the phantom scans. This phantom type is of very limited use for dual-energy CT analyses.

KEYWORDS
15 quantitative imaging/analysis, 18 image reconstruction, 26 biomarkers, 3 IM-CT, 8 phantoms, physical

1 | INTRODUCTION

Computed tomography (CT) is an important modality for clinical imaging. It has become the gold standard for the evaluation of acute conditions in emergency settings[1] and it is the main workhorse in cancer surveillance. However, the use of ionizing radiation in CT may increase the patient's risk of cancer.[2–4] Hence, in contrast to magnetic resonance imaging, volunteer measurements for research purposes only are usually not possible. Thus, anthropomorphic phantoms are widely used instead. Specific application scenarios are acquisition protocol optimizations for dose reduction, task-based image quality assurance, validation of image classification, or reproducibility of biomarker extraction of lesions, as in radiomics.[5] However, most current phantoms lack the fine texture and true anatomic representation of patients, which would be desirable for radiomics analysis.

Several techniques are used to manufacture CT phantoms, such as ex-vivo animal specimens,[6,7] liquid-based phantoms,[8,9] or solid phantoms with homogeneous regions of resins, plastics or waxes.[10] While ex vivo animal specimens have a limited durability, liquid-based and homogeneous solid phantoms do not exhibit realistic texture. A promising technique for CT phantom manufacturing is 3D-Printing. The most commonly used 3D-Printing technique is fused deposition modeling (FDM).[11,12] However, the CT texture and microscopic detail of FDM 3D-printed models is severely limited.[13] Other, higher cost 3D-printing techniques such as polyjet or stereolithography (SLA) printing may be used to achieve finer structural detail.[14,15]

Another approach for generating precise anatomical structures is 3D-printing using iodine-containing ink on paper. This technique was first published by Theodorakou et al. in 2004.[16] Other groups have also shown the feasibility of iodine-ink printing technology for 2D and 3D X-ray imaging.[17,18] It has in the meantime become commercially available (PhantomX, Berlin, Germany).[19–21] Showing high detail in both anatomical structure as well as CT texture it is a promising tool for radiomics feature analysis. It has been used to study radiomics features obtained with different CT acquisition and reconstruction parameters, using real CT phantom scans[22] and simulated virtual phantom images.[23]

The aim of this study is to analyze the quantitative characteristics of such an iodine-ink based 3D-printed phantom in detail. A comparison of the input dataset and the printed phantom was performed in terms of attenuation values, radiomics features, as well as for the frequency distribution. Design considerations and the manufacturing process are described. Potential pitfalls for radiomics testing for such phantoms are presented.

2 | MATERIAL AND METHODS

2.1 | Design, manufacturing and testing workflow

Figure 1 shows the workflow of phantom design, production and analysis.

Suitable datasets for the phantom need to be selected (Step 1) and compiled into a single dataset (Step 2) for phantom production using the iodine-ink based 3D-printing process (Step 3). In the next step, CT scan data is acquired (Step 4) for further analysis (Step 5).

2.2 | Dataset selection and compilation

In theory any grayscale dataset can be printed using the iodine-ink 3D-printing technique. The dataset can be chosen freely and adapted according to the setting in which the phantom is to be used.

In this study three datasets were combined in the phantom: (1) IBSI (Image Biomarker Standardisation

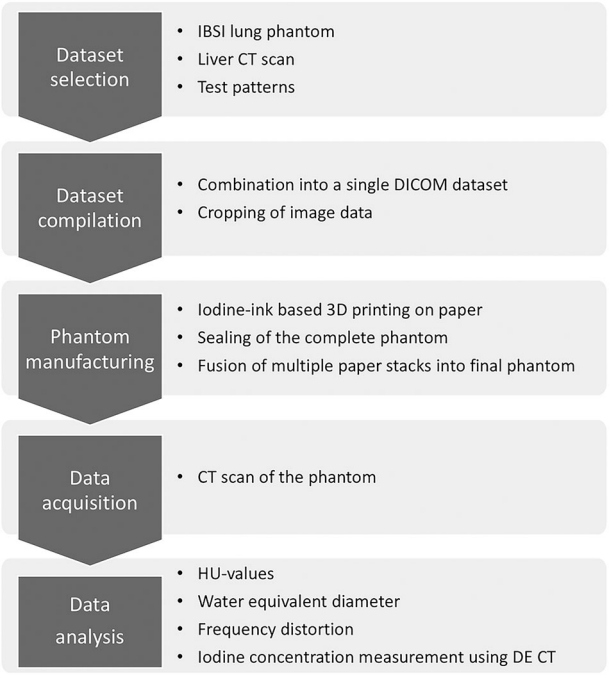

FIGURE 1 Workflow of phantom design, production and analysis. IBSI (= Imaging Biomarker Standardisation Initiative) lung phantom: published lung phantom dataset containing a lung tumor.

Initiative) lung dataset,[24,25] (2) liver dataset, (3) test patterns. Dataset 1) and 2) are based on real patient CT scans, resulting in phantom sections with realistic anatomical detail and real texture features of healthy and diseased tissue. Dataset 1 is a published dataset of a CT scan of a hemithorax containing both unaffected lung and a lung tumor.

The hemithorax was mirrored at a central plane going through the spine to get a full thorax (Figure 2a). The underlying IBSI dataset has been widely used for radiomics feature analysis. Dataset 2 is the upper part of an abdominal CT scan of a colorectal cancer patient (Figure 2b). Image acquisition and reconstruction parameters for the original patient scan of the liver dataset are provided in the supplemental materials (supplemental Table S1). The liver of the patient in this dataset contained both benign lesions (cysts, hemangioma) and a malignant (pathology proven colorectal cancer metastasis) lesion. This part of the dataset was used to analyze the stability and discriminative power of radiomics features in the phantom by studying the different liver lesions and varying scan parameters.[22] Dataset 3 comprises five artificially generated test patterns: a ring pattern, the Siemens star, white noise, dot patterns, an intensity ramp (Figure 2c). In z-direction the test patterns have an extension of 10 cm. The ring pattern, the white noise and the dot patterns (sinusoids along x, y, and z) vary correspondingly in all three spatial directions. The Siemens star and the intensity ramp do not change in the z-direction. Due to challenges described in detail in the results section, two phantom versions were printed. The intensity ramp of the first phantom represents a linear progression of HU values from 0 to 1000 HU, while in the second version it ranges from −300 to 1000 HU.

The three datasets were combined into one continuous image stack, with dataset 1 on top and dataset 3 on the bottom end (Figure 2d). The data sets were isotropically interpolated to a resolution of 0.15 mm before printing.

2.3 | Phantom production

The phantom was manufactured by PhantomX, Berlin, Germany. Using a conventional inkjet printer, an image of the desired structure is printed on a sheet of paper with aqueous potassium iodide solution ink.[19,20] The different shades of gray correspond to different iodine concentrations, which in turn produce the desired HU values in the CT scan. Three-dimensional structures are built sheet by sheet. The phantom production process was restricted to the paper format DIN A4 (21 × 29.7 cm), the printable in-plane area has a size of 26 cm by 19 cm, which was completely used. The phantom is assembled in the z-direction from multiple individual blocks of paper sheets with a thickness of a few centimeters each. The blocks were formed in three steps[19]: (1) A paper sheet is pulled from the paper tray on the build platform and heat compressed against the layer underneath. (2) The paper is cut with a blade. (3) Glue is deposited in lines on the paper in preparation for the next layer. The final 3D model was assembled by applying a thin layer of glue on top of each block and compressing the blocks for 24 h. The complete dataset covered a total length of 30 cm. The manufacturing process would allow for this to be extended further if needed. After assembly of the small paper stacks, the phantom is sealed with black varnish and markers are added for laser positioning (see Figure 2e).

Since the measured HU values also depend on the geometrical shape of the phantom and the amount and density distribution of the irradiated material, a complete theoretical description is difficult and the empirical generation of calibration curves is necessary, which assigns the desired HU values to the amount of iodine to be printed. Two phantoms with different calibrations were produced (phantom version 1 and 2) because systematic deviations between the HU values in the original data set and the HU values measured in the first printed phantom (phantom 1) were observed as part of this work.

2.4 | Data acquisition

The phantom was imaged with a Siemens SOMATOM Definition Edge (Siemens Healthineers, Erlangen, Germany) CT scanner. The helical pitch factor was 1.0, the rotation time was 0.5 s and the detector

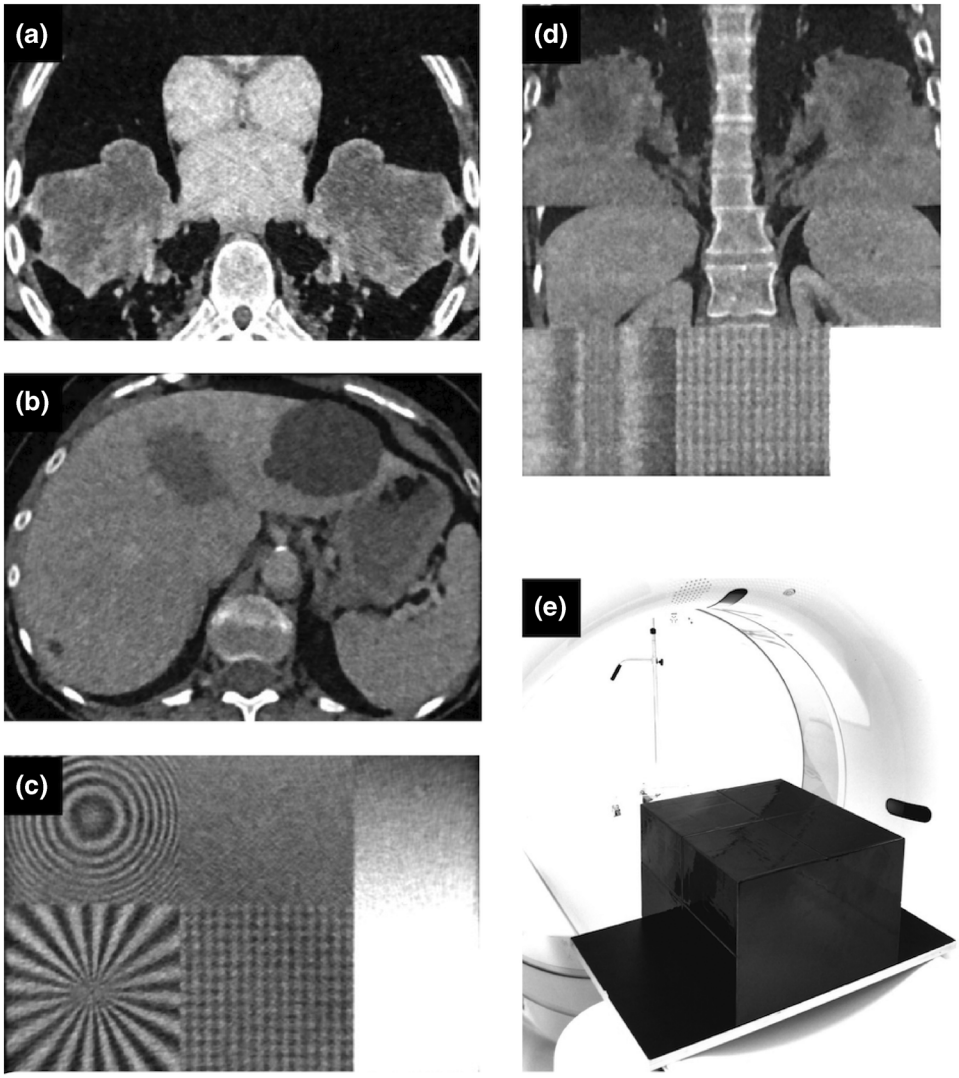

FIGURE 2 CT datasets of the different phantom parts and overall phantom appearance. (a) Transversal CT image of the lung-part, (b) the liver part and (c) the test-patterns. (d) coronal CT image of the phantom. (e) cuboid black phantom positioned on a board in a CT scanner.

configuration was 128 × 0.6 mm (collimation = 38.4 mm) for all phantom acquisitions.

For the analysis of the HU values, phantom CT acquisitions were obtained with different tube voltages of 80/100/120/140 kVp, tube current time product was 318/151/89/58 mAs (no automatic tube current modulation), respectively. This resulted in a volume computed tomography dose index (CTDIvol) of approximately 6 mGy for all acquisitions. Series were reconstructed with filtered back projection (FBP) using a medium smooth kernel (B30f). Slice thickness was 2 mm, slice spacing was 1 mm and the displayed field of view (FOV) was 350 mm.

For the analysis of the water equivalent diameter, phantom acquisitions were performed with a tube voltage of 120 kVp and a tube current time product of 147 mAs (no automatic tube current modulation), resulting in a CTDIvol of approximately 10 mGy. Series were reconstructed with FBP using a medium smooth kernel (B30f). Slice thickness was 2 mm and slice spacing was 1 mm and the FOV was 350 mm.

To compare radiomics features of the original dataset and the phantom, 10 repeated scans on the same CT scanner with the same acquisition and reconstruction parameters as the original patient scan (see Supplemental Table S1) were performed. Only the tube currents and thus the CTDIvol was lower for the phantom because automatic tube current modulation was used and the attenuation of the phantom was smaller compared to the patient due to the restriction to DIN A4 size. As this was a dual-energy CT scan, these images were also used to quantitatively assess iodine concentrations in the phantom.

For the analysis of the frequency distortion, the same acquisition and reconstruction settings as for the analysis of the water equivalent diameter were used. In

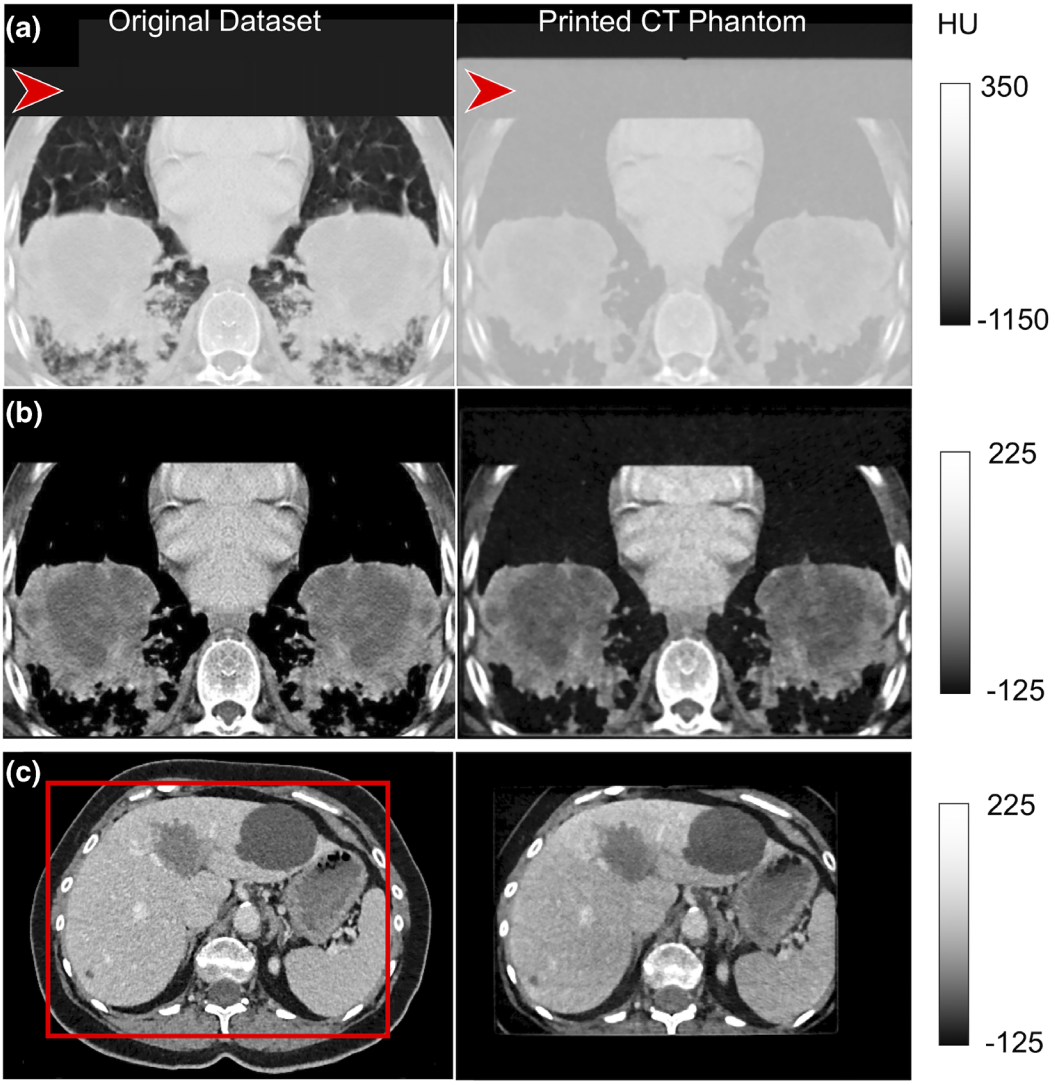

FIGURE 3 Minimum printable HU and maximum phantom size. Left: original CT dataset that serves as a basis for the phantom production. Right: corresponding CT image of the phantom. The windowing is the same for both images. (a) lung dataset, red arrows: in the original dataset the HU above the CT image marked with the red arrow is −1000. This region is represented by pure paper in the phantom. The HU value of the paper is around −100 and defines the minimum HU value that can be reproduced in the phantom. (b) lung dataset shown in a soft tissue window. c) abdominal dataset. The printable area is restricted to 19 × 26 cm2 and is marked by the red rectangle.

addition, series were reconstructed with the Advanced Modeled Iterative Reconstruction (ADMIRE) algorithm at strength level 3 using a medium smooth kernel (I30f). Again, slice thickness was 2.0 mm and slice spacing was 1.0 mm and the FOV was 350 mm.

2.5 | Water equivalent diameter

Since the phantom size was restricted in in-plane dimensions, the phantom was compared to real patients using the water equivalent diameter.

The water equivalent diameter D_W quantifies X-ray attenuation, as it takes into account the size and the composition of objects. It can be calculated from axial CT images using the area (A_{ROI}) and the mean CT number (CT_{ROI}) of a region of interest (ROI)[26]:

$$D_W = 2\sqrt{\left(\frac{1}{1000\ \mathrm{HU}}CT_{ROI} + 1\right)\frac{A_{ROI}}{\pi}}$$

The water equivalent diameter of the two phantoms was determined and compared to water equivalent diameters calculated from patient CT scans. The analysis included contrast-enhanced thoracoabdominal CT scans of 104 patients (57 female, 47 male). The average age was 66 years (19–91 years). The portal vein was chosen as the anatomical landmark where the comparison was performed. A board-certified radiologist (M.M.O.) identified the axial images containing the portal vein in its longest axial course for each scan.

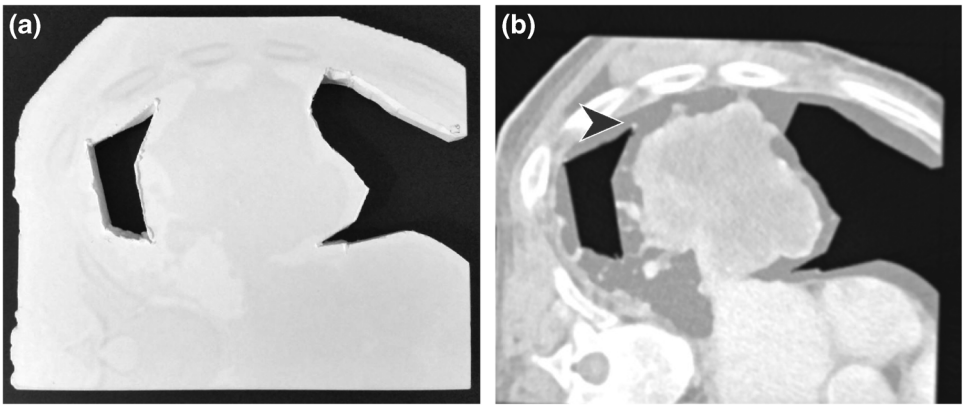

FIGURE 4 Cutting out the lung parenchyma. (a) Test phantom with roughly cut out lung parenchyma parts to achieve more realistic HU values. Only coarse structures can be removed and very sharp edges are created. (b) corresponding CT image. Degradation of cutting edges can lead to hyperdense artifacts (arrowhead).

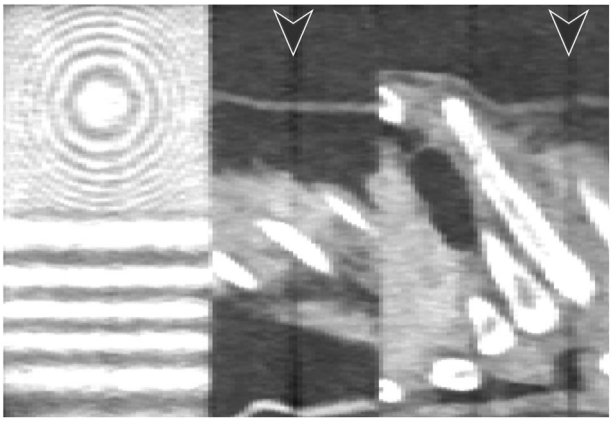

FIGURE 5 Trapped air. Trace amounts of trapped air inside the phantom produce image artifacts at the joints of the phantom parts (arrow heads).

To analyze the water equivalent diameter of the patients, a hand-drawn ROI per patient was used, including the whole patient and excluding surrounding material, in particular the CT scanner table. A linear fit (Python 3.8.8, SciPy 1.6.2) was performed to get a relation between the D_W at the position of the portal vein and the patient weight (or BMI). To analyze the water equivalent diameter of the phantoms, rectangular ROIs were chosen such that only the phantom was included.

2.6 | Radiomics analysis

Radiomics features were extracted in three dimensions from the liver patient scan that was used as input for the phantom production, and from the 10 repetition phantom scans (phantom 2). The open-source PyRadiomics (version 3.0) python package was used with default parameter settings to calculate 86 radiomics features.[27] The features include 18 first order statistics features (e.g., mean, variance, interquartile range), 22 gray level co-occurrence matrix features, 14 gray level dependence matrix features, 16 gray level run length matrix features and 16 gray level size zone matrix features. A list of the individual features can be found in previous work[22] and documentation of feature definitions can be found online (https://pyradiomics.readthedocs.io/). Shape features were not extracted, as the same ROIs of the input dataset and phantom set were compared. The 86 features were each calculated for four different manually segmented 3D-ROIs (normal liver tissue, cyst, hemangioma, liver metastasis).

Features were normalized individually before a principal component analysis (PCA) was performed. Values of individual radiomics features were compared using histogram analysis.

2.7 | Frequency distortion

To compare the frequency content between original data (or reference noise) and the printed, and then CT scanned, phantoms the normalized spectrum modulus was used.[28] This allows estimating systematic biases and frequency distortion introduced by the paper-based phantom and related printing process. The frequency modulus $\hat{f}(\omega)$ was estimated as follows. Given a Bounding Box (BB) of dimension (X,Y,Z), the 2D Fast Fourier Transform (FFT) for each slice was computed, which was interpolated to obtain a square array of maximum dimension between X and Y. The null frequency modulus $|\hat{f}(0)|$ was set to 0. Then the modulus $|\hat{f}(\omega)|$ was cumulated across all Z slices to reduce estimation uncertainty. The 2D modulus map is further transformed into a 1D signal by using only the spectrum profiles along x and y, which we cumulate to approximate an isotropic spectrum. Preliminary investigations showed that 1D spectra along x and y were very similar. Finally,

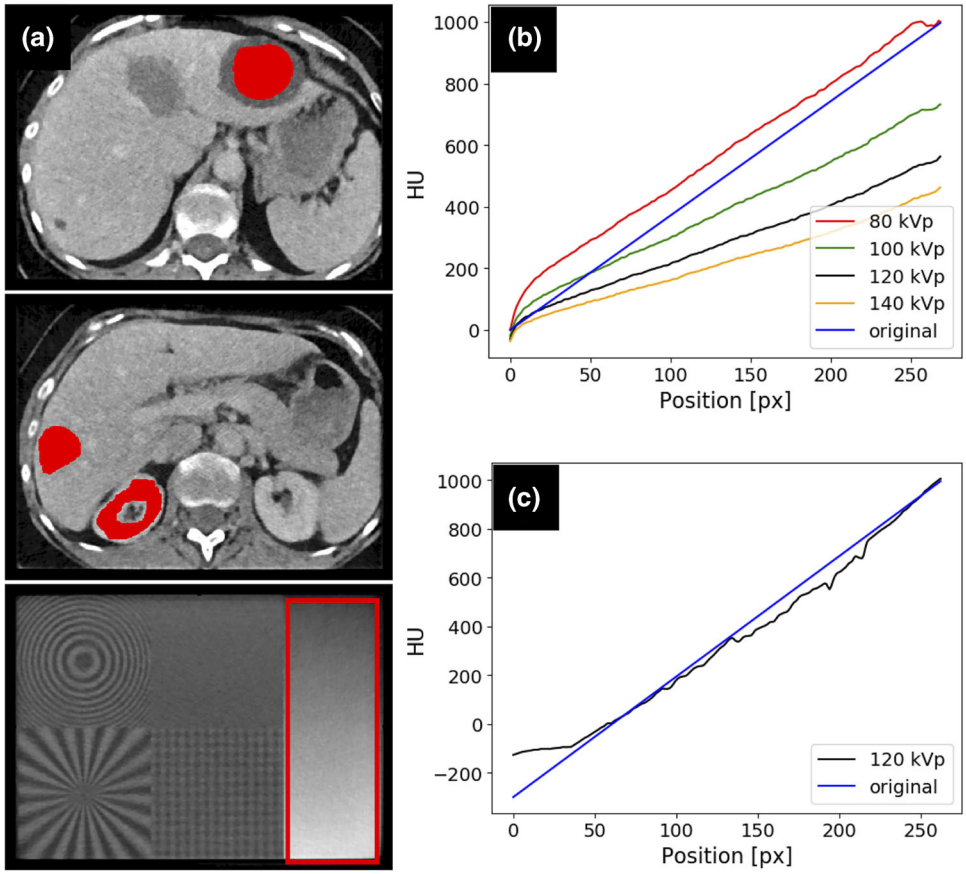

FIGURE 6 Original and printed/measured HU values. (a) ROIs used for the comparison of measured HU values in the phantom versus HU values in the original dataset. (b) HU values along the intensity ramp (see red rectangular ROI in a, bottom)) for phantom version 1. There are strong deviations from the expected HU values (blue line) for all different tube voltages. (c) phantom version 2: the new empirical calibration leads to a better agreement between expected (blue line) and measured HU values (black line).

TABLE 1 HU values (mean ± standard deviation, minimum/maximum) in the three ROIs (liver cyst/healthy tissue, kidney) for the original dataset and the two printed phantoms

ROI	Original	Phantom 1	Phantom 2
Liver cyst	6 ± 14, −44/49	9 ± 12, −32/54	−1 ± 13, −41/41
Liver healthy tissue	122 ± 15, 78/178	73 ± 12, 29/118	94 ± 12, 45/132
Kidney	160 ± 20, 94/220	94 ± 18, −5/148	134 ± 21, 30/187

Note: The agreement of the measured HU values with the original values is improved in phantom version 2.

the spectrum is normalized by a division with the maximum of $|\hat{f}(\omega)|$.

3 | RESULTS

3.1 | Minimum printable HU and maximum phantom size

The CT scans of the respective iodine-ink printed parts are shown on the right. The windowing is identical for both images in one row. Since paper is the substrate material for the iodine-containing ink to be printed on, the minimum HU value that can be represented in the phantom is determined by the HU value of the paper (−100 HU).

In the lung parenchyma it becomes clear that HU values smaller than −100 cannot be printed. Any HU value that is below this threshold in the original data set will result in a HU value close to −100 HU in a CT scan of the printed phantom. Thus, the entire lung parenchyma cannot be visualized with this technique correctly besides possibly very dense structures such as fibrosis (Figure 3a). HU values above 0 can be

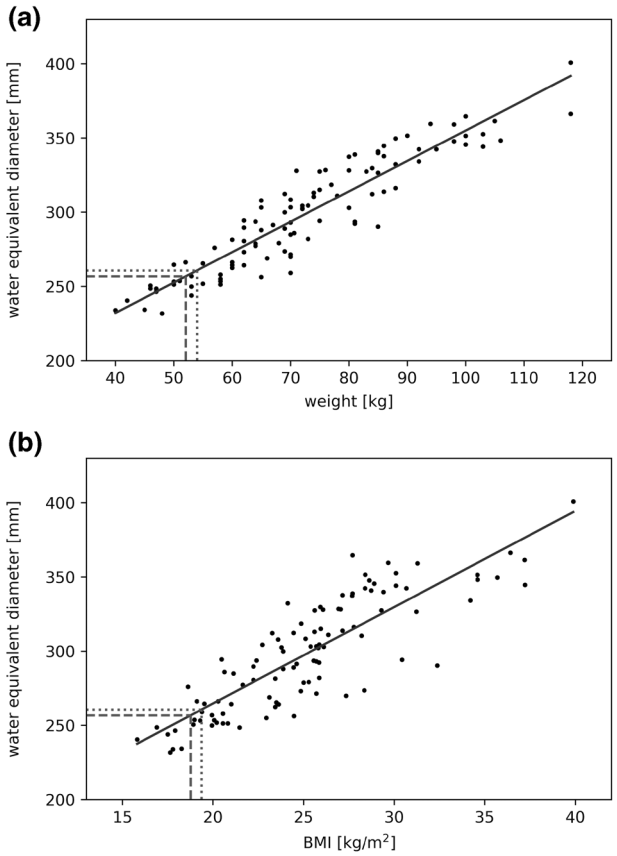

FIGURE 7 Patient water equivalent diameter at the position of the portal vein versus patient weight (a) and patient BMI (b). The solid lines show the best linear fits. The measured water equivalent diameter of phantom 1 (phantom 2) is shown by the dashed (dotted) horizontal lines. The fit equations are used to calculate corresponding weights and BMIs, shown by the vertical lines.

better visualized. For example, the lung tumor is well represented (Figure 3b).

Furthermore, the area of the lung part of the phantom that did not contain any tissue had an HU value of −1000 in the original dataset (Figure 3a, left, red arrow). This area is also represented by the minimum HU value of close to −100 (blank paper) in the phantom (Figure 3a, right, red arrow).

In contrast to this, the abdominal part with typical soft-tissue HU values is realistically represented in the phantom (Figure 3c). The extremely high accuracy in anatomical detail of the printed phantom results in a CT scan which is indistinguishable from a real patient CT scan in the soft tissue window at first glance.

Due to the manufacturing process, the printable area was limited to axial dimensions of 19 cm × 26 cm. This is slightly smaller than an entire abdomen or thorax of an average adult patient. In Figure 3c on the left, the printable area of the liver dataset is marked by the red rectangle.

To potentially overcome the limitation of increased HU-values in air-containing structures additional testing was performed. A thin (2 cm z-direction length) test phantom representing only part of the lung section was built. To achieve lower, more realistic HU values and thus more realistic overall attenuation properties, part of the paper was cut out in areas of lung parenchyma (Figure 4a).

However, the cutting creates very sharp edges with large HU intensity differences of about 1000 HU (Figure 4b). Furthermore, only very coarse structures can be removed by the subsequent cutting-out. Delicate structures cannot be represented. Another pitfall when removing paper parts of the phantom is heating of the cut edges of the remaining phantom. Degradation of these edges can lead to hyperdense artifacts in the images (Figure 4b, red arrow). Because of these limitations this approach was not further pursued.

3.2 | Trapped air

As described above, the phantom is printed in blocks that are about four centimeters thick. These individual blocks are then glued together and assembled into the overall phantom. This can lead to trapped air on the gluing surface and thus small artifacts. The trapped air is not visible on axial slices with reconstruction thicknesses larger than 1 mm. However, it can be visualized as black stripes running through the phantom in coronal or sagittal reconstructions (Figure 5, red arrows). Trapped air mainly occurs in the first two centimeters from the edge of the phantom. Towards the center of the phantom, no trapped air could be detected.

3.3 | Original and printed/measured HU values

The measured HU values in the phantom were compared to the HU values of the original dataset. Three ROIs in the anthropomorphic part of the phantom as well as the intensity ramp (Figure 6a, red areas) were analyzed.

In the anthropomorphic part of the phantom the ROIs contain a liver cyst, healthy liver tissue, and healthy kidney tissue; Figure 6a top/middle). The mean and standard deviation of the HU values within these ROIs are compared between the original dataset and the datasets acquired with the two printed phantoms. The results are shown in Table 1. The agreement between measured and original HU values is better for phantom 2. In all ROIs, the HU values measured in the phantom 2 are lower than in the original data.

The expected HU values of the intensity ramp, as provided in the dataset for printing, are shown by the blue line in Figure 6b-c. The measured values depend on the tube voltage. The phantom manufacturer calibrates the HU values for a tube voltage of 120 kVp. Four different

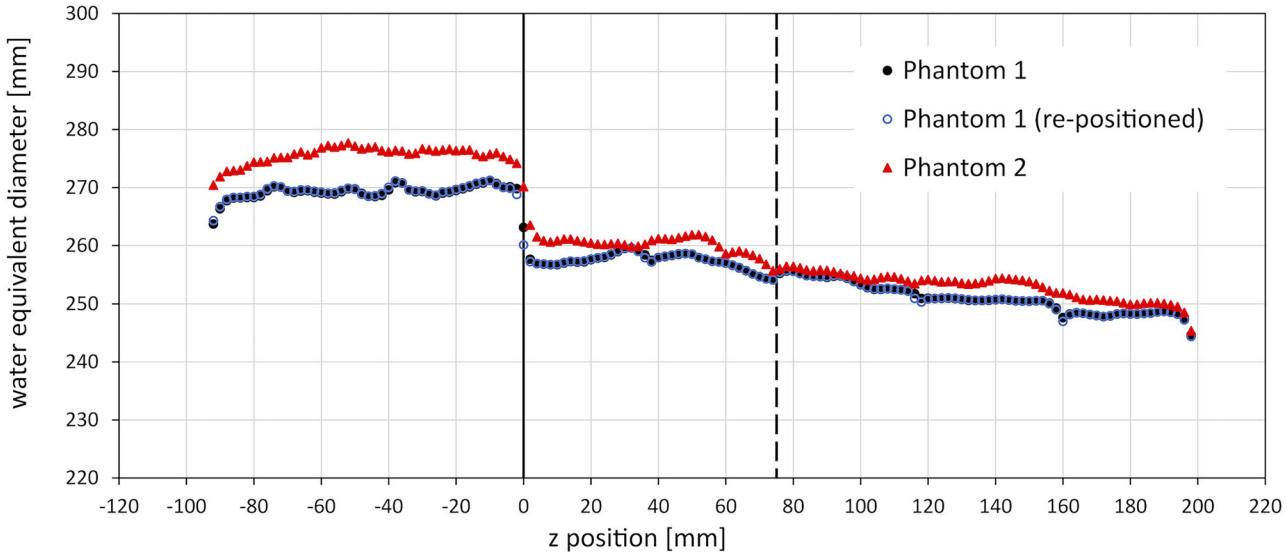

FIGURE 8 Water equivalent diameter along the z-axis of the phantoms. Both phantoms were scanned with the same CT scanner with the same acquisition parameters. The first scan of phantom 1 (black filled circles) was repeated on a different day (re-positioned phantom, blue empty circles). The red triangles show the phantom 2 data. The vertical black solid and dashed lines show the positions where the test pattern part ($z < 0$ mm) and the liver part ($0\ \text{mm} \leq z \leq 75$ mm) end, respectively.

voltages were analyzed (80, 100, 120, and 140 kVp). For all tube voltages, the differences between measured and original HU are apparent (Figure 6b). These large discrepancies over nearly the whole covered HU range led to a new empirical calibration performed by the phantom manufacturer and the printing of a new phantom.

In this second phantom, the HU range of the intensity ramp of the input dataset was also changed. It now starts at – 300 HU in contrast to the previous phantom version, where it starts at 0 HU. The highest HU value in both phantom versions is 1000 HU. The comparison of original and measured HU values for the second phantom is shown in Figure 6c. The flat progression at around −100 HU reflects the minimum printable HU. For a tube voltage of 120 kV, the measured HU values are now in much better accordance with the original HU values, especially for the range between 0 and 150 HU which is most relevant in the abdominal part of the phantom.

3.4 | Water equivalent diameter

Figures 7a-b show the water equivalent diameter at the level of the portal vein for the 104 patients versus patient weight and BMI, respectively.

Using the best fit equations, the measured water equivalent diameter of the phantoms, also at the level of the portal vein, can be used to calculate a typical weight and BMI of a patient with the same D_W. Thus, phantom 1 (2) corresponds to a weight of 52 kg (54 kg) and a BMI of 18.8 kg/m2 (19.4 kg/m^2). As expected, due to the restriction to DIN A4 format, the phantom water equivalent diameter corresponds to that of lightweight patients. Nevertheless, it is still within the range observed for adult patients. Figure 8 shows the water equivalent diameter of both phantoms as a function of z position (long axis of the phantoms).

The transition between the phantom part containing the artificially generated test patterns (denser) and the liver part can be clearly seen. Between the liver and the lung parts of the phantoms, there is no jump in the water equivalent diameter. Due to the minimal printable HU of −100 and the constant cross-sectional area, the average D_W in the lung part is not much lower than in the liver part. Some of the positions where phantom parts are glued together can be seen as small kinks in the water equivalent diameter, in particular for phantom 1, probably due to trapped air (see also Figure 5). The D_W values for phantom 2 are slightly higher than for phantom 1, mainly because of the different HU scale calibration. The D_W is smoother for phantom 2 which could be due to improved manufacturing techniques resulting in reduced trapped air.

3.5 | Radiomics analysis

The PCA of the 10 repetition phantom scans is shown in Figure 9 (top). The four different ROIs (normal liver tissue, cyst, hemangioma, liver metastasis) are separated by the PCA.

In Figure 9 (bottom), the radiomics data from the original patient liver scan is also shown, using the feature normalization and the principal components derived from the phantom scans.

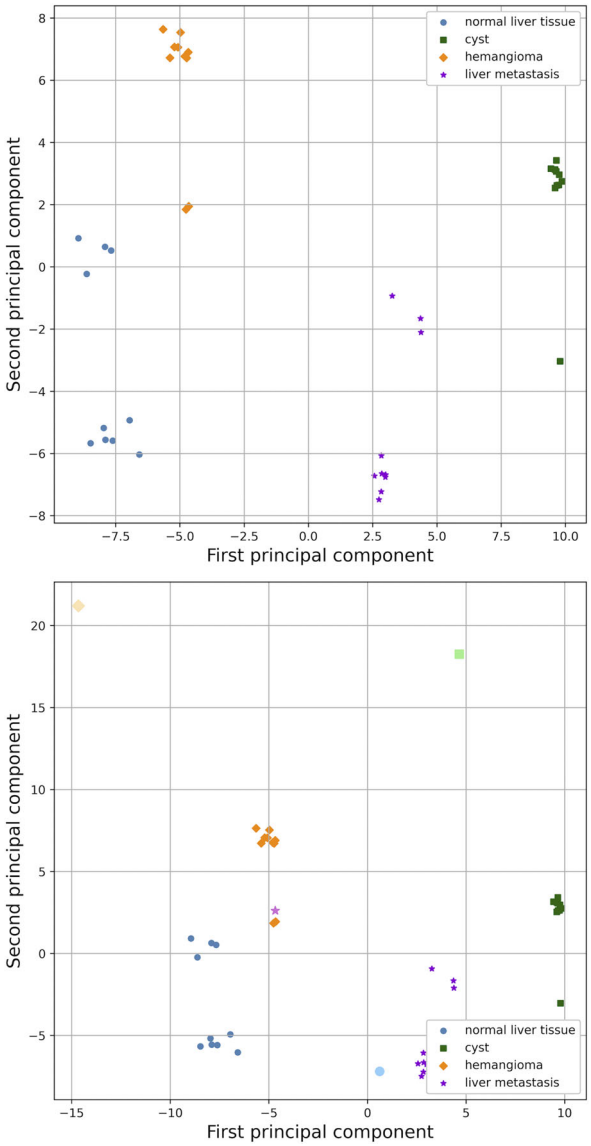

FIGURE 9 Top: Principal component analysis of 86 radiomics features from four different ROIs (normal liver tissue, cyst, hemangioma, liver metastasis) from 10 repetition phantom scans. Bottom: The original patient scan is also displayed (larger symbols, light colors) on top of the phantom data.

Figure 10 shows two examples of individual radiomics features, one first order statistics feature (mean CT number) and one texture feature (gray level non-uniformity). The mean CT number is systematically lower in the phantom scans in all four ROIs. The gray level non-uniformity does not show a systematic shift between the phantom scans and the original patient scan.

3.6 | Iodine measurements

Using dual-energy CT, iodine concentrations measured in the phantom varied between 0 mg/ml and around 47 mg/ml. It should be noted that the attenuation of all structures modeled in the phantom with a density above the minimal printable HU of −100 are due to the presence of iodine ink. This means that structures such as uncomplicated cysts, which in a regular CT do not enhance, show an intrinsic iodine density in the phantom, see Figure 11.

3.7 | Frequency distortion

In a first step the synthetic reference white noise of the input dataset was compared to the corresponding test pattern of the scanned phantom 2 (see Figures 12 and 2c).

For the latter, both FBP with a B30f reconstruction kernel and iterative reconstruction with an I30f reconstruction kernel and 2 mm slice thickness was used. While high frequencies are reduced in both phantom scan reconstructions compared to the reference white noise, this behavior appears to be similar between the two different reconstruction kernels.

Second, the frequency distortion between the original dataset of the liver part and the printed-scanned phantom (FBP, B30f kernel only, 2 mm slice thickness) was compared for the two versions of the phantom. To this end, a rigid registration approach from the Elastix toolbox[29] between the two image series was used and the spectrum modulus for various ROIs was compared. HU distributions and spectrum modulus are compared in Figure 13 for both normal liver tissue and a liver metastasis.

For the two tissue types, we observe a systematic shift in terms of HU distribution and a good correspondence between the frequency spectrums. A systematic boost is observed in the phantom scans for all frequencies. The differences between the HU distributions and spectrum moduli between both phantoms are small. The frequency boost is highlighted in low/mid-frequencies by comparing the spectral signature of one portion of the phantom v2 with pure paper versus pure air (see Figure 14).

4 | DISCUSSION

In this study, a 3D-printed iodine-ink paper phantom is assessed, which allows generating a “patient scan-like” anthropomorphic CT image impression. Not only are solid organs and soft tissues realistically depicted, but also focal benign and malignant lesions, such as cysts, hemangiomas or metastases (that include textures), can be incorporated into the phantom. This extends the use of the phantom beyond that of standard calibration phantoms, since it allows the testing of algorithms for reconstruction, segmentation and classification of anatomical structures and lesions, as is desired for radiomics analyses.

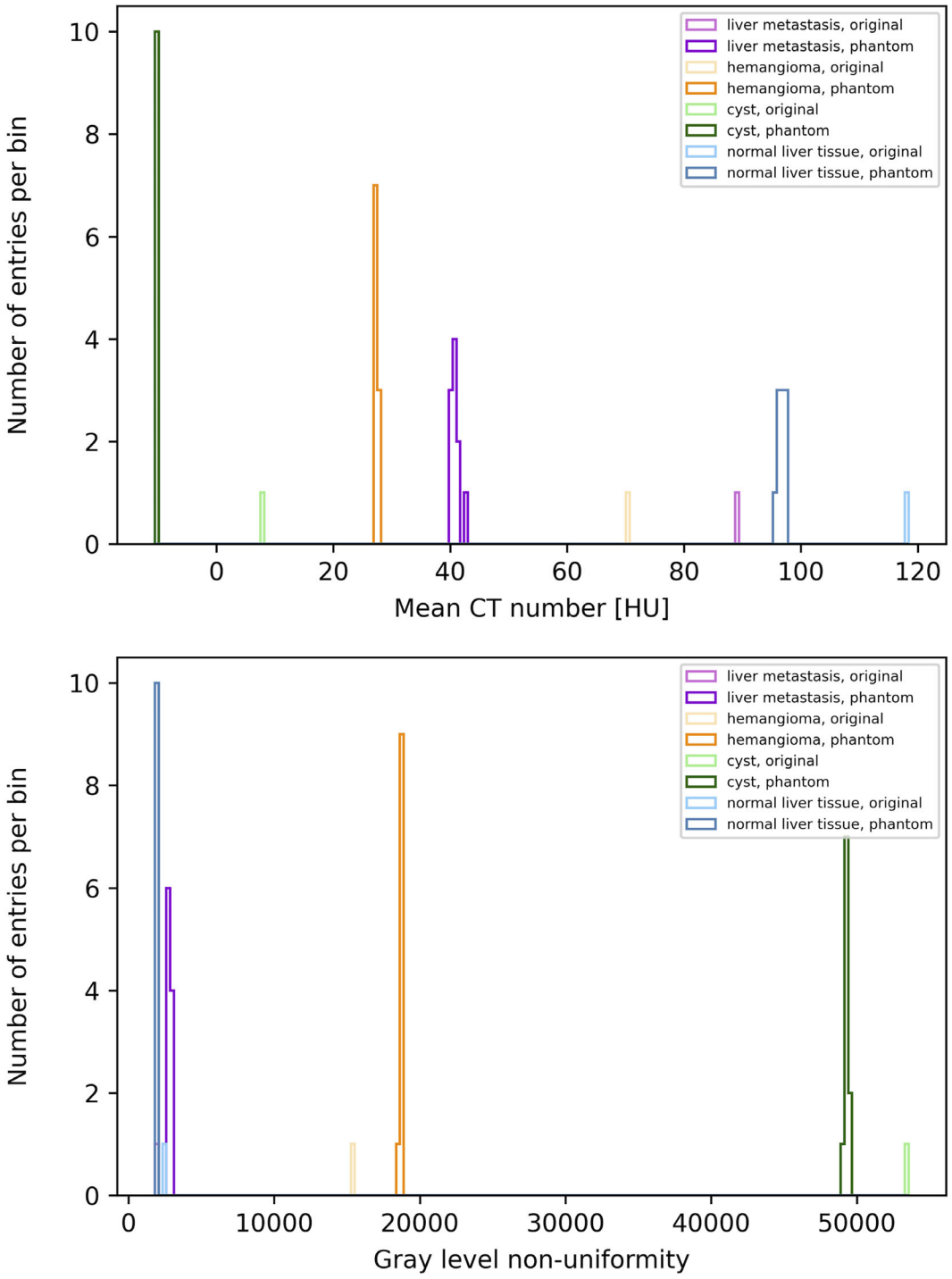

FIGURE 10 Mean CT number (top) and gray level dependence matrix gray level non-uniformity (bottom) of four different ROIs (normal liver tissue, cyst, hemangioma, liver metastasis) for 10 repetition phantom scans (dark colors) and the original patient scan (light colors).

This study has revealed several peculiarities and limitations inherent to the technique of printing iodine-containing ink on paper to create an anthropomorphic CT phantom.

The anatomical depiction of structures is excellent; the geometric resolution of the printable structures is determined by the resolution of the printer. According to the manufacturer the printing resolution is 0.25 mm in-plane (one sheet of paper) and 0.07 mm through plane. The minimum printable differences in HU are about 5 HU. The minimal printable HU is determined by the substrate material, the paper used has an attenuation of around −100 HU. Therefore, structures below the minimum attenuation value of −100 HU cannot be

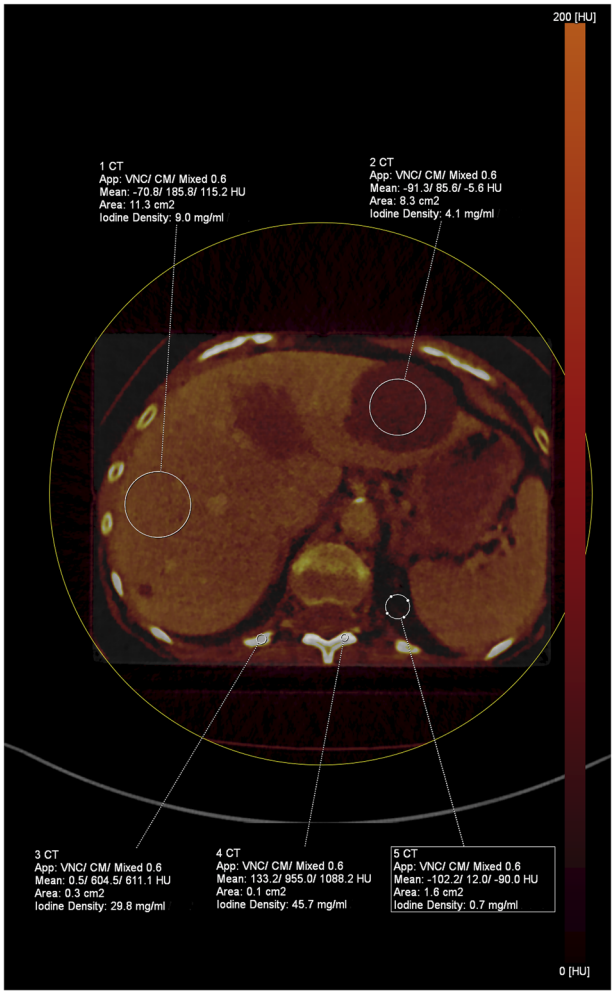

FIGURE 11 Dual-Energy CT scan of the phantom. Iodine concentration measurements in the phantom show that all structures above −120 HU are achieved using the iodine ink. Structures such as the cyst in liver (ROI 2) with an attenuation of −5 HU show an iodine concentration of 4.1 mg/ml. Using higher iodine concentrations, denser structures such as cortical bone (ROI 4) can be modeled with HU values exceeding 1000 HU.

printed with the correct HU values. This is one of the main limitations of this phantom production technique, since structures such as lung parenchyma cannot be reproduced adequately.

Even in the HU domain greater than −100 HU, in which the HU values can in principle be well modulated by the printing technique, deviations occur between the actually measured values and the values of the original data set. One cause is an imperfect calibration. The calibration was carried out empirically. This may further be aggravated by the fact that different CT scanners were used for phantom calibration by the manufacturer and the scans performed in this study. While CT scanners are calibrated to not deviate more than 2 HU from the reference value of 0 HU (i.e., for water), higher density structures may vary quite a bit more.[30] Another cause for deviations between ground-truth HU and printed or measured HU is that the measured HU values depend on the density of the material that is radiated (beam-hardening). That is, the calibration curve may fit better for one area of the phantom than for another area with a different density. The maximum deviation we observed was around 30 HU (e.g., healthy liver tissue original: 122 ± 15 HU, printed and measured 94 ± 12 HU) for the second phantom with improved calibration. It should be emphasized here that the measured HU values also depend on the tube voltage used (see Figure 6b). Since the attenuation in the phantom above −100 HU is mainly determined by the iodine concentration, the dependence of the measured HU values in the phantom on the tube voltage is more pronounced than is the case in biological tissues. The HU values in the phantom presented here are calibrated to a tube voltage of 120 kV.

This also has implications for the use of dual-energy CT in this type of phantoms. Iodine concentrations can be measured and quantified. However, because all structures contain iodine, HU values for both high and low energy acquisitions will be different from an original input dataset. Furthermore, HU values of any 120kVp-equivalent reconstructions (blended average or virtual monoenergetic image) will be different from any 120kVp(-equivalent) input dataset. Therefore, this type of phantoms is very limited for any use with dual-energy CT.

The printable in-plane area corresponds to 19 cm × 26 cm (approximately DINA4 format). This is slightly smaller than an average adult human chest or abdomen cross-section. Therefore, the original data set may need to be cropped. In z-direction, the phantom is assembled from individual blocks of approximately 3 cm thickness. This can result in trapped air which then leads to artifacts in the form of black stripes on the CT scan.

With regard to analyses concerning the radiation dose, it is important that the X-ray attenuation of the phantom is within the range of values observed for adult patients. By analyzing the water equivalent diameter, it was shown that the phantom corresponds approximately to a patient with a bodyweight of 54 kg and a BMI of 19.4 kg/m2. Thus, although the phantom cross-sectional area is restricted, the X-ray attenuation is high enough to realistically simulate a CT of a slim adult patient. In particular, studying imaging biomarker variation at different radiation dose levels can be expected to give realistic results. As the maximum printable size is used for the phantom, using a cropped image set, the resulting shape is a cuboid, which might generate peripheral reconstruction artifacts, not evident in visual inspection.

Differences in the HU scale between the phantom scans and the input dataset are expected, as discussed above. These effects can explain differences between the phantom and the original patient scan observed in some of the first order statistics radiomics features, for example, the mean CT number (see Figure 10, top).

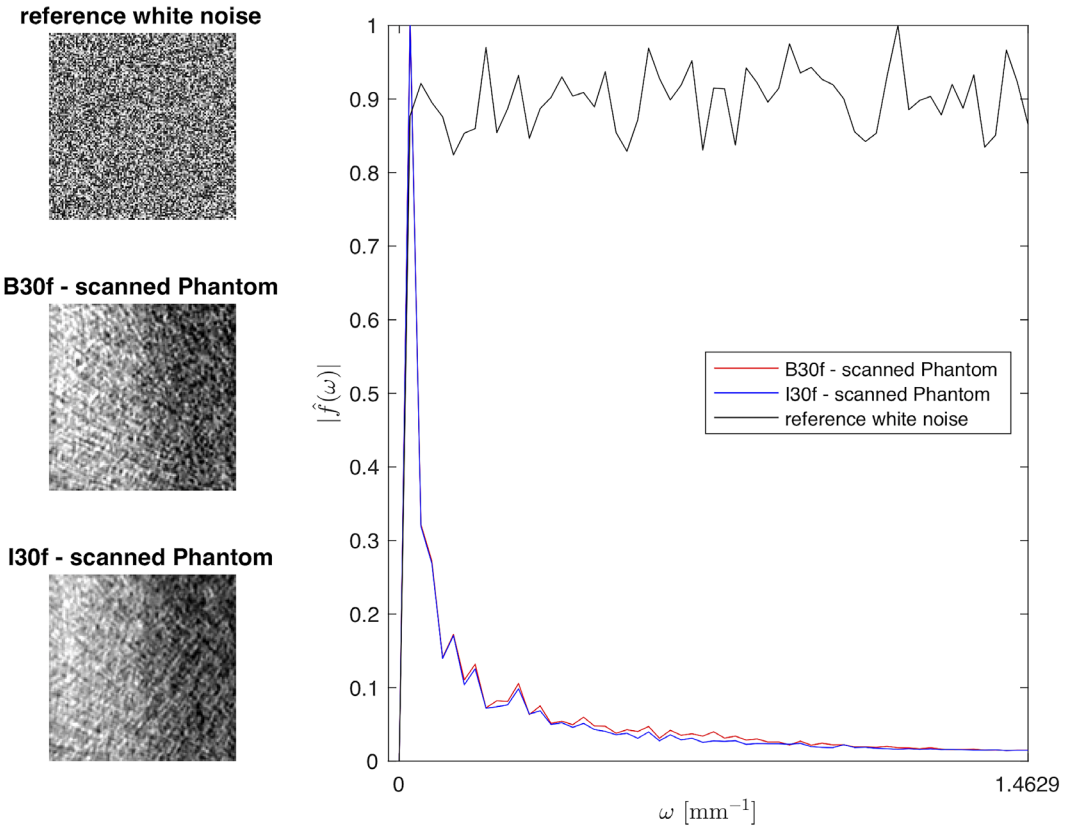

FIGURE 12 Spectrum modulus with reference white noise. The size of the BB is 118 × 118 × 76.

However, systematic shifts between the phantom HU values and the input data HU value do not exclude the use of HU-related first order features in Radiomics studies of the phantom. For example, relative differences between scanners etc. can still be studied, even if there is a systematic HU shift compared to an input dataset. Using a PCA analysis on different tissue types in the phantom and the original dataset, shows persistent clustering of these different tissue types in the PCA graph.

Although there are some expected differences between the phantom and the original patient scan radiomics features, the phantom presents a clear improvement over phantoms that were previously used for radiomics studies in terms of similarity to a real patient CT scan. There are more complex texture features like gray level dependence matrix gray level non-uniformity (Figure 10, bottom) that show discrimination between the ROIs that is very similar for the original patient scan and the phantom scans.

The study was limited to one phantom with a liver part that was derived from a single patient. Therefore, the results are difficult to generalize to other patients or clinical use. Multiple phantoms or phantom parts would be needed to cover inter-patient variabilities, which was beyond the scope of this study.

The frequency analysis of the phantom scans showed two distinct findings. First, for a random white noise pattern incorporated in the phantom, the frequencies were distorted showing a decrease of high frequencies in two reconstructions of phantom scans, using different convolution kernels, compared to the input dataset. This reduction of high frequencies in the white noise pattern may be attributed to a low-pass filter effect of the CT image reconstruction. Second, for healthy liver tissue and a colon cancer metastasis, the Fourier analysis showed a systematic boost for all frequencies in the phantom compared to the original patient CT scan dataset. This frequency boost may be induced by the structure of the paper substrate (see Figure 11). The original dataset for the liver section is a CT scan of an actual patient. Therefore, it already includes CT image reconstruction effects. In summary, we observed a good correspondence between the frequency spectrums of the original patient CT scan and the phantom scans.

Despite the limitations presented here, the technique nevertheless allows realistic imaging of anatomical structures. Phantoms of this manufacturing type are a valuable addition to conventional phantoms for performing texture analysis in CT imaging. For example, the validation of the stability and discriminative power of radiomics features with different reconstruction

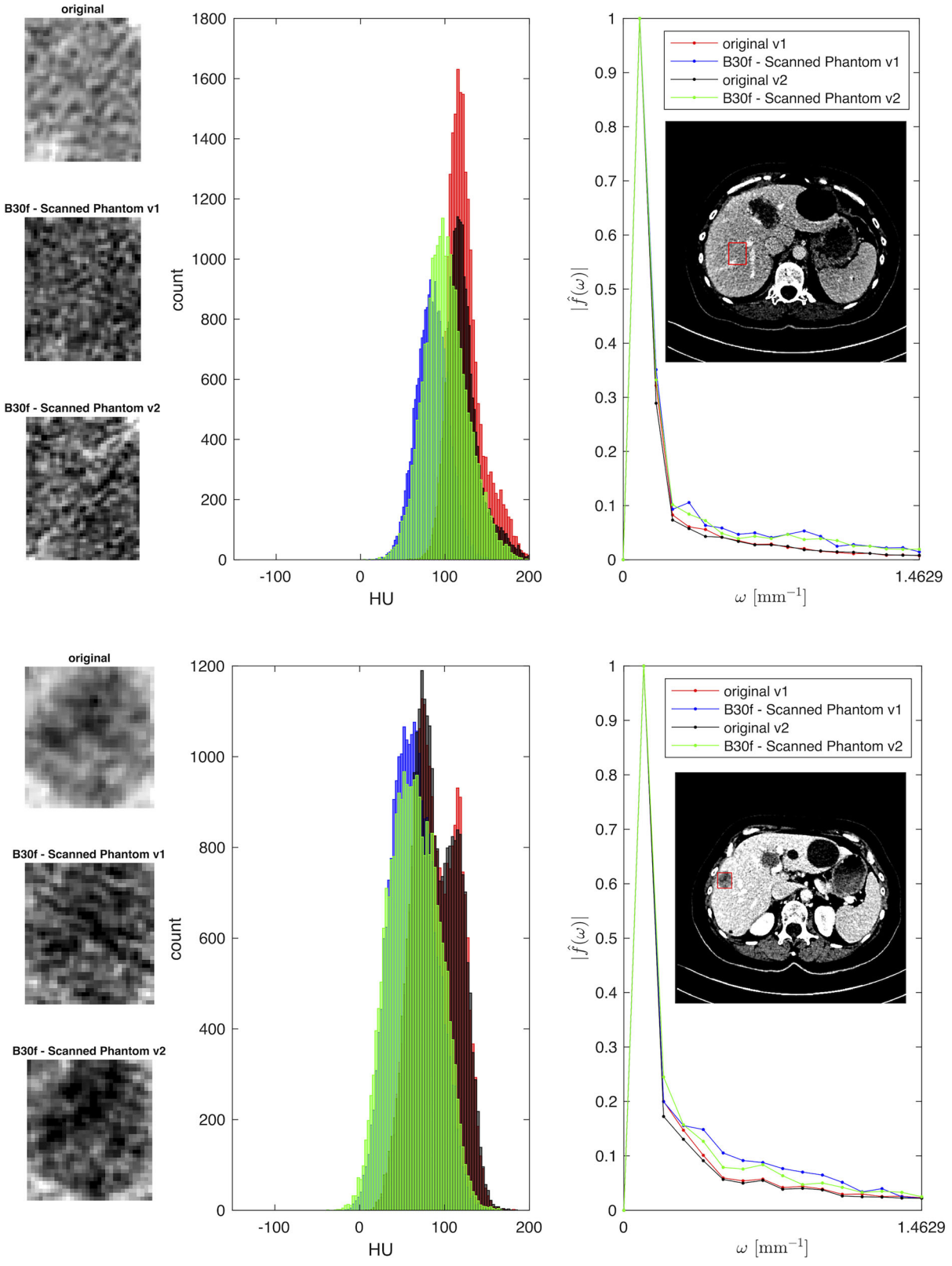

FIGURE 13 HU distribution and spectrum modulus comparison between the original liver dataset and the phantoms. Window level settings used for display are identical between the original and the B30f—scanned phantoms. Top: normal tissue (BB size is 38 × 47 × 12). Bottom: metastasis from colon carcinoma (BB size is 32 × 35 × 22).

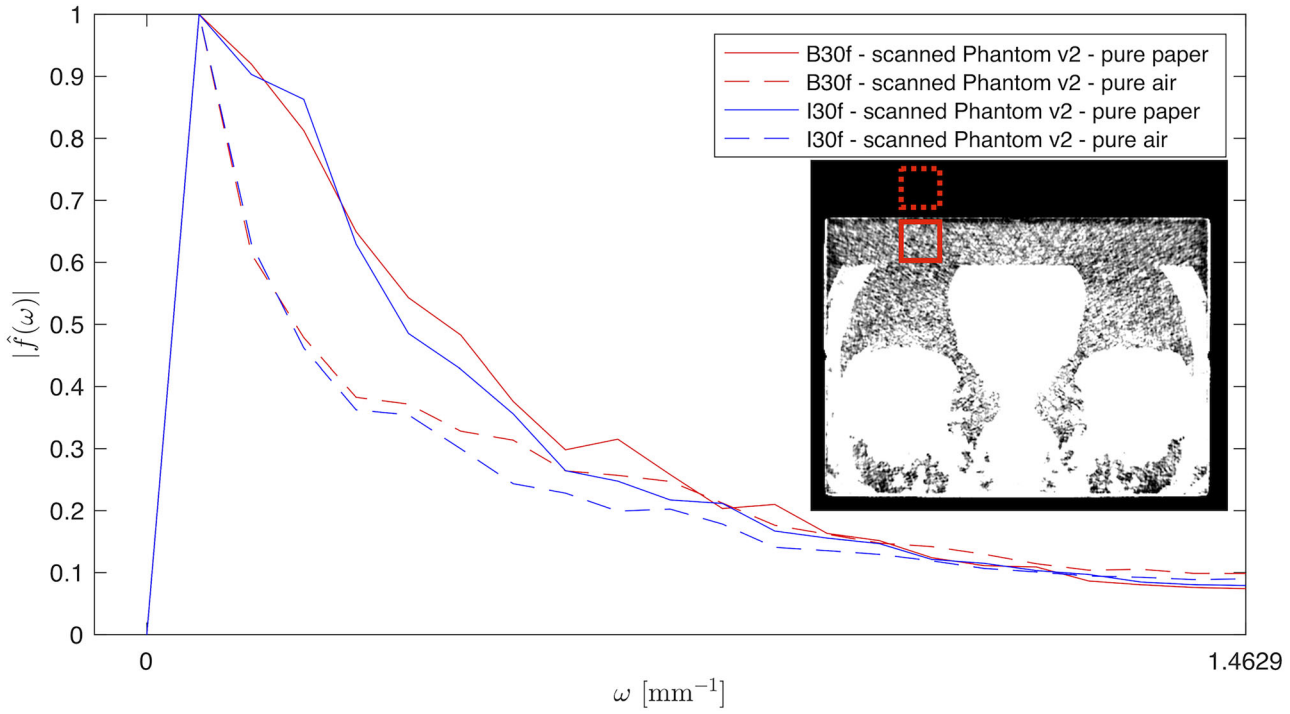

FIGURE 14 Spectrum modulus of one portion of the phantom v2 with pure paper (continuous line) as well as pure air (dashed line). BB sizes are both 43 × 43 × 96.

parameters has been tested using the phantom presented in this work.[22,23] For multicenter studies the influence of different CT scanners on radiomics features may be investigated with such a phantom. Another possible application for the phantom is the evaluation of new reconstruction algorithms at different dose levels, in a more realistic setting, such as for low contrast liver lesions.

The phantom has a very good durability. It is rigid and it is sealed to protect the paper parts from humidity and other factors. Since its purchase 2 years ago, it has been measured at 8 different institutes and no visible changes have occurred so far.

5 | CONCLUSIONS

3D-iodine-ink-printing technology can be used to print anthropomorphic abdominal phantoms with a water equivalent diameter of up to a lightweight adult patient.

Individual radiomics features such as first-order radiomics features can show systematic differences compared to the input dataset, due to differences in HU-values, while some higher order features do not show a systematic bias and exhibit similar values as the input data. In the PCA representation of the radiomics features, the different tissue classes in the phantom scans are separated into similar regions as in the input dataset.

Texture variability in imaging can thus be tested in addition to pure resolution and density. Nevertheless, spatial frequencies are not perfectly preserved by the phantom printing method, as they are influenced by the paper substrate.

Iodine ink printed phantoms are still limited by the manufacturing process in terms of size and small residual air enclosures. As iodine is used to achieve differences in attenuation this phantom type is inherently limited for use with dual-energy CT or different tube voltages.

ACKNOWLEDGMENTS

This work was supported by the Swiss Personalized Health Network (project: QA4IQI).

Open access funding provided by Universitat Basel.

CONFLICT OF INTEREST STATEMENT

The authors have no relevant conflicts of interest to disclose.

REFERENCES

1. Paolantonio P, Rengo M, Ferrari R, Laghi A. Multidetector CT in emergency radiology: acute and generalized non-traumatic abdominal pain. *Br J Radiol*. 2016;89(1061):20150859.
2. Mathews JD, Forsythe AV, Brady Z, et al. Cancer risk in 680 000 people exposed to computed tomography scans in childhood or adolescence: data linkage study of 11 million Australians. *BMJ*. 2013;346(7910):f2360.
3. Pearce MS, Salotti JA, Little MP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet*. 2012;380(9840):499-505.

4. Miglioretti DL, Johnson E, Williams A, et al. The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. *JAMA Pediatr*. 2013;167(8):700-707.
5. Kalendralis P, Traverso A, Shi Z, et al. Multicenter CT phantoms public dataset for radiomics reproducibility tests. *Med Phys*. 2019;46(3):1512-1518.
6. Choi JI, Kim SHH, Kim SHH, et al. Comparison of accuracy and time-efficiency of CT colonography between conventional and panoramic 3D interpretation methods: an anthropomorphic phantom study. *Eur J Radiol*. 2011;80(2):e68-e75.
7. Dachman AH, Lieberman J, Osnis RB, et al. Small simulated polyps in pig colon: sensitivity of CT virtual colography. *Radiology*. 1997;203(2):427-430.
8. Euler A, Solomon J, FitzGerald PF, Samei E, Nelson RC. Can realistic liver tissue surrogates accurately quantify the impact of reduced-kV imaging on attenuation and contrast of parenchyma and lesions? *Acad Radiol*. 2019;26(5):640-650.
9. FitzGerald PF, Colborn RE, Edic PM, Lambert JW, Bonitatibus PJ, Yeh BM. Liquid tissue surrogates for X-ray and CT phantom studies. *Med Phys*. 2017;44(12):6251-6260.
10. Husby E, Svendsen ED, Andersen HK, Martinsen ACT. 100 days with scans of the same Catphan phantom on the same CT scanner. *J Appl Clin Med Phys*. 2017;18(6):224-231.
11. Ma X, Buschmann M, Unger E, Homolka P. Classification of X-ray attenuation properties of additive manufacturing and 3D printing materials using computed tomography from 70 to 140 kVp. *Front Bioeng Biotechnol*. 2021;9:763960.
12. Hazelaar C, Van Eijnatten M, Dahele M, et al. Using 3D printing techniques to create an anthropomorphic thorax phantom for medical imaging purposes. *Med Phys*. 2018;45(1):92-100.
13. Hernandez-Giron I, den Harder JM, Streekstra GJ, Geleijns J, Veldkamp WJH. Development of a 3D printed anthropomorphic lung phantom for image quality assessment in CT. *Phys Medica*. 2019;57:47-57.
14. Ionita CN, Mokin M, Varble N, et al. Challenges and limitations of patient-specific vascular phantom fabrication using 3D Polyjet printing. *Proc SPIE–the Int Soc Opt Eng*. 2014;9038:90380M. doi:10.1117/12.2042266
15. Breitkreutz DY, Bialek S, Vojnovic B, et al. A 3D printed modular phantom for quality assurance of image-guided small animal irradiators: design, imaging experiments, and Monte Carlo simulations. *Med Phys*. 2019;46(5):2015-2024.
16. Theodorakou C, Horrocks JA, Marshall NW, Speller RD. A novel method for producing X-ray test objects and phantoms. *Phys Med Biol*. 2004;49(8):1423-1438.
17. Ikejimba LC, Graff CG, Rosenthal S, et al. A novel physical anthropomorphic breast phantom for 2D and 3D X-ray imaging. *Med Phys*. 2017;44(2):407-416. doi:10.1002/MP.12062
18. Cruz-Bastida JP, Marshall EL, Reiser N, et al. Development of a neonate X-ray phantom for 2D imaging applications using single-tone inkjet printing. *Med Phys*. 2021;48(9):4944-4954.
19. Jahnke P, Limberg FRP, Gerbl A, et al. Radiopaque three-dimensional printing: a method to create realistic CT phantoms. *Radiology*. 2016;282(2):569-575.
20. Jahnke P, Schwarz FB, Ziegert M, et al. A radiopaque 3D printed, anthropomorphic phantom for simulation of CT-guided procedures. *Eur Radiol*. 2018;28(11):4818-4823.
21. Jahnke P, Schwarz S, Ziegert M, Schwarz FB, Hamm B, Scheel M. Paper-based 3D printing of anthropomorphic CT phantoms: feasibility of two construction techniques. *Eur Radiol*. 2019;29(3):1384-1390.
22. Jimenez-del-Toro O, Aberle C, Bach M, et al. The discriminative power and stability of radiomics features with computed tomography variations. *Invest Radiol*. 2021;56(12):820-825. Publish Ah(00).
23. Flouris K, Jimenez-del-Toro O, Aberle C, et al. Assessing radiomics feature stability with simulated CT acquisitions. *Sci Rep*. 2022;12(1):4732.
24. Lambin P. *Radiomics Digital Phantom | CancerData.org*. 2022. Published online 2016:10-11. Accessed. https://www.cancerdata.org/resource/doi%3A10.17195/candat.2016.08.1
25. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749-762.
26. McCollough C, Bakalyar DM, Bostani M, et al. Use of water equivalent diameter for calculating patient size and size-specific dose estimates (SSDE) in CT: the report of AAPM task group 220. *AAPM Rep*. 2014;2014:6-23. Accessed February 22, 2018. http://www.ncbi.nlm.nih.gov/pubmed/27546949
27. Van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77(21):e104-e107.
28. Bracewell RN, The fourier transform and its applications. McGraw-Hill; 1986. https://books.google.ch/books?id=Bk_vAAAAMAAJ
29. Klein S, Staring M, Murphy K, Viergever MA, Pluim JPW. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29(1):196-205.
30. Lamba R, McGahan JP, Corwin MT, et al. CT hounsfield numbers of soft tissues on unenhanced abdominal CT scans: variability between two different manufacturers' MDCT scanners. *AJR Am J Roentgenol*. 2014;203(5):1013-1020.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Bach M, Aberle C, Depeursinge A, et al. 3D-printed iodine-ink CT phantom for radiomics feature extraction - advantages and challenges. *Med Phys*. 2023;50:5682–5697. https://doi.org/10.1002/mp.16373