Received: 18 December 2024 | Revised: 15 April 2025 | Accepted: 2 June 2025

DOI: 10.1002/mp.17950

RESEARCH ARTICLE

MEDICAL PHYSICS

Comparison of three reconstruction algorithms for low-dose phase-contrast computed tomography of the breast with synchrotron radiation

Sandro Donato[1,2,3] | **Simone Caputo**[4] | **Luca Brombal**[5,6] | **Bruno Golosio**[7,8] | **Renata Longo**[5,6] | **Giuliana Tromba**[9] | **Raffaele G. Agostino**[2] | **Gianluigi Greco**[1] | **Benedicta Arhatari**[10] | **Chris Hall**[10] | **Anton Maksimenko**[10] | **Daniel Hausermann**[10] | **Darren Lockie**[11] | **Jane Fox**[12] | **Beena Kumar**[12] | **Sarah Lewis**[13] | **Patrick C. Brennan**[13] | **Harry M. Quiney**[14] | **Seyedamir T. Taba**[13] | **Timur E. Gureyev**[14]

[1]Department of, Mathematics and Computer Science, University of Calabria, Rende, Cosenza, Italy

[2]Department of Physics and STAR-Lab, University of Calabria, Rende, Cosenza, Italy

[3]Division of Frascati, INFN-Laboratori Nazionali di Frascati, Frascati (RM), Italia, Italy

[4]Department of Physics, University of Milano-Bicocca, Milano, Italy

[5]Department of Physics, University of Trieste, Trieste (TS), Italy

[6]Division of Trieste, INFN, Trieste (TS), Italy

[7]Department of Physics, University of Cagliari, Monserrato (CA), Italy

[8]Division of Cagliari, INFN, Monserrato (CA), Italy

[9]Elettra-Sincrotrone Trieste S.C.p.A., Trieste (TS), Italy

[10]Australian Synchrotron, ANSTO, Clayton, Australia

[11]Maroondah BreastScreen, Ringwood East, Australia

[12]Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Australia

[13]Faculty of Health Sciences, The University of Sydney, Lidcombe, Australia

[14]School of Physics, The University of Melbourne, Parkville, Australia

Correspondence
Sandro Donato, Department of Mathematics and Computer Science, University of Calabria, 87036 Rende (CS), Italy.
Email: sandro.donato@fis.unical.it

Funding information
National Health and Medical Research Council, Australia, Grant/Award Number: APP2011204

Abstract

Background: Phase-contrast breast CT imaging holds promise for improved diagnostic accuracy, but an optimal reconstruction algorithm must balance objective image quality metrics with subjective radiologist preferences.

Purpose: This study systematically compares three reconstruction algorithms—filtered back projection (FBP), unified tomographic reconstruction (UTR), and customized simultaneous algebraic reconstruction technique (cSART)—to identify the most suitable approach for phase-contrast breast CT imaging.

Methods: Fresh mastectomy samples were scanned at the Australian synchrotron using monochromatic 32 keV X-rays, a mean glandular dose of 2

mGy, flat-panel detectors with 0.1 mm pixels, and 6-m distance between the rotation stage and the detector. Paganin's phase retrieval method was used in conjunction with all three CT reconstruction algorithms. Objective metrics, including spatial resolution, contrast, signal-to-noise, and contrast-to-noise, were evaluated alongside subjective assessments by seven experienced radiologists. Ratings included perceptible contrast, sharpness, noise, calcification visibility, and overall quality.

Results: cSART excelled in objective metrics, outperforming UTR and FBP. However, subjective evaluations favored FBP due to its higher image contrast, revealing a discrepancy between objective and subjective assessments.

Conclusions: The findings highlight the contrast-focused nature of radiologists' subjective assessments and the potential of cSART for delivering superior objective image quality. These insights inform the development of hybrid evaluation tools and guide clinical translation for future live patient imaging studies.

KEYWORDS
breast CT, synchrotron radiation, reconstruction algorithms

1 | INTRODUCTION

Breast cancer is a complex and significant health issue affecting millions of people worldwide.[1] It is the most common cancer among women globally and can also occur in men, although much less frequently.[2] Breast cancer, like many forms of cancer, arises when cells in the breast tissue begin to grow uncontrollably.[3] These cells can form a tumor, which may be benign (non-cancerous) or malignant (cancerous). Detecting breast cancer early is crucial for improving outcomes. Regular screening mammograms can help detect abnormalities in the breast tissue before symptoms develop.[4] While the implementation of these screening programs has successfully demonstrated a decrease in mortality rates, there is significant potential for further improvement in the diagnostic accuracy of breast imaging. If a suspicious finding is detected, further diagnostic tests, such as ultrasound, MRI, or biopsy, may be performed to confirm the presence of cancer.[5,6] Two-dimensional (2D) mammography has inherent limitations, including low soft tissue contrast and the overlap of different tissues in 2D X-ray projections, which can hinder detecting masses, suspicious lesions or cysts. Overcoming the overlap of diagnostically relevant tissue structures remains a critical challenge in producing clinically valuable breast images. Additionally, there is a potential radiation risk associated with screening mammography,[7] along with significant discomfort for patients due to breast compression.[8] Digital breast tomosynthesis (DBT)[9] and dedicated breast computed tomography (bCT)[10] are newer imaging technologies aimed at addressing the superimposition challenges in 2D mammography. Current data indicate both advantages and disadvantages of DBT and bCT compared to 2D mammography, with no significant dose reduction observed for either technique and limited reduction in discomfort for DBT.

The majority of currently available X-ray-based breast imaging methods primarily rely on X-ray attenuation data, which is based on the differing absorption properties of various soft tissues. However, this approach does not offer substantial contrast for soft tissues due to the minor density variations among different components of breast tissue. Phase-contrast (PhC) imaging represents a more advanced X-ray technique capable of capturing X-ray wave refraction and phase shifts as they pass through objects.[11–13] Particularly with high-energy X-ray beams, phase shifts can provide significantly stronger contrast than attenuation alone. Consequently, retrieving the phase shift of X-ray beams holds great potential for enhancing breast image quality. Various PhC imaging techniques exist, such as propagation-based imaging (PBI),[14,15] analyzer-based imaging,[16] crystal interferometry,[17,18] edge illumination,[19] and grating interferometry.[20] Compared to all other phase-contrast techniques, PBI is experimentally the simplest way to exploit phase shift information because it does not require any X-ray optical elements between the sample and the detector. X-ray PBI CT (PB-CT), which utilizes refraction as well as absorption of X-rays in tissue, shows particular promise due to its superior sensitivity to soft tissues, including tumors. However, PBI requires high spatial coherence of the incident X-ray beam in order to render the phase contrast detectable. In this context, synchrotron facilities have been pivotal in advancing PBI, especially in breast cancer research. Synchrotrons offer unparalleled X-ray brilliance and coherence, enabling high-resolution imaging with exquisite detail. This high

intensity coherent photon flux enhances the detection of subtle tissue density variations, revealing previously elusive levels of detail.[15,21–23] These facilities have fostered collaboration among physicists, pathologists, radiologists, and oncologists, driving breast cancer diagnostics forward.[24–26,15,22,27,13,28]

Ongoing efforts at synchrotrons like Elettra in Italy and Australian Synchrotron focus on refining imaging setups and protocols for potential clinical implementation of PB-CT, aiming for a mean glandular dose (MGD) comparable to clinical mammography or lower.[22,29–31] The requirement of low MGD can be satisfied by either reducing the X-ray fluence per tomographic projection[32,33] or by decreasing the number of projections.[34,35] The former method, while maintaining good angular sampling, results in increased noise in the projection images, leading to a noisier CT image. Conversely, reducing the number of projections significantly below the Nyquist angular sampling criterion introduces significant image artifacts and increased noise when using analytical reconstruction algorithms. Various approaches have been proposed to enhance overall image quality in low-dose CT scans, some of which have been applied to breast CT data, including iterative reconstruction (IR) algorithms.[36–39] Within this framework, optimizing the reconstruction algorithm stands as one of the final stages in clinically implementing PB-CT. The goal is to enhance overall image quality of low-dose CT scans suitable for clinical use, aiming for a total MGD below 5 mGy.

This study aims to compare two novel CT reconstruction algorithms, the Unified Tomographic Reconstruction (UTR)[40] and a customized simultaneous algebraic reconstruction algorithm (cSART),[41] with the well-established filtered back projection (FBP) algorithm,[42] in the context of breast PB-CT. The goal is to evaluate the performance of these algorithms in terms of objective metrics such as contrast, signal-to-noise ratio (SNR), spatial resolution, and the ratio of SNR to spatial resolution, alongside an assessment of their clinical and radiological image quality via a human observer study. Accurate reconstruction is emphasized as essential for diagnostic purposes, impacting the quality and interpretability of medical images. Note that, in principle, "analytical" CT reconstruction algorithms (such as e.g. FBP and UTR) differ significantly in theory and parameter optimization from iterative algorithms (such as cSART). Rather than broadly comparing such algorithms, our study is focused on a specific practical goal: improving the quality of phase-contrast CT reconstruction for breast imaging using near-parallel monochromatic X-ray beams. Conducted on dedicated medical imaging beamlines at synchrotron facilities, our research addresses this issue by evaluating the quality of reconstructed PB-CT images objectively and subjectively.

2 | MATERIALS AND METHODS

2.1 | Breast tissue specimens

Approval was granted by the Human Research Ethics Committee (project number: CF15/3138 -2015001340), and the study utilized 10 fresh mastectomy specimens after having obtained written consent from the patients. Samples were not fixed or preserved, and all were scanned shortly after surgical excision. Although most specimens contained in situ and/or invasive tumors, a few did not exhibit malignancy or contained only benign lesions, as revealed in subsequent pathology examinations. Basic information about each mastectomy sample, including weight, size, and histopathological diagnosis, is provided in Table 1 alongside some key scan parameters.

2.2 | Experimental setup

All tomographic acquisitions for this study were carried out at the Imaging and Medical Beamline (IMBL) of the Australian synchrotron facility in Melbourne. IMBL utilizes a super-conducting wiggler and bent double crystal monochromator system to generate a parallel monochromatic X-ray beam with a cross-section of up to approximately 500 mm (width) $\times$30 mm (height) in the energy range of 20–120 keV, with an energy resolution of $\Delta E/E \cong 10^{-3}$. Two detectors were employed for the scans: (1) a Teledyne-Dalsa Xineos-3030HR flat panel detector with an active area of 296 $\times$296 mm^2 (2988 $\times$ 2988 pixel field of view), a pixel pitch of 99 µm and a frame rate of 40 fps, and (2) a Hamamatsu C10900D CMOS flat panel detector, with an active area of 124.8 $\times$ 124.8 mm^2 (1248 $\times$ 1248 pixel field of view), a pixel size of 100 µm, and a frame rate of 17 fps. The two detectors had very similar performance characteristics in terms of quantum efficiency and spatial resolution. During the scans, samples were placed in a thin-walled plastic cylindrical container measuring 11 cm in diameter. Each scan was conducted at a clinically relevant mean glandular dose of 2 mGy, distributed evenly across either 1200, 1800 or 2400 projections with a uniform angular step of 0.15°, 0.1° or 0.075°, respectively, over 180°, as listed in Table 1. As the mastectomy samples, when placed in the cylindrical container, had the height exceeding that of the incident X-ray beam, the CT scans were performed in several (up to 7) increments ("slabs"), with the height of each slab equal to 2.5 cm and the consecutive slabs overlapping vertically by 1 cm. The overlap was used for subsequent stitching of the slabs in the vertical direction into a single scan with the height of projections exceeding that of the imaged sample. Note that the areas of the overlap (at the top and/or bottom of the slabs) did not receive more incident photons

TABLE 1 The mastectomy samples description.

Sample number	Patient	Breast side	Sample dimension (mm)	Sample weight (g)	Diagnosis	Detector	Number of projections
1	54 yrs female	Left	185 × 190 × 65 (ML x SI x AP).	835	No in situ malignancy; Invasive carcinoma, mixed ductal and lobular features	Hamamatsu	1800
2	40 yrs female	Left	170 × 140 × 25 (ML x SI x AP).	295	High grade DCIS; No invasive tumour	Hamamatsu	1200
3	56 yrs female	Left	230 × 200 × 70 (ML x SI x AP).	1200	Residual intermediate grade DCIS, three foci at previous surgical site, clear of excision margin; No invasive carcinoma	Hamamatsu	1200
4	46 yrs female	Right	220 × 160 × 40 (ML x SI x AP).	773	Negative for malignancy	Hamamatsu	1200
5	53 yrs female	Right	220 × 190 × 40 (ML x SI x AP).	832	High grade DCIS; No invasive tumour	Hamamatsu	1200
6	54 yrs female	Left	200 × 180 × 55 (ML x SI x AP).	906	Post neoadjuvant chemotherapy of invasive carcinoma, no special type: No residual in-situ or invasive disease	Hamamatsu	1200
7	39 yrs female	Right	160 × 140 × 45 (ML x SI x AP).	368	Post chemotherapy and radiotherapy: intermediate grade DCIS; No residual invasive tumour	Hamamatsu	1200
8	44 yrs female	Right	120 × 120 × 20 (MLx SI x AP).	305	Residual invasive lobular carcinoma (post chemo- and radio-therapy)	Xineos	2400
9	53 yrs female	Left	140 × 145 × 27 (SI x ML x AP).	342	Residual tumour bed, No residual invasive tumour	Xineos	2400
10	49 yrs male	Left	128 x 25 x 25 (ML x SI x AP).	234	Invasive carcinoma, no special type, Grade 2	Xineos	2400

Abbreviations: AP, anterior to posterior; DCIS, ductal carcinoma in situ; ML, medial to lateral; SI, superior to inferior; yrs, years old.

on average than the "central" areas of the slabs. The extra exposure in the overlap areas was compensated by the lower intensity, due to the roll-off of the incident beam which had an approximately Gaussian profile with the standard deviation of ∼1 cm in the vertical direction.

The scans employed quasi-plane monochromatic X-rays with an energy of $E = 32$ keV. The distance between the sample and the detector was 6 m, while the source-to-detector distance was 143 m. Dark-current images (with no beam) and flat-field images (with the beam, but without the imaged sample) were collected immediately before and after each CT scan. For all CT scans presented in our paper, ion chamber readings were recorded during acquisition. The chamber was positioned in the X-ray beam upstream of the sample, at a fixed distance from the source. These readings were converted into photon fluence rate and corresponding air kerma rate using a well-established standard protocol. Since the exposure time for each CT projection was known, we calculated the incident photon fluences in the sample plane per projection. This calculation accounted for X-ray beam attenuation due to the air gap between the ion chamber and the sample, as well as the geometrical magnification of the beam between the chamber and the detector. MGD was subsequently calculated using the measured air kerma and the conversion coefficients obtained by Monte Carlo simulations, based on the theoretical framework of Johns and Yaffe.[43] These simulations utilized cylindrical numerical phantoms of different diameters representing breasts of different sizes and glandularity, surrounded by a 5 mm thick layer of tissue simulating the skin.[44] Our modeling and calculations indicate that glandularity has only a minor effect on the MGD conversion coefficients, so we assumed a constant glandularity of 50% for all samples. On the other hand, sample size has a significant impact, therefore, we explicitly accounted for it using a range of different-sized phantoms in our Monte Carlo simulations of the conversion coefficients.

Note that the spatial resolution of the two detectors used in this study was about 150 μm, which was larger than the width of the first Fresnel fringe,

$\sqrt{R'\lambda} \cong \sqrt{5.75m \times 0.3875\text{Å}} \cong 15\mu\text{m}$. Although it may seem counter-intuitive, in-line phase contrast can still be useful in the situations where the detector point-spread function is broader than the width of the first Fresnel fringe – see for example, Ref.[45] In the so-called near-Fresnel regime (in which the vast majority of biomedical in-line phase-contrast imaging experiments have been carried out to date), the "Fresnel number" $N_F = 2\pi\sigma_{sys}^2/(R'\lambda)$ is always larger than one, that is, the spatial resolution of the imaging system, $2\sigma_{sys}$ (which includes the source size and the detector resolution as key components), is always larger than the width of the Fresnel fringe. However, the in-line phase contrast, $C_{PBI} = 0.25\Delta\varphi/N_F$, can still be detectable in this near-Fresnel regime, for example, at the interfaces of different types of tissues, if the ratio of the phase shift, $\Delta\varphi$, at the interface to the Fresnel number is sufficiently large. As a typical phase shift at an adipose-glandular interface in breast tissue, at $E = 32$ keV, can be of the order of many radians, the in-line contrast can be detectable even for $N_F \sim 100$ or more. We certainly observed the in-line phase contrast, in the form of black–white fringes, at the tissue interfaces in the CT projections and in the reconstructed CT slices (without phase retrieval) in the PB-CT scans used in the present paper.

2.3 | Image reconstruction

Image reconstruction was performed using three different algorithms: FBP, UTR and cSART. Before image reconstruction, all projection images were pre-processed through conventional flat field and dark current corrections. Phase-retrieval was also performed using the well-known Paganin's Homogeneous Transport of Intensity Equation (TIE-Hom) algorithm.[46] The TIE-Hom algorithm functions as a low-pass filter controlled by a single parameter, denoted as γ. To ensure accurate phase retrieval, eliminating diffraction fringes at material boundaries, the value of γ should match the ratio δ/β, representing the real decrement to the imaginary part of the relative complex refractive index between the two materials (e.g., glandular and adipose tissue, in the case of breast tissue samples) at a given X-ray energy. Adjusting γ allows for trade-offs between signal-to-noise ratio (SNR) and spatial resolution[47]; increasing γ can improve SNR, albeit at the expense of spatial resolution, while decreasing γ enhances spatial resolution but reduces SNR accordingly. In this study, we employed a "half phase retrieval" approach, setting γ to approximately one-half of the theoretical δ/β value for glandular tissue relative to blood at 32 keV, which is equal to 275. This choice of γ was motivated by our previous studies on optimization of image quality in breast PB-CT.[31] A Hamming filter was employed for FBP reconstructions.

2.4 | The unified tomographic reconstruction algorithm

The UTR method for three-dimensional reconstruction of objects from transmission images collected at multiple illumination directions was described in Ref. 40 The key features of the UTR algorithm are as follows.

1. The UTR method is applicable to experimental conditions relevant to absorption-based, phase-contrast or diffraction imaging using X-rays, electrons and other forms of penetrating radiation or matter waves.
2. It unifies the conventional, phase-contrast and diffraction CT models by intrinsically incorporating both the phase retrieval and the correction for the Ewald sphere curvature (in the cases with a shallow depth of field and significant in-object diffraction).
3. The numerical algorithm implementing UTR, as used in this study, is based on three-dimensional gridding, allowing for fast computational implementation, including parallel processing of multiple input projection images. In principle, this algorithm can be used with any scanning geometry involving plane-wave illumination.

The software code, implementing the UTR algorithm, that was used in the present study, is publicly available.[48] The only non-trivial user-defined parameter of the UTR algorithm relevant to the present study was the "noise-to-signal ratio", which corresponded to the inverse of the SNR in input projections. This parameter was set to 0.05 (corresponding to 5% noise) in all the reconstructions. Increasing the value of this parameter results in stronger low-pass filtering and consequent noise suppression in the reconstructed images, at the expense of spatial resolution.

2.5 | The custom SART algorithm

As discussed in Ref. 41 the main features of the cSART algorithm can be summarized as follows:

1. A relaxation factor, denoted as η, is employed to adjust the iterative corrections. This factor aims to reduce image noise during the reconstruction process. In our approach, η gradually increases from zero to a maximum value over the initial angular steps, then decreases linearly with both the number of iterations and angular steps until it reaches zero at the final angular step of the last iteration.
2. Projections corresponding to various angles are utilized in a random order scheme.
3. Additionally, a bilateral 3D filter is periodically applied to the reconstructed image during the iterative

process. This filter replaces each pixel's content with a weighted average considering both the 3D Euclidean distance and the gray-level difference of neighboring pixels.
4. The optimization process involves adjusting four parameters: the number of iterations, spatial width of the filter ($\sigma_{xy,z}$), pixel intensity difference width (σ_v), and a weighting factor (w).

For optimizing the cSART parameters, reconstructions of a 1 mm thick slice of the tissue were generated using various combinations of the algorithm's parameters. This involved adjusting σ_{xy} and σ_z within the range of 1 to 10 pixels with a step of 1 pixel, σ_v within the range of 0.01 to 0.20 with a step of 0.01, and w within the range of 0.04–0.20 with a step of 0.02. This process resulted in a total of 1800 reconstructions. The number of iterations remained fixed at 5, consistent with a typical SART reconstruction, while the regularization filter was applied every 100 randomly ordered angular steps. Following the optimization discussed in Ref. 41 a subset of optimal parameters was selected based on a threshold value for the frequency peak of the 1D noise power spectrum compared to the equivalent FBP one, evaluated in uniform region of interests within the adipose tissue. Reconstructions falling within this threshold (difference less than 15%) were then compared in terms of signal-to-noise ratio and spatial resolution. The subset of parameters that yields the best values for both metrics was then chosen to perform the full volume reconstruction. This optimization process was repeated three times, once for each different number of acquired projections. For datasets with 1200 and 1800 projections, the optimal subset of parameters was: $w = 0.04$, $\sigma_{xy,z} = 10$, $\sigma_v = 0.20$. For datasets with 2400 projections, the optimal parameters were: $w = 0.06$, $\sigma_{xy,z} = 10$, $\sigma_v = 0.07$.

2.6 | Two types of reconstructed slices

The reconstructed slices produced by the three different algorithms were initially produced in the planes corresponding to the coronal view in mammography. Note that within the theoretical framework of TIE-Hom (Paganin's method) utilized in this paper (and in many other publications), $\delta(\mathbf{r})$ and $\beta(\mathbf{r})$ are always related via a simple multiplicative constant, γ, making it trivial, in principle, to calculate one distribution from the other. However, if the assumption of homogeneity is violated in some places of the sample, then, while the reconstruction of β by the TIE-Hom method remains quantitatively accurate everywhere outside a narrow vicinity of sample edges and interfaces between different tissues (which can be either over-smoothed or retain traces of Fresnel fringes), the distribution of δ may become qualitatively incorrect.[49] Therefore, it is generally preferable to report the results of a TIE-Hom reconstruction in terms of β, rather than δ. Many researchers working with CT data tend to report the results in the form of the linear attenuation coefficient, μ, which is directly related to β, $\mu = (4\pi/\lambda)\beta$. However, we prefer to report the results in terms of β, because it is dimensionless, unlike μ.

For radiological assessments, we digitally reoriented the images into the axial view, representing the craniocaudal view commonly used in breast imaging. While maintaining the original in-plane resolution of 100 μm, we employed 30-pixel binning to create slices measuring 3 mm in thickness, for assessment purposes. The latter binning was performed as follows. Each "column" consisting of 30 pixel values at a fixed transverse (x,y), location inside a stack of 30 adjacent original 100 μm slices was sorted into two bins, the "lower" bin containing all pixel values lower or equal to the selected threshold value $\beta_{Thresh} = 2 \times 10^{-10}$ and the "upper" bin containing the pixels values higher than the threshold value. This particular threshold value was chosen on the basis of analysis of many reconstructions of mastectomy samples, collected at IMBL with plane monochromatic X-rays with $E = 32$ keV, as an optimal threshold between the β values of soft (adipose and glandular) tissues and calcifications. If the upper bin was empty (indicating the absence of calcification at this location), the values in the lower bin were averaged, producing a single "denoised" soft-tissue value of β for the resultant 3-mm thick slice. If the upper bin was not empty (i.e., a calcification was present), the output "denoised" calcification value of β was made equal to the average pixel value in the upper bin only. This procedure was developed in collaboration with radiologists and was employed in order to avoid averaging of β values of microcalcifications with those of z-adjacent pixels containing soft tissues. The 3 mm axial slices were produced with a 1.5 mm step, that is, with a 15-pixels overlap of the consecutive 30-slice stacks of original thin slices. This overlap was created in order to reduce apparent "jumps" in the appearance of consecutive 3 mm thick slices during the subjective radiological evaluation. For the evaluation, the 3 mm axial slices were saved in DICOM files, together with sufficient information about the sample, to enable convenient examination of these files on medical PACS systems. For the DICOM files, the data was also converted from the original 32-bit floating-point to a 12-bit integer format by a linear mapping,

$$I_{out} = I_{out,min} + \left(I_{out,max} - I_{out,min}\right) \frac{\left(\beta_{in} - \beta_{in,min}\right)}{\left(\beta_{in,max} - \beta_{in,min}\right)} \quad (1)$$

with fixed mapping parameters equal to $\beta_{in,min} = 5 \times 10^{-11}$, $\beta_{in,max} = 7 \times 10^{-10}$, $I_{out,min} = 0$ and, $I_{out,max} = 2^{12} - 1 = 4095$ with the last two parameters representing the lowest and highest pixel intensity values in the output rescaled 12-bit images. All input values smaller

than $\beta_{in,min}$ were converted to $\beta_{in,min}$, and all input values larger than $\beta_{in,max}$ were converted to $\beta_{in,max}$, prior to the linear mapping. The values $\beta_{in,min} = 5 \times 10^{-11}$ and $\beta_{in,max} = 7 \times 10^{-10}$ were chosen to be just below the lowest possible value of β for soft breast tissues (at $E = 32$ keV) and above the highest possible value of β for calcifications, respectively. Therefore, any pixel value outside the range $[\beta_{in,min}, \beta_{in,max}]$ could appear in the original reconstructed floating-point images only due to noise, artefacts or extraneous inclusions, such as, for example, surgical clips. Choosing this input range for the linear mapping allowed us to effectively maximize the contrast of the essential features in the output 12-bit images, without losing any important information.

2.7 | Objective assessment

Objective assessment of CT-reconstructed coronal and axial slices was performed using the contrast, spatial resolution and signal-to-noise ratio (SNR) as measures of image quality. It was shown in Refs.[50,51] that, in the case of CT imaging at a fixed radiation dose, the ratio of SNR^2 to spatial resolution in the third power is proportional to Shannon's information capacity of the imaging system. In other words, this ratio reflects the capacity of the system to extract information from each detected photon about the 3D distribution of the refractive index on which the photons were scattered. It is also known that, when an image is post-processed by means of linear filtering (e.g., convolution or deconvolution), its SNR and spatial resolution change. For example, low-pass filtering typically increases the SNR, but spoils the spatial resolution. However, the ratio of SNR to spatial resolution (in the appropriate power, depending on the dimensionality of the image) remains constant after such image filtering operations[50] and can be changed only by means of non-linear processing such as, for example, machine learning or other methods utilizing a priori information about the imaging system or the imaged object.

The contrast was measured in the images by selecting rectangular regions across boundaries between the adjacent adipose and glandular tissue areas, 1D-averaging the pixel values along the shorter dimension of the rectangle and creating a histogram from the resultant averaged 1D profile along the longer dimension of the rectangle. The histogram was created by dividing the range of all 1D-averaged pixel values into five equal intervals and creating five bins containing the pixels with values in the corresponding intervals. The contrast was then defined as

$$C = \frac{\bar{\beta}_{max} - \bar{\beta}_{min}}{\bar{\beta}_{max} + \bar{\beta}_{min}} \tag{2}$$

where $\bar{\beta}_{max}$ and $\bar{\beta}_{min}$ were the average values of the top and bottom bins, respectively, of the 5-bin histogram. Note that in Equation (2) and in subsequent related formulae below, we used the notation corresponding to the original reconstructed coronal slices which contained 2D distributions of 32-bit floating-point values of the imaginary part of the refractive index, β. When similar measurements were performed on the 3 mm thick axial slices, then, instead of β, the relevant formulae involved the corresponding 12-bit pixel intensity values, I, obtained according to Equation (1).

The SNR of a stochastic function, such as, for example, reconstructed distribution of refractive index, $\beta(\mathbf{r})$ at a point $\mathbf{r}$, was defined as

$$\mathrm{SNR} = \frac{\bar{\beta}(\mathbf{r})}{\sigma_\beta(\mathbf{r})} \tag{3}$$

where $\bar{\beta}(\mathbf{r})$ was the statistical mean ("the signal") and $\sigma^2{}_\beta(\mathbf{r}) = < [\beta(\mathbf{r}) - \bar{\beta}(\mathbf{r})]^2 >$ was the (noise) variance (the angular brackets denote the ensemble average). In the present study, practical measurements of SNR were performed in "flat" regions of reconstructed slices. In this context, the flatness of an image inside a certain region means that the reconstructed values of $\beta(\mathbf{r})$ were approximately constant in that region. We generally assumed that the "spatial ergodicity" hypothesis was satisfied in our images, implying that the ensemble averages, involved in Equation (2) and elsewhere, could in practice be substituted by spatial averages within such flat areas.[52] This allowed us to evaluate the SNR from single stochastic images instead of ensembles of such images. Evaluating the SNR in flat areas of the reconstructed slices also allowed us to minimize the contribution of the natural variations of the tissue density to the measured values of SNR.

The contrast-to-noise (CNR), was defined according to the following formula:

$$\mathrm{CNR} = C \times \mathrm{SNR} = \frac{\bar{\beta}_{max} - \bar{\beta}_{min}}{2\sigma_\beta(\mathbf{r})} \frac{\bar{\beta}(\mathbf{r})}{\left(\bar{\beta}_{max} + \bar{\beta}_{min}\right)/2} \tag{4}$$

Note that the CNR defined in this way was dimensionless and invariant with respect to linear scaling of β. Furthermore, while the contrast C in Equation (4) was measured as described in conjunction with eq. (2) above, the SNR was measured outside the selected rectangle crossing the boundary between adjacent glandular and adipose tissue regions (in which the contrast was evaluated). This was done in order to avoid the influence of the strong tissue density variation inside the selected rectangle on SNR. The square region where the SNR was evaluated (according to Equation (3)) in conjunction with Equation (4) was always selected immediately adjacent to the top or the left side of the rectangle, while ensuring that the region where SNR was measured lied in a flat area of the image.

Since we are interested in estimating the spatial resolution in imaging, we focus on the spatial resolution of the corresponding class of 3D computational imaging systems that is relevant to the present study. Such imaging systems incorporate illumination of a sample with coherent monochromatic incident X-ray beam, transmission (scattering) of the beam through the sample, free-space propagation of the transmitted beam from the sample to the detector, image acquisition by the detector and, finally, a CT reconstruction of the 3D distribution of refractive index in the sample. The spatial resolution of such composite hardware-software imaging systems is defined as the width of the 3D point-spread function (PSF) of the imaging system, that is, the width of response of the imaging system to a delta-function-like input (with an idealized infinitely-narrow 3D point-like sample). Assuming that the imaging system is linear and shift-invariant, any real (reconstructed) image is considered to be a convolution of an "ideal reconstruction", $\beta_{ir}(\mathbf{r})$, corresponding to an imaging system with a delta-function PSF, and a real PSF, $P(\boldsymbol{r})$, of the imaging system: $\beta\ (\mathbf{r}) = (\beta_{id} * P) \equiv \int \beta_{id}(\boldsymbol{r}')P(\boldsymbol{r}-\boldsymbol{r}')d\boldsymbol{r}'$. In turn, the width of the 3D PSF is defined via its second spatial integral moment:

$$\text{Res} = \left(\frac{4}{3}\frac{\int |\boldsymbol{r}-\boldsymbol{r}'|^2 P(\boldsymbol{r})\, d\boldsymbol{r}}{\int P(\boldsymbol{r})\, d\boldsymbol{r}}\right)^{1/2} \tag{5}$$

where $\bar{\boldsymbol{r}} \equiv \int \boldsymbol{r}\, P(\boldsymbol{r})\, d\boldsymbol{r}$. Regarding the choice of normalization factor, (4/3), included in Equation (5), note that, in the case of n-dimensional Gaussian distributions, $P_{\text{Gauss}}(\boldsymbol{r}) = (2\pi)^{-n/2}\ \sigma^{-n}\ \exp[-|\boldsymbol{r}|^2/(2\sigma^2)]$, we get Res = 2σ.

As mentioned above, an objective imaging quality characteristic,[50] closely related to Shannon's information capacity of the imaging system, is proportional to the ratio of SNR to the appropriate power of the spatial resolution. For a 3D imaging system, such as CT, the relevant ratio is $\text{SNR}^2/\text{Res}^3$[50]. Note that in the case of Poisson photon-counting statistics, the latter ratio corresponds to the number of photons (or, more precisely, to the number of noise equivalent quanta, NEQ[53]) per minimal resolvable volume. It is therefore clear, in particular, that $\text{SNR}^2/\text{Res}^3$ is proportional to the incident photon fluence and, hence, is also proportional to the radiation dose, D, delivered to the sample during imaging. Accordingly, the ratio

$$Q_S \equiv \frac{SNR}{Res^{3/2}D^{1/2}} \tag{6}$$

which is closely related to "intrinsic imaging quality characteristic",[50,51] can be used for objective assessment of performance of an imaging system. This characteristic has been shown to be directly proportional to the Shannon information capacity of an imaging system, quantifying the information (in bits) that each detected photon transmits about the imaged sample (see Gureyev et al.[15,50,51]). Therefore, Qs can be used to measure how effectively a CT reconstruction algorithm utilizes detected photons to maximize information about the sample while minimizing the radiation dose. This allows us to address one of the key questions of the present study, that is, to determine which of the three tested CT reconstruction algorithms objectively provides the most (Shannon) information about the imaged breast samples. As in the present study we compared the quality of three different PB-CT reconstruction algorithms using the data from scans collected at a fixed dose (2 mGy MGD), we ignored the constant dose parameter in our objective comparisons and only measured and reported the ratios of $\text{SNR}/\text{Res}^{3/2}$.

Notably, Equation (6) is generally applicable to photon-counting detectors, whereas our experiments utilized flat-panel detectors. However, in practice, applying Equation (6) to SNR measurements from uniform image regions acquired with non-photon-counting detectors (with the use of dark current subtraction) introduces only a multiplicative constant equal to the DQE(0) (i.e., detective quantum efficiency at zero spatial frequency) of the detector.[53] This constant has no impact on our study's findings, as all algorithms were compared using the same input images, and all sample images were acquired under identical controlled conditions.

Measurements of spatial resolution were also performed in "flat" regions of reconstructed slices. Evaluation of the spatial resolution was based on the effect of PSF on the noise distribution in images.[52] We used a method based on the Fourier transform of the equation $\beta(\boldsymbol{r}) = (\beta_{\text{id}} * P)(\boldsymbol{r})$:

$$\hat{\beta}\ (\boldsymbol{k}) = \hat{\beta}_{id}\,(\boldsymbol{k})\,\hat{P}\,(\boldsymbol{k}) \tag{7}$$

where the overhead hat symbol denotes the Fourier transform, $\hat{f}\ (\boldsymbol{k}) = \int\!\!\int \exp(-2\pi\ \boldsymbol{k} \bullet \boldsymbol{r})\ f(\boldsymbol{r})\ d\boldsymbol{r}$. The noisy photon fluence in the ideal reconstruction was assumed to be uncorrelated between different voxels and having constant mean and variance within flat regions. Then the Fourier transform of the ideal reconstruction was also a flat noisy distribution and the width of the product of the two functions in the right-hand side of Equation (7) was determined primarily by the width of the modulation transfer function (MTF), $|\hat{\text{P}}(\mathbf{k})|$. The width of the MTF was straightforward to measure in practice using Fourier transforms of flat regions of reconstructed coronal slices of $\beta(\mathbf{r})$ in accordance with Equation (5) with $|\hat{\text{P}}(\mathbf{k})|$ in place of $\text{P}(\mathbf{r})$. After that, assuming that the PSF was approximately Gaussian, and hence the MTF was also Gaussian, we applied the known relationship between the widths of a Gaussian distribution and its Fourier transform to evaluate the width of the

PSF[54]:

$$\text{Res}\ [P_{\text{Gauss}}] = 2/\text{Res}\left[\hat{P}_{\text{Gauss}}\right] \quad (8)$$

For each objective metric, statistical tests were conducted to evaluate significant differences among the reconstruction algorithms. A paired sample *t*-test was performed using MATLAB 2020a with the Statistical and Machine Learning Toolbox. To account for multiple comparisons among the three algorithms, the Bonferroni correction was applied. This method adjusts the significance level to control for the increased risk of "False positive" errors when performing multiple tests. Since three pairwise comparisons were conducted (cSART vs. UTR, cSART vs. FFBP, UTR vs. FBP), the significance level was divided by the number of comparisons ($\alpha = 0.05/3$), resulting in an adjusted $\alpha = 0.017$.

2.8 | Subjective assessment

This assessment was performed on 3 mm-thick craniocaudal (axial) slices reconstructed using the three proposed algorithms. Seven assessors independently evaluated the radiological image quality in this study. The panel of assessors comprised breast specialist radiologists, diagnostic radiographers and medical physicists. Conducted in a setup resembling digital mammography reading rooms, the assessments employed a high-specification workstation equipped with a single 12 MP monitor and typical tools such as zooming, panning, and window/level adjustment. The evaluation process involved examining the image quality of three different reconstructed techniques displayed in a three-panel (grids) in a synchronized format. Here, the middle panel served as the reference against which the right and left panels were compared. To minimize order effects and sequence bias, the image sets within the panels were randomly allocated for each sample, with assessors unaware of the sequence of the image sets (reconstruction techniques) in each panel.

Using a five-point rating scale, assessors were tasked with evaluating image quality criteria for each sample. This scale ranged from indicating clearly better (+2) or slightly better (+1) image quality compared to the reference, to equal quality (0), slightly worse (−1), or clearly worse (−2) than the reference. For each sample, assessors were asked to rate the following five Image Quality Criteria.

1. Perceptible contrast: the differences in radiolucency between soft tissue regions, indicating how well soft tissue variations were displayed.
2. Sharpness of tissue interfaces: the clarity of boundaries between different tissue types, measuring how well the image visualized transitions between tissues.
3. Calcification visibility: the visibility and sharpness of micro-calcifications, specifically assessing their clarity and prominence in the image.
4. Image noise: the presence of quantum mottle in the image, with a higher score indicating less noise in the test image compared to the reference.
5. Overall image quality: a holistic assessment of the entire stack of images, considering all aspects of image quality.

The observer study employed a multiple-reader, multiple-case (MRMC) design, where all assessors evaluated all images. Inter-observer agreement regarding image ratings was assessed using the intraclass correlation coefficient (ICC). Utilizing SPSS Statistics v28, a two-way mixed model of ICC was generated based on absolute rating scores. It's widely acknowledged that an ICC below 0.4 suggests poor reliability, while values between 0.4 and 0.6 indicate fair reliability, 0.6 to 0.75 imply good reliability, and anything exceeding 0.75 reflects excellent reliability. We also examined the agreement between assessors by calculating Cronbach's alpha for all assessors and performed sensitivity analyses by individually excluding each assessor to ensure there were no outliers in the panel of assessors.

Following this, image quality underwent analysis through visual grading characteristics (VGC) analysis using VGC Analyzer software v1.0.2. For each image criterion, the cumulative distributions of rating data for the test images were plotted against the reference images, yielding a curve. The area under this curve (AUCVGC) served as a metric for measuring the difference in image quality between the two sets. In the interpretation of VGC analysis results, an AUCVGC of 0.5 denotes equivalence between the image sets, values between 0 and 0.5 signify lower quality and those between 0.5 and 1 indicate higher quality in the test images compared to the reference.

Statistical tests employed a nonparametric approach. To establish the confidence interval (CI) of the AUCVGC and calculate p-values for testing the null hypothesis, bootstrapping was conducted with 2000 resamplings of the rating scores. Given the small number of assessors, the analysis considered a random-observer scenario, incorporating bootstrapping of assessors to ensure results' generalizability to the assessor population.

3 | RESULTS AND DISCUSSION

3.1 | Objective image quality measurements

Figure 1 depicts three coronal slices from sample 5, reconstructed using the three considered algorithms. Signal-to-noise ratio (SNR) and spatial resolution were

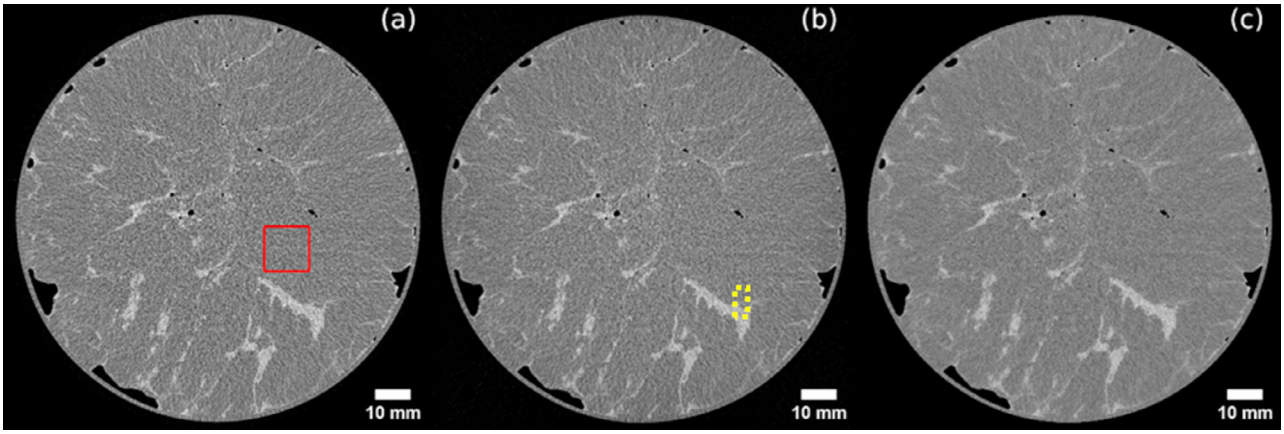

FIGURE 1 Coronal slices from sample 5, reconstructed using the three considered algorithms: (a) the image obtained with FBP, (b) UTR, (c) cSART. The solid red square highlights a region of interest used for the measurement of SNR and spatial resolution, while the yellow dotted rectangle shows a typical selection used for measurement of contrast and CNR.

measured within uniform regions selected in adipose tissue (see Figure 1a). Contrast was measured across fibroglandular and adipose interfaces (see Figure 1b). The objective image quality measurements were performed on two different datasets: the original reconstructions in the form of coronal slices and the axial slices obtained with the 30-pixel binning described in the previous section. For the reconstructions of the coronal slices, which retained the original voxel size, regions of interest (ROIs) of 128 × 128 pixels were used for the measurements of SNR and spatial resolution. For the thick axial slices, we employed ROIs of 64 × 64 pixels, as the selection of uniform regions was more challenging after the binning procedure which increased the apparent presence of glandular tissue. All the measurements were made using the definitions provided in Equations (2)–(8), as implemented in the X-TRACT software.[55] For each sample, three slices were selected to evaluate the objective metrics for each reconstruction algorithm. The primary criterion for selecting the region of interest was the presence of a sufficiently large and uniform area to accommodate the ROIs of the chosen size within adipose tissue. Results of the analysis of thin coronal slices are presented in Figure 2.

Figure 2(a) shows the contrast values, measured according to Equation (2), in thin coronal slices of 10 samples reconstructed using the three algorithms. The contrast was generally highest in the FBP-reconstructed images (p-values against UTR and cSART < 0.001), while UTR and cSART produced lower and statistically indistinguishable contrast (p-value $= 0.14$).

Figure 2(b) presents the SNR values, measured in the same slices according to Equation (3). The results demonstrate that cSART consistently achieved the highest SNR values (minimum = 6.9, maximum = 11.5, average = 8.5), followed by UTR (minimum = 5.1, maximum = 7.2, average = 6.4), with FBP performing the worst (minimum = 4.6, maximum = 6.1, average = 5.6). These differences were statistically significant, as evaluated by paired samples t-tests (all p-values < 0.01). Notably, the minimum, maximum, and average values reported were calculated across all samples, while the individual data points in the plots represent the average values from the three slices per sample.

Figure 2(c) depicts the CNR values, measured according to Equation (4), for the same 10 samples. Similar to the SNR results, cSART yielded the highest CNR values, while UTR and FBP produced comparable results. Statistical analysis confirmed significant differences between cSART and the other algorithms (p-values < 0.01), but no significant difference between UTR and FBP (p-value $= 0.13$).

The spatial resolution, measured using Equations (5), (7), and (8), is shown in Figure 2(d). cSART provided the best spatial resolution (minimum = 170 μm, maximum = 204 μm, average = 182 μm), while FBP exhibited the worst performance (minimum = 205 μm, maximum = 237 μm, average = 218 μm). UTR results fell in between (minimum = 172 μm, maximum = 225 μm, average = 198 μm). Paired samples t-tests produced the following p-values: cSART versus UTR = 0.05 (not significant), cSART versus FBP and UTR versus FBP < 0.001.

Finally, Figure 2(e) presents the calculated proxy for intrinsic imaging quality, as defined by Equation (6), which involves the ratio of SNR to the resolution raised to the power of 1.5. This metric clearly favored cSART, with UTR as the second best and FBP performing the worst (all p-values < 0.001). These results indicate that cSART-reconstructed slices objectively contained more measurable (Shannon) information about the imaged samples compared to slices reconstructed using UTR or FBP.

Figure 3 presents an example of three 3 mm-thick craniocaudal (axial) slices from sample 6, reconstructed using the three proposed algorithms. A large lesion with multiple micro-calcification clusters is clearly visible in the upper-central region of these slices.

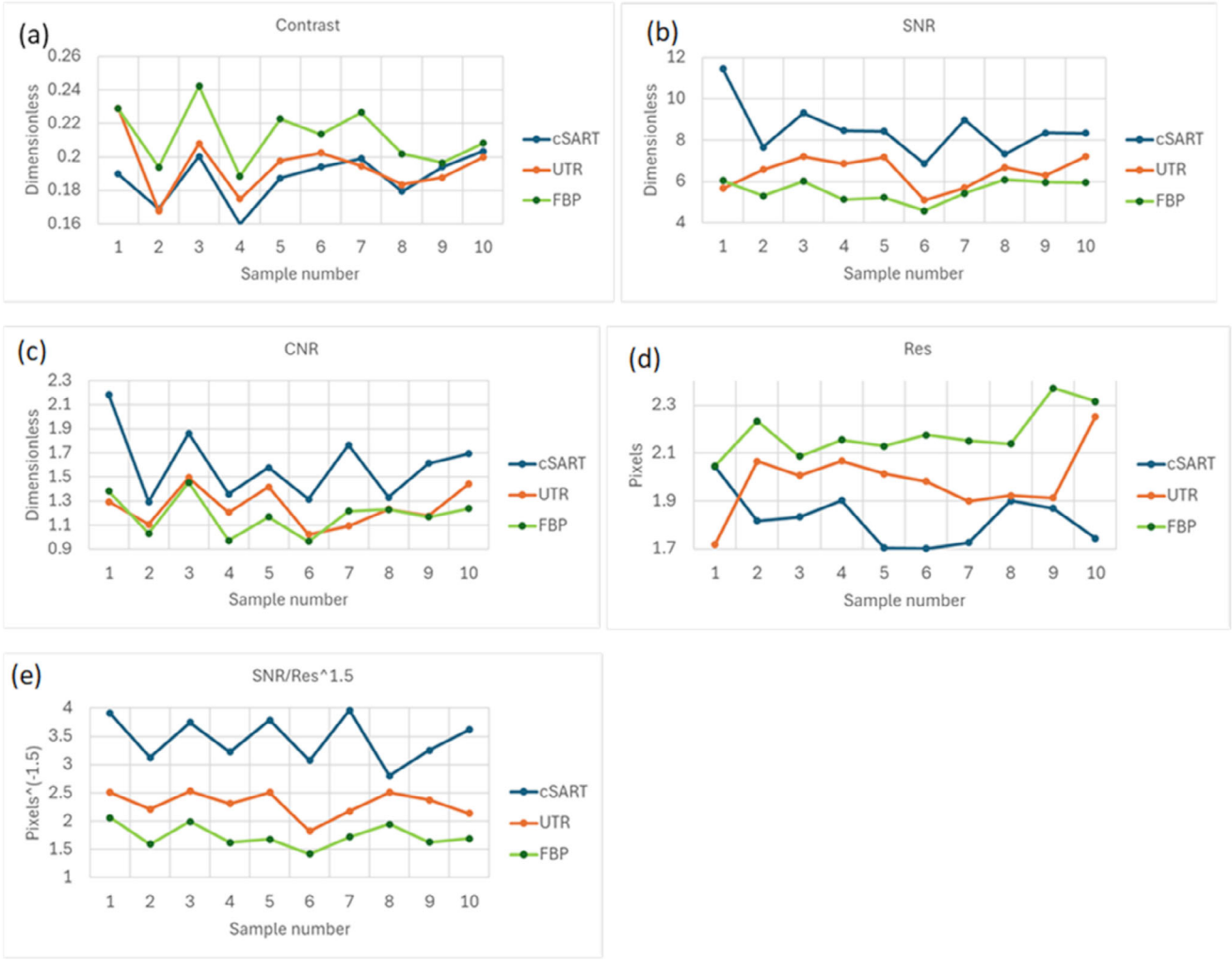

FIGURE 2 Objective measurements of image quality characteristics in coronal slices of 10 different mastectomy samples reconstructed using the cSART, UTR and FBP algorithms from PB-CT scans collected with 32 keV planar monochromatic X-rays and the MGD of 2 mGy. (a) Contrast, (b) SNR, (c) CNR, (d) Spatial resolution (in the detector pixel size), (e) SNR/Res$^{1.5}$ (in the inverse detector pixel size taken to the power 1.5). The individual values shown in this figure, along with their corresponding standard deviations, are reported in Table S1.

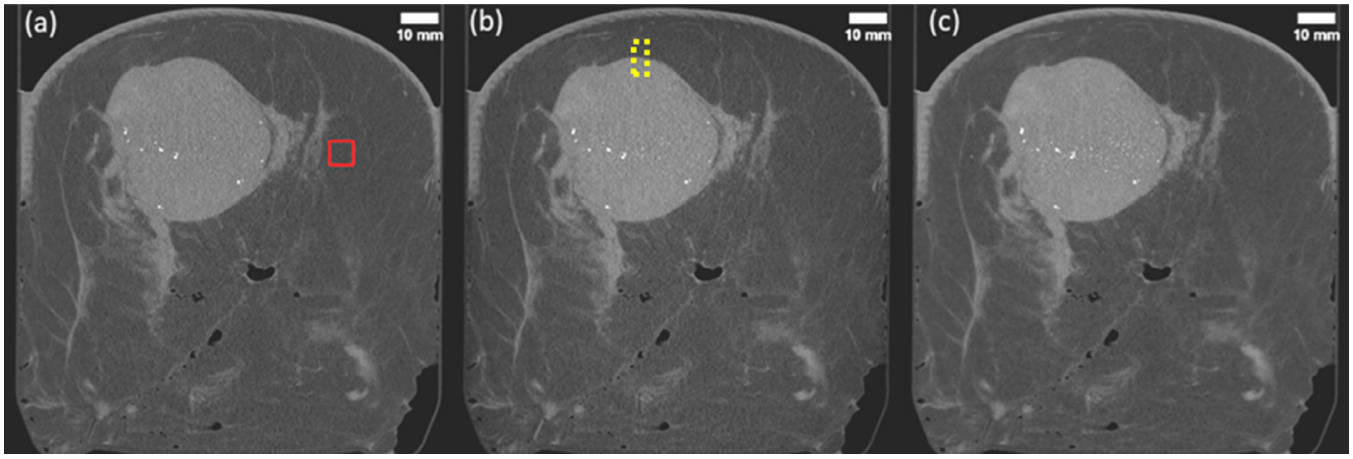

FIGURE 3 Reconstructed 3mm-thick axial slices from sample 6, obtained with the three considered algorithms: (a) the image obtained with FBP, (b) UTR, (c) cSART. The solid red square highlights a region of interest used for the measurement of SNR and spatial resolution, while the yellow dotted rectangle shows a typical selection used for measurement of contrast and CNR.

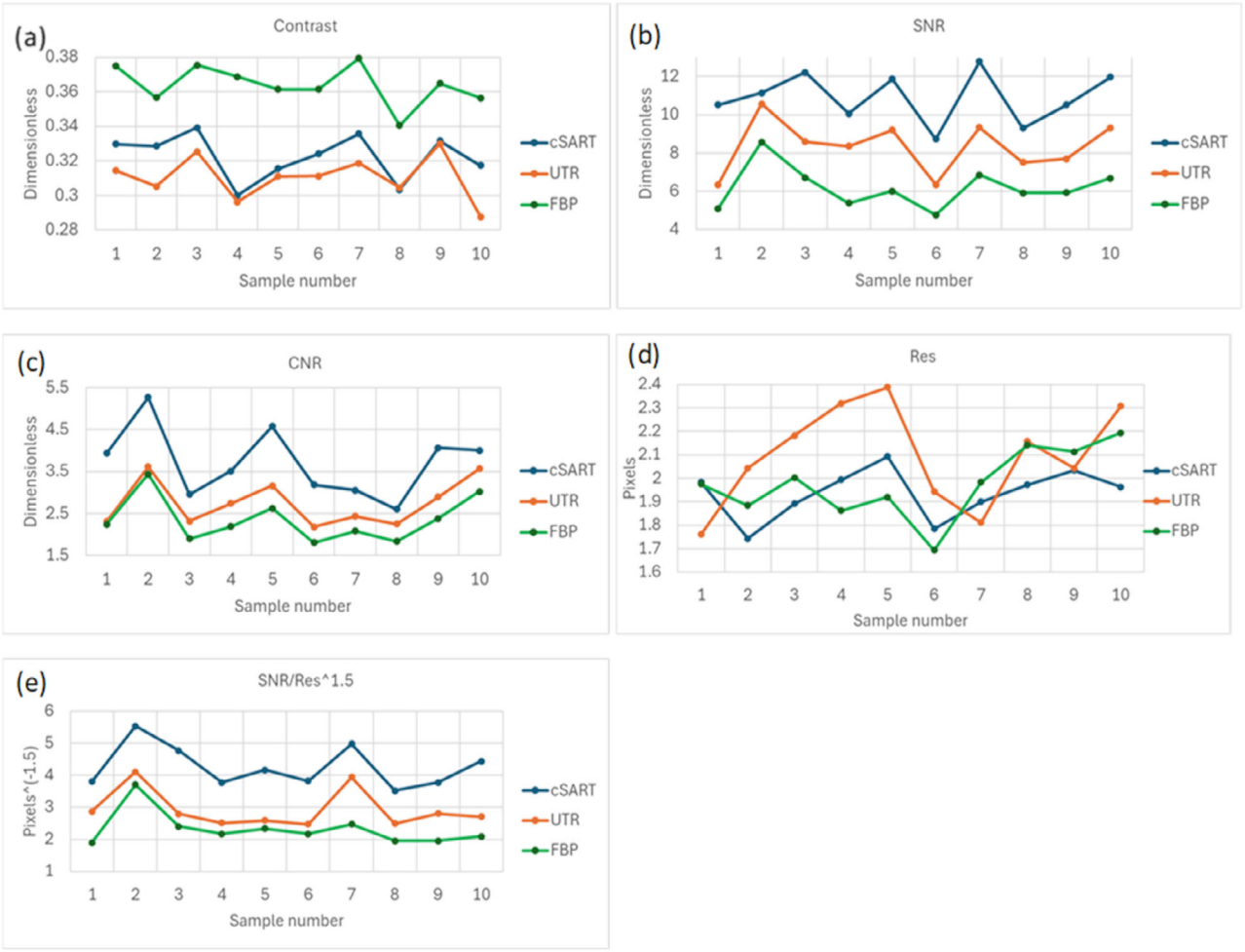

FIGURE 4 Objective measurements of image quality characteristics in 3 mm-thick axial slices of 10 different mastectomy samples reconstructed using the cSART, UTR and FBP algorithms from PB-CT scans collected with 32 keV planar monochromatic X-rays and the MGD of 2 mGy. (a) Contrast, (b) SNR, (c) CNR, (d) Spatial resolution (in the detector pixel size), (e) SNR/Res$^{1.5}$ (in the inverse detector pixel size taken to the power 1.5). The individual values shown in this figure, along with their corresponding standard deviations, are reported in Table S2.

Figure 4 presents the measurements of objective image quality characteristics in 3 mm-thick, 12-bit axial slices. Each data point represents the average value calculated across three selected slices. Overall, the results observed in these thick axial slices closely mirror those obtained for the thin coronal slices presented earlier. Repeating the measurements in the rescaled thick axial slices was crucial, as these slices (rather than the thin coronal ones) were used for the subjective image quality assessments described in the subsequent section.

Figure 4(a) shows that the contrast was significantly highest in the FBP-reconstructed slices (p-values < 0.001). While cSART and UTR produced generally comparable contrast values, the UTR contrast was statistically higher than that of cSART (p-value < 0.01).

Figure 4(b) demonstrates that the SNR results align with those found for the coronal slices. Specifically, cSART consistently achieved the highest SNR values (minimum = 8.7, maximum = 12.8, average = 10.9), followed by UTR (minimum = 6.3, maximum = 10.6, average = 8.3), and FBP yielding the lowest values (minimum = 4.8, maximum = 8.6, average = 6.3). These differences were highly statistically significant, as determined by paired samples t-tests (all p-values < 0.001).

The trends in CNR, shown in Figure 4(c), mirrored those observed for SNR. cSART exhibited the highest CNR values, followed by UTR and FBP, with all comparisons demonstrating statistical significance (p-values < 0.001).

As shown in Equation (4), CNR is inherently proportional to SNR; therefore, their behavior is expected to be similar. However, reporting both metrics remains important for two key reasons. First, confirming the proportionality between measured CNR and SNR for each reconstruction algorithm validates that our implementations of the three reconstruction algorithms functioned correctly using the experimental data. Second, CNR is more directly related to human perception of image quality, making it particularly relevant to readers interested

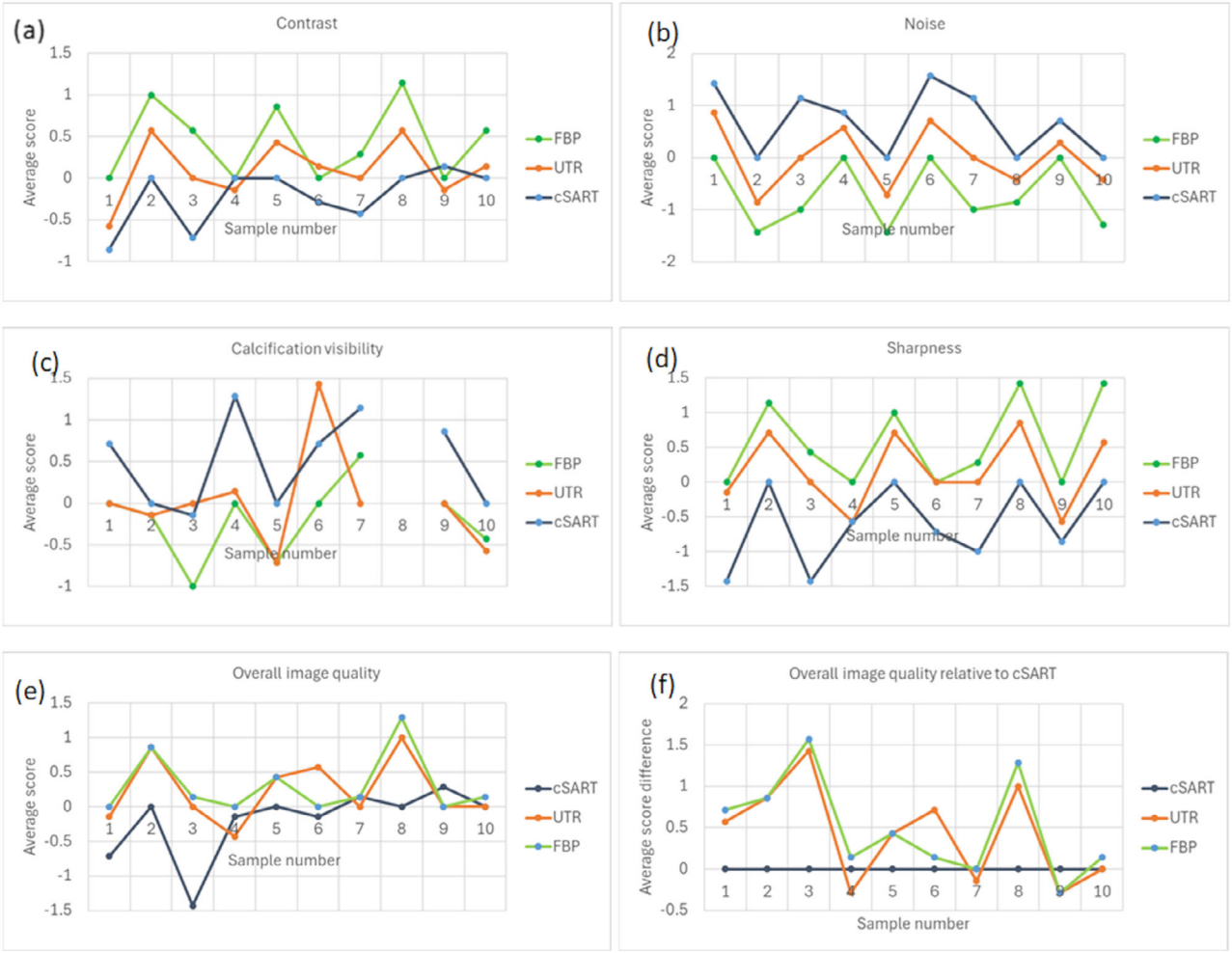

FIGURE 5 Subjective assessment of image quality in 3mm-thick axial slices of 10 different mastectomy samples reconstructed using the cSART, UTR and FBP algorithms from PB-CT scans collected with 32 keV planar monochromatic X-rays and the MGD of 2 mGy. (a) Average (across all assessors) subjective image contrast scores; (b) average image noise scores; (c) average calcification visibility scores (note that sample 8 did not contain any calcifications); (d) average image sharpness scores; (e) average overall image quality scores; (f) differences between the average overall image quality scores for each of the three algorithm and those of cSART.

in visual interpretability. On the other hand, SNR provides insight into the fundamental performance of both the imaging hardware and reconstruction algorithms, reflecting the efficiency of the entire computational imaging pipeline. Moreover, as noted in Equation (6), SNR is directly linked to the intrinsic imaging quality characteristic, making it a crucial metric for assessing imaging system performance.

In contrast, the spatial resolution results measured in the thick axial slices (Figure 4(d)) revealed no statistically significant differences between the three algorithms. The p-values were 0.03, 0.36, and 0.15 for cSART versus UTR, cSART versus FBP, and UTR versus FBP, respectively. However, when the proxy for intrinsic imaging quality—defined as the ratio of SNR to spatial resolution raised to the power of 1.5—was analyzed (Figure 4(e)), the results clearly differentiated the three algorithms. This differentiation was driven primarily by the significant statistical differences observed in SNR (all p-values < 0.001).

3.2 | Subjective assessment of image quality

The results of the subjective image quality assessment are summarized in Figure 5 and Table 2. The calculated ICC for this study indicated excellent agreement among assessors for rating perceptible contrast, sharpness of tissue interfaces, and image noise, with good reliability observed in the ratings of calcification visibility and overall image quality. Note that some features in the images that looked like calcifications were in fact the artefacts from brighter noisy pixels that have been amplified by our thresholding algorithm (see e.g., Figure 3(c)). Despite the appearance of these artefacts in a small number of images, we still consider our current thresholding algorithm (used for conversion from the 0.1 mm thick 32-bit floating-point coronal slices to 3-mm thick 12-bit integer axial slices) the best compromise preventing the calcifications from being "washed out" by the 30-pixel averaging involved in the conversion process.

TABLE 2 Results of assessment of subjective image quality in 3 mm-thick axial slices of 10 different mastectomy samples reconstructed using the cSART, UTR and FBP algorithms from PB-CT scans at $E = 32$ keV X-rays at 2 mGy MGD.

	Inter-observer agreement	Average image quality score			VGCAUC and *p*-value		
Image quality criteria	**ICC**	**FBP**	**UTR**	**cSART**	**UTR against FBP**	**cSART against FBP**	**cSART against UTR**
Perceptible contrast	0.80	0.4	0.1	−0.2	**0.38 (0.01)***	**0.29 (0.01)***	0.39 (0.06)
Sharpness of tissue interfaces	0.90	0.6	0.2	−0.6	**0.38 (0.01)***	**0.17 (<0.001)*****	**0.26 (<0.01)****
Calcification visibility	0.68	−0.2	0	0.5	0.55 (0.35)	**0.68 (0.02)***	0.62 (0.11)
Image noise	0.91	−0.7	0	0.7	**0.70 (<0.01)****	**0.82 (<0.01)****	**0.70 (<0.01)****
Overall image quality	0.62	0.3	0.2	−0.2	0.46 (0.31)	**0.32 (0.02)***	**0.35 (0.03)****

P-values are shown in parentheses, with those below 0.05 significance level indicated in bold. Asterisks *, **, and *** denote significance levels below 0.05, 0.01, and 0.001, respectively.

While the presence of these artefacts may have affected the evaluation of calcification visibility in the subjective assessment, the effect is not expected to be a major one, as only a few images were affected by this problem.

When assessing contrast and sharpness, FBP received the highest scores from the assessors, followed by UTR. FBP demonstrated significantly better performance than both UTR ($p = 0.01$) and cSART ($p = 0.01$) for both criteria, while the difference between UTR and cSART was significant for sharpness ($p < 0.01$) but not for contrast. In terms of calcification visibility, cSART achieved the highest scores, followed by UTR; however, the only significant difference among the three techniques was observed between cSART and FBP. For image noise, cSART was rated significantly better than both FBP ($p < 0.01$) and UTR ($p < 0.01$), while UTR was significantly better than FBP ($p < 0.01$). Regarding overall image quality, which encompasses various image quality criteria, there was no significant difference between FBP and UTR, but both of these techniques were rated significantly better than cSART ($p = 0.02$ and $p = 0.03$, respectively).

4 | DISCUSSION

While the results of the objective image quality measurements clearly show that cSART consistently outperforms both UTR and FBP in terms of various quality characteristics, including spatial resolution, SNR, and CNR, the subjective image quality assessment revealed a general preference for FBP reconstructions. Interestingly, the only objective metric that strongly correlated with the subjective evaluation was the image contrast, which was consistently higher for FBP compared to both UTR and cSART. This outcome suggests that radiologists may place greater importance on image contrast (i.e., the sharpness and visibility of key image features) when subjectively assessing image quality, compared to other factors such as image noise (SNR), which is a major component of objective assessments.

This observation leads to a hypothesis, which could be explored further in future studies: medical imaging specialists, when evaluating medical images, may subjectively assign more weight to contrast, particularly in terms of visual clarity and feature recognition, rather than noise reduction, even though the latter is objectively critical for image quality. The results of the current study indicate that cSART, in terms of objective metrics, is more capable of providing Shannon information about the imaged samples. This suggests that cSART could be a better choice for automated image analysis tools, especially with the rapid growth of AI-assisted diagnostics. However, if subjective preferences for contrast persist in clinical practice, this may influence algorithm selection in certain settings, such as visual assessments by radiologists.

Despite this difference between objective and subjective assessments, we argue that the findings of this study do not diminish the value of the cSART and UTR reconstruction algorithms. Rather, they highlight the complexity of image quality evaluation, where both objective metrics and subjective radiological preferences play significant roles. If automated tools that rely on objective quality metrics gain further acceptance in medical imaging, cSART's superior information content could become a major advantage for applications in breast phase-contrast CT imaging.

Furthermore, beyond these results, the current work presents a robust methodology for medical image quality assessment that extends beyond the specific context of phase-contrast breast CT imaging. We proposed a systematic use of a set of five physics-based objective image quality characteristics: contrast, SNR, CNR, spatial resolution, and Qs, with the latter characteristic particularly effective in quantifying the Shannon information about the imaged sample provided by a computational imaging system. Complementing these objective metrics, we also applied our previously developed methodology for systematic subjective image quality assessment, using five key characteristics—perceptible contrast, sharpness of tissue interfaces, calcification visibility, image noise, and overall image quality—and analyzed them through visual grading characteristics (VGC). This integrated approach, which has been developed in collaboration with practicing

radiologists and medical imaging specialists, provides a convenient way to assess medical image quality and to explore the relationships between objective and subjective image quality evaluations.

By combining both objective and subjective assessments, this study allows for an in-depth analysis of the strengths and limitations of each of the three PCT reconstruction approaches and highlights the potential for these findings to inform the development of automated image analysis tools in the future.

ACKNOWLEDGMENTS

The experimental part of this research was undertaken on the Imaging and Medical beamline at the Australian Synchrotron, part of ANSTO. The following funding is acknowledged: National Health and Medical Research Council, Australia (APP2011204). The authors would like to acknowledge staff from the University of Melbourne, Monash University, Australian Synchrotron, CSIRO and the University of Sydney for their contribution toward the collection of the original image data and the radiological assessment. Sandro Donato is supported by the project Tech4You- Technologies for climate change adaptation and quality of life improvement (C.I. ECS 00000009, CUP H23C22000370006)

Open access publishing facilitated by Universita della Calabria, as part of the Wiley - CRUI-CARE agreement.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data are available from request to the authors.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209-249. doi:10.3322/caac.21660
2. Fox S, Speirs V, Shaaban AM, Male breast cancer: an update. *Virchows Arch*. 2022;480:85-93. doi:10.1007/s00428-021-03190-7
3. Sledge GW, Miller KD, Exploiting the hallmarks of cancer. *Eur J Cancer*. 2003;39:1668-1675. doi:10.1016/S0959-8049(03)00273-9
4. Autier P, Boniol M, Mammography screening: a major issue in medicine. *Eur J Cancer*. 2018;90:34-62. doi:10.1016/j.ejca.2017.11.002
5. Jafari SH, Saadatpour Z, Salmaninejad A, et al., Breast cancer diagnosis: imaging techniques and biochemical markers. *J Cell Physiol*. 2018;233:5200-5213. doi:10.1002/jcp.26379
6. Vaughan CL, Novel imaging approaches to screen for breast cancer: recent advances and future prospects. *Med Eng Phys*. 2019;72:27-37. doi:10.1016/j.medengphy.2019.09.001
7. Feig SA, Hendrick RE, Radiation risk from screening mammography of women aged 40–49 years. *JNCI Monogr*. 1997;1997, 119-124. doi:10.1093/jncimono/1997.22.119
8. Boone JM, Nelson TR, Lindfors KK, Seibert JA, Dedicated breast CT: radiation dose and image quality evaluation. *Radiology*. 2001;221:657-667. doi:10.1148/radiol.2213010334
9. Chong A, Weinstein SP, McDonald ES, Conant EF, Digital breast tomosynthesis: concepts and clinical practice. *Radiology*. 2019;292:1-14. doi:10.1148/radiol.2019180760
10. Sarno A, Mettivier G, Russo P, Dedicated breast computed tomography: basic aspects: breast computed tomography with dedicated scanners. *Med Phys*. 2015;42:2786-2804. doi:10.1118/1.4919441
11. Bronnikov AV, Theory of quantitative phase-contrast computed tomography. *J Opt Soc Am A*. 2002.19:472. doi:10.1364/JOSAA.19.000472
12. Momose A, Takeda T, Itai Y, Hirano K, Phase–contrast X–ray computed tomography for observing biological soft tissues. *Nat Med*. 1996;2:473-475. doi:10.1038/nm0496-473
13. Taba ST, Gureyev TE, Alakhras M, Lewis S, Lockie D, Brennan PC, X-ray phase-contrast technology in breast imaging: principles, options, and clinical application. *Am J Roentgenol*. 2018;211:133-145. doi:10.2214/AJR.17.19179
14. Brombal L, Donato S, Dreossi D, et al., Phase-contrast breast CT: the effect of propagation distance. *Phys Med Biol*. 2018;63:24NT03. doi:10.1088/1361-6560/aaf2e1
15. Gureyev TE, Nesterets YaI, Baran PM, et al., Propagation-based x-ray phase-contrast tomography of mastectomy samples using synchrotron radiation. *Med Phys*. 2019;46:5478-5487. doi:10.1002/mp.13842
16. Keyriläinen J, Fernández M, Bravin A, et al., Comparison of *in vitro* breast cancer visibility in analyser-based computed tomography with histopathology, mammography, computed tomography and magnetic resonance imaging. *J Synchrotron Radiat*. 2011.18:689-696. doi:10.1107/S090904951102810X
17. Davis TJ, Gao D, Gureyev TE, Stevenson AW, Wilkins SW, Phase-contrast imaging of weakly absorbing materials using hard X-rays. *Nature*. 1995;373:595-598. doi:10.1038/373595a0
18. Ingal VN, Beliaevskaya EA, X-ray plane-wave topography observation of the phase contrast from a non-crystalline object. *J Phys Appl Phys*. 1995;28:2314-2317. doi:10.1088/0022-3727/28/11/012
19. Olivo A, Edge-illumination x-ray phase-contrast imaging. *J Phys Condens Matter*. 2021;33:363002. doi:10.1088/1361-648X/ac0e6e
20. Hellerhoff K, Birnbacher L, Sztrókay-Gaul A, et al., Assessment of intraductal carcinoma in situ (DCIS) using grating-based X-ray phase-contrast CT at conventional X-ray sources: an experimental ex-vivo study. *PLOS ONE*. 2019.14:e0210291. doi:10.1371/journal.pone.0210291
21. Baran P, Mayo S, McCormack M, et al., High-resolutionX-ray phase-contrast 3-D imaging of breast tissue specimens as a possible adjunct to histopathology. *IEEE Trans Med Imaging*. 2018;37:2642-2650. doi:10.1109/TMI.2018.2845905
22. Longo R, Arfelli F, Bonazza D, et al., Advancements towards the implementation of clinical phase-contrast breast computed tomography at Elettra. *J Synchrotron Radiat*. 2019.26:1343-1353. doi:10.1107/S1600577519005502
23. Arana Peña LM, Donato S, Bonazza D, et al., Multiscale X-ray phase-contrast tomography: from breast CT to micro-CT for virtual histology. *Phys Med*. 2023.112:102640. doi:10.1016/j.ejmp.2023.102640
24. Castelli E, Tonutti M, Arfelli F, et al., Mammography with synchrotron radiation: first clinical experience with phase-detection technique. *Radiology*. 2011.259:684-694. doi:10.1148/radiol.11100745
25. Donato S, Arana Peña LM, Arfelli F, et al., Integrating X-ray phase-contrast imaging and histology for comparative evaluation of breast tissue malignancies in virtual histology analysis. *Sci Rep*. 2024.14:5831. doi:10.1038/s41598-024-56341-6
26. Donato S, Bonazza D, Advancing breast cancer diagnosis in 3D: the transformative power of X-ray phase-contrast microtomography for virtual histology. *TrAC Trends Anal Chem*. 2024.180:117943. doi:10.1016/j.trac.2024.117943

27. Pacilè S, Dullin C, Baran P, et al., Free propagation phase-contrast breast CT provides higher image quality than cone-beam breast-CT at low radiation doses: a feasibility study on human mastectomies. *Sci Rep.* 2019.9:13762. doi:10.1038/s41598-019-50075-6
28. Zhao Y, Brun E, Coan P, et al., High-resolution, low-dose phase contrast X-ray tomography for 3D diagnosis of human breast cancers. *Proc Natl Acad Sci.* 2012.109:18290-18294. doi:10.1073/pnas.1204460109
29. Baran P, Pacile S, Nesterets YI, et al., Optimization of propagation-based x-ray phase-contrast tomography for breast cancer imaging. *Phys Med Biol.* 2017.62:2315-2332. doi:10.1088/1361-6560/aa5d3d
30. Oliva P, Di Trapani V, Arfelli F, et al., Experimental optimization of the energy for breast-CT with synchrotron radiation. *Sci Rep.* 2020.10:17430. doi:10.1038/s41598-020-74607-7
31. Tavakoli Taba S, Baran P, Lewis S, et al., Toward improving breast cancer imaging: radiological assessment of propagation-based phase-contrast CT technology. *Acad Radiol.* 2019.26:e79-e89. doi:10.1016/j.acra.2018.07.008
32. McCollough CH, Primak AN, Braun N, Kofler J, Yu L, Christner J, Strategies for reducing radiation dose in CT. *Radiol Clin North Am.* 2009.47:27-40. doi:10.1016/j.rcl.2008.10.006
33. Gunn MLD, Kohr JR, State of the art: technologies for computed tomography dose reduction. *Emerg Radiol.* 2010.17:209-218. doi:10.1007/s10140-009-0850-6
34. Lu B, Lu H, Palta J, A comprehensive study on decreasing the kilovoltage cone-beam CT dose by reducing the projection number. *J Appl Clin Med Phys.* 2010.11:231-249. doi:10.1120/jacmp.v11i3.3274
35. Sidky EY, Chartrand R, Boone JM, Pan Xiaochuan, Constrained ${\rm T}p{\rm V}$ minimization for enhanced exploitation of gradient sparsity: application to CT image reconstruction. *IEEE J Transl Eng Health Med*. 2014;2:1-18. doi:10.1109/JTEHM.2014.2300862
36. Bian J, Yang K, Boone JM, Han X, Sidky EY, Pan X, Investigation of iterative image reconstruction in low-dose breast CT. *Phys Med Biol.* 2014.59:2659-2685. doi:10.1088/0031-9155/59/11/2659
37. Greffier J, Macri F, Larbi A, et al., Dose reduction with iterative reconstruction: optimization of CT protocols in clinical practice. *Diagn Interv Imaging.* 2015.96:477-486. doi:10.1016/j.diii.2015.02.007
38. Makeev A, Glick SJ, Investigation of statistical iterative reconstruction for dedicated breast CT: statistical iterative reconstruction for dedicated breast CT. *Med Phys.* 2013.40:081904. doi:10.1118/1.4811328
39. Pacilè S, Brun F, Dullin C, et al., Clinical application of low-dose phase contrast breast CT: methods for the optimization of the reconstruction workflow. *Biomed Opt Express.* 2015.6:3099. doi:10.1364/BOE.6.003099
40. Gureyev TE, Brown HG, Quiney HM, Allen LJ, Unified fast reconstruction algorithm for conventional, phase-contrast, and diffraction tomography. *J Opt Soc Am A.* 2022.39:C143. doi:10.1364/JOSAA.468350
41. Donato S, Brombal L, Arana Peña LM, et al., Optimization of a customized simultaneous algebraic reconstruction technique algorithm for phase-contrast breast computed tomography. *Phys Med Biol.* 2022.67:095012. doi:10.1088/1361-6560/ac65d4
42. Natterer F, The mathematics of computerized tomography. *Soc Ind Appl Math*. 2001;doi:10.1137/1.9780898719284
43. Johns PC, Yaffe MJ, Theoretical optimization of dual-energy x-ray imaging with application to mammography. *Med Phys.* 1985.12:289-296. doi:10.1118/1.595766
44. Nesterets YI, Gureyev TE, Mayo SC, et al., A feasibility study of X-ray phase-contrast mammographic tomography at the imaging and medical beamline of the Australian synchrotron. *J Synchrotron Radiat*. 2015;22:1509-1523. doi:10.1107/S160057751501766X
45. Gureyev TE, Nesterets YI, Stevenson AW, Miller PR, Pogany A, Wilkins SW, Some simple rules for contrast, signal-to-noise and resolution in in-line x-ray phase-contrast imaging. *Opt Express.* 2008.16:3223. doi:10.1364/OE.16.003223
46. Paganin D, Mayo SC, Gureyev TE, Miller PR, Wilkins SW, Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object. *J Microsci.* 2002.206:33-40. doi:10.1046/j.1365-2818.2002.01010.x
47. Gureyev TE, Nesterets YI, Kozlov A, Paganin DM, Quiney HM, On the "unreasonable" effectiveness of transport of intensity imaging and optical deconvolution. *J Opt Soc Am A.* 2017.34:2251. doi:10.1364/JOSAA.34.002251
48. Gureyev TE, n.d. Unified Tomographic Reconstruction algorithm, https://github.com/timg021/UTR_public
49. Gureyev T, Mohammadi S, Nesterets Y, Dullin C, Tromba G, Accuracy and precision of reconstruction of complex refractive index in near-field single-distance propagation-based phase-contrast tomography. *J Appl Phys.* 2013.114:144906. doi:10.1063/1.4824491
50. Gureyev T, Nesterets Y, De Hoog F, Spatial resolution, signal-to-noise and information capacity of linear imaging systems. *Opt Express.* 2016.24:17168. doi:10.1364/OE.24.017168
51. Gureyev TE, Mayo SC, Nesterets YI, et al., Investigation of the imaging quality of synchrotron-based phase-contrast mammographic tomography. *J Phys Appl Phys.* 2014.47:365401. doi:10.1088/0022-3727/47/36/365401
52. Goodman JW, 2015;*Statistical optics*, Second ed. ed, Wiley series in pure and applied optics. John Wiley & Sons Inc.
53. Cunningham IA, Shaw R, Signal-to-noise optimization of medical imaging systems. *J Opt Soc Am A.* 1999.16:621. doi:10.1364/JOSAA.16.000621
54. Gureyev TE, Paganin DM, Quiney HM, Signal-to-noise and spatial resolution in in-line imaging. 1. Basic theory, numerical simulations and planar experimental images. *J Synchrotron Radiat.* 2024.31:896-909. doi:10.1107/S1600577524003886
55. Gureyev TE, Nesterets Y, Ternovski D, et al., Toolbox for advanced x-ray image processing, in: Sanchez Del Rio M, Chubar O (Eds.), *Presented at the SPIE Optical Engineering + Applications*, 2011, p. 81410B. doi:10.1117/12.893252

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: Donato S, Caputo S, Brombal L, et al. Comparison of three reconstruction algorithms for low-dose phase-contrast computed tomography of the breast with synchrotron radiation. *Med Phys*. 2025;52:e17950. https://doi.org/10.1002/mp.17950