IPEM
Institute of Physics and
Engineering in Medicine

PAPER • OPEN ACCESS

Task-based detectability in anatomical background in digital mammography, digital breast tomosynthesis and synthetic mammography

To cite this article: P Monnin *et al* 2024 *Phys. Med. Biol.* **69** 025017

View the article online for updates and enhancements.

You may also like

Physics in Medicine & Biology

PAPER

Task-based detectability in anatomical background in digital mammography, digital breast tomosynthesis and synthetic mammography

P Monnin[1,*], J Damet[1], H Bosmans[2,3] and N W Marshall[2,3]

[1] Institute of radiation physics (IRA), Lausanne University Hospital (CHUV) and University of Lausanne, Rue du Grand-Pré 1, 1007 Lausanne, Switzerland
[2] UZ Gasthuisberg, Department of Radiology, Herestraat 49, 3000 Leuven, Belgium
[3] Medical Imaging Research Center, Medical Physics and Quality Assessment, Katholieke Universiteit Leuven, 3000 Leuven, Belgium
[*] Author to whom any correspondence should be addressed.

E-mail: pascal.monnin@chuv.ch

Keywords: detectability, mammography, digital breast tomosynthesis, model observer, anatomical noise

Abstract

Objective. Determining the detectability of targets for the different imaging modalities in mammography in the presence of anatomical background noise is challenging. This work proposes a method to compare the image quality and detectability of targets in digital mammography (DM), digital breast tomosynthesis (DBT) and synthetic mammography. *Approach.* The low-frequency structured noise produced by a water phantom with acrylic spheres was used to simulate anatomical background noise for the different types of images. A method was developed to apply the non-prewhitening observer model with eye filter (NPWE) in these conditions. A homogeneous poly (methyl) methacrylate phantom with a 0.2 mm thick aluminium disc was used to calculate 2D in-plane modulation transfer function (MTF), noise power spectrum (NPS), noise equivalent quanta, and system detective quantum efficiency for 30, 50 and 70 mm thicknesses. The in-depth MTFs of DBT volumes were determined using a thin tungsten wire. The MTF, system NPS and anatomical NPS were used in the NPWE model to calculate the threshold gold thickness of the gold discs contained in the CDMAM phantom, which was taken as reference. *Main results.* The correspondence between the NPWE model and the CDMAM phantom (linear Pearson correlation 0.980) yielded a threshold detectability index that was used to determine the threshold diameter of spherical microcalcifications and masses. DBT imaging improved the detection of masses, which depended mostly on the reduction of anatomical background noise. Conversely, DM images yielded the best detection of microcalcifications. *Significance.* The method presented in this study was able to quantify image quality and object detectability for the different imaging modalities and levels of anatomical background noise.

1. Introduction

The development of digital breast tomosynthesis (DBT) has paved the way to diverse implementations of pseudo-volumetric imaging in mammography (Sechopoulos 2013a, 2013b). This has also led to the development of synthetic mammography (SM) imaging, in which a 2D image is generated from DBT projections and/or reconstructed planes, with the aim of substituting digital mammography (DM) images (Durand 2018). The development of breast phantoms that include anatomical background structures to assess the imaging performance of DM, DBT and SM devices aims to provide a direct and relevant evaluation of the detectability of structures such as microcalcifications and mass-like lesions (Hadjipanteli *et al* 2017 and 2019, Tanguay *et al* 2019, Marshall and Bosmans 2022). Breast phantoms which mimic anatomical breast noise and some low-frequency and high-frequency structures for task-based performance evaluation have been recently developed. The phantoms used in these studies attempt to match breast anatomical noise in terms of magnitude

and correlation of structures (Kiarashi *et al* 2015, Cockmartin *et al* 2017, Ikejimba *et al* 2017, Glick and Ikejimba 2018).

The use of Fourier-based metrics such as presampling modulation transfer function (MTF), noise power spectrum (NPS), noise equivalent quanta (NEQ) and detective quantum efficiency (DQE) (ICRU Report 54 1996, Cunningham 2000) to quantify detector or system imaging properties is well-established (Zhao *et al* 2009 and 2017, Marshall *et al* 2011, Marshall and Bosmans 2012). The metric of NEQ and model observers in the spatial frequency domain have been generalized to include the influence of anatomical noise and scattered radiation for CBCT and tomosynthesis (Tward and Siewerdsen 2008, Gang *et al* 2010, Reiser and Nishikawa 2010, Prakash *et al* 2011). Detectability of low-frequency and high-frequency objects calculated from Fourier-based metrics have been shown to correlate well with observer performance for a wide range of imaging parameters in DBT (Gang *et al* 2011). In previous work, we have used cascaded linear system theory to include the scatter fraction and the anti-scatter device in a global system DQE and describe in this way the SNR transfer through the complete imaging system (Monnin *et al* 2017). Recent work has applied this method to the evaluation of global system performance of several DBT systems with different image reconstruction algorithms (Monnin *et al* 2020), yielding a methodology to assess Fourier metrics and detection performance of DBT systems using a non-prewhitening with eye filter (NPWE) model observer.

This work extends the method described in Monnin *et al* (2020) to calculate a detectability index which includes a structured anatomical noise term. Fourier-based image quality metrics were calculated for the DM, DBT and SM imaging modes of three mammography systems at three breast equivalent thicknesses. The threshold thicknesses of the CDMAM phantom discs were compared to those given by the NPWE model. The good agreement obtained between the observer model and the contrast-detail analysis established a threshold detectability index that pertained for all the imaging modalities and conditions considered in this study. This threshold index was then used to estimate the threshold diameter of spherical microcalcifications and masses for different imaging levels of system noise (detector air kerma) and anatomical background noise.

2. Materials and methods

2.1. Mammography systems and imaging setup

Three flat panel-based mammography systems were involved in this study: a GEHC Senographe Pristina (GE HealthCare, Chicago, USA), a Hologic Selenia Dimensions (Hologic, Massachusetts, USA) and a Siemens Revelation (Siemens Healthineers, Germany). See table 1 for their technical characteristics. For each system, a phantom made of 180 mm × 240 mm plates of poly(methyl) methacrylate (PMMA) was imaged in three thicknesses: 30 mm, 50 mm and 70 mm. The tube voltage, anode/filter (A/F), tube current–time product (mAs) and grid configuration of these three PMMA thicknesses imaged under automatic exposure control (AEC) are shown in table 2. Spacers were used when establishing these technique factors so that the total thickness (PMMA + spacers) was equal to the breast equivalent thickness for a given PMMA thickness (Dance *et al* 2000). For DM mode, 'For Processing' and 'For Presentation' images were acquired, referred to respectively as 'raw' and 'processed' DM images. DBT scans were performed and DBT planes and SM images were generated using the standard clinical image reconstruction and processing (table 1). A 1 mm reconstructed DBT plane spacing was used for all systems.

As described in Monnin *et al* (2020), a phantom made of PMMA plates (the 'NPWE phantom') was imaged in three thickness configurations (30, 50 and 70 mm) to measure the 2D MTF and the 2D NPS required in the NPWE observer model to calculate the detectability index. The NPWE phantom contained a 0.2 mm thick aluminium disc of 50 mm diameter placed on top of 20 mm of PMMA at the reference point (van Engen 2013), 6 cm from chest wall side and laterally centred. Four DM and DBT acquisitions were made for each of the three NPWE phantom thicknesses, with the phantom slightly moved between each acquisition to produce four different samples of the disc. An additional acquisition of the 50 mm NPWE phantom placed side by side with a structured phantom (the 'L1 phantom') was made. The L1 phantom is a 48 mm thick semi-cylindrical PMMA box filled with water and acrylic beads of different diameters that generates images containing structured noise similar to DM and DBT images of breasts (Cockmartin *et al* 2017, Vancoillie *et al* 2021).

Images of the CDMAM phantom were acquired at the same thicknesses and conditions used for the NPWE phantom for the DM and DBT imaging modes. Attenuation of the CDMAM phantom is approximately equivalent to that of 10 mm PMMA (CDMAM manual, Artinis), and therefore a 10 mm PMMA plate was replaced by CDMAM. The CDMAM phantom was always positioned on top of 20 mm of PMMA, ensuring a fixed height above the detector for all acquisitions. The phantom was slightly moved between each of eight similar acquisitions.

3

Table 1. Characteristics of the mammography systems.

System	Detector type	Pixel pitch (μm) DM/DBT/SM	DBT			
			X-ray tube motion	Number of projections	Angular range	Reconstruction method
GEHC Senographe Pristina	CsI/a-Si TFT switch	100/100/-	step-and-shoot	9	$\pm12.5°$	iterative
Hologic Selenia Dimensions	a-Se/TFT switch	70/100/106	continuous	15	$\pm7.5°$	FBP with iterative optimization
Siemens Revelation	a-Se/TFT switch	85/85/85	continuous	25	$\pm25°$	EMPIRE (iterative)

4

Table 2. Acquisition parameters and x-ray beams characteristics.

Modality	AEC mode	Grid	Phantom thickness (mm)	kV	mAs	Anode	Filter (mm)	K_{in} (μGy)	DAK (μGy)	φ (mm^{-2} μGy^{-1})	$\frac{\Delta q_{IRF}}{\bar{q}}$	SF_{in}
GEHC Senographe Pristina												
DM/DBT	STD	In	30	26	40	Mo	Mo (0.03)	206.1	96.9	4844	0.217	0.360
			50	34	36	Rh	Ag (0.03)	247.9	104.1	7353	0.137	0.504
			70	34	90	Rh	Ag (0.03)	245.7	88.5	7979	0.126	0.589
Hologic Selenia Dimensions												
DM	Auto filter	In	30	26	92	W	Rh (0.05)	285.9	157.2	5776	0.177	0.334
			50	31	176	W	Rh (0.05)	334.5	160.6	6783	0.147	0.462
			70	34	198	W	Ag (0.05)	338.6	145.6	8680	0.113	0.555
DBT	Auto filter	Out	30	28	43	W	Al (0.7)	400.1	400.1	6630	0.157	0.291
			50	33	60	W	Al (0.7)	492.5	492.5	9149	0.109	0.460
			70	42	70	W	Al (0.7)	822.2	822.2	15013	0.064	0.568
Siemens Revelation												
DM	Opdose	In	30	27	50	W	Rh (0.05)	189.6	104.3	5866	0.174	0.355
			50	30	100	W	Rh (0.05)	180.3	90.2	6593	0.150	0.483
			70	32	200	W	Rh (0.05)	162.5	73.1	7593	0.132	0.568
DBT	Opdose	Out	30	27	100	W	Rh (0.05)	351.6	351.6	5866	0.174	0.355
			50	30	180	W	Rh (0.05)	300.5	300.5	6593	0.150	0.483
			70	32	320	W	Rh (0.05)	240.7	240.7	7593	0.132	0.568

2.2. X-ray beams, radiant contrast and grid characteristics

We consider imaging systems composed of an anti-scatter grid (if used), a detector, image reconstruction (DBT and SM) and image processing stages (processed DM, DBT and SM). The values of all quantities are expressed in the detector (image) plane considered as the reference plane. The air kerma, the mean photon fluence (photons/mm^2) and the scatter fraction (SF) are denoted as K, $\bar{q}$ and SF, respectively. When necessary, the subscripts *in* or *out* specify whether these quantities correspond to values at the grid *input* or *output*, respectively. For the modalities which do not use an anti-scatter grid, the quantities with the subscripts *in* and *out* are therefore equal.

For air kerma measurements, the relevant PMMA phantom was held in the compression plate at a height of 2 cm above the breast support table (see figure 2 in Monnin *et al* 2020). A calibrated ionization chamber was positioned at the reference point (van Engen 2013) above the system (grid) entrance plane, 1 cm below the PMMA/compression plate. The measured air kerma includes primary and scattered radiation at the system entrance (grid entrance if used for imaging). An inverse square distance correction was applied to express the input air kerma at the detector (image) plane, noted K_{in} in table 2. The detector air kerma (DAK) is the product between K_{in} and the total grid transmission. DAK is equal to K_{in} for the case without grid. The product of K_{in} with the photon fluence per air kerma unit (Boone 1998), noted φ in table 2, gave the input photon fluence $\bar{q}_{in}$. The photon fluence per air kerma unit φ depends only on the x-ray spectrum and is the same for the configurations grid in or grid out.

The NPWE phantom contains a thin 0.2 mm Al disc used to calculate the impulse response function (IRF) of the imaging system. This disc produces a radiant contrast $\Delta \bar{q}_{IRF}$ (zero-frequency signal), which is the difference in photon fluence measured with and without an additional 0.2 mm Al plate fixed at the tube output. We assume the relative disc contrast $\frac{\Delta \bar{q}_{IRF}}{\bar{q}}$ depends on the disc thickness (T_{Al}) and on the linear attenuation coefficient of aluminium (μ_{Al}) as given in equation (1), neglecting the effect of the thin 0.2 mm Al disc on the SF

$$\frac{\Delta \bar{q}_{IRF}}{\bar{q}} = 1 - \exp\left(-\mu_{Al} T_{Al}\right). \tag{1}$$

For a given disc, $\frac{\Delta \bar{q}_{IRF}}{\bar{q}}$ varies with the x-ray beam energy, determined by the tube voltage, the anode/filter (A/F) combination and the phantom thickness, but neither with the SF produced in the phantom nor with the total grid transmission.

SFs were measured for 30, 50 and 70 mm PMMA using the beam stop method described in Monnin *et al* (2017). Lead discs with radii between 1.5 mm and 6 mm were positioned on the PMMA thickness at the reference point and imaged under the same conditions as the NPWE and CDMAM phantoms (table 2). The linear response function between 'For Processing' pixel values and DAK was established from the raw projection images for all the DM and DBT beams (van Engen 2013). Pixel values measured within the disc in the 'For Processing' projection images were converted to DAK values using the linear response functions. The SFs measured with grid in and out gave SF_{out} and SF_{in}, respectively. The total grid transmission (T_g) was calculated from the ratio of DAK measured using a 5×5 mm^2 ROI at the reference point in the linearized projection images, acquired with grid in and grid out. The primary grid transmission T_p, scatter grid transmission T_s and grid DQE (DQE_{grid}) were determined from T_g, SF_{in} and SF_{out} using equations (2), (3) and (4), respectively (Monnin *et al* 2017)

$$T_p = \frac{1 - SF_{out}}{1 - SF_{in}} \bullet T_g \tag{2}$$

$$T_s = \frac{SF_{out}}{SF_{in}} \bullet T_g \tag{3}$$

$$DQE_{grid} = \frac{T_p^2}{T_g} = \left(\frac{1 - SF_{out}}{1 - SF_{in}}\right)^2 \bullet T_g. \tag{4}$$

The SF increases with the x-ray path length in the phantom, and thus with the projection angle for DBT. The x-ray path length increases with the inverse of the cosine of the projection angle, leading to a 10% increase for the maximum projection angle of 25° (Siemens Revelation). The difference in SF between the central and the outer projections in DBT therefore remains below a few percent and was not considered in this study.

2.3. In-plane Fourier metrics: IRF, MTF, NPS, NEQ and system DQE

The use of Fourier-based metrics to characterize properties of medical imaging devices requires the approximation of a stationary and spatially invariant imaging system (Metz and Doi 1979, Cunningham 2000). We consider that pixel values d are spatially invariant and depend linearly on the photon fluence per pixel q, with a gain $\partial d / \partial q$. This local linear relationship was considered for all image types: raw and processed DM images, reconstructed DBT planes and SM images. Although image reconstruction and image processing algorithms use

logarithmic transforms that are nonlinear processes, processed DM images, reconstructed DBT planes and SM images were considered linear in a small-signal approximation for a small range of signal variations (contrast) in the images (Tward and Siewerdsen 2008, Zhao and Zhao 2008). The 0.2 mm aluminium disc in the NPWE phantom was used to produce a signal (radiant contrast) sufficiently small (between 6% and 22%, table 2) so that the system gain could be considered constant, and the log-normalization performed in the DBT (and by implication in the SM reconstructions) could be considered linear in a small-signal approximation. The impulse response function (IRF), NPS, NEQ and system detective quantum efficiency (system DQE) were therefore calculated in the Fourier space to characterize the image quality parameters of DM, DBT and SM images, following the calculation methodology detailed in previous work (Monnin *et al* 2020), and summarized below.

In this study, the following coordinate system was used: x for the left-right direction, y for the front-back direction and z for the vertical direction. DBT produces a 3D IRF, which was expressed as the product of two separable functions governed by different physical parameters in the in-plane xy- and z-components according to equation (5)

$$IRF_d(f_x, f_y, f_z) = IRF_{d,xy}(f_x, f_y) \cdot MTF_z(f_z). \tag{5}$$

IRF_d denotes the system IRF measured with scatter, whose magnitude is expressed in pixel values d. $IRF_{d,xy}$ is the IRF component in the xy plane and MTF_z the in-depth MTF of DBT reconstructions characterized using the thin wire. The 2D 'in-plane IRF' was measured in the in-focus DBT plane of the thin 0.2 mm Al disc of the NPWE phantom. It is denoted $IRF_{d,ip}$ and is equal to the 3D IRF_d integrated over the z-frequency bandwidth (Zhao and Zhao 2008).

$$IRF_{d,ip}(f_x, f_y) = \int_{-f_{z,Nyq}}^{f_{z,Nyq}} IRF_d(f_x, f_y, f_z) df_z = IRF_{d,xy}(f_x, f_y) \int_{-f_{z,Nyq}}^{f_{z,Nyq}} MTF_z(f_z) df_z. \tag{6}$$

The case for 2D (DM and SM) images is obtained by setting $f_z = 0$ in equations (5) and (6), where the in-plane IRF reduces to the xy-component of the IRF. The in-plane IRF was measured in the image of the 0.2 mm Al disc of the NPWE phantom with a radial version of the angled edge method (Samei, Flynn and Reimann 1998). Extension of the method to radial coordinates is detailed in Monnin *et al* (2016). Radial edge spread functions (ESF) originating from the disc centre were plotted every 2°, in a square ROI of size 100×100 mm^2, each of them covering a 4° angular aperture. Each radial ESF produced a radial IRF. The in-plane 2D IRF was calculated from the 180 radial IRFs using the angular interpolation method described in Monnin *et al* (2016). The in-plane system MTF was defined as the in-plane system IRF normalized to 1.0 at its maximum value. Unlike the IRF, the MTF does not vary with $\bar{q}$, system gain or disc contrast, and was used to compare the signal frequency transfer between the different systems and acquisition modes.

The 3D NPS of DBT stacks were measured from two homogeneous volumes of interest (VOIs) of 60×60 mm^2 placed in the NPWE phantom, on both sides of the aluminium disc, at 60 mm from the disc centre and 60 mm from the chest wall side (figure 1 in Monnin *et al* 2020). The first and last images of the stack were excluded from the VOIs. Computation was made without detrending correction, leading to low-frequency peaks in the NPS due to signal trends present in the VOIs. The in-plane 2D NPS, denoted $NPS_{d,ip}$, was obtained by integrating the 3D NPS over the z-frequency bandwidth (Siewerdsen *et al* 2002, Zhao and Zhao 2008).

$$NPS_{d,ip}(f_x, f_y) = \int_{-f_{z,Nyq}}^{f_{z,Nyq}} NPS_d(f_x, f_y, f_z) df_z. \tag{7}$$

The in-plane 2D NPS for DM and SM images were calculated from 2D regions of interest (ROIs) instead of VOIs.

The NEQ and the system DQE (DQE_{sys}) have been shown to quantify image quality and the detective efficiency of the imaging system including the scattered radiation, the anti-scatter grid, the reconstruction stages and the image processing (Monnin *et al* 2017, 2020). These metrics were calculated from the in-plane IRF and NPS according to equations (8) and (9), respectively

$$NEQ(f_x, f_y) = \frac{IRF_{d,ip}^2(f_x, f_y)}{\left(\frac{\Delta \bar{q}_{IRF}}{\bar{q}}\right)^2 \cdot NPS_{d,ip}(f_x, f_y)} \tag{8}$$

$$DQE_{sys}(f_x, f_y) = \frac{IRF_{d,ip}^2(f_x, f_y)}{\left(\frac{\Delta \bar{q}_{IRF}}{\bar{q}}\right)^2 \cdot (1 - SF_{in})^2 \cdot \bar{q}_{in} \cdot NPS_{d,ip}(f_x, f_y)}. \tag{9}$$

The mean radial MTF, NEQ and DQE_{sys} curves are radial averages of the corresponding 2D metrics, excluding the 0° and 90° axial values.

2.4. In-depth resolution (MTF$_z$)

In-depth resolution of DBT reconstructions was characterized using MTF$_z$ and measured using the thin wire method (Li *et al* 2007, Hu *et al* 2008, Marshall and Bosmans 2012). A thin tungsten wire of diameter 20 μm was stretched vertically in air from the compression plate to the breast table, tilted by 20° in the front-back y-direction and by 2–3° in the left-right x-direction. The wire was imaged with the same kV and A/F used for the phantoms (table 2), but with a low mAs value to avoid saturation of pixel values. For each image in the stack (z-position), the voxel with the maximum intensity gave the (x, y) coordinates of the wire position in the DBT planes. A linear least squares regression of x-coordinates to y-coordinates of the wire gave the azimuthal angle of the wire relative to the image matrix. A further linear curve fit of the z-coordinates to azimuthal positions of the wire gave the polar angle (inclination) of the wire in a spherical coordinate system. The intensity of all the voxels of the reconstructed volume having the same azimuthal coordinates as the wire were plotted as a function of their vertical distance (z-position) to the wire to obtain the oversampled PSF$_z$. The amplitude of the Fourier transform of the PSF$_z$, was normalized to 1.0 at the zero frequency to give MTF$_z$.

2.5. Anatomical breast noise

In order to obtain the anatomical background NPS, denoted NPS$_{d,a}$, the L1 and the 50 mm NPWE phantoms were rotated by 90°, positioned adjacent to each other at the chest wall edge and imaged together. These two phantoms yielded equivalent attenuation and hence the same mean pixel value and system noise level (quantum, fixed pattern and electronic noises). Acquisition factors relevant to each imaging system for 50 mm phantom thickness were set and three images acquired for each system. The in-plane NPS of system noise was measured in the NPWE phantom, and was subtracted from the in-plane NPS measured in the L1 phantom to give the anatomical in-plane NPS (equation (10))

$$\mathrm{NPS}_{d,a}(f_x, f_y) = \mathrm{NPS}_{d,ip,L1}(f_x, f_y) - \mathrm{NPS}_{d,ip,\mathrm{NPWE}}(f_x, f_y). \tag{10}$$

The measured anatomical NPS was then fitted to the empirical power law relationship used to describe the power spectrum of anatomical structures (Gang *et al* 2010, Reiser and Nishikawa 2010, Cockmartin *et al* 2013) given in equation (11), with the coefficients κ and β quantifying the magnitude and correlation of the structured noise

$$\mathrm{NPS}_{d,a}(f_x, f_y) = \kappa \cdot (f_x^2 + f_y^2)^{-\beta/2}. \tag{11}$$

We are therefore assuming that the structures in the L1 phantom model the structures in a typical breast to some acceptable degree (Cockmartin *et al* 2013). Images of anatomical background noise were generated and added to CDMAM images to quantify its deleterious effect on object detectability. For this purpose, a homogeneous image with a white Gaussian noise of random phase was generated, Fourier-transformed and filtered by $(f_x^2 + f_y^2)^{-\beta/4}$ (Bochud *et al* 1995, Båth *et al* 2005a and 2005b, Reiser and Nishikawa 2010). The NPS magnitude of the generated noise was then adjusted to correspond to that measured on the images of the L1 phantom. An inverse Fourier transform yielded a simulating structure that was mathematically equivalent to the image of anatomical noise. An anatomical NPS with a different random noise texture was added to each of the 8 CDMAM images of equivalent PMMA thickness 50 mm, as shown in figure 1. For the two other thicknesses of 30 and 70 mm, the magnitude of the anatomical NPS measured on the 50 mm thick phantom was converted to another thickness T using equation (12)

$$\mathrm{NPS}_{d,a,T}(f_x, f_y) = \left(\frac{\bar{d}_T}{\bar{d}_{50}} \bullet \frac{\Delta\mu_T}{\Delta\mu_{50}} \bullet \frac{\Delta z_T}{\Delta z_{50}} \right)^2 \bullet \mathrm{NPS}_{d,a,50}(f_x, f_y). \tag{12}$$

The magnitude coefficient κ was considered proportional to the square of the mean pixel value $\bar{d}$, of the effective slab thickness Δz, and of the difference between the linear attenuation coefficients of the breast glandular and adipose tissues $\Delta\mu = \mu_{\mathrm{glandular}} - \mu_{\mathrm{adipose}}$, as reported in previous studies (Mainprize and Yaffe 2010, Gang *et al* 2012, Hill *et al* 2013)
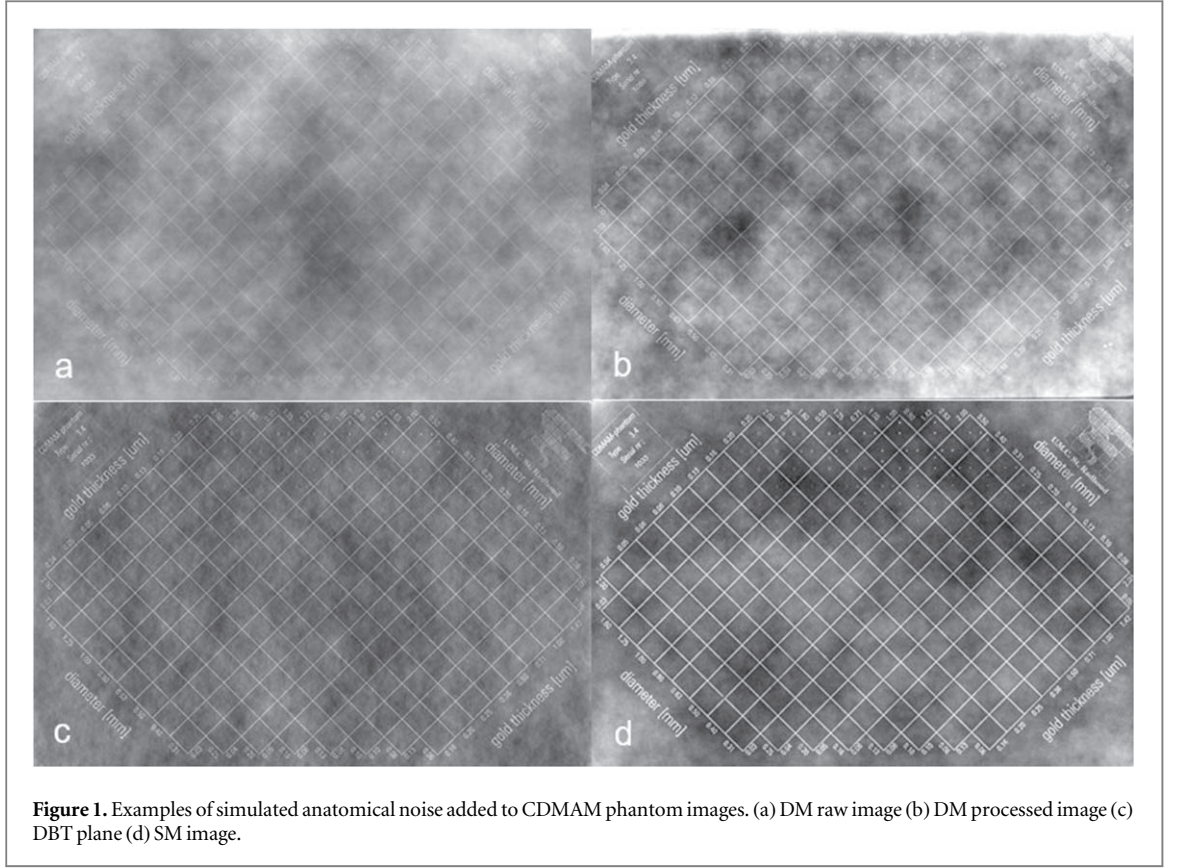
$$\kappa \propto (\Delta\mu \bullet \Delta z \bullet \bar{d})^2. \tag{13}$$

The effective slab thickness Δz is the integral of the in-depth PSF$_z$ over the breast thickness T and is equal to T for DM and SM images for which PSF$z = 1.0$

$$\Delta z = \int_{-T/2}^{T/2} \mathrm{PSF}_z(z) \, dz. \tag{14}$$

2.6. NPWE observer model

We consider the detectability of a spherical object of radius R embedded in a homogeneous thickness, with a 3D shape spectrum $S_{\mathrm{obj}}(R, f_x, f_y, f_z)$. The NPWE detectability index d'_{3D} for a 3D object uses the 3D IRF$_d$ and

Figure 1. Examples of simulated anatomical noise added to CDMAM phantom images. (a) DM raw image (b) DM processed image (c) DBT plane (d) SM image.

NPS_d (equation (10b) in Monnin *et al* 2020)

$$d'_{3D} = \frac{\Delta q_{obj}/\bar{q}}{\Delta \bar{q}_{IRF}/\bar{q}} \bullet k(R). \tag{15}$$

Where $k(R)$ represents the NPWE detectability index value for an object that produces a radiant contrast equal to that of the 0.2 mm Al disc of the NPWE phantom:

$$k(R) = \frac{\int_{-f_{x,Nyq}}^{f_{x,Nyq}} \int_{-f_{y,Nyq}}^{f_{y,Nyq}} VTF^2(f_x,f_y) \cdot \left(\int_{-f_{z,Nyq}}^{f_{z,Nyq}} IRF_d(f_x,f_y,f_z) \bullet S_{obj}(R,f_x,f_y,f_z) df_z \right)^2 df_x df_y}{\sqrt{\int_{-f_{x,Nyq}}^{f_{x,Nyq}} \int_{-f_{y,Nyq}}^{f_{y,Nyq}} VTF^4(f_x,f_y) \cdot \left(\int_{-f_{z,Nyq}}^{f_{z,Nyq}} IRF_d(f_x,f_y,f_z) \bullet S_{obj}(R,f_x,f_y,f_z) df_z \right)^2 \cdot NPS_{d,ip}(f_x,f_y) df_x df_y}}. \tag{16}$$

Assuming separability of the *xy*- and *z*-components of the 3D IRF, equation (16) transforms to equation (17)

$$k(R) = \frac{\int_{-f_{x,Nyq}}^{f_{x,Nyq}} \int_{-f_{y,Nyq}}^{f_{y,Nyq}} VTF^2(f_x,f_y) \bullet IRF_{d,xy}^2(f_x,f_y) \cdot \left(\int_{-f_{z,Nyq}}^{f_{z,Nyq}} MTF_z(f_z) \bullet S_{obj}(R,f_x,f_y,f_z) df_z \right)^2 df_x df_y}{\sqrt{\int_{-f_{x,Nyq}}^{f_{x,Nyq}} \int_{-f_{y,Nyq}}^{f_{y,Nyq}} VTF^4(f_x,f_y) \bullet IRF_{d,xy}^2(f_x,f_y) \cdot \left(\int_{-f_{z,Nyq}}^{f_{z,Nyq}} MTF_z(f_z) \bullet S_{obj}(R,f_x,f_y,f_z) df_z \right)^2 \cdot NPS_{d,ip}(f_x,f_y) df_x df_y}}$$
$$\tag{17}$$

Using equation (6), equation (18) gives a practical form of equation (17) that involves the in-plane IRF, in-plane NPS, and MTF_z measured in the images

$$k(R) = \frac{\left(\int_{-f_{z,Nyq}}^{f_{z,Nyq}} MTF_z(f_z) df_z \right)^{-1} \bullet \int_{-f_{x,Nyq}}^{f_{x,Nyq}} \int_{-f_{y,Nyq}}^{f_{y,Nyq}} VTF^2(f_x,f_y) \bullet IRF_{d,ip}^2(f_x,f_y) \cdot \left(\int_{-f_{z,Nyq}}^{f_{z,Nyq}} MTF_z(f_z) \bullet S_{obj}(R,f_x,f_y,f_z) df_z \right)^2 df_x df_y}{\sqrt{\int_{-f_{x,Nyq}}^{f_{x,Nyq}} \int_{-f_{y,Nyq}}^{f_{y,Nyq}} VTF^4(f_x,f_y) \bullet IRF_{d,ip}^2(f_x,f_y) \bullet \left(\int_{-f_{z,Nyq}}^{f_{z,Nyq}} MTF_z(f_z) \bullet S_{obj}(R,f_x,f_y,f_z) df_z \right)^2 \bullet NPS_{d,ip}(f_x,f_y) df_x df_y}}. \tag{18}$$

For DM and SM images, the slice projection theorem gives the projection obtained by setting $f_z = 0$ in equation (18). The 3D NPWE model for projection images reverts to the 2D form given in equation (19)

$$k(R) = \frac{\int_{-f_{x,\mathrm{Nyq}}}^{f_{x,\mathrm{Nyq}}} \int_{-f_{y,\mathrm{Nyq}}}^{f_{y,\mathrm{Nyq}}} \mathrm{VTF}^2(f_x, f_y) \bullet S_{\mathrm{obj}}^2(R, f_x, f_y) \bullet \mathrm{IRF}_{d,ip}^2(f_x, f_y) \, df_x \, df_y}{\sqrt{\int_{-f_{x,\mathrm{Nyq}}}^{f_{x,\mathrm{Nyq}}} \int_{-f_{y,\mathrm{Nyq}}}^{f_{y,\mathrm{Nyq}}} \mathrm{VTF}^4(f_x, f_y) \bullet S_{\mathrm{obj}}^2(R, f_x, f_y) \bullet \mathrm{IRF}_{d,ip}^2(f_x, f_y) \bullet \mathrm{NPS}_{d,ip}(f_x, f_y) \, df_x \, df_y}}. \tag{19}$$

The visual transfer function (VTF) of the human eye used in this study is an approximation of the Barten contrast sensitivity function of the human eye (Barten 1990) with a general form given in equation (20) (Burgess 1994, Segui and Zhao 2006, Gang *et al* 2011, Prakash *et al* 2011)

$$\mathrm{VTF}(f_x, f_y) = (f_x^2 + f_y^2)^{n/2} \bullet \exp(-c(f_x^2 + f_y^2)^{a/2}). \tag{20}$$

2.7. CDMAM phantom and NPWE model: threshold detectability index

The detectability index *d'* of the discs of different diameters and contrasts of the CDMAM phantom was calculated using equations (15), (18) and (19) of the NPWE model. The Fourier spectrum S_{obj} of a disc of radius *R* and thickness *T* is given in equation (21), where J_1 is the Bessel function of the first kind

$$S_{\mathrm{obj}}(R, f_x, f_y, f_z) = \frac{R}{\sqrt{f_x^2 + f_y^2}} \cdot J_1(2\pi R \sqrt{f_x^2 + f_y^2}) \cdot \mathrm{sinc}(\pi T f_z). \tag{21}$$

All the CDMAM discs are much thinner than the DBT plane thickness T_z, leading to a sinc function that is approximately 1.0 below the Nyquist frequency $f_z = 0.5/T_z$. As a result, the discs can be treated as 2D objects

$$S_{\mathrm{obj}}(R, f_x, f_y) \cong \frac{R}{\sqrt{f_x^2 + f_y^2}} \bullet J_1(2\pi R \sqrt{f_x^2 + f_y^2}). \tag{22}$$

The term $\Delta \bar{q}_{\mathrm{obj}}/\bar{q}$ in equation (15) represents the relative radiant contrast of the CDMAM gold discs and is proportional to the disc thickness *T* when $\Delta\mu \bullet T < 0.1$, which is the case for all the CDMAM discs

$$\frac{\Delta \bar{q}_{\mathrm{obj}}}{\bar{q}} = 1 - e^{-\Delta\mu \bullet T} \cong \Delta\mu \cdot T. \tag{23}$$

The term $\Delta\mu$ is the difference in mean linear attenuation coefficients between the gold disc and the PMMA + Al base of the CDMAM phantom for a given x-ray spectrum

$$\Delta\mu = \mu_{\mathrm{gold}} - \mu_{\mathrm{PMMA+Al}} \tag{24}$$

Equations (15), (19) and (23) were used to determine the threshold gold disc thickness (T_T) that corresponds to a threshold detectability index d_T'.

$$T_T \cong \frac{\Delta \bar{q}_{\mathrm{IRF}}}{\bar{q}} \bullet \frac{d_T'}{\Delta\mu \bullet k(R)} \tag{25}$$

2.8. Threshold diameter of a spherical object

We consider spherical objects of radius *R* embedded in a homogeneous thickness (equation (26))

$$S_{\mathrm{obj}}(r) = \begin{cases} 1 - \left(\frac{r}{R}\right)^2 & \text{if} \quad r < R \\ 0 & \text{if} \quad r > R \end{cases} \tag{26}$$

The object shape spectrum corresponds to equation (27) (Bateman and Erdélyi 1954)

$$S_{\mathrm{obj}}(R, f_x, f_y, f_z) = \frac{R^{3/2}}{\pi} \bullet \frac{J_{3/2}(2\pi R(f_x^2 + f_y^2 + f_z^2)^{1/2})}{(f_x^2 + f_y^2 + f_z^2)^{3/4}} \tag{27}$$

The maximum radiant contrast (x-ray attenuation) Δq_{obj} occurs along the diameter of the sphere

$$\frac{\Delta q_{\mathrm{obj}}}{\bar{q}} = 1 - e^{-2 \bullet \Delta\mu \bullet R}, \tag{28}$$

where $\Delta\mu$ is the difference between the mean linear attenuation coefficients of the spherical object and the background mammography tissue.

$$\Delta\mu = \mu_{\mathrm{object}} - \mu_{\mathrm{mammo}}. \tag{29}$$

Microcalcifications and masses were considered as spheres made of calcium oxalate (CaC_2O_4) and glandular tissue, respectively. The background mammographic tissue was considered as a mixture of adipose and glandular tissues whose glandular fraction (by weight) depended on the thickness: 67%, 20% and 4% glandularity for the PMMA thicknesses 30, 50 and 70 mm, which are equivalent in attenuation to breast thicknesses 32, 60 and 90 mm, respectively (Dance *et al* 2000). The threshold object diameter of a spherical object is twice the threshold radius R_T that corresponds to a threshold value of the detectability index in the NPWE model observer, noted d_T'. R_T was calculated numerically using equation (30)

$$(1 - e^{-2\bullet\Delta\mu\bullet R_T}) \bullet k(R_T) = \frac{\Delta\bar{q}_{\text{IRF}}}{\bar{q}} \bullet d_T' \tag{30}$$

R_T for the masses was calculated as a function of the level of the anatomical NPS by adjusting the magnitude coefficient in equation (11) from 0 to 2κ, i.e. from the absence of anatomical noise until twice the level measured in the L1 phantom. R_T for the microcalcifications was calculated as a function of K_{in}, between 100 and 1000 μGy. For this calculation, the impulse response for a linear system was considered proportional to the air kerma at the system input K_{in} (equation (31))

$$\text{IRF}_d(f_x, f_y, K_{\text{in}}) = \text{IRF}_d(f_x, f_y, 1) \bullet K_{\text{in}}. \tag{31}$$

Considering quantum limited systems, the system NPS ($NPS_{d,s}$) is proportional to K_{in} to a good approximation, and the anatomical NPS ($NPS_{d,a}$) is proportional to K_{in}^2 (Mainprize and Yaffe 2010). The variation of the in-plane NPS in the image was therefore calculated as a function of K_{in} according to equation (32)

$$\text{NPS}_{d,ip}(f_x, f_y, K_{in}) \cong \text{NPS}_{d,s}(f_x, f_y, 1) \bullet K_{\text{in}} + \text{NPS}_{d,a}(f_x, f_y, 1) \bullet K_{\text{in}}^2. \tag{32}$$

2.9. CDMAM scoring and validation of NPWE model against human observers

The CDMAM images were scored using the CDCOM module and contrast-detail (c-d) curves were generated using the standard processing method described by Young *et al* (2006). The cases where the in-focus DBT plane was not obvious, planes around the plane considered to have the best focus were scored and the plane with the lowest contrast thresholds was taken to be the in-focus plane. Uncertainty on the threshold gold thickness was estimated using a bootstrap method.

The predictive value of the CDMAM results has not yet been established against human observers in the presence of anatomical background noise. The agreement between CDCOM scoring, the NPWE model and human performance was therefore evaluated. Six human observers read ten series of images of the CDMAM phantom, five with system noise only and five with anatomical background noise. Software developed in-house was used to crop and rotate each CDMAM square to prevent the observers learning the disc positions, which would bias the results. Each observer read two images in every series, giving 20 images in total for each reader. A Barco 10-megapixel radiology display (model MDMC-12133) (Barco, Kortrijk, Belgium) was used to display the images. During reading, the room lighting was subdued, giving an ambient light level of approximately 6 lux. A typical viewing distance of approximately 65 cm was suggested, although the readers were free to zoom the image content while scoring. The number of reading sessions ranged from 2 to 4 sessions (median value of 3), spread over a period of approximately 3 weeks, depending on the availability of the observer at the hospital. The reading time was approximately 19 min per image, averaged over all images and observers. The series included homogeneous background (raw and processed DM, DBT and SM images for the Siemens and DBT images for the Hologic) and anatomical background (processed DM, DBT and SM images for the Siemens, and DBT and SM images for the Hologic). The average from 12 readings from each series gave the threshold thickness of human reading for each disc diameter. The uncertainty in the threshold thickness was quantified using two times the standard error of the mean (SEM).

3. Results

3.1. Acquisition parameters and x-ray beam characteristics

The AEC settings obtained for the three PMMA thicknesses on the three mammography units, along with the measured values of the photon fluence per unit air kerma, the relative radiant contrast for the 0.2 mm aluminium thickness and SF_{in} are given in table 2. The 0.2 mm aluminium disc in the NPWE phantom produced a signal (radiant contrast) between 6.4% and 21.7% (table 2). We assume that this is sufficiently small for the assumptions of constant system gain and small signal linearity of the log-normalization stage in the DBT and SM reconstructions to be valid. The air kerma at the system input (K_{in}) obtained under AEC for the homogenous phantom varied widely among the systems and imaging modes, from 163 μGy for the Siemens Revelation (DM, 70 mm) to 822 μGy for the Selenia Dimensions (DBT, 70 mm). SF_{in} varied between 0.291 and 0.360 for the thickness 30 mm, between 0.460 and 0.504 for 50 mm and between 0.555 and 0.589 for 70 mm.

This amounts to variations of less than 7% measured for the different imaging modes and systems and can be explained by differences in x-ray spectra, geometry and compression plate size and composition.

3.2. In-plane Fourier metrics: MTF, NPS, NEQ and system DQE

The mean radial system MTF of raw DM images decreased to zero around 11 mm^{-1} on the Selenia Dimensions and Revelation, and around 6 mm^{-1} on the Pristina (figures 2(a)-(c)). The useful signal bandpass in the digital images is however limited by the Nyquist frequency determined at the sampling stage by the pixel pitch, resulting in spatial frequencies of 7.14, 5.88 and 5.00 mm^{-1} for the Selenia Dimensions, Revelation, and Pristina, respectively. For all three systems, the image processing or reconstruction algorithm gave the system MTF a similar peak shape that contributed to edge enhancement on processed DM, DBT and SM images. The system MTFs of processed images peaked between 0.7 and 1.3 mm^{-1} for the three systems. These results are consistent with in-plane MTFs previously measured in DM and DBT images (Zhao *et al* 2009 and 2017, Marshall and Bosmans 2022). Unlike the other systems, the image processing applied on DM and DBT images increased the MTF compared to raw DM images on the Pristina system.

The mean radial NEQ curves of raw DM images describe the ability of the system to transfer signal and noise, and show that the imaging systems act as low-frequency filters (figures 3(a)-(c)). The NEQ of raw DM images increased with the number of detected primary photons, and therefore increased with the DAK chosen by the AEC for a given system, the primary fraction at the detector and the detector DQE. For processed DM, DBT and SM images, image processing or reconstruction cause the NEQ to differ from the original low-frequency shape determined by the detector MTF and NPS. The processing increased the NEQ of DM images acquired on the Selenia Dimensions and the Revelation systems, but not for the Pristina. This result shows that the image processing modifies the signal and noise transfers differently and is amplifying the signal content and decreasing the noise to some extent. The Selenia Dimensions gave the highest NEQ of the studied systems for the DBT planes for the thickness 70 mm. The high exposure delivered by the AEC (table 2) contributes to this result. The NEQ frequency bandpass of DBT planes was reduced compared to that of DM images on the Selenia Dimensions and Revelation systems but was the same on the Pristina. The Revelation system gave the lowest NEQ of the study for the Insight 2D images for the thickness 70 mm. The NEQ of the Insight 2D images peaked at frequencies between 2 and 3 mm^{-1} for the thicknesses 30 and 50 mm but was poor at lower and higher frequency. This behaviour arises from strong noise filtration that dramatically reduces the high-frequency NPS. The NEQ of the C-View images was systematically lower than that of the corresponding DM images and DBT planes, especially at high spatial frequency.

Mean radial system DQE results are visualized in figures 4(a)-(c). The system DQE measured on raw DM images is the product of the detector and the grid DQE. The raw DM images had a maximum system DQE between 0.50 and 0.65 for the three systems and three thicknesses. The grid DQE, and hence the system DQE, increased with SF_{in} and with the phantom thickness. Image reconstruction and processing algorithms can strongly modify the system DQE, confirming that signal and noise can be handled differently in these processes. The Pristina system gave similar DQE while the Selenia Dimensions and Revelation gave very different DQE for the different image types. The processed DM images gave the highest system DQE over the widest frequency range for the Selenia Dimensions and the Revelation systems. While image processing increased the system DQE of DM images to approximately 0.8 on the Selenia Dimensions and Revelation systems, the low-frequency DQE of processed DM images was decreased for the Pristina system. The DBT planes gave a lower DQE over a reduced bandwidth, compared to DM images. The DQE of DBT planes reached up to 5 mm^{-1} on the Pristina and Selenia Dimensions but fell to zero at 3 mm^{-1} on the Revelation. The SM images had a reduced frequency bandwidth that differed strongly for the three thicknesses on the Revelation and Selenia Dimensions systems. The comparison in DQE between the processed against the raw images shows the importance of the image processing and reconstruction algorithm for overall system efficiency.

The grid DQE ranged between 0.96 and 1.10 for the thickness 30 mm, between 1.23 and 1.35 for the thickness 50 mm and between 1.37 and 1.58 for the thickness 70 mm (table 3). The grid DQE curves plotted as a function of SF_{in} show that the cellular grid of the Selenia Dimensions system had a DQE between 7% and 15% higher than the linear grids of the Pristina and Revelation systems (figure 5).

3.3. In-depth resolution (MTF$_z$)

Although the DBT plane spacing was 1 mm for the three systems, the different systems and reconstructions gave large differences in in-depth resolution (figure 6). The MTF$_z$ of the Siemens system reached zero at the z-frequency of 0.49 mm^{-1} for an angular span 50°. This cut-off frequency corresponds to an effective plane thickness of 2.04 mm, comparable to the FWHM of the measured PSF$_z$ (2.05 mm). The Pristina gave the narrowest effective plane thickness of 1.56 mm, while the narrowest angular acquisition of 15° of the Selenia Dimensions gave a thicker effective plane of 2.95 mm.
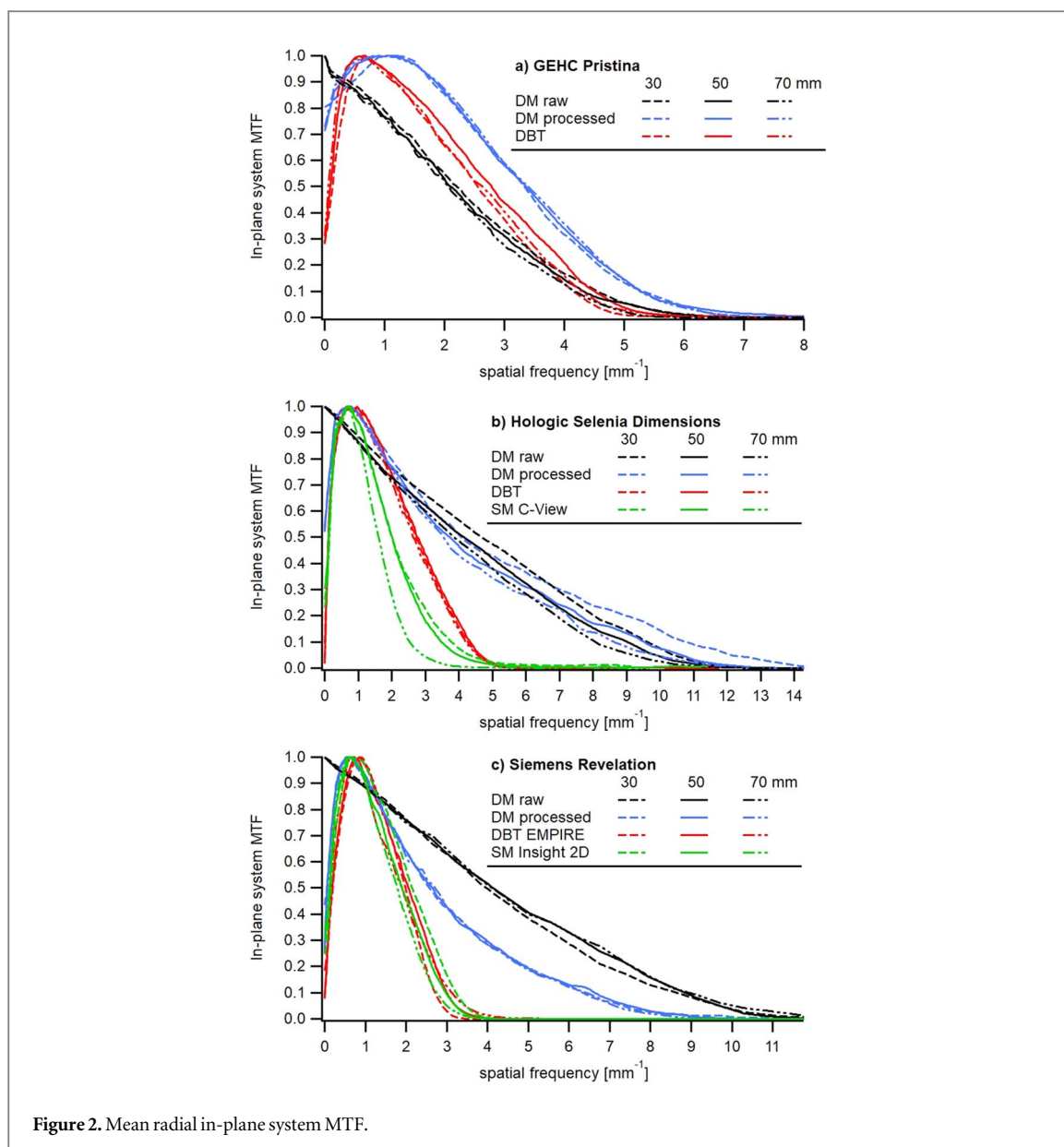
Figure 2. Mean radial in-plane system MTF.

3.4. Background anatomical breast noise

The large differences in NPS magnitudes measured in the L1 phantom are mostly due to differences in the signal gains used for the different image types (figures 7(a)-(c)). The power-law noise parameters κ and β fitted to the anatomical NPS measured in the L1 phantom are shown in table 4. The power exponent β varied between 2.8 and 3.6 for the different systems, imaging modes and image types, in agreement with β values around 3.0 measured in mammograms (Chen *et al* 2012, Mainprize *et al* 2012, Cockmartin *et al* 2013, Hill *et al* 2013). The power law parameters κ and β were the same for the raw and processed DM images. This result shows that the image processing applied on DM images did not significantly modify the frequency composition of the low-frequency anatomical structured noise. DBT planes gave lower anatomical noise magnitude κ and power term β than DM images, in agreement with theoretical prediction (Metheany *et al* 2008) and previous measurements (Engstrom *et al* 2009). This result confirmed that DBT reduces the amount of structured noise in the reconstructed DBT planes compared to DM projection images. The κ and β attributes of SM images lie somewhere between DM and DBT images, with an anatomical NPS magnitude higher than DBT but reduced compared to DM. The increased angular span of the DBT acquisitions on the Siemens system reduced the noise magnitude as quantified by κ, showing that a wider angular DBT scan improves the out-of-plane clutter rejection and reduces the superposition of structured noise on the reconstructed planes (Yoon *et al* 2009).

The simulated anatomical NPS was able to match the NPS measured in the L1 phantom for all the systems and imaging modalities (figures 7(a)-(c)). This simulated anatomical noise was added to the CDMAM phantom images to study the effect of this noise on detectability. It is important to note that the aim of this study was not a

Figure 3. Mean radial NEQ.
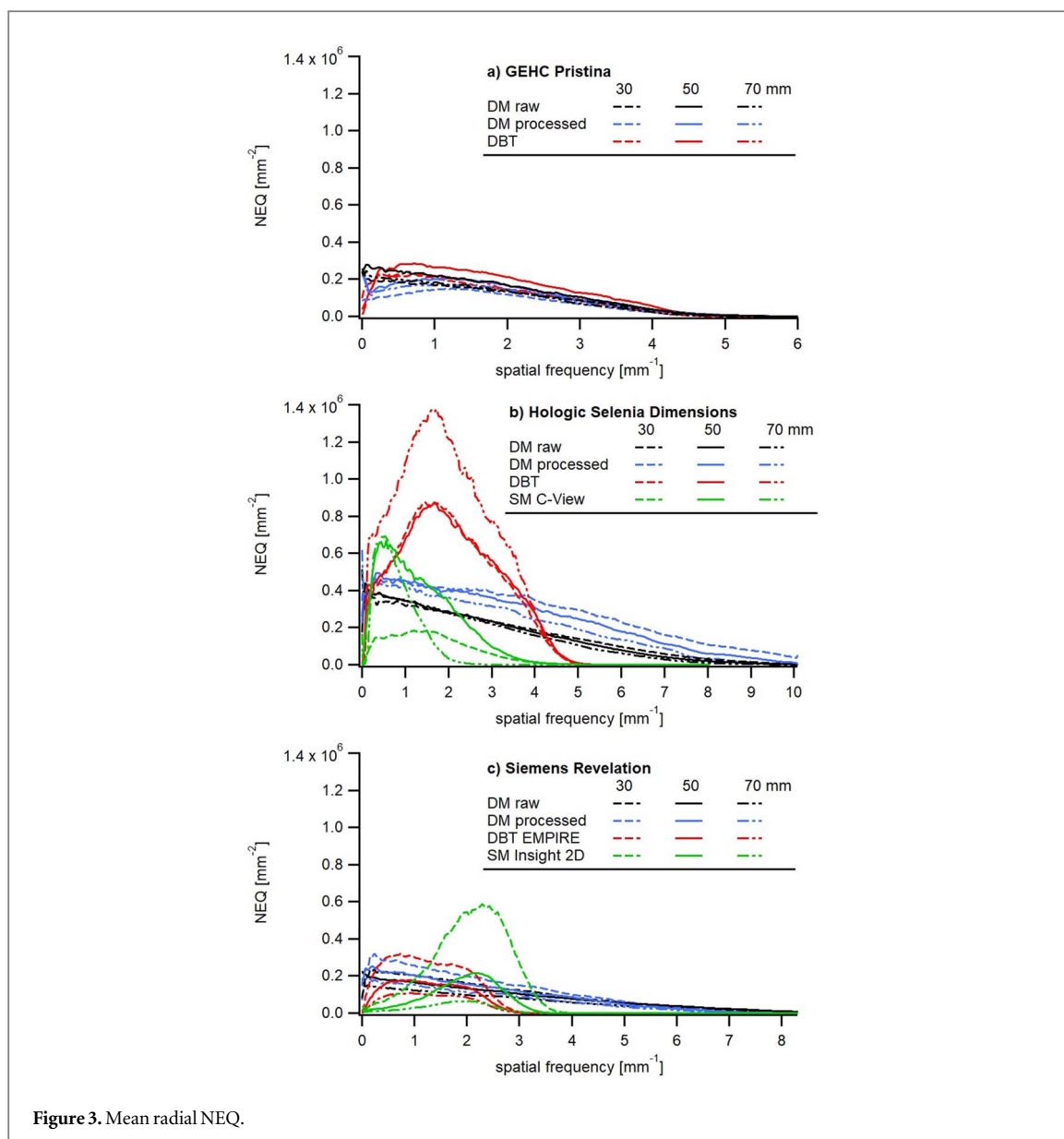
precise determination of κ and β, but an investigation of their effect on the detectability given by the CDMAM and the NPWE model. Only a rough similarity between the simulated background NPS and the anatomical NPS magnitude measured in the L1 phantom was therefore considered in this study.

3.5. Validity of the NPWE model compared against the CDMAM phantom

The three free parameters of the VTF used in the NPWE model were adjusted to find the best correlation between the threshold gold thickness (T_T) of the CDMAM discs given by the CDCOM software (van Engen 2013) and by the NPWE model. The threshold gold thickness for CDMAM discs of different diameters corresponds to the smallest disc thickness visible in 62.5% of cases by a human observer (Young *et al* 2006). The parameters $n = 0.65$, $c = 0.0012$ cycle/° and $a = 2.0$ were used in equation (20), with a viewing distance of 400 mm without magnification. This VTF covered a wide frequency bandwidth until 10 mm^{-1} and gave a maximum response at 2.36 mm^{-1} (figure 8).

The T_T values obtained from the CDMAM images were plotted against the ratios T_T/d' calculated with the NPWE model (figures 9(a)–(c)). A log–log linear correlation was found for all the systems and imaging conditions and for the three thicknesses. This correlation also held both for images containing anatomical background noise and those with a homogeneous noisy background. The least squares linear regression to all the CDMAM discs between 0.1 and 2.0 mm gave a Pearson linear correlation coefficient (PCC) of 0.980 for a slope set to 1.0 (black line in figures 9(a)–(c)). The NPWE observer model accurately predicted T_T values over a broad range of conditions, suggesting that this model is suitable for image quality characterization over the range of
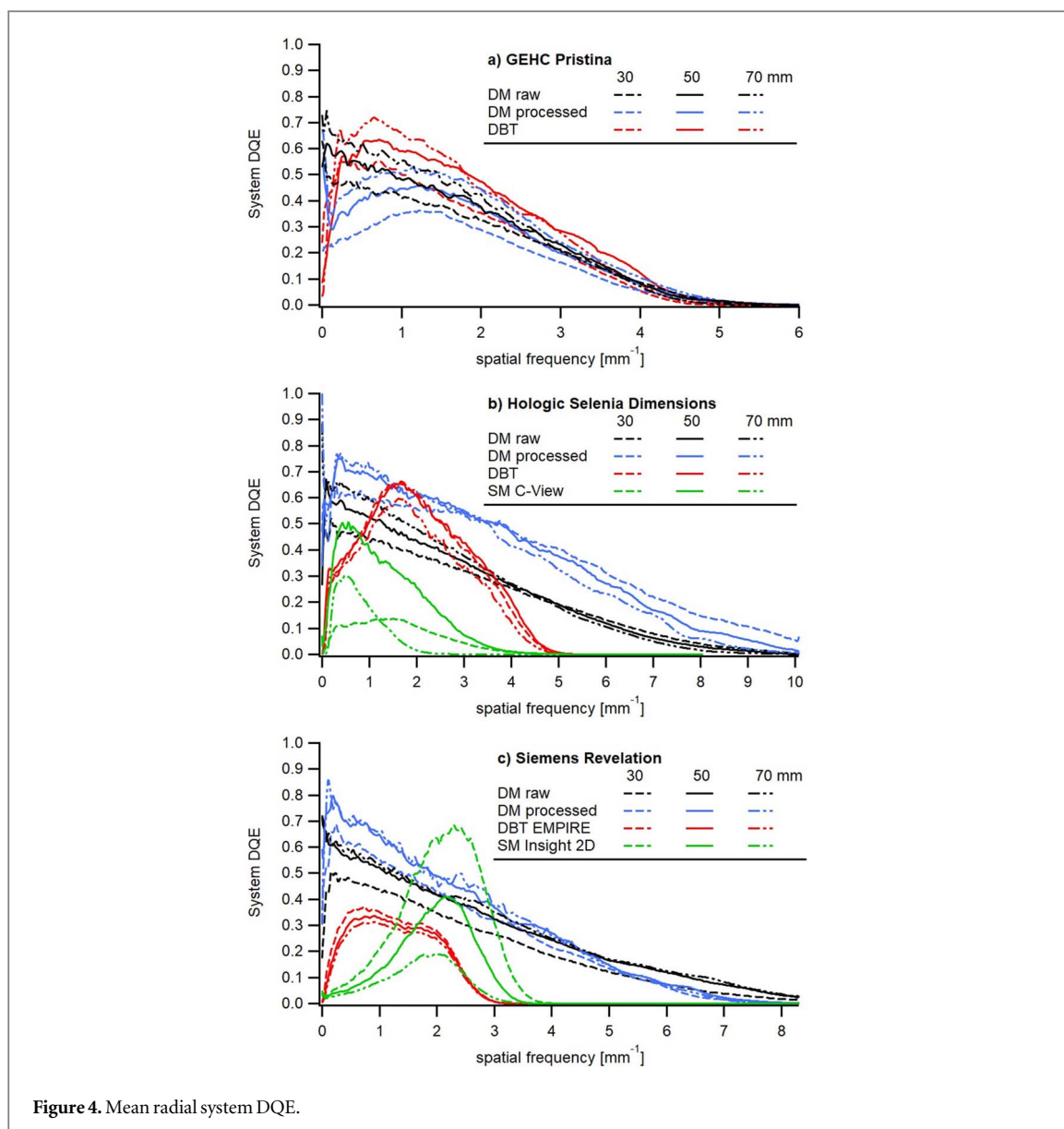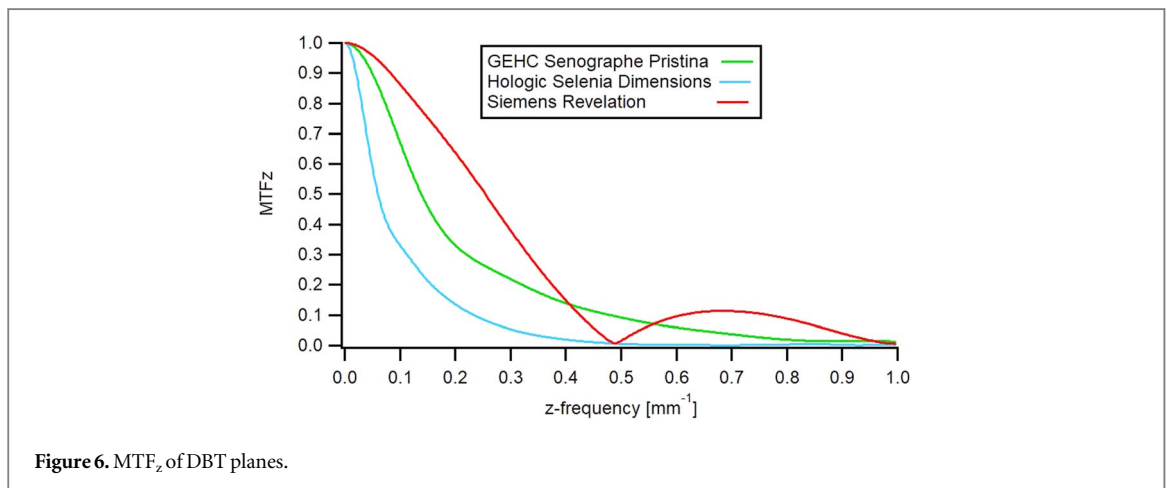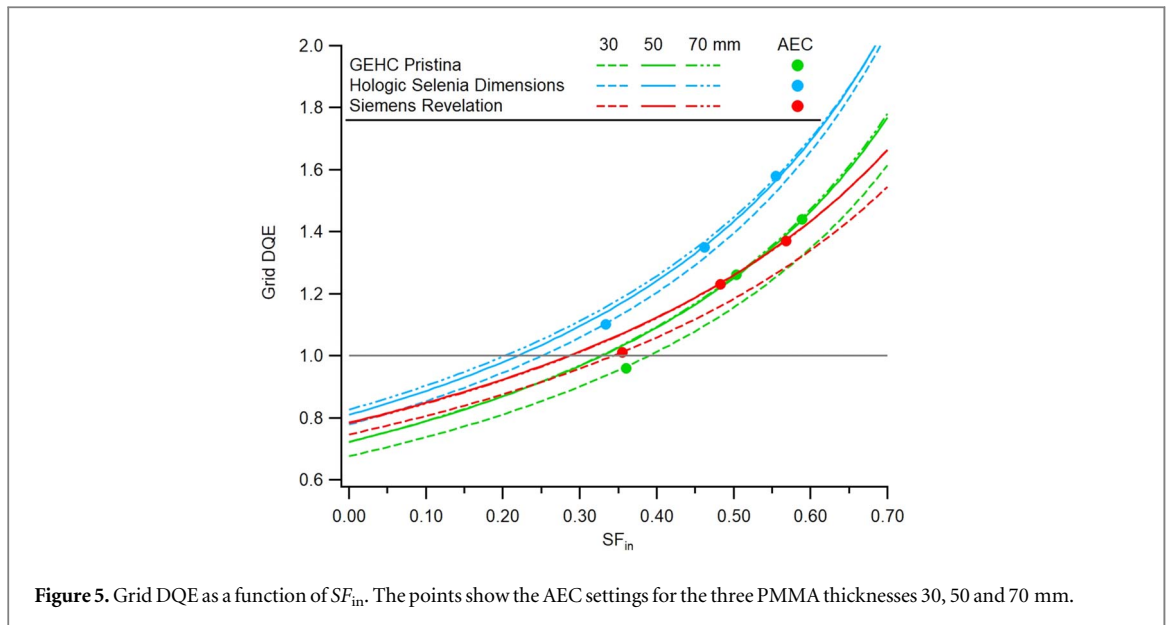
Figure 4. Mean radial system DQE.

Table 3. Grid characteristics.

System	Modality	Phantom thickness (mm)	Tube voltage (kV)	T_p	T_s	T_g	Grid DQE	SF_{out}
GEHC Senographe Pristina	DM & DBT	30	26	0.68	0.11	0.47	0.96	0.087
		50	34	0.72	0.11	0.42	1.26	0.136
		70	34	0.72	0.11	0.36	1.44	0.178
Hologic Selenia Dimensions	DM	30	26	0.78	0.09	0.55	1.10	0.055
		50	31	0.81	0.11	0.48	1.35	0.100
		70	34	0.83	0.12	0.43	1.58	0.151
Siemens Revelation	DM	30	27	0.75	0.19	0.55	1.01	0.129
		50	30	0.78	0.19	0.50	1.23	0.186
		70	32	0.78	0.19	0.45	1.37	0.242

thicknesses, background noise and imaging modalities considered, i.e. for DM raw and processed images, DBT planes and reconstructed SM images. The best numerical correspondence in threshold gold disc thickness (T_T) between the CDMAM and the NPWE model given by the linear log–log correlation was obtained for a threshold detectability index $d_T' = 2.92 \pm 0.11$ (uncertainty 2 sigma).

Figure 5. Grid DQE as a function of SF_{in}. The points show the AEC settings for the three PMMA thicknesses 30, 50 and 70 mm.

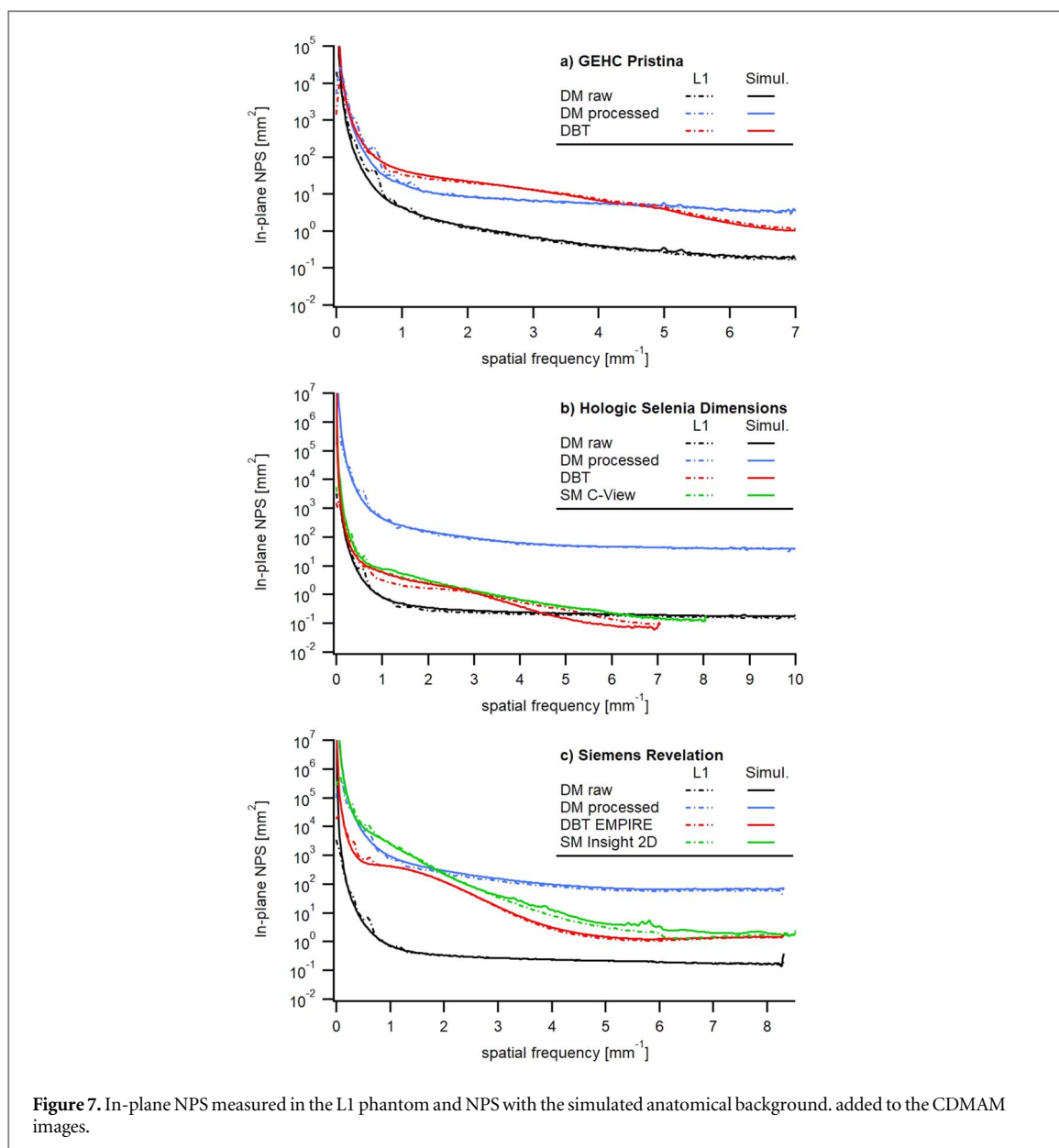**Figure 6.** MTF_z of DBT planes.

3.6. Validity of the CDMAM phantom and NPWE model compared against human observers

The human reading results of the CDMAM phantom were in good agreement with CDCOM scoring for the plain background images (figure 10(a)). The averaged ratio of human to CDCOM threshold thicknesses for all disc diameters was 0.961. The human and CDCOM threshold thicknesses were less consistent in the anatomical background. Slightly better performance for human observers in the anatomical backgrounds was found with a ratio of human to CDCOM results equal to 0.819. The error bars encompass the identity line except for the Insight 2D images, for which the human readers performed significantly better than CDCOM.

The human observers gave results in good agreement with those obtained with the NPWE model (figure 10(b)). As expected, the anatomical background resulted in a greater relative increase of threshold gold thickness for the large diameter discs compared to the smaller discs. The averaged ratios of human to NPWE threshold thicknesses for all disc diameters were 0.992 for the plain background images and 1.070 in the anatomical background. Human observer performance was slightly poorer in the anatomical background than the NPWE model. All the error bars cut the line fitted between NPWE and CDCOM values, except for the five largest discs of the Insight 2D images with plain background.

3.7. Threshold diameter of microcalcifications (high-frequency tasks)

The threshold diameter D_T of spherical microcalcifications (MCs) is the diameter of a sphere made of calcium oxalate (CaC_2O_4) that corresponded to the threshold detectability index $d_T' = 2.92$ calculated using the 3D NPWE model observer. The detectability index of small objects, which are considered to be high-frequency tasks, depends mostly on the amount of system noise and has little dependence on low-frequency anatomical background noise (Ruschin *et al* 2007, Gang *et al* 2010, Vancoillie *et al* 2021). This is largely governed by the DAK. D_T was therefore determined without anatomical noise as a function of K_{in}, for K_{in} values between 100 and

Figure 7. In-plane NPS measured in the L1 phantom and NPS with the simulated anatomical background. added to the CDMAM images.

1000 μGy (figures 11(a)-(c)). As expected, D_T increased with the phantom thickness and decreased with K_{in} for all the imaging conditions. For the three thicknesses, the processed DM images gave the smallest D_T values for the Hologic and Siemens systems, whereas DBT planes gave slightly better detectability than DM images for the Pristina. SM images systematically gave the largest D_T values. The ranking given by the NPWE model in this study is consistent with recent studies where SM images gave the poorest detectability of small details, whereas DM gave a better detection than DBT images (Hadjipanteli *et al* 2017, Mackenzie *et al* 2021, Vancoillie *et al* 2021). In Ikejimba *et al* (2021), MC detectability was highest for DBT and DM images compared to SM, with slightly better detectability for DBT images.

The D_T values obtained for MCs under AEC conditions depended on the DAK chosen by the AEC and varied for the different imaging modalities between 0.21 and 0.27 mm for the thickness 30 mm, 0.23–0.29 mm for the thickness 50 mm and 0.26–0.36 mm for the thickness 70 mm (table 4). The DM processed images acquired on the Selenia Dimensions gave the smallest D_T values, consistent with the highest DAK values of the three systems for the three thicknesses. The threshold diameters calculated in our study correspond to a threshold detectability index d_T' = 2.92 and to a percentage of correct (PC) detection of 62.5% established on the CDMAM phantom. The threshold diameters of MCs calculated in our study are larger than those found in Ikejimba *et al* (2021) and in Vancoillie *et al* (2021). The D_T values in our study correspond to the detection of a single dot whereas the detection task of the other authors consisted of a cluster of several MCs, which can be an easier detection task. The composition of MCs was also different, and this will influence the detection threshold value. Ikejimba *et al* (2021) used a mixture of calcium hydroxyapatite and epoxy, whose attenuation coefficient is not explicitly given,

Table 4. Parameters κ and β of the anatomical breast noise measured on the L1 phantom, and threshold diameter of masses and microcalcifications (in mm) obtained under the AEC (in bold the best result for each system).

| System	Modality	κ	β	MCs ($\kappa = 0$)			Masses ($\kappa = 0$)			Masses ($\kappa \times 1$)			Masses ($\kappa \times 2$)		
				30	50	70	30	50	70	30	50	70	30	50	70 mm
GEHC Senographe Pristina	DM raw	4e-6	−3.2	0.251	0.281	0.298	**2.4**	**2.0**	**2.0**	**9.9**	7.7	8.7	**12.7**	9.9	11.2
	DM processed	4e-6	−3.2	0.272	0.284	0.298	3.0	2.2	2.2	10.4	7.3	8.3	13.6	9.8	11.1
	DBT	2e-6	−2.8	**0.250**	**0.273**	**0.292**	3.4	2.5	2.4	10.6	**5.2**	**4.8**	14.9	**6.7**	**6.1**
Hologic Selenia Dimensions	DM raw	3e-6	−3.4	0.224	0.237	0.260	2.2	1.7	1.8	9.2	6.9	7.9	11.9	9.1	10.4
	DM processed	3e-6	−3.4	**0.210**	**0.225**	**0.255**	**2.2**	**1.7**	**1.8**	9.8	6.6	8.1	13.6	9.6	11.6
	DBT	3e-7	−3.1	0.222	0.247	0.275	2.5	2.0	2.1	**5.0**	**2.9**	**2.7**	**6.3**	**3.5**	**3.2**
	SM C-View	4e-7	−3.4	0.265	0.258	0.314	2.9	1.8	2.1	5.5	3.6	4.6	7.9	5.0	6.4
Siemens Revelation	DM raw	6e-6	−3.4	0.244	0.263	0.287	2.5	2.0	2.0	11.8	8.9	10.2	15.2	11.8	13.3
	DM processed	6e-6	−3.4	**0.239**	**0.257**	**0.280**	**2.5**	**1.9**	**2.0**	11.6	8.0	9.8	17.3	11.6	14.7
	DBT	5e-7	−3.0	0.250	0.276	0.328	3.3	2.6	2.9	**5.2**	**3.2**	**3.3**	**6.1**	**3.8**	**3.7**
	SM Insight 2D	2e-6	−3.6	0.254	0.287	0.364	3.2	2.8	3.2	19.6	7.5	7.7	25.9	12.2	12.2
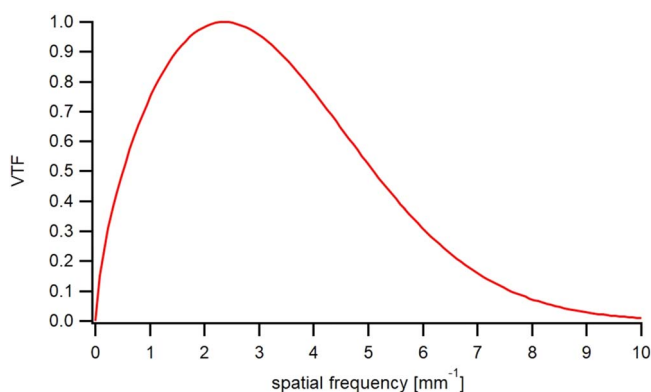
Figure 8. Visual transfer function (VTF) used in the NPWE model.

however the attenuation coefficient of the calcium hydroxyapatite component used in Ikejimba *et al* (2021) is approximately twice as large as that of the calcium oxalate used in our study. The calcium carbonate used in the L1 phantom-based study of Vancoillie *et al* (2021) has an attenuation coefficient that is 40% to 60% larger higher than that of calcium oxalate.

3.8. Threshold diameter of masses (low-frequency tasks)
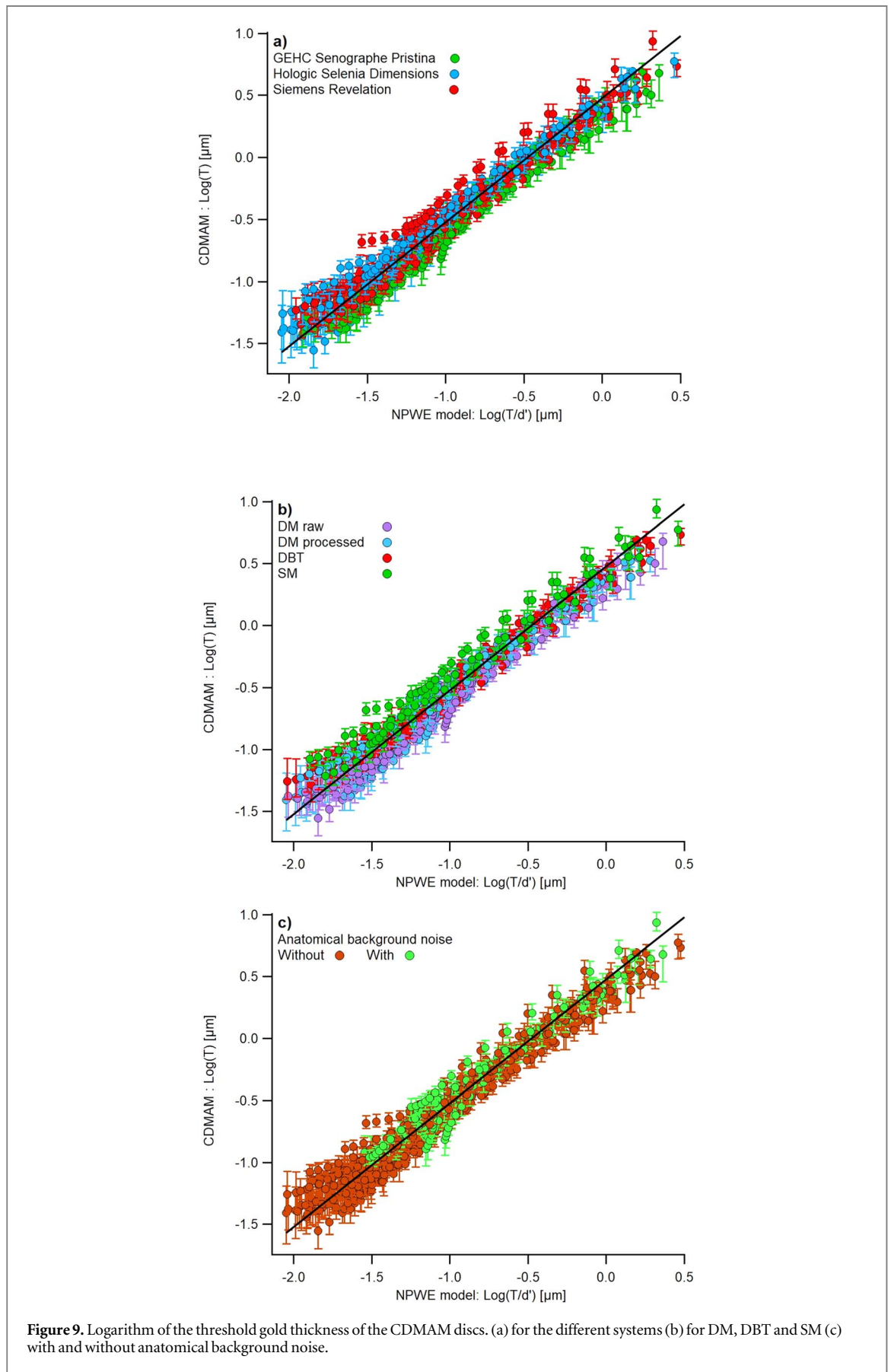
Masses were modelled as glandular spheres. As for MCs, the threshold diameter D_T corresponded to the threshold detectability index $d_T' = 2.92$ obtained from the 3D NPWE model observer. The detection of large objects (low-frequency tasks) depends mostly on the amount of low-frequency anatomical background noise, and has little dependence on exposure level (Ruschin *et al* 2007, Gang *et al* 2010, Vancoillie *et al* 2021). D_T was therefore calculated as a function of added anatomical noise, for κ values between 0 and 2κ, with κ being the level of background noise measured on the images of the L1 phantom (figures 12(a)-(c)). D_T increased with κ for all the imaging conditions. A smaller thickness was associated with a higher glandularity and a lower contrast, resulting in higher D_T values. For the three systems, DM images gave the smallest D_T values in the absence of anatomical background noise ($\kappa = 0$), whereas DBT planes gave better detectability when anatomical noise was present (table 4). Except for the Insight 2D images which gave particularly poor detectability for the thickness 30 mm, SM images gave D_T values intermediate between DM and DBT. These results are consistent with recent technical image quality studies (Hadjipanteli *et al* 2019, Ikejimba *et al* 2021, Vancoillie *et al* 2021).

D_T values obtained under the AEC conditions varied for the different imaging modalities within the range 1.7–3.4 mm for $\kappa = 0$, 2.7–19.6 mm for κ x 1 and 3.2–25.9 mm for κ x 2 (table 4). The smallest D_T values were obtained on DM processed images of the Selenia Dimensions for $\kappa = 0$ and on DBT planes of the Selenia Dimensions for κ x 1 and κ x 2. The threshold diameters of masses calculated in our study for the phantom thckness 50 mm are consistent with those found in Ikejimba *et al* (2021) and in Vancoillie *et al* (2021). Masses of diameter 5 mm in Ikejimba *et al* (2021) gave a PC between 0.6 and 0.8. Human observers in Vancoillie *et al* (2021) found threshold diameters between 2 and 6 mm (PC 62.5%). For the same PC of 62.5%, Hadjipanteli *et al* (2019) obtained D_T values around 10 mm for DM and 6 mm for DBT on simulated images with anatomical background, slightly larger than our results for the thickness 50 mm.
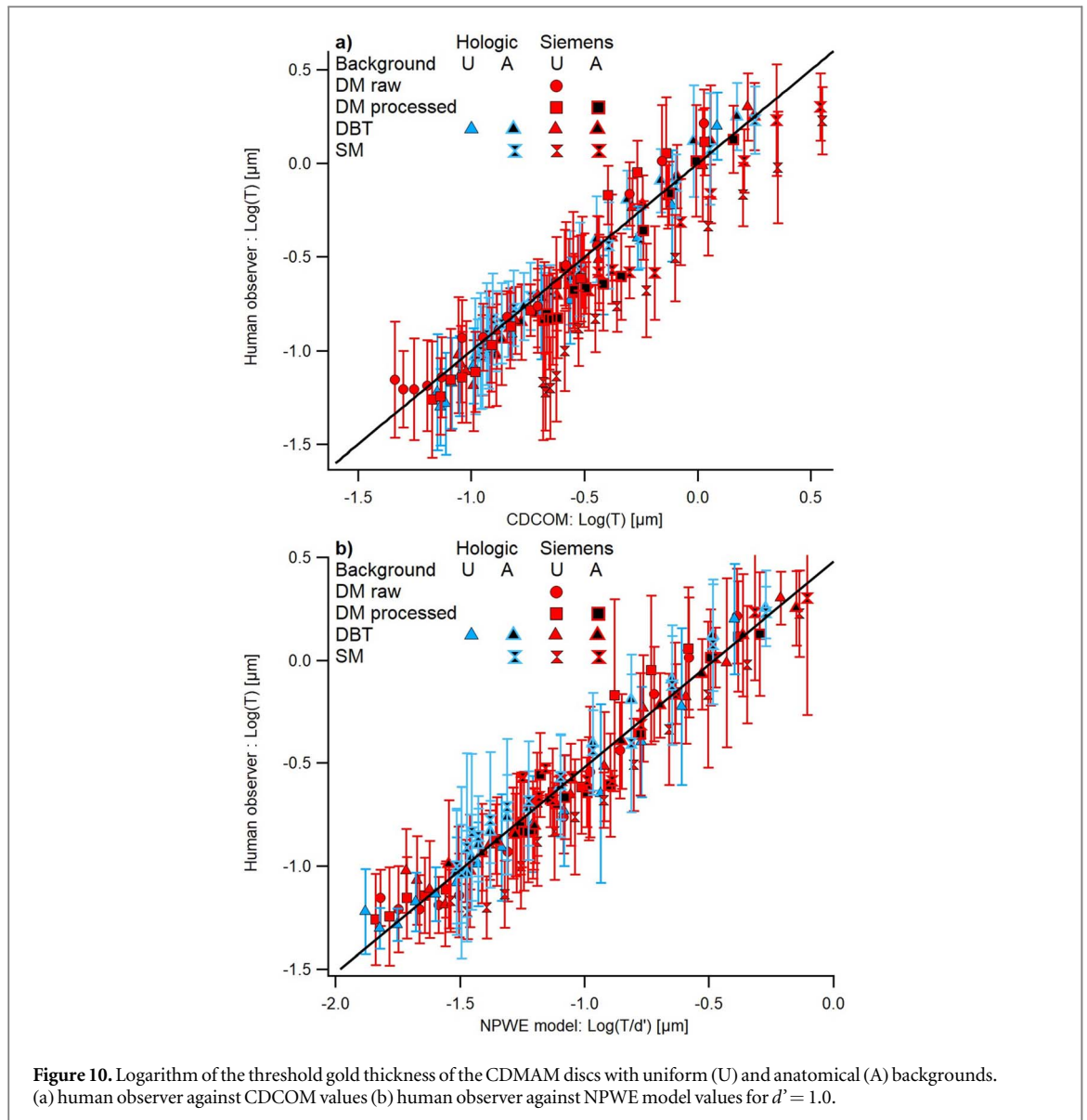
4. Discussion

This study has established a link between the CDMAM phantom, human reading and the NPWE model to study detectability in DM, DBT and SM. The NPWE model gave coherent and robust results compared to the CDMAM phantom and human reading for a range of discs between 0.1 and 2.0 mm in diameter, for PMMA thickness between 30 and 70 mm, and for DM, DBT and SM images with and without anatomical background noise. These results valid for a wide range of x-ray beam energies, SF levels and noise characteristics extended those published in previous work for DBT planes (Monnin *et al* 2020) and for raw DM images in a homogeneous noisy background (Monnin *et al* 2011).

The discs in the CDMAM phantom are considerably thinner than the reconstructed DBT plane spacing, allowing the detection of MCs to be treated as a 2D problem. In this study, we further considered masses whose threshold diameter exceeded the size of the CDMAM gold discs and was not negligible with regards to the effective DBT plane thickness, leading to the use of a 3D NPWE model. The 3D NPWE model may thus be less accurate or diverge for objects larger than 2.0 mm that cover a lower frequency range. Thick objects can also give

Figure 9. Logarithm of the threshold gold thickness of the CDMAM discs. (a) for the different systems (b) for DM, DBT and SM (c) with and without anatomical background noise.

a high signal that is not compatible with the assumption of small signal linearity required when applying this model. As a result, nonlinear iterative image processing can lead to sharpness and noise levels dependent on the contrast of the imaged structures. In this study, the assessment of detectability performance using transfer

Figure 10. Logarithm of the threshold gold thickness of the CDMAM discs with uniform (U) and anatomical (A) backgrounds. (a) human observer against CDCOM values (b) human observer against NPWE model values for *d'* = 1.0.

functions in the 3D NPWE model is a linear approximation to nonlinear systems and must be ideally used under task-based conditions. The MTF and NPS used in the NPWE model were calculated under small contrast condition provided by the 0.2 mm Al disc. For thick objects with high contrast, the results could potentially differ from those obtained in this study, which only holds for low-contrast objects. Nevertheless, the extension of the NPWE model to a 3D formulation for tomosynthesis has been shown to have a reasonable correspondence with human observer performance over a broad range of imaging tasks and conditions in the presence of low-frequency background noise (Gang *et al* 2011). The 3D NPWE model has therefore been successfully used to explore the imaging performance and optimize the imaging parameters of CBCT and tomosynthesis for different imaging tasks like the detection of punctual (small), spherical, Gaussian objects (Gang *et al* 2010, Hu and Zhao 2011). The threshold diameters of masses for human observers in recent studies follow similar trends and are compatible with our results (Ikejimba *et al* 2021, Vancoillie *et al* 2021). The accuracy of our model must nevertheless be studied for large objects, and this is a limitation of the present study.

The hierarchy of detection performance for large (low-frequency) tasks like masses depended strongly on the amount of anatomical noise added to the images. We conclude that the use of detectability indices or threshold thicknesses calculated in homogeneous background is, as might be expected, not representative of clinical situations for comparing the performance of DM and DBT for low-frequency tasks. The anatomical noise reduced the detectability indices of masses, and the magnitude of this reduction depended on the in-depth resolution, which governs the ability of the system to reduce or restrict propagation of out of plane anatomical texture to the image plane containing the targets of interest. In our study, the system with a narrow angular span (Selenia Dimensions) could have similar or even superior detectability to the wide-angle systems for masses (Pristina and Revelation). This implies that other factors such as the reconstruction algorithm, in addition to
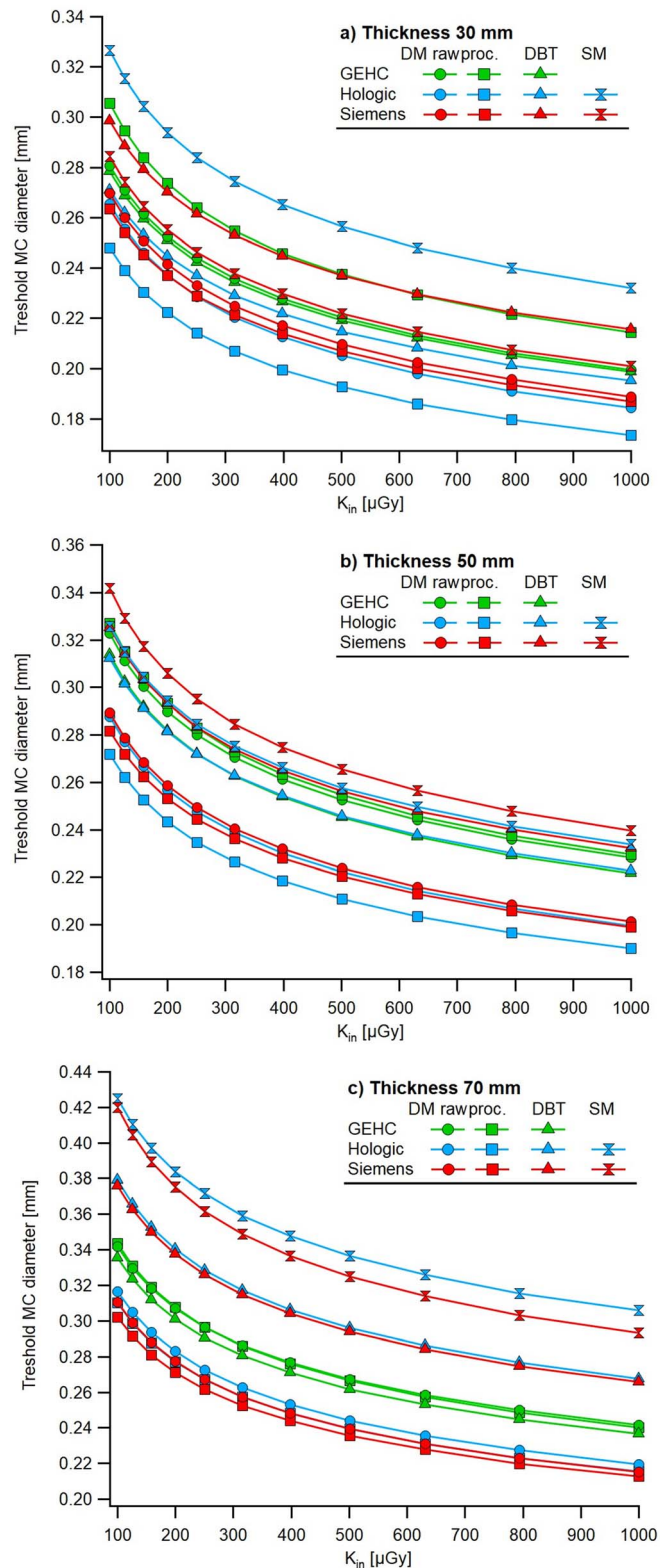
Figure 11. Threshold diameter of spherical microcalcification as a function of K_{in} for three PMMA thicknesses. (a) 30 mm (b) 50 mm (c) 70 mm.

angular range, have an important role in determining in-depth resolution and threshold detectability in anatomical background.

Conversely, detection of small (high-frequency) tasks like microcalcifications depends little on low-frequency anatomical noise, mostly on the system (quantum and electronic) noise and hence on the DAK and primary x-ray detector efficiency. Our study showed that DBT did not improve the detection performance of microcalcifications, consistent with data in the literature (Gang *et al* 2010, Hadjipanteli *et al* 2017, Mackenzie *et al* 2021, Ikejimba *et al* 2021, Vancoillie *et al* 2021). Detection performance of small objects can therefore be
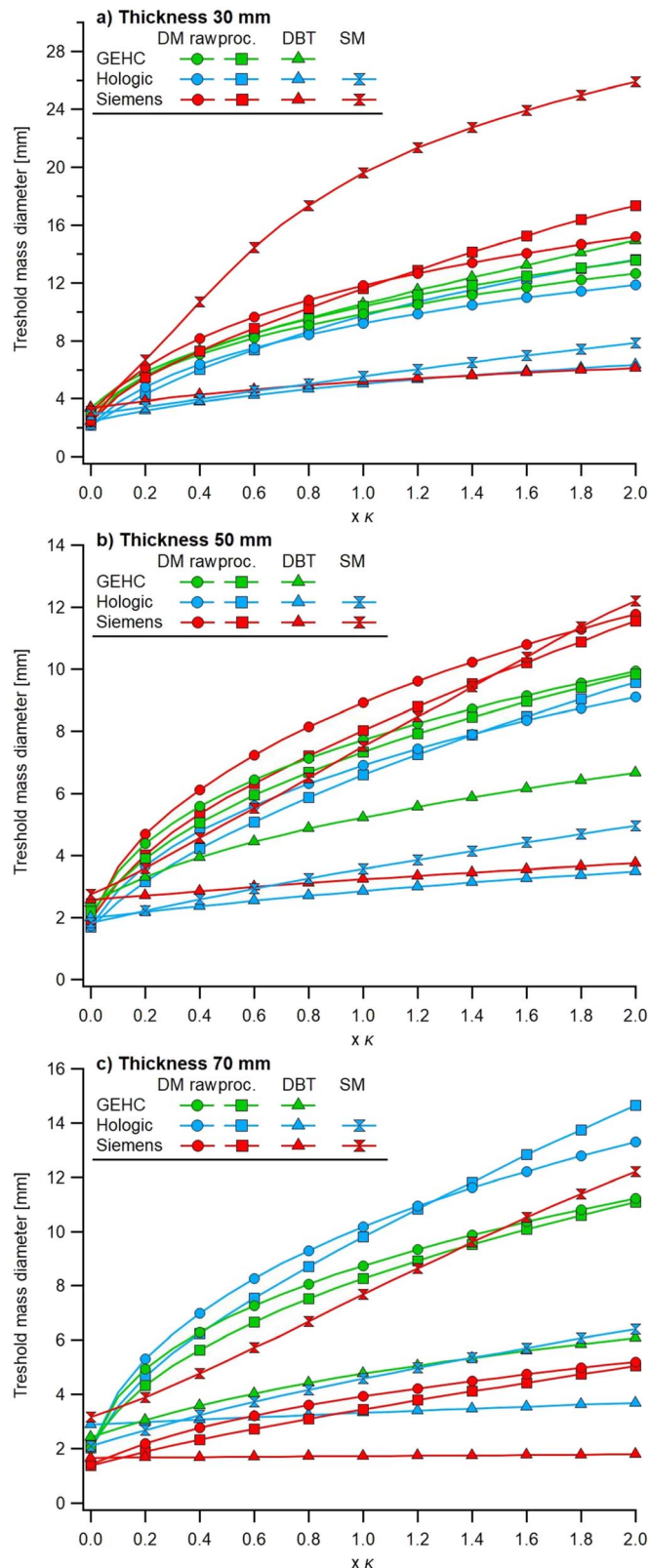
Figure 12. Threshold diameter of spherical masses as a function of κ for three PMMA thicknesses. (a) 30 mm (b) 50 mm (c) 70 mm.

characterized with homogeneous test-objects like the CDMAM phantom, if used with caution. The test-object must remain parallel to the reconstructed planes and obviously the details must remain fully in the plane that is evaluated. The effective thickness of DBT planes measured in our study (between 2.05 and 2.68 mm) was much thicker than the CDMAM discs and insured to fully include the discs in a single DBT plane. A further assumption is that CDCOM is functioning as expected on the image type assessed, and if doubted this must be checked via manual scoring.

It is important to note that current test-objects dedicated to measure physical image quality parameters were not designed for the evaluation of SM images. Phantoms with improved anthropomorphic structures might be necessary with for this evaluation. SM images are the result of computational algorithms that have evolved over time and are different between manufacturers. Various methods will be used, involving, for example, processing steps that identify slices and regions containing mass-like and MC-like features and that may adapt to some degree to breast tissue patterns. Adaptive enhancing filters can therefore result in SM images that look quite different and may give different results when applied to test-objects. The 3D-structured L1 phantom used in this study produces background structure whose power spectrum is close to that of breast images (Cockmartin *et al* 2017), but the structure is not anthropomorphic. Thus, the conclusions of this study cannot be generalized beyond the L1 test-object and SM algorithms used in our study. However, the detection results for SM images obtained in this study using the L1 phantom are consistent with those found by Ikejimba *et al* (2017), using a phantom that more closely resembles breast images. Further research using phantoms with improved anthropomorphic structures and objects resembling MCs or masses is needed to confirm the extent to which our results for SM images are representative of patient imaging.

In this work, the VTF parameters were optimized to predict the CDMAM performance for a wide range of discs sizes, imaging modalities and background noise types. VTF parameters were adjusted to improve the agreement between human and NPWE performance in the presence of background noise, as described in other studies (Bouwman *et al* 2016). The resulting VTF used in this work covered a wide range of spatial frequencies, between 0 and 10 mm^{-1}. The predictive value of the CDMAM results was established against human observers in the presence of anatomical background noise. Human reader performance was generally in good agreement with CDCOM scoring and with the NPWE model, with better agreement for plain images, and slightly lower agreement for the images with anatomical background simulating structure. Only the SM images gave inconsistent results for some disc diameters. The NPWE model gave reasonably close agreement with human readings (averaged ratio 0.992 and 1.070 for images without and with structured anatomical background, respectively). This is consistent with earlier studies comparing contrast-detail detectability in plain images (Marshall 2006, Segui and Zhao 2006, Monnin *et al* 2011). The extensive application of the NPWE model in this work shows that this model can predict target detectability over a wide range of imaging conditions. Further research is however necessary to validate the conditions for which the NPWE model can be used to match the CDMAM results in the presence of low-frequency, structured background noise.

5. Conclusion

This study compared the quantitative image quality metrics and detectability of DM, DBT and SM images based on the measurement of image quality indices (MTF, NPS and NEQ) and system efficiency (system DQE) for the situation of cluttered backgrounds. Used in the NPWE observer model, these metrics correctly predicted the detectability of the thin CDMAM discs for a wide range of beam energies, scatter fractions, image processing and reconstruction algorithms of three mammography device vendors. Furthermore, these predictions held for images with and without low-frequency structured background noise. The quantitative data on threshold diameters are potentially valuable technical performance indicators and it may be possible to explain or predict these results from the different components making up the NPWE model. The analysis method used in this work is proposed as a means of including the influence of anatomical structured noise on image quality and detectability in mammography, for a range of imaging tasks and imaging conditions.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary information files).

ORCID iDs

N W Marshall ⓘ https://orcid.org/0000-0001-5549-5133

References

Barten P G J 1990 Evaluation of subjective image quality with the square-root integral method *J. Opt. Soc. Am.* **7** 2024–31
Båth M, Håkansson M, Börjesson S, Kheddache S, Grahn A, Bochud F O, Verdun F R and Månsson L G 2005a Nodule detection in digital chest radiography: part of image background acting as pure noise *Radiat. Prot. Dosim.* **114** 102–8

Båth M, Håkansson M, Tingberg A and Månsson L G 2005b Method of simulating dose reduction for digital radiographic systems *Radiat. Prot. Dosim.* **114** 253–9

Bateman H and Erdélyi A 1954 *Tables of Integral Transforms, Vol. 1 and Vol. 2* ed A Erdélyi (New York: McGraw-Hill) (https://doi.org/10.2307/3608613)

Bochud F O, Verdun F R, Hessler C and Valley J-F 1995 Detectability of radiological images: the influence of anatomical noise *Proc. SPIE 2436, Medical Imaging 1995: Image Perception* (https://doi.org/10.1117/12.206845)

Bouwman R W, van Engen R E, Broeders M J M, den Heeten G J, Dance D R, Young K C and Veldkamp W J H 2016 Can the non-prewhitening model observer, including aspects of the human visual system, predict human observer performance in mammography? *Phys. Med.* **44** 1559–69

Burgess A E 1994 Statistically defined backgrounds: performance of a modified nonprewhitening observer model *J. Opt. Soc. Am.* A **11** 1237–42

Chen L, Abbey C K, Nosrateih A, Lindfors K K and Boone J M 2012 Anatomical complexity in breast parenchyma and its implications for optimal breast imaging strategies *Med. Phys.* **39** 1435–41

Cockmartin L *et al* 2017 Design and application of a structured phantom for detection performance comparison between breast tomosynthesis and digital mammography *Phys. Med. Biol.* **62** 758–80

Cockmartin L C, Bosmans H and Marshall N W 2013 Comparative power law analysis of structured breast phantom and patient images in digital mammography and breast tomosynthesis *Med. Phys.* **40** 081920

Cunningham I A 2000 Applied linear-systems theory *Medical Imaging Volume 1. Physics and Psychophysics.* ed J Beutel, H L Kundel and R L Van Metter (Bellingham, WA: SPIE) pp 79–156

Dance D R, Skinner C L, Young K C, Beckett J R and Kotre C J 2000 Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol *Phys. Med. Biol.* **45** 3225–40

Durand M 2018 Synthesized Mammography: clinical evidence, appearance, and implementation *Diagnostics* **8** 22

Engstrom E, Reiser I and Nishikawa R 2009 Comparison of power spectra for tomosynthesis projections and reconstructed images *Med. Phys.* **36** 1753–8

Gang G J, Tward D J, Lee J and Siewerdsen J H 2010 Anatomical background and generalized detectability in tomosynthesis and cone-beam CT *Med. Phys.* **37** 1948–65

Gang G J, Lee J, Stayman J W, Tward D J, Zbijewski W, Prince J L and Siewerdsen J H 2011 Analysis of Fourier-domain task-based detectability index in tomosynthesis and cone-beam CT in relation to human observer performance *Med. Phys.* **38** 1754–68

Gang G J, Zbijewski W, Webster Stayman J and Siewerdsen J H 2012 Cascaded system analysis of noise and detectability in dual-energy cone-beam CT *Med. Phys.* **39** 5145–56

Glick S J and Ikejimba L C 2018 Advances in digital and physical anthropomorphic breast phantoms for x-ray imaging *Med. Phys.* **45** e870–85

Hadjipanteli A, Elangovan P, Mackenzie A, Looney P T, Wells K, Dance D R and Young K C 2017 The effect of system geometry and dose on the threshold detectable calcification diameter in 2D-mammography and digital breast tomosynthesis *Phys. Med. Biol.* **62** 858–77

Hadjipanteli A, Elangovan P, Mackenzie A, Wells K, Dance D R and Young K C 2019 The threshold detectable mass diameter for 2D-mammography and digital breast tomosynthesis *Phys. Med.* **57** 25–32

Hill M L, Mainprize J G, Carton A-K, Muller S, Ebrahimi M, Jong R A, Dromain C and Yaffe M J 2013 Anatomical noise in contrast-enhanced digital mammography: I. Single-energy imaging *Med. Phys.* **40** 051910

Hu Y-H, Zhao B and Zhao W 2008 Image artifacts in digital breast tomosynthesis: investigation of the effects of system geometry and reconstruction parameters using a linear system approach *Med. Phys.* **35** 5242

Hu Y-H and Zhao W 2011 The effect of angular dose distribution on the detection of microcalcifications in digital breast tomosynthesis *Med. Phys.* **38** 2455–66

Ikejimba L C, Graff C G, Rosenthal S, Badal A, Ghammraoui B, Lo J Y and Glick S J 2017 A novel physical anthropomorphic breast phantom for 2D and 3D x-ray imaging *Med. Phys.* **44** 407–16

Ikejimba L C, Salad J, Graff C G, Goodsitt M, Chan H-P, Huang H, Zhao W, Ghammraoui B, Lo J Y and Glick S J 2021 Assessment of task-based performance from five clinical DBT systems using an anthropomorphic breast phantom *Med. Phys.* **48** 1026–38

Kiarashi N, Nolte A C, Sturgeon G M, Segars W P, Ghate S V, Nolte L W, Samei E and Lo J Y 2015 Development of realistic physical breast phantoms matched to virtual breast phantoms based on human subject data *Med. Phys.* **42** 4116–26

Li B, Avinash G B, Uppaluri R, Eberhard J W and Claus B E H 2007 Optimization of slice sensitivity profile for radiographic tomosynthesis *Med. Phys.* **34** 2907–16

Mackenzie A, Kaur S, Thomson E L, Mitchell M, Elangovan P, Warren L M, Dance D R and Young K C 2021 Effect of glandularity on the detection of simulated cancers in planar, tomosynthesis, and synthetic 2D imaging of the breast using a hybrid virtual clinical trial *Med. Phys.* **48** 6859–68

Mainprize J G, Tyson A H and Yaffe M J 2012 The relationship between anatomic noise and volumetric breast density for digital mammography **39** 4660–8

Mainprize J G and Yaffe M J 2010 Cascaded analysis of signal and noise propagation through a heterogeneous breast model *Med. Phys.* **37** 5243–50

Marshall N W 2006 A comparison between objective and subjective image quality measurements for a full field digital mammography system *Phys. Med. Biol.* **51** 2441–63

Marshall N W, Monnin P, Bosmans H, Bochud F O and Verdun F R 2011 Image quality assessment in digital mammography: I. Technical characterization of the systems *Phys. Med. Biol.* **56** 4201–20

Marshall N W and Bosmans H 2012 Measurement of system sharpness for two digital breast tomosynthesis systems *Phys. Med. Biol.* **57** 7629–50

Marshall N W and Bosmans H 2022 Performance evaluation of digital breast tomosynthesis systems: physical methods and experimental data *Phys. Med. Biol.* **67** 22TR03

Metheany K G, Abbey C K, Packard N and Boone J M 2008 Characterizing anatomical variability in breast CT images *Med. Phys.* **35** 4685–94

Metz C E and Doi K 1979 Transfer function analysis of radiographic imaging systems *Phys. Med. Biol.* **24** 1079–106

Monnin P, Marshall N W, Bosmans H, Bochud F O and Verdun F R 2011 Image quality assessment in digital mammography: II. NPWE as a validated alternative for contrast detail analysis *Phys. Med. Biol.* **56** 4221–38

Monnin P, Bosmans H, Verdun F R and Marshall N W 2016 A comprehensive model for quantum noise characterization in digital mammography *Phys. Med. Biol.* **61** 2083–108

Monnin P, Verdun F R, Bosmans H, Rodriguez Pérez S and Marshall N W 2017 A comprehensive model for x-ray projection imaging system efficiency and image quality characterization in the presence of scattered radiation *Phys. Med. Biol.* **62** 5691–722

Monnin P, Verdun F R, Bosmans H and Marshall N W 2020 In-plane image quality and NPWE detectability index in digital tomosynthesis *Phys. Med. Biol.* **65** 095013

Prakash P, Zbijewski W, Gang G J, Ding Y, Stayman J W, Yorkston J, Carrino J A and Siewerdsen J H 2011 Task-based modeling and optimization of a cone-beam CT scanner for musculoskeletal imaging *Med. Phys.* **38** 5612–29

Reiser I and Nishikawa R M 2010 Task-based assessment of breast tomosynthesis: effect of acquisition parameters and quantum noise *Med. Phys.* **37** 1591–600

Ruschin M *et al* 2007 Dose dependence of mass and microcalcification detection in digital mammography: free response human observer studies *Med. Phys.* **34** 400–7

Samei E, Flynn M J and Reimann D A 1998 A method for measuring the presampled MTF in digital radiographic systems using an edge test device *Med. Phys.* **25** 102–13

Sechopoulos I 2013a A review of breast tomosynthesis: I. The image acquisition process *Med. Phys.* **40** 014301–1

Sechopoulos I 2013b A review of breast tomosynthesis: II. Image reconstruction, processing and analysis, and advanced applications *Med. Phys.* **40** 1–17

Segui J A and Zhao W 2006 Amorphous selenium flat panel detectors for digital mammography: validation of a NPWE model observer with CDMAM observer performance experiments *Med. Phys.* **33** 3711–22

Siewerdsen J H, Cunningham I A and Jaffray D A 2002 A framework for noise-power spectrum analysis of multidimensional images *Med. Phys.* **29** 2655–71

Tanguay J, Lalonde R, Bjarnason T A and Yang C-Y J 2019 Cascaded systems analysis of anatomic noise in digital mammography and dual-energy digital mammography *Phys. Med. Biol.* **35** 5511–29

Tward D J and Siewerdsen J H 2008 Cascaded systems analysis of the 3D noise transfer characteristics of flat-panel cone-beam CT *Med. Phys.* **35** 5511–29

van Engen R E 2013 *Digital Mammography Update. European Protocol for the Quality Control of the Physical and Technical Aspects of Mammography Screening* (Luxembourg: European Commission, Office for Official Publications of the European Union)

Vancoillie L, Cockmartin L, Marshall N and Bosmans H 2021 The impact on lesion detection via a multi-vendor study: a phantom-based comparison of digital mammography, digital breast tomosynthesis, and synthetic mammography *Med. Phys.* **48** 1–23

Yoon S, Gang J G, Tward D J, Siewerdsen J H and Fahrig R 2009 Analysis of lung nodule detectability and anatomical clutter in tomosythesis imaging of the chest *Proc. SPIE* **7258** 72581M

Young K C, Cook J J H, Oduko J M and Bosmans H 2006 Comparison of software and human observers in reading images of the CDMAM test object to assess digital mammography systems *Proc. SPIE* **6142** 39–51

Zhao B, Zhou J, Hu Y H, Mertelmeier T, Ludwig J and Zhao W 2009 Experimental validation of a three-dimensional linear system model for breast tomosynthesis *Med. Phys.* **36** 240–51

Zhao B and Zhao W 2008 Three-dimensional linear system analysis for breast tomosynthesis *Med. Phys.* **35** 5219–32

Zhao C, Vassiljev N, Konstantinidis A C, Speller R D and Kanicki J 2017 Three-dimensional cascaded system analysis of a 50 μm pixel pitch wafer-scale CMOS active pixel sensor x-ray detector for digital breast tomosynthesis *Phys. Med. Biol.* **62** 1994–2017