Can different Catphan phantoms be used in a multi-centre audit of radiotherapy CT image quality?

Anne T. Davis[a,b]*, Antony L. Palmer[a,b], Andrew Nisbet[a]

[a.] Department of Medical Physics and Biomedical Engineering, University College London, UK
[b.] Department of Medical Physics, Portsmouth Hospitals NHS Trust, Portsmouth, UK

*Corresponding author.
Department of Medical Physics, Portsmouth Hospitals NHS Trust, Portsmouth, PO6 3LY, UK
E-mail address: Anne.Davis2@porthosp.nhs.uk (Anne Davis)

Abstract

Purpose

To determine the variation between Catphan image quality CT phantoms, specifically for use in a future multi-centre image quality audit.

Method

14 Catphan phantoms (models 503, 504 and 604) were scanned on a Canon Aquilion Prime CT scanner using a single scan protocol. Measurements were made of noise in the uniformity section, visibility of low contrast targets and contrast, x-ray attenuation and CT number for 5 materials in the sensitometry section. Scans were also acquired using one phantom and varying reconstruction field of view, image slice thickness, effective tube-current-time product and iterative reconstruction settings to determine how the degree of inter-phantom variability compared with the magnitude of changes from scan parameter alteration.

Results

Across all phantoms the mean CT value in the uniformity section was 7.0 (SD 0.9) range: 4.9 - 8.1 HU. For the different materials the CT numbers were air: -1004 $\pm$5, Polymethylpentene: -190 $\pm$2, Polystyrene: -42 $\pm$2, Delrin: 321 $\pm$5 and Teflon: 898 $\pm$8 HU. Consistency of low contrast targets through visual scoring was good. Measured contrast was lower (p <0.001) with more variability for 504 versus 604 models. All phantoms produced identical tube current settings with x-ray tube current modulation, indicating no x-ray attenuation differences. The degree of change in image quality metrics between phantoms was small compared with results when scan parameters were varied.

Conclusion

Catphan phantoms model 604 showed minimal differences and will be used for multi-centre inter-comparison work, with the consistency between phantoms appropriate for measuring possible variations in image quality.

Introduction

The Catphan phantom (The Phantom Laboratory, Salem, NY, USA) is a test tool designed for the measurement of image quality metrics on a CT scanner [1-3]. It can be used to monitor the constancy of CT scanner performance over time, performance against specification or to assess the impact of changes in clinical scan protocols on image quality [4]. It is widely available in radiotherapy departments as it is often supplied with linear accelerators for quality control testing of on-board cone-beam imaging systems. This study undertook foundational work to support the future use of Catphan phantoms as tools to evaluate radiotherapy planning CT scan protocols for adult imaging. There was a particular emphasis on tests which assess low contrast detail visibility. The work investigated the differences between a collection of Catphan phantoms so as to establish whether the phantoms were interchangeable and could be used in a future multicentre image quality audit, and more generally for comparisons between departments. The degree of difference in measured image quality metrics when using various phantoms was also compared against the magnitude of variation in results arising from changes in CT scan protocol parameters. If the variation of results arising from different phantoms was much smaller than the variation from changing scan parameters then the use of multiple Catphans would be feasible.

It is important to establish the degree of any difference between phantoms in measuring image quality. Historically, in other imaging modalities, inconsistent manufacturing of image quality phantoms has made image quality inter-comparison difficult when using multiple phantoms of the same type [5, 6]. This work investigates, for the Catphan phantom, the basic premise that any groups of test tools used must be consistent if such a multi-centre audit is to be meaningful.

Materials and Methods

The Catphan phantom is modular in construction with each module designed for a specific set of tests. The model number of the phantom defines the type of modules within the phantom. 14 phantoms were included in this study of which two were model 503, six were model 504 and six were model 604. Table 1 shows the modules within the various models of phantom [1-3]. The parts of the phantoms intended for high resolution and geometry tests were not assessed as part of the study. Model 604 is an updated version of model 504. The differences, based on specification information from the manufacturer, between the 503/504 phantoms and 604 phantoms are given in Table 2. The parameters to be measured were: X-ray attenuation (based on indicated CT dose index); low contrast levels based on measurement and visibility of targets; uniformity of CT number and Hounsfield unit (HU) values for the materials in the sensitometry section.

Table 1. The module within the various phantoms

	Catphan models and module reference numbers		
Module type	Model 503	Model 504	Model 604
Low contrast resolution	✗	✓CTP515	✓CTP730
Uniformity	✓CTP 486	✓CTP 486	✓CTP729
Sensitometry and geometry	✓CTP 404	✓CTP 404	✓ CTP732
High resolution	✓CTP 528	✓CTP 528	✓ CTP732

Table 2. Specification of the model 503, 504 and 604 phantoms

Parameter	Model 503/504	Model 604	Tolerances
Low contrast targets	For model 504 only: 3 groups: nominal 1%, 0.5% and 0.3% contrast. 9 targets in each group with diameters between 15 and 2 mm. All 40 mm in length.	3 groups: nominal 1%, 0.5% and 0.3% contrast. 9 targets in each group with diameters between 15 and 2 mm. All 40 mm in length.	None given
	3 groups: sub-slice lengths of 7 mm, 5 mm and 3mm. 4 targets in each group with diameters between 9 and 3 mm.	Not present.	None given
Uniformity	Typical values of HU between 5 and 18 HU. 15 cm diameter section	Typical values of HU between 5 and 18 HU. 20 cm diameter section	Within 2% (or 20 HU) of water. i.e. 0$\pm$20 HU
Sensitometry	7 targets: Air Polymethylpentene Low density polyethylene Polystyrene Acrylic - Delrin - Teflon	9 targets: Air Polymethylpentene Low density polyethylene Polystyrene Acrylic Bone 20% Delrin Bone 50% Teflon	Typical HU values: -1046 to -986 -220 to -172 -121 to -87 -65 to -29 92 to 137 211 to 263 344 to 387 667 to 783 941 to 1060
High resolution: Line pair grating MTF bead MTF wire	1 to 21 lp/cm 0.28 mm -	1 to 15 lp/cm 0.18 mm 0.05 mm	None given

Comparison of phantoms

The phantoms were scanned on a Canon Aquilion Prime CT scanner and all scans were acquired in one session with the aim of ensuring consistent scanner performance. The scanner was subject to daily quality control testing to confirm constancy. The scan parameters used were: helical scanning, 120 kV, 0.5 x 40 collimation, 0.625 pitch, 220 mm diameter reconstruction and acquisition field of view (FOV), FC21 reconstruction kernel, and image slice width and slice interval were set at 2 mm. The image matrix was 512 by 512 pixels. Iterative reconstruction (IR) was switched off. The 120 kV setting was chosen since that is the primary kilovoltage used in UK radiotherapy departments [7]. The phantom was aligned so that it was centred in the scanner aperture and the whole phantom

length scanned. Initially 10 scans were acquired for one phantom. Once scanning consistency was established 3 identical scans were acquired for all the other phantoms.

(i) Hounsfield unit uniformity

The tube rotation time was set to 1 s, the tube current fixed at 400 mA and the scans acquired. Macros developed in ImageJ (National Institute of Health, USA) were used to analyse the images produced for this test and those described in sections (ii) and (iii). ImageJ is available open-source. A region of interest (ROI) of approximately half the phantom diameter was placed on the images of the uniform section of the phantom, see Figure 1(a) [8]. A macro was used to ensure consistency of ROI size and position on all images. Measurements of HU and standard deviation for noise were taken from three adjacent images and from three scans for each phantom. Additional measurements were made using 4 ROIs placed at the edges of the image and one in the centre, see Figure 1(b). The difference in CT number obtained from ROIs at the image centre and edges was calculated [9].

(ii) Low contrast target visibility and measured contrast

Quantitative measurement of contrast was made. Using the same scans acquired for the uniformity measurement, images from the centre of the low contrast target section were selected. ROIs were placed on the 15 mm and 8 mm diameter targets in the 1% contrast group and the 15 mm diameter targets in the 0.5% and 0.3% contrast groups. ROIs of approximately 12 mm diameter were placed on the background material on the image next to the targets as shown in Figure 1(c). The size of the ROI was chosen to fit in the space immediately adjacent to the target but large enough to minimise the impact of pixel value variation arising from image noise. The larger targets were chosen to ensure ROIs could be accurately positioned within the circular target and because variation of CT number due to noise will be greater with smaller ROIs leading to greater uncertainties in the results. After initial measurements, a second method was used, Figure 1(d). Two ROIs were placed on the background, on the outer and inner edges of the low contrast target. The average value of these two background measurements was calculated. Contrast was calculated by subtracting HU of the background from the HU of the low contrast circular target. One of the 604 phantoms had a rotated module so was not included in the measurements. Contrast values arising for the 504 and 604 model phantoms were compared using the student t-test to assess significance.

Visual scoring of the low contrast targets was then performed. Viewing conditions were standardised with low light levels to improve visibility. Brightness and contrast levels were chosen for optimal viewing and set at the same level for all images [10]. The image was displayed to fill the whole screen on a 56 cm viewing monitor. The image from the centre of the scan of the contrast section and the two images immediately adjacent were selected. Each of the nominal 1%, 0.5% and 0.3% contrast groups were scored by counting how many targets could be seen. Half scores were allocated where part of the target was seen. With 9 targets in each group the maximum possible score per group was 9. Three people scored the images, scoring a total of 9 images for each phantom. The order of scoring was randomised.   The consistency of the scoring methodology was checked using repeated scoring.

(a)

(b)

(c)

Nominal 1% contrast

Nominal 0.5% contrast

Nominal 0.3% contrast

(d)

Nominal 1% contrast

Nominal 0.5% contrast
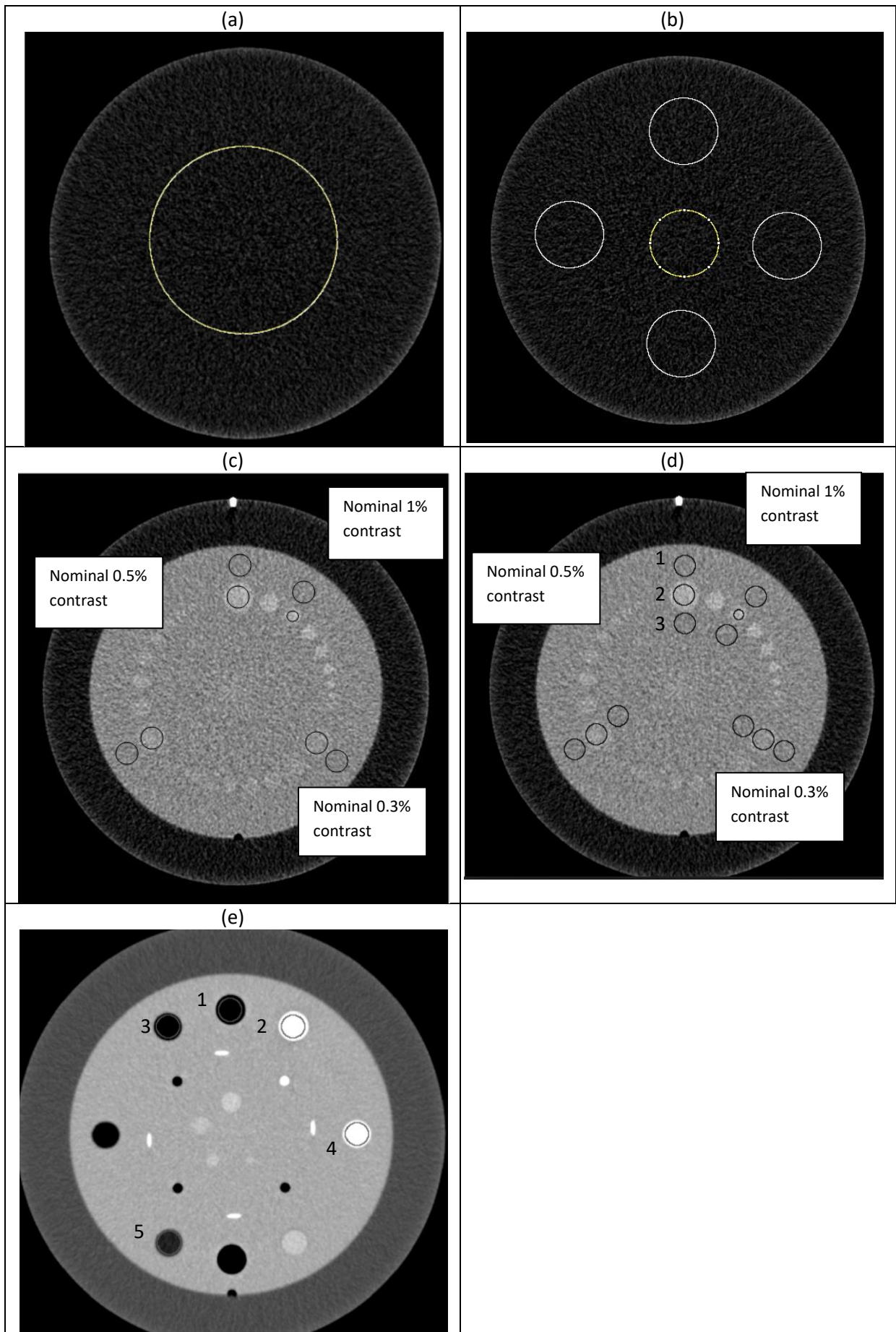
1
2
3

Nominal 0.3% contrast

(e)

1
3
2
4
5

Figure 1. (a) ROI placed on a uniformity image; (b) Whole image uniformity check; (c) Measurement of contrast in a 604 model; (d) Method of measuring contrast using two ROIs to obtain the background HU value in a 604 model;(e) Measurements in sensitometry section of 504 model Catphan.

(iii)   Hounsfield units scale in the sensitometry section

In the sensitometry section there are some differences between the 503/504 models and the 604 models. The 604 phantoms contain additional inserts representing two types of bone material. Using the images from the centre of the sensitometry section, ROIs were placed on the inserts for air, Teflon, Polymethylpentene (PMP), Delrin and Polystyrene, as indicated by numbers 1 to 5 on Figure 1(e). These materials were chosen to cover the full range of CT numbers typically found in human tissue. The Hounsfield unit expected range based on the manuals for the Catphans models is given in Table 4 [1-3]. 3 images from each scan were used, with 3 scans, giving 9 measurements per phantom. Care was taken to ensure the ROIs were centred within the inserts and all ROIs were 25 pixels in diameter.

(iv)   X-ray attenuation of the phantom

The tube current modulation system, Sure Expose, was enabled on the CT scanner with the noise index set to 'standard' [11]. Tube rotation time was 0.75 s. Lateral and anterior-posterior view scout scans were acquired initially as required when using modulation. After the helical scan acquisitions, the images of the low contrast section of the phantom were reviewed and the tube current value from the centre slice noted. This region was chosen since comparability of the low contrast section of the phantoms was deemed particularly important in the future image quality audit work. As, however, the model 503 phantoms did not have a low contrast section, the tube current value from the image at the centre of the uniformity section was also recorded and compared. Additionally, the CT dose index from the whole scan was also recorded. On the Canon Aquilion Prime scanner the CT dose index is an average value from the images acquired over the whole scan length.

Phantom sensitivity to scan parameter changes

These measurements aimed to establish the degree of change in measured image quality metrics caused by changing parameters in the CT scan protocol. A single 604 model phantom was selected. The same CT scanner was used as for the phantom inter-comparison, and one scan parameter varied at a time, as shown in Table 3. The scan parameters changed were (i) effective tube current product (mAs), (ii) image slice thickness, (iii) reconstruction field of view and (iv) IR setting. These parameters are those which are generally varied within scan protocols used for radiotherapy planning CT [12]. Three identical scans were acquired for each set of parameters.

Table 3. Scan protocol settings when assessing the impact of scan parameter changes

Scan parameter	Variation with effective mAs	Variation with reconstruction FOV	Variation with slice thickness	Variation with iterative reconstruction (IR) strength setting
kV	120	120	120	120
Pitch	0.825	0.825	0.825	0.825

Collimation	0.5 x 40	0.5 x 40	0.5 x 40	0.5 x 40
Reconstruction kernel	FC21	FC21	FC21	FC21
Acquisition FOV (mm)	200	550	200	200
Reconstruction FOV (mm)	200	Varied: 200, 300, 400, 550	200	200
IR strength	Off	Off	Off	Varied: Off, mild, standard, strong
Image slice thickness (mm)	2	2	Varied: 1,2,3,5	2
Effective mAs	Varied: 242 364,485, 606	441	441	441

Following the same techniques as mentioned in the method section where phantoms were compared, uniformity, noise, low contrast and low contrast target visibility were measured. Viewing conditions were optimised and standardised as before. When scoring visibility of low contrast targets in the images with the larger reconstructed FOVs, zoom was applied to the image to match the image size of the 200 mm diameter reconstructed FOV. In each scan set, 3 images were processed from each of the uniformity and low contrast resolution modules.

Results

Results from phantom comparison

(i)   Hounsfield unit uniformity.

The measured HU and noise values, as measured by standard deviation of HU within the ROI placed on the images of the uniformity section, were very similar for all phantoms, see Figures 2 and 3. The results of a one-way analysis of variance (ANOVA) statistical test on the results from the three models of phantom gave a p value of 0.93 indicating no significant difference between phantom models since the test of significance difference is for $p < 0.05$. The variation for HU value was very much lower than the 0 ± 20 HU range given in the Catphan user manuals for this parameter, see Table 2. CT numbers varied between 4.9 and 8.1 HU. The mean value was 7.0, SD 0.9 HU. CT number uniformity values, the difference between the centre and edge of the phantom, are shown in Figure 4. All values were less than 1.5 HU and there were no differences in the range of values seen between the phantom models.
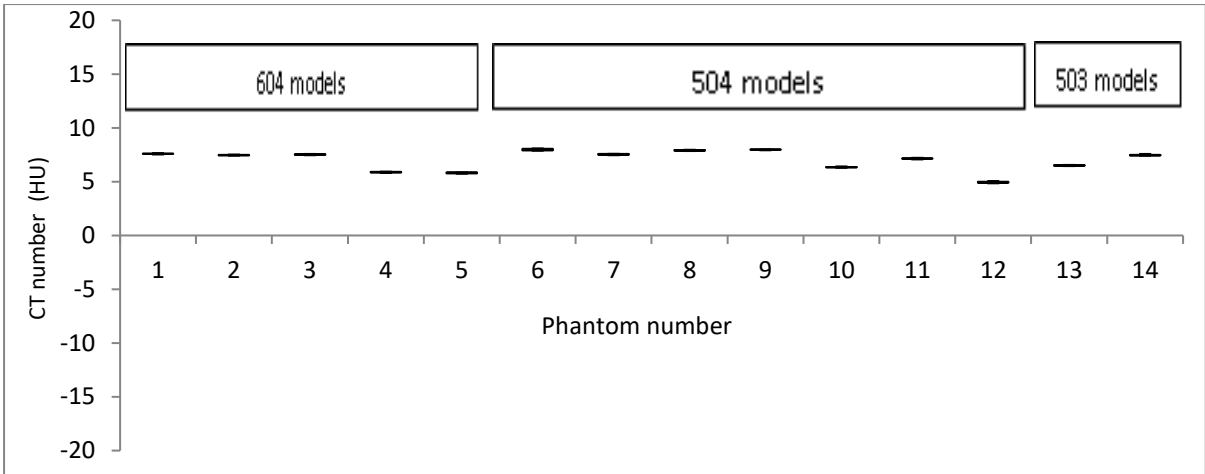
Figure 2. CT number for the uniformity section of different phantoms
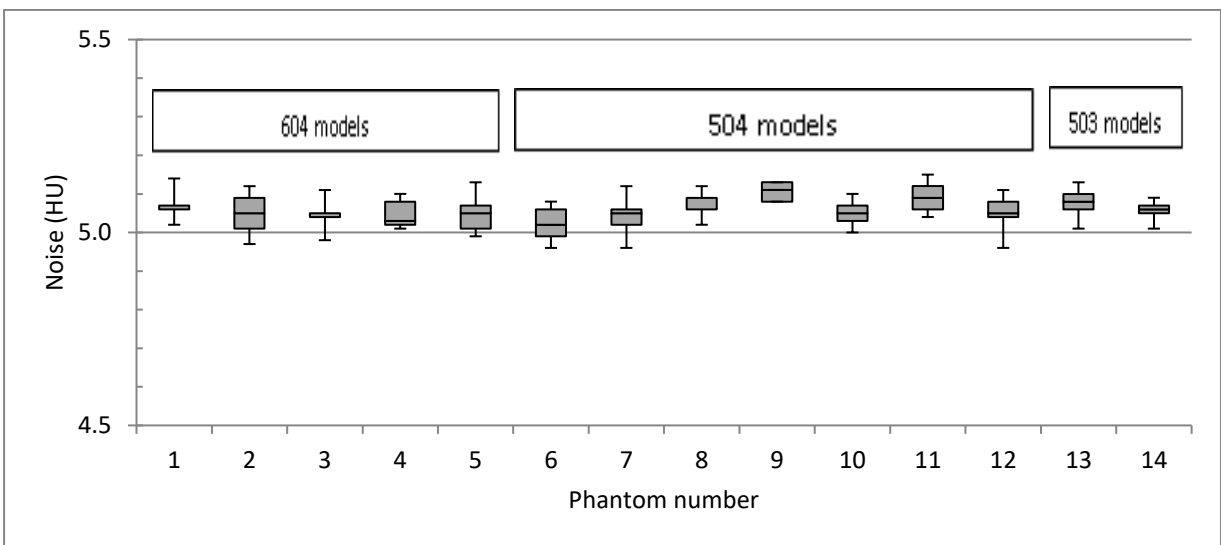

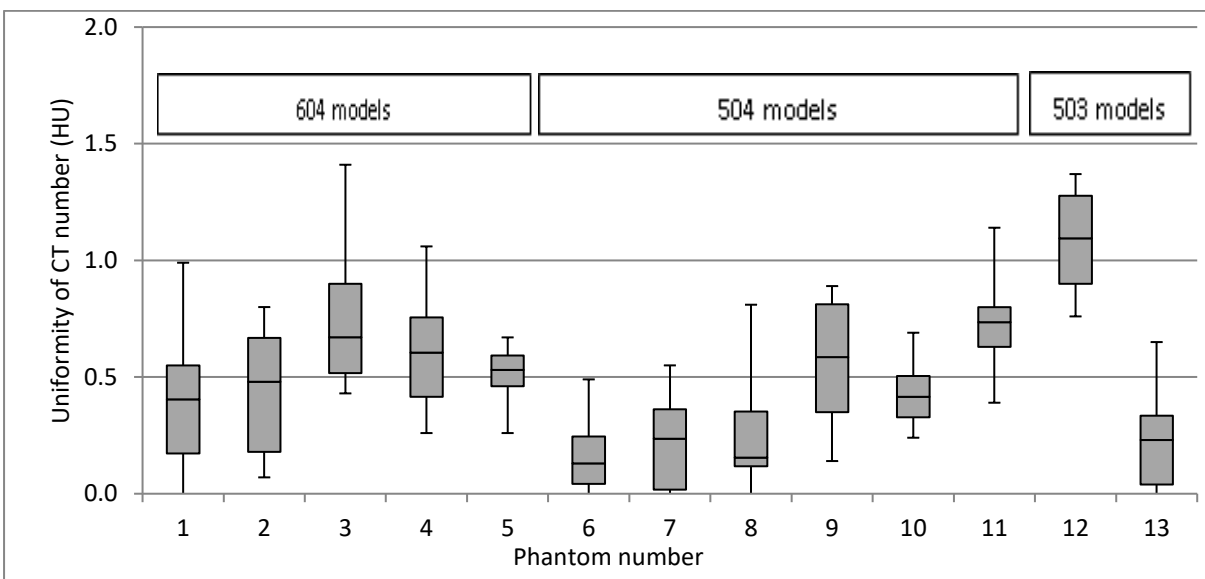Figure 3. Measured noise from the uniformity section of different phantoms

Figure 4. Uniformity for different phantoms, difference between phantom centre and edge

(ii)   Low contrast target visibility and measured contrast

Results for measured contrast (15 mm diameter target, nominal 1%), using the method where 2 ROIs measured the average background HU value, are shown in Figure 5. Results generally were at the expected level since the HU value of the low contrast target needs to be 10 HU to give a contrast of 1%. It can be seen that the mean contrast values for the 504 model phantoms were lower than those for the 604 phantoms. Contrast values were, for model 604: mean 1.12 and SD 0.05, and for model 504: mean 0.99 and SD 0.11, p value <0.001. This difference between the phantom models was also seen for other targets of different size diameters and other nominal contrast values. The 503 model phantoms do not contain a low contrast target module.

A review of the HU values measured on the contrast targets and the background identified that the background ROI HU values closest to the edge of the phantom, position 1 shown in Figure 1 (c), were higher for the 504 model phantoms (typical value 47 HU) than for position 3 in the 504 phantoms and positions 1 and 3 for 604 model phantoms (typical value 45 HU). With HU values of the low contrast targets measured in the region of 56 HU for both phantom models, the 2 HU difference of measured background accounts for the contrast difference between the 604 and 504 model phantoms. The initial use of a single background ROI, and previous to that, smaller ROIs on both the contrast targets and background, produced greater variability of results between all phantoms. This was noticeable on the measured contrast values for the 8 mm diameter target where the smaller ROI meant increased variability due to the influence of noise. This was evident by a greater spread of results for individual phantoms. The results of 10 scans on a single phantom, phantom number '0'on Figure 5, showed a similar range of results as for 3 scans on the same phantom. This indicated that the choice to collect results from 3 scans on each phantom was appropriate.

With visual scoring of low contrast targets this difference in contrast between the phantom models was not evident, see Figure 6.
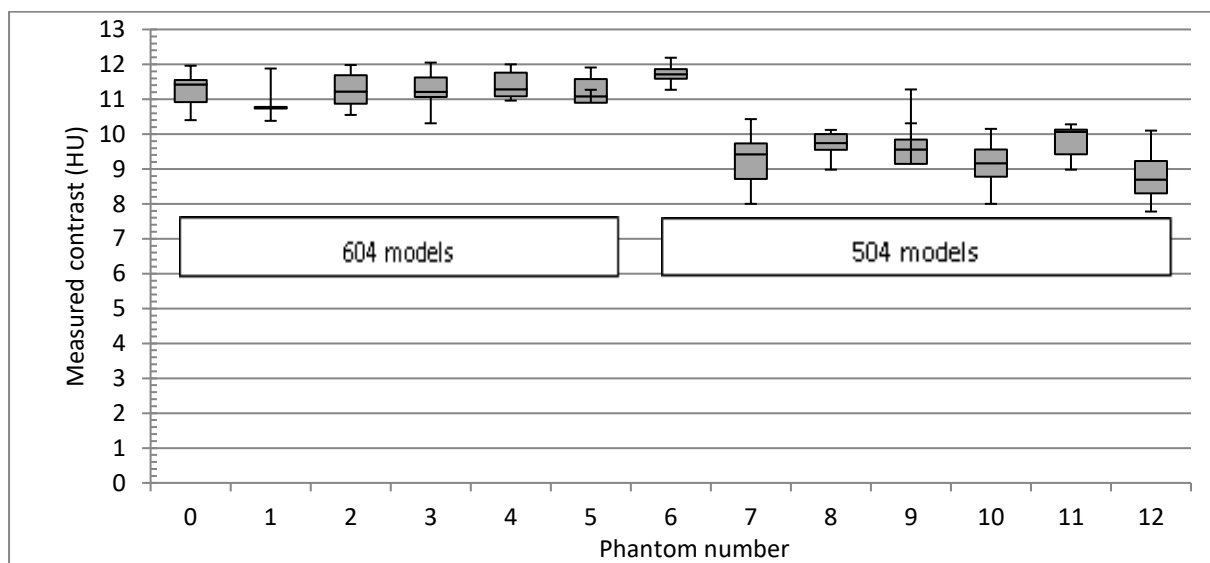

Figure 5. Measured contrast for the 15 mm diameter target of nominal 1% contrast for different phantoms.
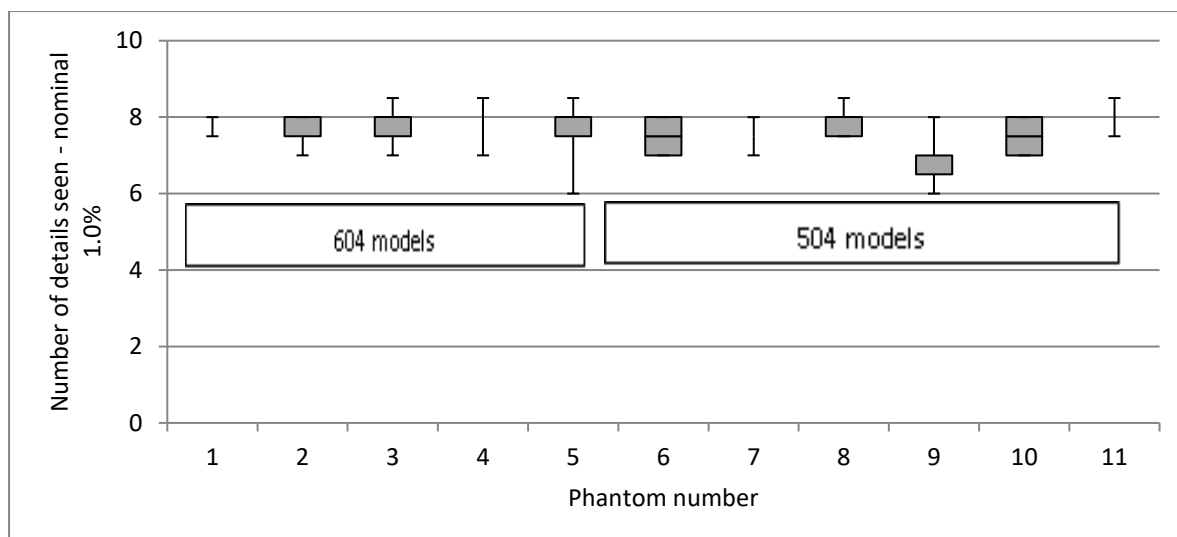
Figure 6. Visual scoring of low contrast targets, nominal 1% contrast group, for different phantoms.

(iii) Hounsfield unit scale

The results from measurements on the various inserts in the sensitometry section from all phantoms are shown in Table 4. The range of CT numbers was up to $\pm$ 8 HU of the average value for the materials with larger measured values for air, Delrin and Teflon. For the PMP and Polystyrene inserts the range of measured values was small at $\pm$ 2 HU. There was no difference noted between the different models of phantom.

Table 4. CT numbers measured for materials in the sensitometry sections of all the phantoms

Material	Typical values quoted by Phantom Laboratory	Average CT number, all images	Range of CT numbers, all images
Air	-1046 to -986	-1004	-1008 to -997
PMP	-220 to -172	-190	-192 to -188
Polystyrene	-65 to -29	-42	-43 to -39
Delrin	344 to 387	321	316 to 326
Teflon	941 to 1060	898	890 to 905

(iv) X-ray attenuation of the phantoms

The tube current values for the images at the centre of the low contrast section were all 180 mA when the tube current modulation system on the scanner was switched on. The tube current values for the uniformity section were also all identical for the different phantoms including the 503 models. The average CTDI dose index resulting from scans of the whole phantom varied from 58.6 to 62.3 mGy across the set of 14 phantoms. These differences can be accounted for by slight unintended differences in scan lengths and a consequential slight increase in tube current where the metal support bracket was erroneously included at the end of the scan on some occasions.

Results from variation of scan parameters

(i) Effective tube-current product setting

When the effective tube-current product (effective mAs) was increased the measured noise in the uniformity section reduced as expected, see Figure 7. The measured contrast did not vary as effective mAs increased, as shown in Figure 8, although the spread of the results reduced slightly as image noise reduced for higher set mAs values. Results for visual scoring of the low contrast targets, see Figure 9, clearly showed the reduced visibility of low contrast targets for lower effective mAs values and that the higher contrast targets (nominal 1% group) were easier to see than the lower contrast targets (nominal 0.3% group).
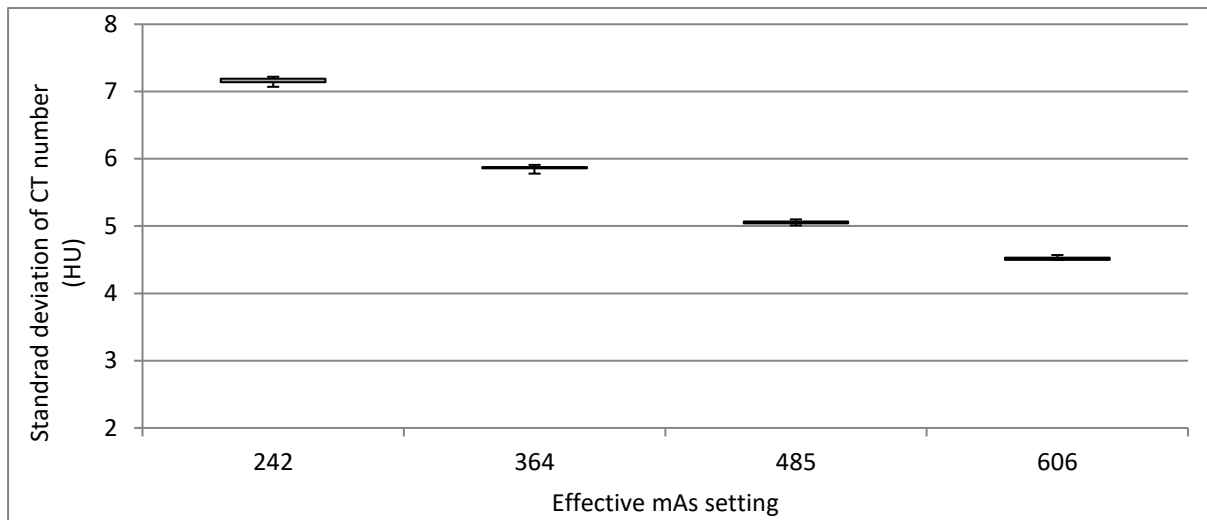
Figure 7. Variation of noise in the uniformity module versus effective mAs
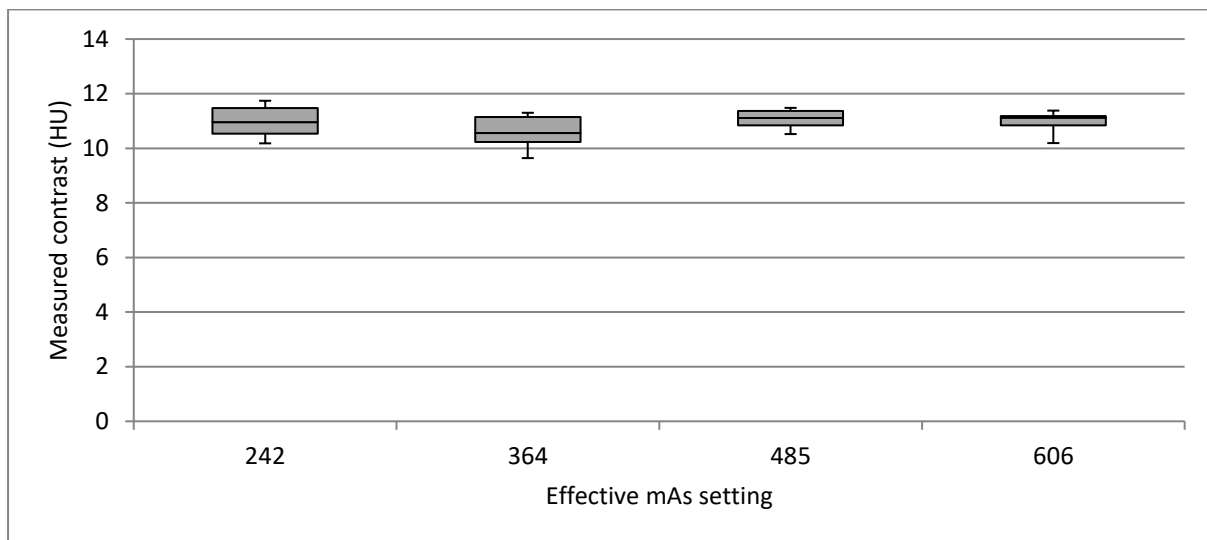
Figure 8. Contrast values for the 15 mm diameter target (nominal 1%) with various effective mAs settings.
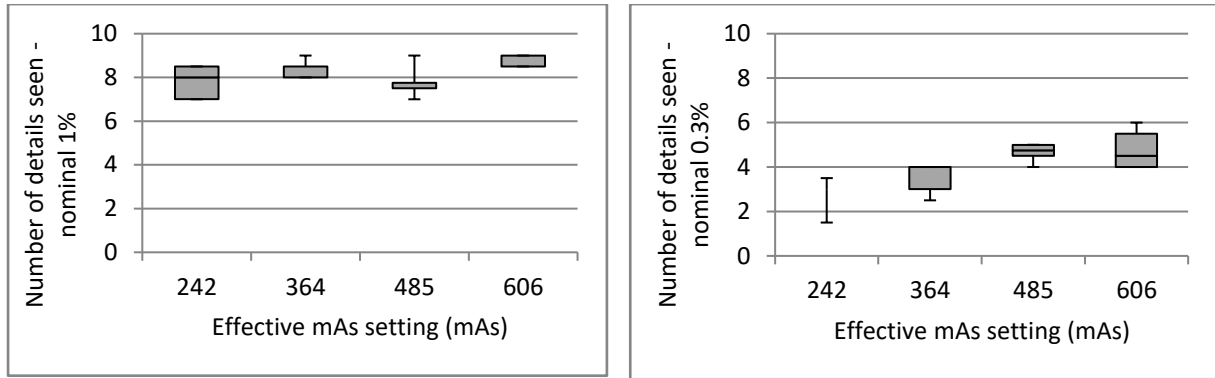
Figure 9. Visibility of low contrast targets with different effective tube-current-time settings with details of (a) nominal 1% contrast and (b) nominal 0.3% contrast

(ii) Image slice thickness

With increased image slice thickness the noise in the image reduced, as shown in Figure 10. Whilst the measured contrast values (nominal 1%) were relatively unchanged for different image slice thickness settings, the reduction in noise for larger slices did result in a reduction in the spread of results, see Figure 11. For visual scoring, however, more of the low contrast targets could be seen with the thicker slices as shown in Figure 12. This is due to the fact that both contrast and noise levels affect target visibility [13]. The lowest contrast group (0.3%) had the lowest results and the widest spread of results. The difference in scores between the 5 mm slice and the 1 mm slice was most significant for the lowest 0.3% contrast target group with approximately 7 targets seen for the 5 mm slice and 3 targets for the 1 mm slice.
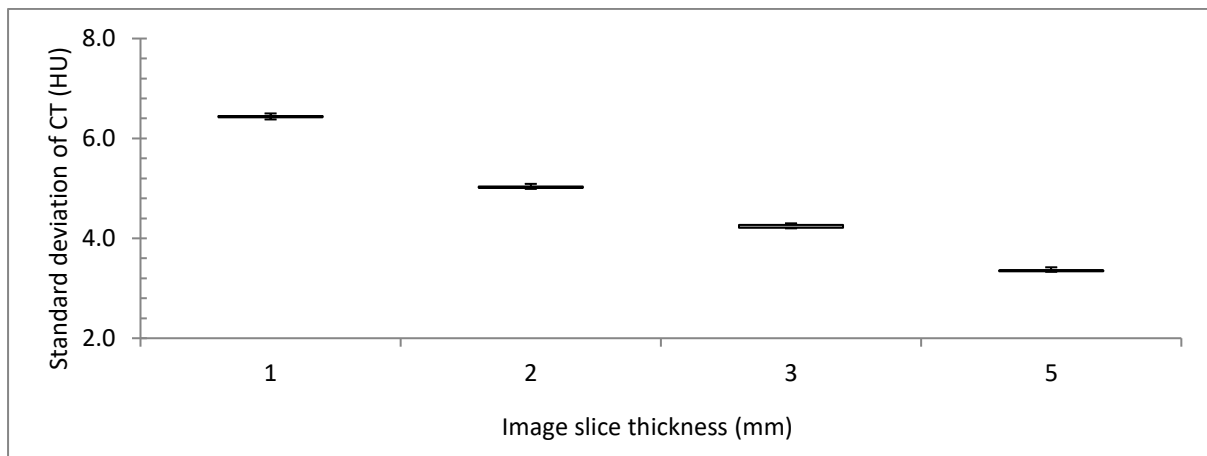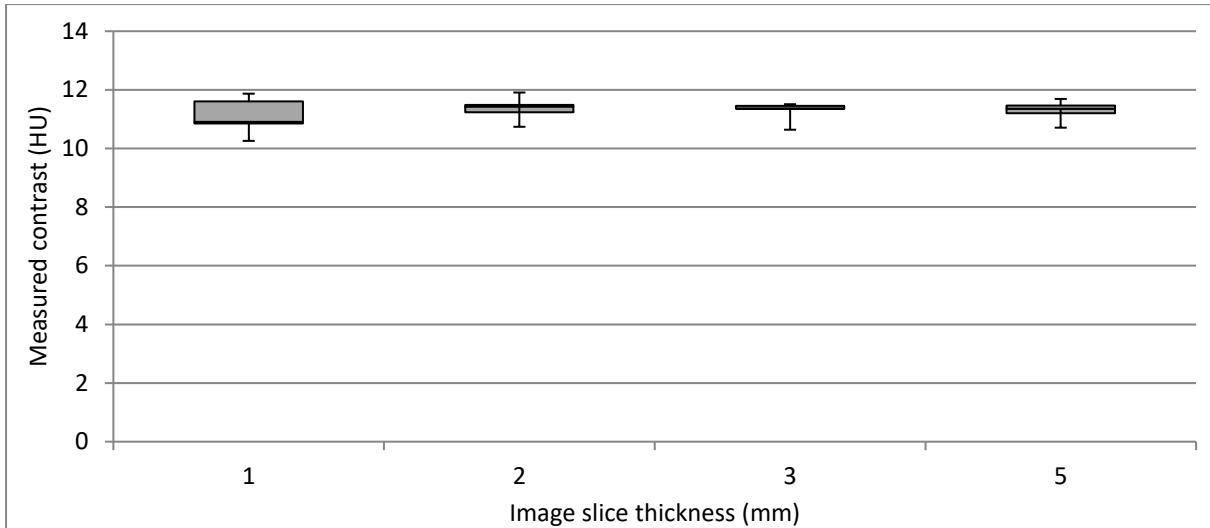

Figure 10. Measured noise versus slice thickness

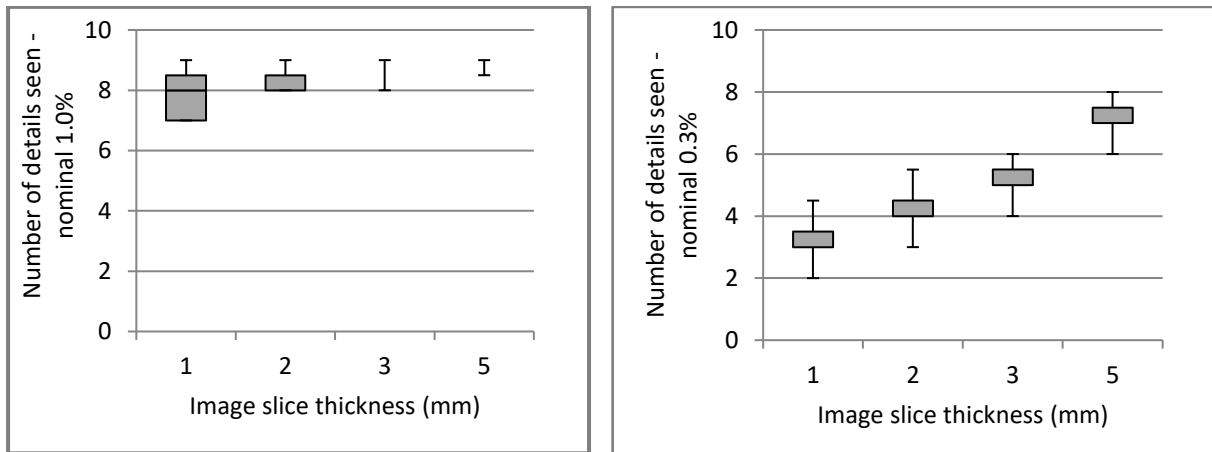Figure 11. Measured contrast for 15 mm target (nominal 1% contrast) versus slice thickness

Figure 12. Score for low contrast target groups with different slice thickness settings with details of (a) nominal 1% contrast and (b) nominal 0.3% contrast

(iii) Reconstruction field of view diameter

Changes in image noise and measured contrast (for nominal 1%, 15 mm target) were minimal when varying reconstruction field of view. The visibility of the low contrast targets did vary with the smaller, lower contrast targets being less visible with the largest FOVs, see Figure 13. This is in line with research findings by other authors which have identified that to maintain visibility by the human observer, contrast must increase as target size reduces [10]. The larger FOVs effectively reduce the size of targets within the image which makes the smallest targets within the 0.5% and 0.3% contrast groups invisible, reducing the overall scores.
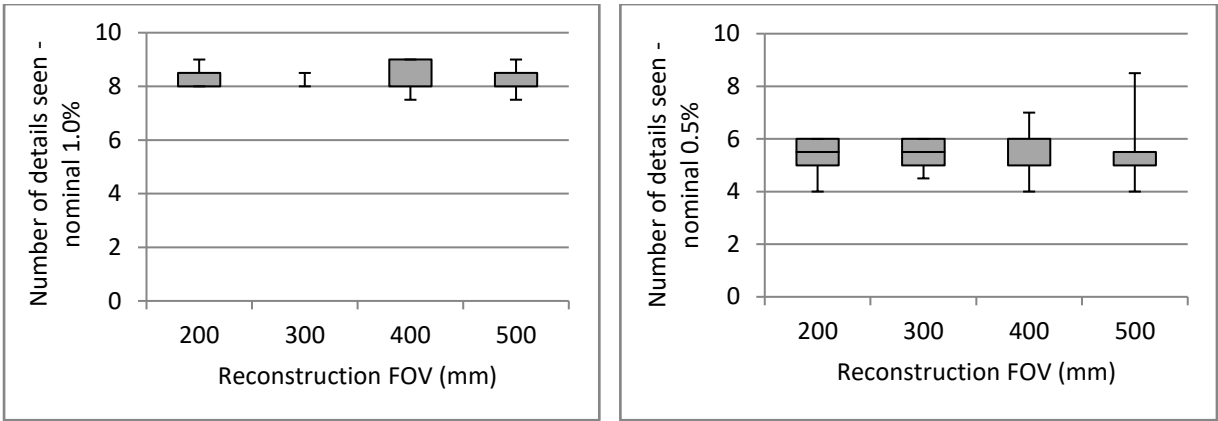
Figure 13. Score for low contrast target groups with different reconstruction FOV settings with details of (a) nominal 1% contrast and (b) nominal 0.3% contrast

(iv) Iterative reconstruction setting

Figure 14 shows that measured noise reduced as the IR setting is increased, with the greatest difference seen between the 'mild' setting and AIDR switched off. The measured contrast, for the 15 mm nominal contrast 1% target circle, did not change with IR setting, see Figure 15. The visual scoring of low contrast targets, see Figure 16, showed that the visibility of targets did improve slightly with increased strength of IR, especially for the lowest contrast (nominal 0.3%) target group where image noise is particularly detrimental to target visualisation [10, 13, 14].
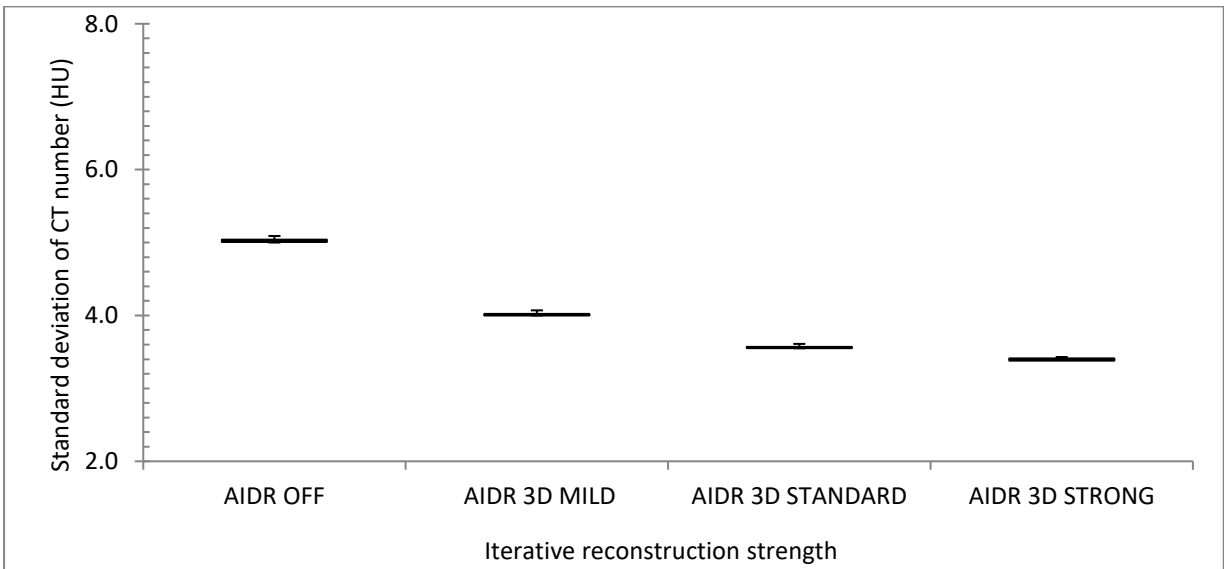

Figure 14. Variation of noise in the uniformity module from variation of IR strength
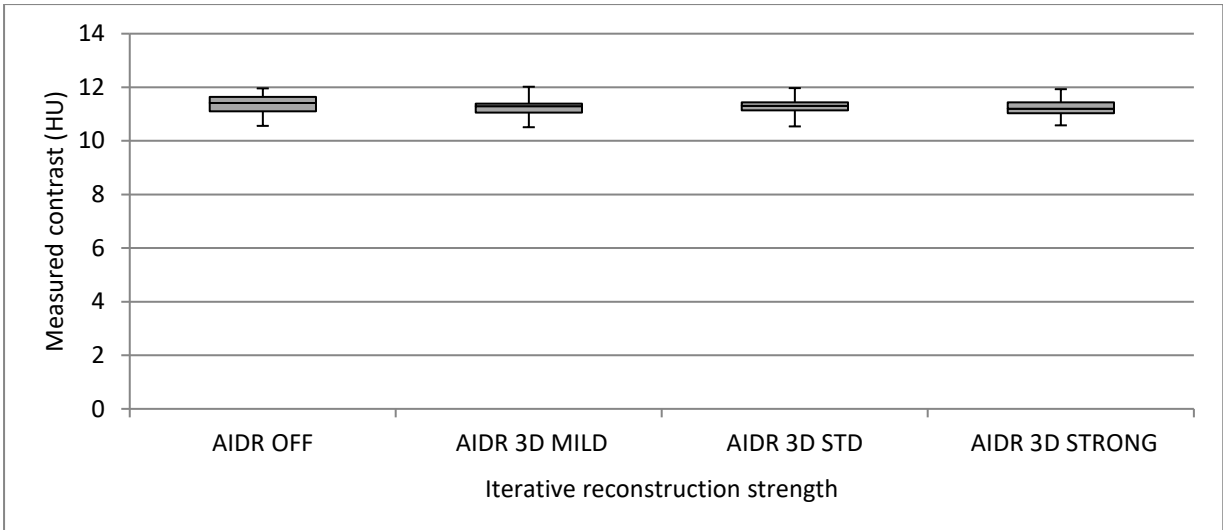
Figure 15. Measured contrast in the low contrast target module with varying IR strength
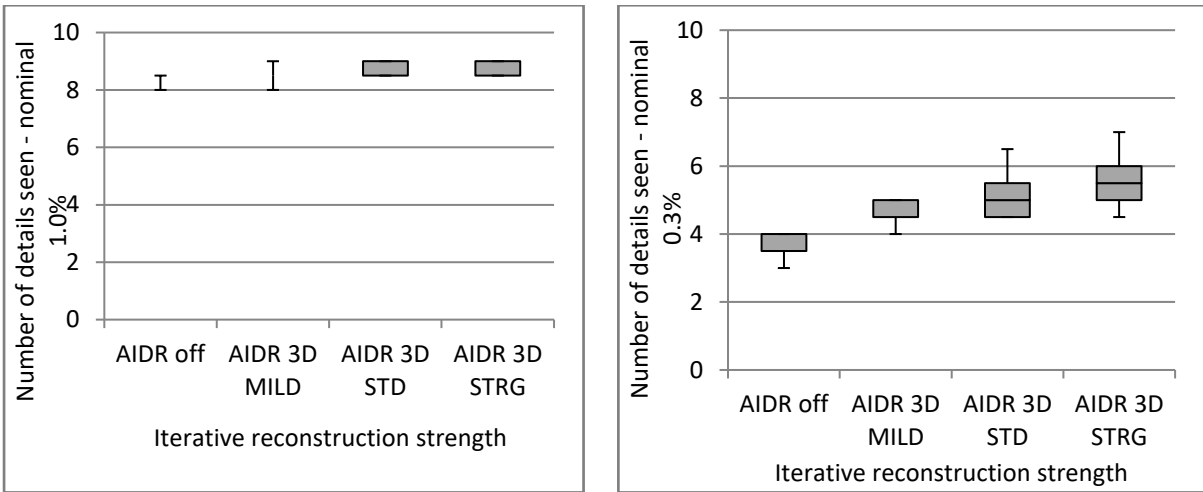

Figure 16. Number of targets visualised in the low contrast target module with varying IR strength with details of (a) nominal 1% contrast and (b) nominal 0.3% contrast

Discussion

The uniformity sections for all the phantoms were shown to be well matched in terms of measured CT number and noise. HU values, at 4.9 to 8. 1, were well within the 0 ± 20 HU quoted in the phantom technical specifications[1-3]. The variation of measured noise for different phantoms was $\pm$ 0.1 HU, which is small when compared to the changes of 1 or more for measured noise when varying scan parameters as shown when comparing Figures 3, 7, 10 and 14. This indicates that all phantoms could be used interchangeably when comparing the effects of scan parameter changes on noise.

Low contrast for the nominal 1% contrast targets, as measured with placement of ROIs, did show differences between phantoms. In general the 604 model phantoms had contrast values in the region of 1.1% compared to 1.0% for the 504 models. This was due to higher background HU values measured in the centre of the 504 model phantoms. The Phantom Laboratory were contacted and asked for comment. Their suggestion was that the different construction method of the two

phantom models, in particular the differences in the outer casing that holds the modules, contribute differently to X-ray beam hardening which affects the CT number of the material. The modules in the 604 models also do not have the rows of inner contrast targets which is present in 503 models. All CT numbers are in part determined by the materials around the point of measurement so HU in the centre is likely to be affected by the presence or absence of the inner groups of targets [14]. The Phantom Laboratory also uses a technique which places an annulus shaped ROI around a low contrast target to measure average background HU for contrast measurement. That will improve the averaging of HU values for that measurement. Additionally, it was clear from the early results within this work that 2 large background ROIs on the inner and outer edges of the circular targets gave less variable results than a single background ROI or smaller diameter ROIs. This is because variability of measured HU, due to image noise, is reduced with larger ROIs. Hence, consistency of measured contrast does depend on measurement technique used, with larger ROIs being preferable. The results from the 604 phantom models within this group of phantoms appeared more consistent than those for the 504 models (Figure 5). Results were close to the nominal values though the variability of results indicates that error bars of approximately 10 % need to be applied i.e. 11 $\pm$ 1 HU contrast for the nominal 1.0% target for the 604 model phantoms.

Variation of the chosen scan parameters produced no changes in measured contrast. Scan parameters that were not tested here which do have the potential to affect contrast are reconstruction kernel and tube kilovoltage [4, 14, 15]. These are not routinely varied for a given examination type in radiotherapy CT and so have not been investigated.

Visual scoring of low contrast targets was less sensitive to the background HU differences in the 504 and 604 phantom models and there were no notable differences between any of the phantoms for this parameter. The spread of results when scoring images from a single phantom, Figure 6, indicates error bars of $\pm$ 0.5 a target to be appropriate for three scorers undertaking repeated measurements. Visual scoring of images obtained when varying scan parameters showed reductions in target visibility in excess of 0.5 of a target. Reduced mAs, reduced nominal contrast, reduced slice thickness and increased FOV all showed a reduction in scores of between 1 and 4 targets depending on the parameter and the extent of the change. This confirms that different phantoms can be used to assess differences in low contrast target visibility provided suitable error bars, in this case, $\pm$ 0.5 of a target, are applied.

The Hounsfield units measured in the various inserts in sensitometry section were in good agreement across the set of phantoms. The range was small, less than $\pm$5% of the average HU values of the insert material. The HU values for Delrin and Teflon were noted to be outside the quoted typical values from the Catphan technical specification, see Tables 2 and 4. CT numbers, whilst material dependent, are also a product of the scanner kilovoltage, beam shaping filter, reconstruction kernel, volume and type of material surrounding the insert. Typical values quoted by the Phantom Laboratory can be considered indicative but not nominal values. Reconstruction kernels, in particular on Canon CT scanners, are known to affect the HU values [4]. Other publications have identified that HU does not vary significantly with mAs, slice thickness, reconstruction FOV or iterative reconstruction strength, other than a small change in variability due to noise changes [16].

The check of X-ray attenuation based on scanning the phantoms with automatic tube current modulation showed there to be no significant differences in attenuation across the 14 phantoms.

To summarise, the measured differences between Catphan phantoms, are given in Table 5.

Parameter	Measured variation
CT number in uniform section (HU)	± 1.6
Noise of uniformity section (HU)	± 0.07
Uniformity (edge vs centre) (HU)	± 0.15
Measured contrast - nominal 1%	± 0.1
Visual scoring, nominal 1% – number of details	± 0.5
CT numbers (HU) : air	± 5
PMP	± 2
Polystyrene	± 2
Delrin	± 5
Teflon	± 8
X-ray attenuation - based on indicated tube current (mA)	± 1

Table 5. Tolerances which should be applied to measured values when using a selection of different Catphan phantoms.

Conclusion

The results from 14 Catphan phantoms of models 503, 504 and 604 were found to be very similar for most measured parameters. The only parameter where there was an apparent notable difference was the low contrast target section. If contrast measurements are to be made, these results suggest using the 604 model phantoms would be preferable to reduce variability caused by phantom differences. Results show the degree of difference between phantoms for the low contrast, noise and sensitometry sections to be small compared to changes seen when varying CT scan protocol parameters. This work indicates the size of error bars which should be applied to results if using a range of different Catphan phantoms for comparison of image quality between centres.

Differences between phantoms are much smaller than might be suggested by the technical specification in the different Catphan manuals. It would be useful if The Phantom Laboratory could indicate in the user manual the typical manufacturing tolerances achieved so that users would know what magnitude of difference to expect when using multiple phantoms of the same model interchangeably.

Acknowledgements

References

[1] The Phantom Laboratory. The Catphan 504 Manual. 2013.

[2] The Phantom Laboratory. The Catphan 604 Manual. 2015.

[3] The Phantom Laboratory. The Catphan 503 Manual. 2016.

[4] Davis AT, Palmer AL, Pani S, Nisbet A. Assessment of the variation in CT scanner performance (image quality and Hounsfield units) with scan parameters, for image optimisation in radiotherapy treatment planning. Phys Med. 2018;45:59-64.

[5] Underwood AC, Law J, Clayton C. Reproducibility of Leeds TOR(MAM) mammographic test object plates. Br J Radiol. 1997:186-91.

[6] Fabiszewska E, Grabska I, Pasicz K. The threshold contrast thickness evaluated with different CDMAM phantoms and software. Nukleonika. 2016;61:53-9.

[7] Wood TJ, Davis AT, Earley J, Edyvean S, Findlay U, Lindsay R, et al. IPEM topical report: the first UK survey of dose indices from radiotherapy treatment planning computed tomography scans for adult patients. Phys Med Biol. 2018;63:185008.

[8] Rasband WS. ImageJ. Bethesda, Maryland, USA: U. S. National Institutes of Health; 1997-2018.

[9] Edyvean S, Lewis M, Keat N, Jones A. IPEM Report no. 32. Measurement of the Performance Characteristics of Diagnostic X-ray Systems used in Medicine. Part III Computed Tomography X-ray Scanners. York, UK: Institute of Physics and Engineering in Medicine; 2003.

[10] Wang J, Langer S. A brief review of human perception factors in digital displays for picture archiving and communications systems. Journal of Digital Imaging. 1997;10:158-68.

[11] Martin C, Sookpeng S. Setting up computed tomography automatic tube current modulation systems. J Radiol Prot. 2016;36:R74-R95.

[12] Wood TJ, Davis AT, Earley J, Edyvean S, Findlay S, Lindsay R, et al. IPEM topical report 2: the first UK survey of dose indices from radiotherapy treatment planning computed tomography scans for adult patients. Phys Med Biol. 2018;107287.R1.

[13] Goldman LW. Principles of CT: radiation dose and image quality. J Nucl Med Technol. 2007;35:213-25.

[14] Kalender WA. Computed Tomography: Fundamentals, System Technology, Image Quality, Applications. 3rd ed. Erlangen: Publicis; 2011.

[15] Tang K, Wang L, Li R, Lin J, Zheng X, Cao G. Effect of low tube voltage on image quality, radiation dose, and low-contrast detectability at abdominal multidetector CT: phantom study. J Biomed Biotechnol. 2012;2012:130169.

[16] Davis AT, Palmer AL, Nisbet A. Can CT scan protocols used for radiotherapy treatment planning be adjusted to optimize image quality and patient dose? A systematic review. Br J Radiol. 2017;90:20160406.