Improving the quality of head and neck radiotherapy CT images using a second image reconstruction set

Anne T. Davis[a,b], David Nash[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 University NHS Trust, Portsmouth, UK

Abstract

Purpose: To improve the quality of radiotherapy head and neck CT images through use of an additional image set reconstructed from the raw data of the primary scan, thus allowing parameters such as reconstruction field-of-view (FOV) and kernel to be optimised without impacting on the images used for treatment planning dose calculations.

Methods: Using a Catphan image quality phantom and a Toshiba Aquilion LB CT scanner, qualitative and quantitative measurements were made for different reconstruction kernels and FOV diameters. The preferred FOV diameter and kernels were selected. Clinical images from six patients were reconstructed using those kernels (FC13, FC41, FC44, FC64) and the chosen FOV, 200 mm. The images were ranked to choose the kernel which gave best image quality for organ delineation. The scanner workflow was adjusted to produce for every scan a second image set using the chosen kernel and FOV. Finally, for 10 patient scans, image quality was compared for the two reconstructed images.

Results: The second image set, was produced using kernel FC44 and 200 mm FOV. The primary image set using 550 mm FOV and FC13 was unchanged and contours from the second image set merged onto the first. Oncologists reported increased confidence in contouring in all cases using the new procedure.

Conclusion: Production of a second image set, using a reduced reconstruction FOV and a kernel which optimises contrast and sharpness, significantly improves the quality of head and neck CT images for contouring, and avoids any dose increase.

Key words: radiotherapy, CT, image, quality

Highlights

- Quality improvement of head and neck radiotherapy CT images with no additional dose
- A second image reconstruction set with optimised reconstruction FOV and kernel
- No adverse impact on images used for treatment planning dose calculations
- Improved image contrast and sharpness and oncologist confidence when contouring
- Method can be implemented on all makes and models of CT scanner

Introduction

The objective of radiotherapy CT scanning is to produce images which can be used for a) delineation of the target and organs at risk, b) for dose calculation within the radiotherapy treatment planning system (TPS), and c) as a reference for treatment set-up consistency. There are often constraints when setting up the CT imaging protocols that relate to the requirements of the treatment planning system [1-3]. For photon radiotherapy, the TPS will contain a calibration curve (CT-to-density table), which converts the CT values for different tissue types to electron or physical density. When acquiring the data for the TPS calibration curve(s), through use of a specialist phantom, the scan parameters must match those for the clinical scan protocols to ensure that the CT values within the calibration curve are representative of those seen in clinical images. Some scan parameters can change the CT value for a given tissue type when adjusted, though this is dependent on the make and model of the scanner [4, 5]. Many radiotherapy centres choose to have only one or two TPS calibration curves to avoid the risk of the wrong one being selected. Consequently, the different CT protocols for scanning various body regions may all contain very similar scan parameters. Tube current is commonly varied through use of the scanner modulation system but for other scan parameters, such as reconstruction kernel, image slice thickness and kilovoltage, centres often restrict selection such that parameters are set the same across all protocols. For example, scan protocol information submitted to a UK audit showed that 57% (30 of 53) of scanners used the same reconstruction kernel for their head and neck, lung, prostate and brain protocols [6]. Restricting the selection of scan parameters is contrary to the concept of image optimization and has the potential to result in sub optimal clinical image quality.

With head and neck imaging, where delineated anatomical details are generally small (for example the optic chiasm is typically 4 mm in diameter), an added difficulty is the need to ensure the image field of view (FOV) is large enough to accommodate the shoulders for inclusion in the treatment plan. The reconstructed FOV diameter for a head and neck scan is typically between 350 and 700 millimetres (data from 52 scanners in UK centres) [6]. A large diameter FOV results in a large image pixel size and indistinct detail when viewing the image for delineation. Additionally, contrast between different types of soft tissue in this region is low making delineation difficult. Examples of organs difficult to delineate include the spinal cord, pituitary gland, optic nerve and chiasm [7-9].

This paper presents a method to improve the details seen in the brain region of radiotherapy head and neck CT images to aid delineation. Changes to two scan parameters, namely reconstruction kernel and reconstruction FOV were investigated. A simple methodology to determine the optimum settings is presented and the modifications to the imaging and planning workflows that were made for seamless implementation are explained. The work utilises the production of a second reconstructed image data set from the raw data of the planning CT scan. The use of additional reconstructions is common place in diagnostic CT scanning but they are not routinely used within the context of CT imaging for radiotherapy treatment planning currently.

Methods

This study was carried out on a Toshiba (Tochigi Prefecture, Japan) Aquilion LB scanner, software version 3.35, used for radiotherapy CT scanning in a UK oncology centre. The parameters for the original head and neck scan protocol are shown in Table 1. Reconstruction FOV diameter was selected on a per patient basis depending on patient size. The patient's shoulders and the CT couch had to be visible in the images for treatment planning purposes. The reconstruction method was filtered-back projection with no iterative reconstruction available on this scanner. The image matrix size was 512

by 512 pixels and all scans used helical mode. The reconstruction kernel, FC13, was selected by the scanner manufacturer's clinical application specialist when the scanner was installed. It is understood that their rational was to minimise CT number variation for different scan protocols and hence all scan protocols were allocated FC13 which is a kernel intended for body imaging. Tube current was fixed, not modulated, as advised by the clinical applications specialist when the protocol was developed. In line with national professional guidance, intravenous contrast was administered to all patients with lymph node involvement unless contraindicated on medical grounds [10]. The protocol for contrast administration was: Omnipaque (GE Healthcare, Chicago, USA) contrast, concentration 300 mg I/ml, flow rate 1 mL/s, max volume 95mls and the contrast amount was the weight of the patient in kg + 10 mls, 90 second delay between contrast administration and the start of the CT scan. The role of IV contrast is generally to enhance the visibility of areas of blood flow and to support detection and delineation of the lymph nodes and the target volume.

Table 1 Parameters for the original head and neck planning CT scan

Kilovoltage	Effective mAs	Collimation	Pitch	Image slice thickness (mm)	Reconstruction FOV (mm)	Reconstruction kernel	$CTDI_{vol\ 16\ cm}$ (mGy)
120	267	1 x 16	0.938	2	400 to 700	FC13	75.8

With the initial workflow, the acquired images were sent to Prosoma (Medcom, France) for organ and target delineation and subsequently to Philips (Amsterdam, Netherlands) Pinnacle version 9.6 TPS system for plan development and dose calculation. For this study, a second set of images were reconstructed from the raw data of the planning CT scan so that the first set of images could be used for dose calculation and the second set for delineation. Since both image sets originated from the same raw data set there were no concerns about movement or misregistration and there would be no adverse effects on the treatment dose accuracy related to any CT number changes caused by changing reconstruction parameters, nor any additional dose to the patient.

The choice of the reconstruction parameters for the second image set for delineation was carried out in stages, namely selection of the optimal diameter for the reconstruction FOV, next selection of the optimal reconstruction kernel based on both phantom and then clinical images. Once the reconstruction parameters were selected and images produced, the fourth step was implementation of the new workflow to utilise the second image set and the final step was to compare the clinical images using both reconstructions to determine whether image quality had improved.

(i) Selection of the reconstruction FOV using a phantom

Assuming no change to the image matrix size, as the reconstruction FOV is reduced the pixel size will also reduce in a linear manner, resulting in improved resolution and sharpness [11]. Image noise will increase according to an inverse square relationship, meaning that if the reconstruction FOV is reduced by a factor of two, the pixel area reduces by a factor of four and the noise will increase by a factor of two. Therefore very small FOVs might detrimentally affected image quality by increasing image noise to an unacceptable level. The aim was therefore to use a reduced FOV which balanced

improved resolution without a significant increase in noise and provided sufficient coverage of the required anatomy. Improving small organ visibility in the brain was the priority so a default 200 mm diameter reconstruction FOV was chosen as the likely preferred FOV since it allowed the whole head to fill the image. Adjustment could be made by the operator, on a patient-by-patient basis, allowing for head size. Reconstruction FOVs smaller than 200 mm could have been used to further improve resolution but this would have both increased noise and necessitated additional reconstructions for full coverage of the anatomy required for contouring. There was a limitation in the number of image sets which were considered practical for workflow management and for handling by the radiotherapy TPS, hence the decision to have two image sets in total with the second reconstruction utilising a 200 mm reconstruction FOV giving both improved image sharpness and adequate coverage.

To test whether reduced FOV did deliver an improvement in image quality, a phantom, Catphan model 503 (The Phantom Laboratory, USA), was scanned, aligned with the low contrast targets, see Figure 1. Scan parameters used were those in Table 1 but with reconstruction FOV diameters between 550 and 200 mm. 550 mm was chosen as it was the most common largest FOV for clinical imaging, with 700 mm used only rarely. FC23 (Head—with Beam Hardening Correction; fine grain size) was the reconstruction kernel. Viewing conditions were standardised through adjustment of display window width and level so that brightness and contrast looked similar on all images. Widow width and level could not be set the same for all images as the different kernels gave different displayed brightness and contrast under those conditions. Electronic zoom was applied so that the size of the targets was the same on each of the images. A simple scoring system was used, counting the number of targets seen within a group. Scores could range from 9 (all targets from 15 to 2 mm diameter seen) to 0 (no targets seen). Two observers (Physicists) who were experienced at reviewing Catphan low contrast images undertook the scoring, providing four sets of scores for each image. All images were scored and averaged for a given FOV.

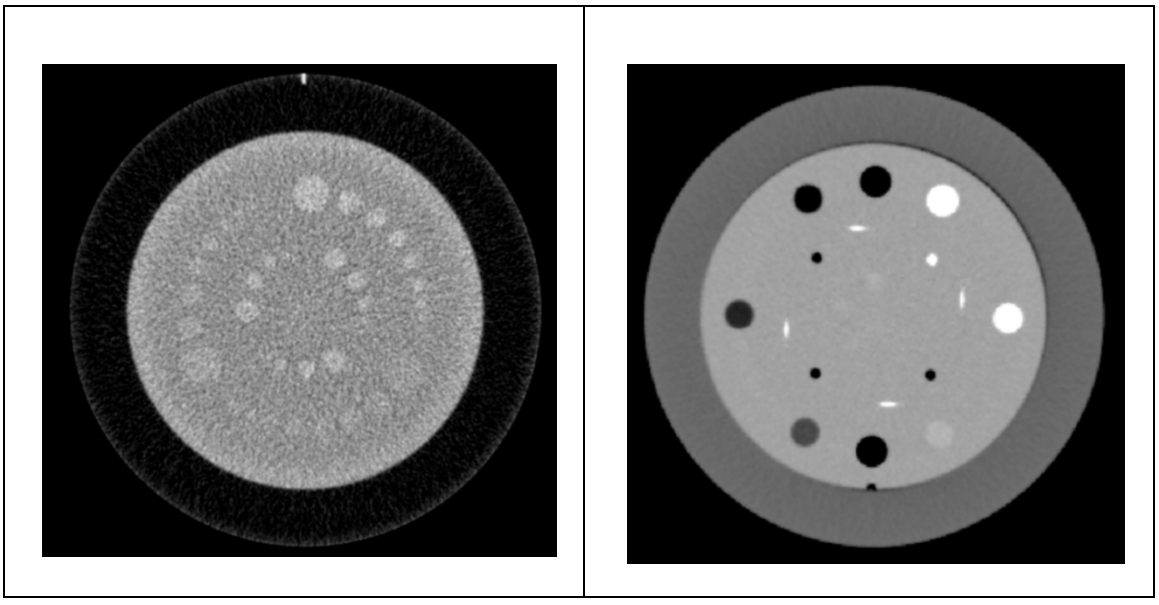


Figure 1 Images of the Catphan phantom showing the low contrast targets (left) and the sensitometry section (right)

A second assessment was made to determine the visibility of the edge of the 15 mm diameter target, see Figure 2. This method allocated a score on a scale between 1 and 5 to the edge and noted the percentage of the edge that was visible at that score. The scoring system 1 to 5 related to the following definitions: 1: edge not visible, 2: edge fuzzy or clumpy, 3: edge defined but lacking sharpness, 4: edge reasonably well defined and 5: edge sharply defined. Scoring was applied to the whole edge and could result in different scores allocated to different proportions of the edge. The overall score for the visibility of the target edge would then be calculated by multiplying each score by the associated percentage and adding together all edge scores for that target. For example, a single target edge might score: 4 (reasonably well defined) x 40% + 3 (defined but lacking sharpness) x 40% + 2 (fuzzy or clumpy) x 20%. Adding the individual scores of 160, 120 and 40 for this case would produce a total edge score for that target of 320. The maximum possible score was 500, related to a score of 5 given to 100% of the edge. To validate this method of scoring edges, a further test was added which was simply to compare and rank the order of low contrast target images in terms of best to worst visibility for the 15 mm diameter target. The best image was given the score 1. Where multiple sets of images were scored for the same parameters, ranks were added together to give total score for the ranking. Lowest value indicated best performance.

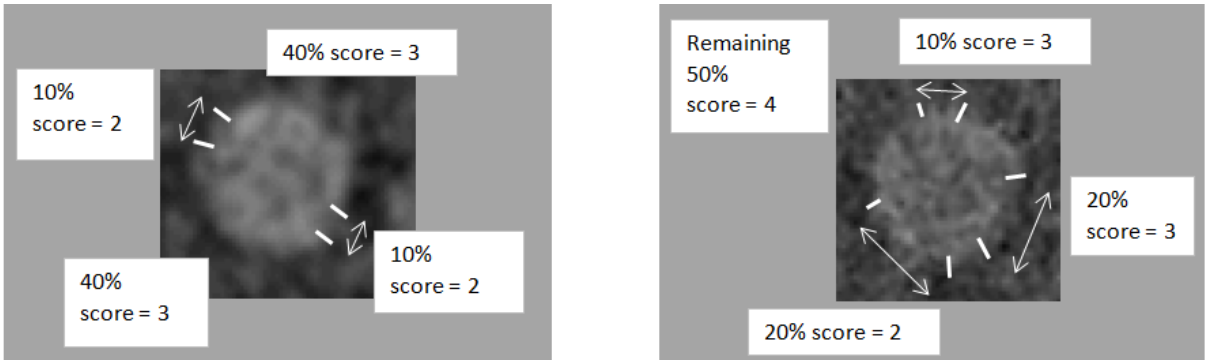


Figure 2 Two examples of scoring the visibility of the edge of the 15 mm diameter, nominal 1% contrast target. The left image (kernel FC64) scores 20%x2 + 80%x3 = 280; The right image (kernel FC13) scores 20%x2 + 30%x3 + 50%x4 = 330, Score 4 is 'edge reasonably well defined' 3 is 'defined but lacking sharpness', score 2 is 'fuzzy or clumpy.

(ii) Selection of the reconstruction kernel using a phantom

A large number of reconstruction kernels were available on the scanner and those intended for head and body imaging are shown in Table 2 [12]. The reconstruction kernels selected for investigation were FC13 (original), FC08, FC17, FC23, FC41, FC44 and FC64. The choice of kernels from within the different groups was guided by technical information provided by Toshiba in their scanner operator manual which recommended a specific kernel for a particular anatomical region. Whilst the majority of kernels selected were for head imaging, three (FC13, FC08 and FC17) were for body imaging. These were included in the investigation since the original kernel FC13 selected by the clinical applications specialist was a body kernel. FC17 and FC44 were in the same groups as FC13 and FC41 respectively but were added to enable examination of the impact of sharper kernels from within the same group

as another kernel. The numbering of the kernels is such that, within a particular group, the higher number indicates a sharper kernel. Sharper kernels will potentially both enhance the visibility of edges and increase noise. Within a group some also have increased contrast such as FC08 and FC17. It was also known that on this scanner model changing the reconstruction kernel could change the CT values and therefore might impact on image contrast [4].

Table 2 Reconstruction kernels available on the Toshiba Aquilion LB scanner

Kernel group	Description of kernel group
FC01 to FC09	Body algorithms with BHC (FC7 to FC9 have increased contrast)
FC11 to FC19	Body algorithms without BHC (FC17 to FC 19 have increased contrast)
FC20 to FC26	Head algorithms with BHC (fine grain)
FC41 to FC44	Head algorithms without BHC
FC 62 to FC68	Head algorithm with BHC (coarse grain)

BHC is beam hardening correction

Preliminary work was undertaken to establish what level of radiographic contrast was typical in the original images produced using the FC13 kernel. Clinical staff identified the spinal cord as the structure which was routinely most difficult to see. For 20 clinical head and neck image sets (10 with intravenous contrast and 10 without), one image from each set was selected at the level of vertebra C4. A measure of image contrast of the spinal cord was made by placing a region of interest (ROI) in the centre of the cord and a second ROI in homogenous tissue adjacent to the cord. Care was taken, when placing the ROIs, to avoid crossing organ boundaries and any artefacts. Contrast was calculated by subtracting the CT number in the cord ROI from that in the adjacent ROI [13].

Six sets of scans were then taken of the Catphan phantom using the scan parameters given in Table 1 but with various reconstruction kernels. The reconstructed image FOV was 200 mm. Each of the six kernels under investigation was chosen in turn and three identical scans acquired. Images were selected showing the sections containing the low contrast targets and the sensitometry inserts, see Figure 1 **Error! Reference source not found.**. Two images of each section were collected from each of the scan sets, giving six images for each section and reconstruction kernel. To investigate the impact of any CT number change with position within the FOV, further scans of the Catphan were made with the phantom off-set laterally by 13 cm, positioning the phantom towards the outer edge of the scan FOV. 13 cm was chosen as a reasonable representation of the position of the edge of the shoulder on a head and neck scan. The reconstruction FOV was 550 mm diameter.

Within the groups of the low contrast targets, circular details ranged in diameter from 15 to 2 mm. The three groups had nominal contrast values of 1%, 0.5 % and 0.25%. This equates to a CT number difference, compared to the background phantom material, of approximately 10, 5 and 2.5 HU. The group of targets chosen was that with the nominal contrast closest to that of the contrast of the spinal cord measured from the 20 clinical images sets. The same scoring methods were used as for the images with the different reconstruction FOV, with the number of targets seen counted and the sharpness of the edge of the 15 mm diameter target scored.

For the sensitometry section, the eight inserts of different materials provided a range of CT numbers ranging from approximately -1000 to +1000 HU. Average CT numbers for the various inserts were compared against CT numbers for different clinical tissue types and those most closely matching soft

tissue were determined to be polystyrene, low density polyethylene (LDPE), acrylic and the 'background', the unspecified material that the sensitometry inserts are embedded in, see Table 3. For each of the reconstruction kernels, CT values were measured by placing ROIs on each insert and the contrast between the different materials was calculated. The process was repeated for the images acquired when the phantom was off-set laterally.

Table 3 CT values for inserts in the Catphan and for various types of clinical soft tissue

Material of insert	Typical CT number	Clinical tissue type	Clinical CT number range
Sensitometry section background	80	Muscle, soft tissue, grey / white matter	5 to 80
Background & acrylic	80 & 100	Blood	50 to 90
Polystyrene & LDPE	-50 & -100	Fat	-30 to -200

Finally, the diameter of the polystyrene insert was measured for all kernels to establish whether the choice of kernel had any impact on the apparent edge position of the insert. The method chosen for this was to plot a profile across the insert on the image to obtain an inverted 'top hat' shaped profile of the CT numbers. The width of the insert was defined by measuring across the profile at half the maximum height. The widths of the insert from the profile plots for the various kernels were then compared.

(iii) Selection of the reconstruction kernel using clinical images

The four top scoring kernels, based on phantom measurements, were selected and head and neck scans from six patients were reconstructed using those kernels, producing four image sets for each patient. These images were reviewed by three reviewers (one oncologist, one radiotherapy physicist and one imaging physicist) who ranked the images in order of their preference for image quality and suitability for delineation, thereby choosing the kernel they judged gave the best performance. Assessment of image quality was made by reviewing visibility of the parotids, spinal cord, brain stem, arteries and musculature, with the spinal cord generally being the most challenging anatomy to visualise.

(iv) Quantitative phantom measurements for different kernels

Further images were collated to allow calculation of both noise-power spectra (NPS) and task or target transfer function (TTF) for each of the reconstruction kernels. The NPS is a measure of noise levels at different frequencies within the image and will reflect the texture of the noise pattern seen in the image. An image with coarser grain noise pattern will have a lower frequency for the peak of the NPS than an image where the noise pattern is finer. Images were acquired using the standard clinical head scan protocol with a 200 mm reconstruction FOV set. NPS and TTF were calculated using Quantitative IQ software coded using ImageJ (National Institutes of Health, USA) which had been validated against other open source software which calculate NPS and TTF [14-16]. The NPS for each different kernel was produced using an average of 20 images with 16 ROIs of 64 pixel square positioned at 50 mm from the centre of the image, see Figure 3 [17]. The peak and average frequency, and overall noise level were recorded for each reconstruction kernel. The TTF was calculated using the polystyrene insert and an average of 9 images from each scan[17]. This insert was chosen since the contrast level (relative to

the phantom background material) was in the soft tissue Hounsfield unit range and was sufficiently high to avoid producing a noisy TTF profile. Frequency values for the normalised TTF at 50% and 10% of the profile were recorded for each kernel. The contrast values for acrylic versus background were used for a contrast-noise (CNR) calculation. Acrylic was chosen for the contrast measure since it had the subtlest contrast of all the materials measured. TTF, NPS, noise and CNR results were compared against the scores for visual ranking by the observers of the clinical images and edge visibility of the low contrast targets to determine whether those parameters could be used to predict the choice of kernel based on visual scoring methods.

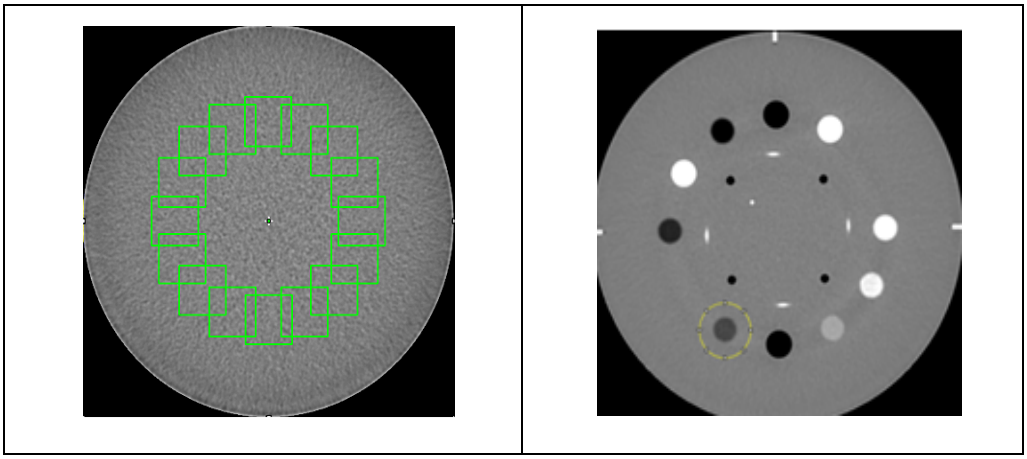


Figure 3 Images from the Catphan phantom used for calculation of the NPS (left) and TTF (right).

(v) Implementation of the new image workflow

Additional reconstructions are relatively easy to implement and can be set up by scanner operators or with support from CT scanner applications specialists. Following review of the phantom and clinical images, the scan protocol was changed to routinely produce a second set of images reconstructed with the chosen reconstruction kernel and the reduced diameter FOV of 200 mm. This second image set showed a magnified view of the central region of the slice. The reconstruction was generally centred but could be moved off-centre if required, such as for a patient with a large neck tilt. The planning CT scan parameters and the first image set for the TPS were unchanged. Both sets of images were exported from the CT scanner into the Prosoma (Medcom, France) virtual simulation system where they were fused with no adjustment needed. They were made available to the oncologists for delineation and once contours had been added were sent for treatment planning. All delineations were attached to the planning images. The planning images and contours were subsequently sent to Pinnacle for treatment planning and dose calculation as per normal protocol. This included sending the radiotherapy dose distribution back to Prosoma for approval, where if desired the oncologist could use the reduced FOV images.

(vi) Comparison of clinical images from original and new protocols

Using the chosen reconstruction kernel and with the selected diameter for the reduced reconstruction FOV, ten more sets of clinical images were reconstructed. Two experienced head and neck oncologists then compared each of the new image sets against the original image sets. They were asked to answer a set of questions when undertaking the comparison. The questions asked are shown in Table 4.

Table 4 Assessment of clinical image quality for second reconstruction image set compared to first image set

Question	Possible answers
Was the image contrast enhanced?	Yes / no
How does image quality compare to the original image set?	Superior/Equivalent/Inferior
Was contouring time reduced?	Yes/no/no change
Was your confidence in contouring increased?	Yes/no/no change

Lastly, measurements of contrast for the spinal cord were made on the first and second reconstruction image sets for those ten patients (no IV contrast). A two sample paired student t-test was performed on the results to determine whether there was a difference in measured contrast values for the two image sets.

Results

(i) Selection of the reconstruction FOV

The preliminary measurements of the contrast level for the spinal cord from the clinical images gave results of 21 HU (SD 10) and 20 HU (SD 8) respectively for scans with and without IV contrast. This equates to a 2% contrast level against the background. The contrast levels available in the Catphan phantom low contrast target groups were 1, 0.5 and 0.25%. Lower numerical contrast values equate to harder to see targets. So, use of the 0.5% and 0.25% contrast groups for this work was discounted as the targets were more subtle and difficult to see than the contrast of the spinal cord in clinical images. The 1% (~10 HU) contrast group was used for image quality evaluation, accepting that it was not an exact match for the spinal cord contrast of approximately 2%.

Table 5 shows the scores from the 1% group of low contrast targets in the Catphan phantom for different diameter reconstruction FOVs. The smallest target seen was 4 mm in diameter (corresponding to a score of 7) and approximately the same score was obtained for all set values of FOV diameter. The score for the edge visibility of the 15 mm 1% contrast target shows higher scores for reducing FOV diameter. The 200 mm diameter FOV has therefore improved the image compared to the larger FOVs.

Table 5 Scores from the Catphan images acquired with different diameters reconstruction FOVs

Diameter of reconstruction FOV (mm)	Average score for number of targets seen in the 1% contrast group	Score for edge definition of the 1% contrast 15 mm diameter target	Ranking of 15 mm target (lowest is top score)
550	7.5	193	12

400	7.2	250	8
200	7.5	283	4

(ii) Selection of the reconstruction kernel using a phantom

The values of contrast between different inserts in the Catphan sensitometry section, for different reconstruction kernel, are shown in Table 6 with a clear range of different contrast values produced by the different kernels. Kernels FC44 and FC41 delivered the highest contrast values for the range of materials which are considered soft tissue equivalent and FC23 and FC08 the lowest.

Table 6 Contrast values from different insert materials in the Catphan for various reconstruction kernels with the phantom centred in the FOV. The 'top three' best performing kernels for each measure are highlighted in bold.

	Contrast (HU)						
Materials	FC44	FC41	FC13	FC17	FC64	FC23	FC08
Acrylic – polystyrene	**176**	**174**	163	160	**165**	155	150
Background – polystyrene	**106**	**104**	99	96	**105**	92	92
Acrylic – background	**70**	**70**	**64**	63	59	60	58
Acrylic – LDPE	**240**	**236**	**222**	220	211	214	205

The results obtained from the measurements on the images where the phantom had been offset and positioned towards the outer edge of the scan FOV showed that whilst CT numbers could change considerably with position within the FOV, the differences in HU values between different materials did not change significantly. For example, CT number contrast differences were measured as up to 15 HU (7%) different for the highest contrast value, acrylic and LDPE. The relative ranking, however, of the results from the different kernels when the phantom was off-set laterally were the same as when the phantom was centred.

Table 7 shows the results from scoring the Catphan 1% contrast group of targets. FC13, FC44 and FC41 kernels scored the highest on edge definition and target ranking, FC64 scored the lowest. For the number of targets visible within the 1% contrast group, results for FC44, FC41, FC13, FC64 and FC08 were similar, and scores for FC17 and FC23 were slightly lower. It is interesting to note that the visibility of the edge of the 15 mm target cannot be predicted by the scores from counting the number of different diameter targets visible within the group. There are only small differences in performance of the various kernels when counting the number of visible targets.

Table 7 Scores for low contrast targets (nominal 1% group) for various reconstruction kernels. The 'top three' best performing scores for each measure are marked in bold.

Image quality measure	FC44	FC41	FC13	FC17	FC64	FC23	FC08
Average number of low contrast targets visible	8.2	**8.7**	8.2	7.8	**8.3**	7.9	**8.5**
Standard deviation	*0.8*	*0.5*	*0.8*	*0.4*	*0.5*	*0.2*	*0.5*

Edge definition score, 15 mm diameter low contrast target	**330**	**330**	**338**	308	257	303	310
Standard deviation	21	20	8	18	38	17	24
Ranking of 15 mm target (lowest value is top score)	**16**	**14**	**13**	19	42	20	33
Standard deviation	*6*	*5*	*4*	*6*	*14*	*7*	*11*

For the next stage, the aim was to shorten the evaluation list to four kernels. FC23 and FC08 were excluded based on low contrast values. FC13, FC41 and FC44 were best performers for both contrast and edge definition scores so were retained. Of the remaining kernels (FC17 and FC64) FC17 was excluded on the basis that the contrast was virtually the same as for FC13 but the edge definition worse than FC13. FC17 and FC13 were also from within the same kernel group as shown in Table 2. FC64 was left in the evaluation process since whilst it performed worst at edge definition scores it had high contrast scores for some parts of the contrast range. Hence, the short list for kernel evaluation with clinical images was FC44, FC41, FC13 and FC64.

The results from measurement of the diameter of the polystyrene insert for all the reconstruction kernels were within 12.3 $\pm$ 0.1 mm indicating no significant variation in edge position on the different images, arising from kernel selection.

(iii) Selection of the reconstruction kernel using clinical images

The results from the review of six clinical image sets each produced with the four selected kernels FC41, FC44, FC13 and FC64 ranked the kernels (best to worst). When the ranking scores from the three reviewers were grouped together the results were FC44 = first, FC41 and FC13 = joint second and FC64 = fourth. The reviewers commented that kernel FC44 produced a clear visual improvement to the images compared to those produced with other kernels. FC44 was therefore chosen as the preferred kernel. FC44 was then combined with the chosen 200 mm diameter reconstruction FOV to produce the second set of reconstructed images.

It is of interest to note that the chosen kernel, FC44, is recommended for head imaging, rather than FC13 originally in the scan protocol which is for body imaging. Neither FC44 nor FC13 kernels correct for beam hardening. Beam hardening artefacts in the head, when they occur, are generally seen as bright or dark streaks in the vicinity of high attenuation material such as bone. For example, there may be a thin dark band in the section of the soft tissue immediately adjacent to the bone of the skull [18] . This was not judged a particular problem for contouring.

(iv) Quantitative phantom measures for different kernels

The results from the measures of NPS, TTF and noise are shown in Table 8. For the images in this study, the level of noise measured did not predict which kernel would be selected during visual assessment. For example, it can be seen that FC64 had the lowest noise value but also had low visual scores from visual assessment of phantom edge definition and clinical image ranking. FC44, however, has the highest noise value of all the kernels assessed. For this reason CNR was also not a good predictor of kernel selection. The NPS results, which describe the texture of the noise, indicate that FC64 has the lowest frequency for both the average and the peak of the NPS, whilst FC44 has the highest values.

The high frequencies indicate in a fine texture noise pattern whilst the low frequencies suggest a coarse grain pattern. Visual review of the noise pattern on the phantom images confirmed this. TTF, the measure of resolution and image sharpness, was highest for FC44 and lowest for FC64. It is interesting to compare the results for FC41 and FC13 which were scored second and third in the clinical image review and first (FC13) and joint second (FC41) for phantom edge scores. FC13 generally scores higher for TTF and NPS results than FC41 but contrast is lower than for FC41. This suggests that maximising radiographic contrast is most important for optimising clinical head and neck image quality for contouring, with low frequency of noise (fine grain) and high frequency of TTF (sharp image) next most important. Overall, noise levels are less important than the texture of the noise when visualising the defined edge of an organ or a phantom target.

Table 8 Results from measures of TTF, NPS, noise and CNR. The 'top three' best performing scores for each measure are marked in bold.

Measured parameter	FC44	FC41	FC13	FC17	FC64	FC23	FC08
TTF50 (per mm)	**0.33**	0.26	0.29	**0.32**	0.23	0.29	**0.34**
TTF10 (per mm)	**0.57**	0.46	**0.55**	0.51	0.39	**0.55**	0.52
Frequency of peak of NPS (per mm)	**0.19**	0.17	0.17	**0.20**	0.14	0.17	**0.24**
Average frequency of NPS (per mm)	**0.27**	0.19	**0.26**	0.23	0.16	**0.26**	0.25
Noise	2.2	**1.3**	1.7	1.7	**1.0**	**1.6**	1.7
Contrast (Acrylic-background)	**70**	**70**	**64**	63	59	60	58
CNR	32	**54**	38	37	**59**	38	34

(v) Implementation of the new image workflow

There were no problems experienced with use of the new workflow. The oncologists could switch between viewing the central portion of the image FOV using either the first image set (planning images) or the second image set (reduced diameter FOV images). Delineation for organs outside of the reduced diameter FOV was carried out on the planning image set but was generally only the most inferior nodal chain since most organs could be seen on the 200 mm FOV images. All contours were successfully fused onto the planning images.

(vi) Comparison of clinical images from original and new protocols

The results from the oncologists' evaluation of the ten sets of images for the second reconstruction set (FC44 and 200 mm FOV) compared with first image set (FC13 and 550 FOV) are shown in Table 8. The second reconstruction was considered to have delivered an improvement in image quality for delineation or contouring in all cases, positively impacting on either the time taken or the confidence of the oncologist in accuracy.

Table 8 Results from the oncologists' evaluation of new reconstructions image (200 mm FOV and FC44 kernel) versus the original image sets (550 FOV and FC13 kernel) for ten scans

Question	Answers and number of responses
Was the image contrast enhanced for the new reconstruction?	7: Yes, 3: No
How does image quality compare to the original image set?	9: Superior; 1: Equivalent

Was contouring time reduced?	8: Reduced; 2: No change
Was your confidence in contouring increased?	10: Yes

The results from the measurements of contrast on the spinal cord for the two kernels were: average contrast 30.0 HU (SD 8.5) and 20.3 HU (SD 7.9) for FC44 and FC13 respectively. The result from the two sample paired student T-test was *t (19) = 3.3 and p <0.01* indicating significance and confirming that the contrast values were different for the two kernels. Clinical observers also confirmed that the spinal cord was generally more visible with the FC44 kernel.

The example images from a clinical head scan given in Figure 4 visually show the quality improvement that was achieved across the whole image through production of the second reconstruction image set using 200 mm diameter FOV and kernel FC44. For comparison purposes, the first reconstruction image set was electronically zoomed to match that of the second image set. Organ edges are much more clearly defined.

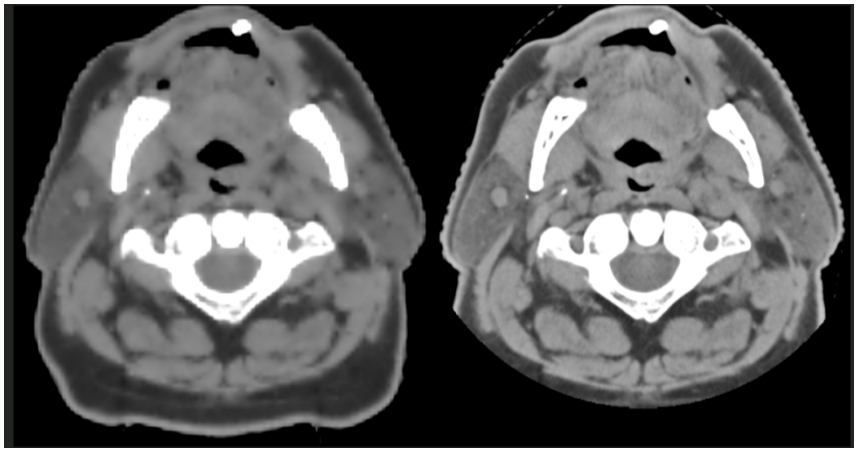
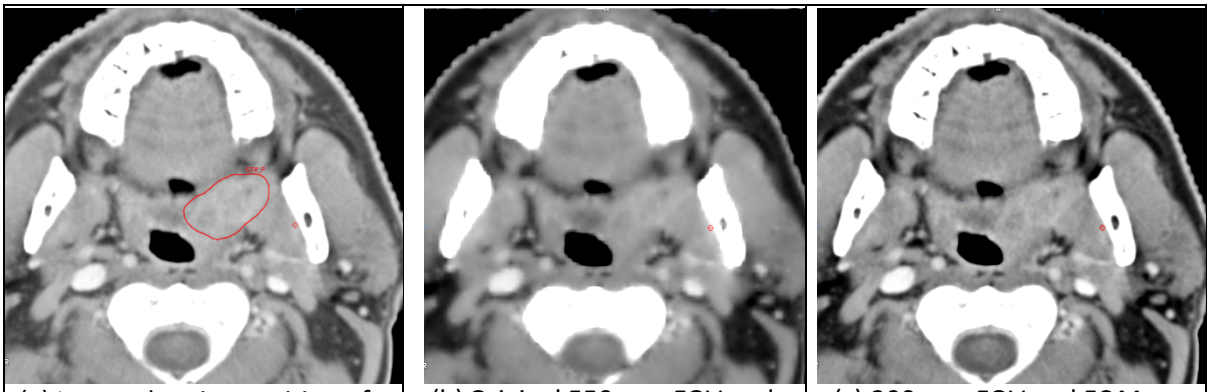


Figure 4 Images from a CT head and neck scan showing (left) the original 550 mm diameter reconstruction FOV and reconstruction kernel FC13 and (right) the improved image quality achieved by use of a 200 mm diameter reconstruction FOV and reconstruction kernel FC44.

Images in Figure 5 show the central portion of a treatment plan image with the gross target volume (GTV) identified. Tumours / target regions can be generally more difficult to visualise than organs at risk due to a lack of the clear anatomical boundaries. Whilst the differences are subtle, there is an improvement in the visibility of the target due to improved radiographic contrast.



(a) Image showing position of GTV	(b) Original 550 mm FOV and FC13 image	(c) 200 mm FOV and FC44 image

Figure 5 An image from a head and neck scan showing the treatment planning Gross Target Volume GTV marked on image (a), the poor visibility of the GTV region with the 550 mm diameter FOV and FC13 reconstruction kernel in image (b) and the improved contrast and boundary of the GTV region seen with the 200 mm diameter FOV and reconstruction kernel FC44 in image (c).

Discussion

The comparison of phantom scores against the ranked clinical images, when choosing the reconstruction kernel for head and neck CT, indicated a clinical preference for kernels which gave both high contrast in the soft tissue CT number range and sharp edges. FC44 was one of the two kernels which scored highest for both contrast and image sharpness from the phantom measures and was chosen as the preferred kernel for the clinical images. It is notable that FC64 which was included in the clinical image review due to the high contrast results for some materials, had the lowest score for sharpness and ranked lowest in the clinical preference shortlist. NPS results identified that the noise texture was important, with fine noise pattern (higher frequency) being preferable. This highlights that high scores for contrast, sharpness and NPS frequency were important predictors of clinical choice for the preferred kernel.

The scoring of the 15 mm diameter target to determine the visibility of the edge was a good indicator of low contrast sharpness and the two scoring methods used (ranking and allocating scores for whole edge visibility) were effective at providing an initial selection of kernels to test on clinical images. The scoring of the visibility of different diameter low contrast targets within the nominal 1% contrast group was a relatively insensitive test of kernel performance and reconstruction FOV diameter, probably because a target is generally marked as seen even if the edge is poorly defined.

The choice of the Catphan phantom inserts to use when measuring contrast must be relevant to the radiographic contrast levels of interest in the clinical images. There can be differences in how the various kernels perform across different parts of the CT number scale. In this head and neck example, where soft tissue contrast definition is important, and where one of the objectives was to improve

spinal cord visibility, none of the inserts exactly matched the CT number range seen when measuring the contrast of the spinal cord (typical CT number range 20 to 80 HU). The phantom background material with the acrylic insert (nominal CT values 60 HU and 120 HU respectively) and the acrylic with the polystyrene inserts (nominal CT values 120 HU and -50 HU respectively) provided contrast values closest to that of the spinal cord. The results of both these contrast measures predicted the increase in the spinal cord contrast seen in the clinical images for FC44 compared to FC13.

The improvement in edge visibility with reduced diameter FOV was as expected since pixel size decreased. For a 200 mm FOV, pixel size was reduced to 0.39 mm compared to 1.07 and 1.37 mm with 550 and 700 mm FOVs respectively. As with kernel variation, the low contrast target visibility scores showed little variation with FOV diameter. This was thought to be because the primary indicator of low contrast visibility is the level of image noise, and noise changes are minimal with reconstruction FOV changes.

Although this work has been carried out on one particular make/model of CT scanner, all scanners have a range of reconstruction kernels available which produce different levels of edge enhancement, smoothing or noise reduction depending on the kernel selected. Siemens for example have several dozen kernels available which are intended for use with imaging the head, ranging from the very smooth H10s to the very sharp H90s kernel. GE and Philips both have fewer than ten available kernels, but that may still provide scope for improvement of image quality through choice of reconstruction kernel. All scanners have flexibility on the reconstruction FOV diameter which can be selected. The method of image optimisation is therefore applicable to all CT scanners, irrespective of make and model.

Conclusion

The use of a reduced diameter image reconstruction FOV clearly improves the quality of head and neck images for delineation in radiotherapy treatment planning. This improvement can be easily implemented through production of a second set of reconstructed images for use when delineating the brain region, alongside the larger FOV image set used for treatment planning. Both image sets can be produced from the same single scan raw data set so there are no concerns about image registration or additional CT radiation dose to the patient. The reconstruction kernel used for the second image set can also be optimised to improve image contrast since there is no impact on the primary image set used for planning the treatment.

This paper has presented a method to choose the optimal reconstruction kernel, reconstruction FOV diameter and described the workflow for clinical implementation. Improving image contrast and sharpness are important when optimising head and neck images. Even in centres where it is not possible to undertake the phantom based work, clinical image quality can be compared for a collection of patient images reconstructed with different kernels and FOVs to determine the optimal parameters. The advantage of the preliminary measurements with the phantom is that it is quicker to target possible best reconstruction parameters and therefore reduce the image evaluation work required by busy clinical staff.

Of the 46 centres which sent CT dose data and scan parameter information to a UK audit, no other centres declared use of a second reconstruction image set for head and neck imaging [6]. This is despite the reconstruction FOVs in use for this examination being large. The method outlined here has not been documented in the scientific literature and some guidance documents continue to reference only the use of large FOVs for radiotherapy head and neck imaging despite the detrimental impact on image quality. Use of a second reconstructed image set at reduced FOV, particularly where the very largest FOVs are in use, would universally improve the quality of head and neck CT images. Production of multiple image reconstructions is possible on all CT scanners, even older and lower specification scanners. It is particularly beneficial for delivering high quality head and neck imaging where it is not possible to use other imaging modalities such as MRI or PET to support pre-treatment imaging. Finally, the use of a reduced FOV image set has the potential for use when imaging other body regions and in the centre involved in this study has subsequently been implemented for use in gynaecological planning CT imaging.

References

[1] International Atomic Energy Agency. IAEA Human Health Series No 19, Quality Assurance for Computed Tomography: Diagnostic and Therapy Applications. International Atomic Energy Agency; 2012.
[2] Bolsi A, Placidi L. Optimal CT Protocols for CT-Guided Planning Preparation in Radiotherapy. Imaging and Interventional Radiology for Radiation Oncology: Springer; 2020. p. 27-45.
[3] Mutic S, Palta JR, Butker EK, Das IJ, Huq MS, Loo LN, et al. Quality assurance for computed-tomography simulators and the computed-tomography-simulation process: report of the AAPM Radiation Therapy Committee Task Group No. 66. Med Phys. 2003;30:2762-92.
[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] Davis AT, Palmer AL, Nisbet A. Can different Catphan phantoms be used in a multi-centre audit of radiotherapy CT image quality? Phys Med. 2020;78:38-47.
[6] 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.
[7] Brouwer CL, Steenbakkers RJ, Bourhis J, Budach W, Grau C, Gregoire V, et al. CT-based delineation of organs at risk in the head and neck region: DAHANCA, EORTC, GORTEC, HKNPCSG, NCIC CTG, NCRI, NRG Oncology and TROG consensus guidelines. Radiother Oncol. 2015;117:83-90.
[8] Breunig J, Hernandez S, Lin J, Alsager S, Dumstorf C, Price J, et al. A system for continual quality improvement of normal tissue delineation for radiation therapy treatment planning. Int J Radiat Oncol Biol Phys. 2012;83:e703-8.
[9] Zhu M, Bzdusek K, Brink C, Eriksen JG, Hansen O, Jensen HA, et al. Multi-institutional quantitative evaluation and clinical validation of Smart Probabilistic Image Contouring Engine (SPICE) autosegmentation of target structures and normal tissues on computer tomography images in the head and neck, thorax, liver, and male pelvis areas. Int J Radiat Oncol Biol Phys. 2013;87:809-16.
[10] McNair H, Franks K, van Herk M. On Target 2: Updated Guidance for Image-guided Radiotherapy. 2022.
[11] Kalender WA. Dose in x-ray computed tomography. Phys Med Biol. 2014;59:R129-50.
[12] Toshiba. Operation manual for Toshiba scanner Aquilion TSX 201A. Toshiba Medical Systems Corporation; 2008.

[13] Dance D, Christofides S, Maidment A, McLean I, Ng K. Diagnostic radiology physics: A handbook for teachers and students. Endorsed by: American Association of Physicists in Medicine, Asia-Oceania Federation of Organizations for Medical Physics, European Federation of Organisations for Medical Physics. Austria: IAEA; 2014.
[14] Platten DJ. Quantitative IQ ImageJ plugin. 2020.
[15] Wasbo E. ImageQC. https://github.com/EllenWasbo/ImageQC2022.
[16] Solomon JB. imQuest. 7.1 ed. https://deckard.duhs.duke.edu/.
[17] The International Commission on Radiation Units and Measurements. Radiation dose and image quality assessment in computed tomography - Report 87. Oxford: Oxford University Press; 2013.
[18] Barrett JF, Keat N. Artifacts in CT: recognition and avoidance. Radiographics. 2004;24:1679-91.