Optimisation of chest Computed Tomography using a phantom: impact of mAs and reconstruction techniques on Image Quality

C.S. Reis[1]; T. Faqir[2]; V. Harsaker[3]; P. Hogg [2]; L. Kristoffersen[3]; I.L. van Rein[4]; K. Stancombe[2]; N.C. Warmerdam[4]; C. Wergeland[3]

1   Escola Superior de Tecnologia da Saúde de Lisboa/Lisbon School of Health Technology (ESTeSL), Lisbon, Portugal
2   University of Salford, Salford, United Kingdom
3   Oslo and Akershus University College of Applied Sciences, Oslo, Norway
4   Department of Medical Imaging and Radiation Therapy, Hanze University of Applied Sciences, Groningen, the Netherlands

Abstract:

Objectives: To verify if the mAs and reconstruction techniques affect the visualisation of relevant structures in lung Computed Tomography (CT) using a phantom.

Methods: Images were acquired using various mAs and reconstruction techniques. Image quality (IQ) was analysed applying two approaches: perceptual, using 5 observers and objective (edge gradient calculation) to verify the sharpness of the structures. Dose was recorded. Wilcoxon Signed Rank test was used to compare the data from the perceptual image analysis. *P*-values were calculated (Bonferroni-Correction method) to compare reconstruction techniques and mAs. A Kappa Test with linear weighting was performed to calculate the level of agreement between observers.

Results: The Wilcoxon-Signed-Rank-Test showed no significant difference between the reconstruction techniques tested (p<0.05). In addition, the test showed no significant difference between any of the mAs values with a Bonferroni correction (p = 0.0167). For 10 mAs the observers scored differently, depending on which structures they were looking at. The overall IQ was acceptable and the nodules were well defined. The agreement for visualising the range of anatomical regions (Kappa test linear-weighting) suggests that observer 2 and 3 had a poor agreement level (0-0.366) and observer 1,4 and 5 had moderate agreement (0.5714-0.751).

Conclusion: The visual measures of IQ were largely unaffected by reconstruction techniques or mAs values. However, further work is needed for a better understanding of visual and clinical value of reconstruction techniques at lower doses.

Keywords: Lungs CT, reconstruction techniques, mAs, Image Quality, Optimisation.

Introduction

According to the Eurostat Database and the UK National Health Service, Computed Tomography (CT) is the radiological examination with the highest growth showing an increase of 10.3% in the UK alone for 10 consecutive years (1,2). The requests for CT scans has increased over time due to the improvements in detection of many pathologies (3). For this reason CT is used in screening programs such as lung and colon cancer detection, where asymptomatic patients are examined and early detection can be made (4). This increase in use has made optimisation a major topic. CT scans are associated with high radiation doses with an effective dose ranging from 2 to 16 mSv (5). These examinations may be associated with an increase in the risk of developing cancer, with a chance of approximately 1 in 2000 (6). In comparison, conventional radiography has a lower effective dose, ranging from 0.001 to 8 mSv for the more extensive exams (5). The increase in number of CT scans performed with the associated increase in risk is becoming a public health issue and for that reason it is important to reduce these risks by optimising the examinations according to the principle of 'As Low As Reasonably Possible/Practicable' (ALARP). Therefore, it is necessary to reduce dose while maintaining diagnostic image quality (IQ).

Manufacturers have implemented several techniques using both hardware and software in order to reduce dose without compromising IQ (7)we investigated whether images reconstructed using filtered back projection (FBP. One of the most recent strategies is the use of reconstruction techniques to improve the quality of images acquired with lower radiation dose. Filtered back projection (FBP) is frequently used for modern CT systems. FBP assumes the data is exact, but the projection data is noisy. The filter amplifies the noise and enhances or diminishes details on the image (8). This technique is considered an adequate method for reconstruction; however low doses or morbidly obese patients affect the performance of FBP, as they can promote artefacts. An alternative to FBP is iterative reconstruction (IR). Although this technique is not new, CT technology did not have the computational power to run this software until recently. IR can reduce dose by using algebraic reconstruction and is expected to allow imaging with similar noise levels and IQ as FBP (9).

There are several IR software solutions available and SAFIRE (Sinogram-Affirmed Iterative Reconstruction; Siemens Medical Solutions) is one of the most recent. SAFIRE is a hybrid technique that combines FBP and IR. Previous studies have shown that SAFIRE is capable of a 65% dose reduction without losing diagnostic information (10). The objectives of this study were to verify if the mAs and the reconstruction techniques affect the visualisation of anatomical details in lung CT exams using a phantom.

Methods
Image Acquisition
A multipurpose chest phantom (N1 "LUNGMAN"; Kyoto Kagaku) was used to produce the images (11). The phantom was positioned supine, head-first into the CT gantry and remained untouched during all acquisitions.

A Siemens Somatom Definition AS 128 slice CT scanner was used to acquire the images (12). The scanner was located at University Medical Centre in Groningen (UMCG). The scanner was warmed up and calibrated. All equipment used was subjected to the manufacturer specification for quality controls to ensure accuracy of the results. Six sets of 560 images were acquired (table1).

For each acquisition the Dose Length Product (DLP) was recorded. From the six sets provided, IQ analysis was only carried out on the three lower mAs values (10, 20, and 30 mAs). This was to verify if the observers could visualise various anatomical structures at a low mAs, which in turn meant a lower dose to the patient.

Exposure Parameters	Values
mAs	10, 20, 30, 40, 50, 66
kVp	120
Pitch	1.2
Slice Thickness	0.6mm
Matrix	512 x 512
Reconstruction Techniques	FBP, SAFIRE level 1, 3, 5
Body Kernel(13)	B31f, I31f
Reconstruction Plans	Axial, Coronal

Table 1 Exposure parameters used for image acquisition and reconstruction

Criteria	Likert scale used for each parameter
Lung edge	1 - It is not visible
Borders of larger vessels	2 - I can see it partially
Calcification in right main bronchi	3 - I can see it
Border of nodule	4 - It is clearly defined
Overall noise	1 - very poor: excessive noise or poor vessel wall definition
	2 - poor: poor vessel wall definition and prominent image noise
	3 - adequate: some image noise, vessel walls definition is minimal
	4 - good: minimal image noise definition of vessel walls are visible
	5 - very good: excellent definition of vessel walls, limited perceptual image noise
Overall image quality	1 - very poor: poor IQ due to artefacts, no definition between anatomical structures
	2 - poor; prominent artefacts, minimal definition between anatomical structures
	3 - adequate: minor artefacts present, definition between anatomical structures
	4 - good: no perceptual artefacts present, clear definition between anatomical structures
	5 - very good: no perceptual artefacts present, total definition between anatomical structures

Table 2 Criteria analysed by the observers and Likert scales provided

Perceptual IQ Analysis

The same axial and coronal slices were selected for each data set and analysed according to anatomical criteria provided by European guidelines (14), as well as for noise and overall IQ (table 2). Both axial and coronal slices were randomised, anonymised and four repeats were present in both axial and coronal data sets to determine the intra-observer-reliability. Slice selection was performed considering the anatomical details presented in each image.

A blind analysis of all images was undertaken by 5 qualified radiographers ranging in age of 31-58 years, with 5-32 years experience. Questionnaires were provided to all the observers to check whether they have had their eyesight tested within the last 12 months, if their eyesight was compromised and whether they wore glasses or contact lenses to correct it. The observers were trained using a presentation to show which relevant structures they had to analyse (figure 1 and 2). The images were randomised and the observers had to verbalise their answers. Three researchers were present at the time of scoring; one to train the observer and select the images, a second to manually enter the data from the observers and a third to monitor the two researchers to minimise error.
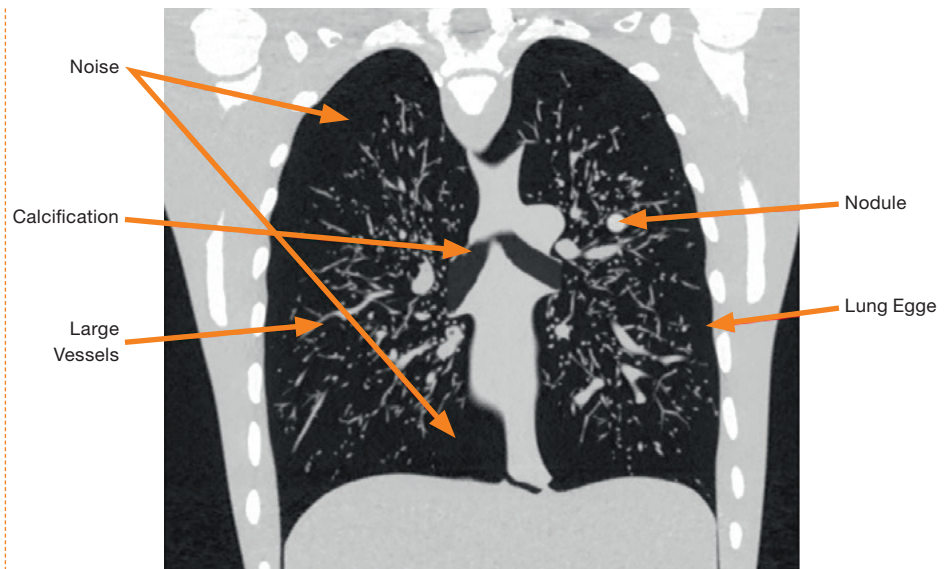

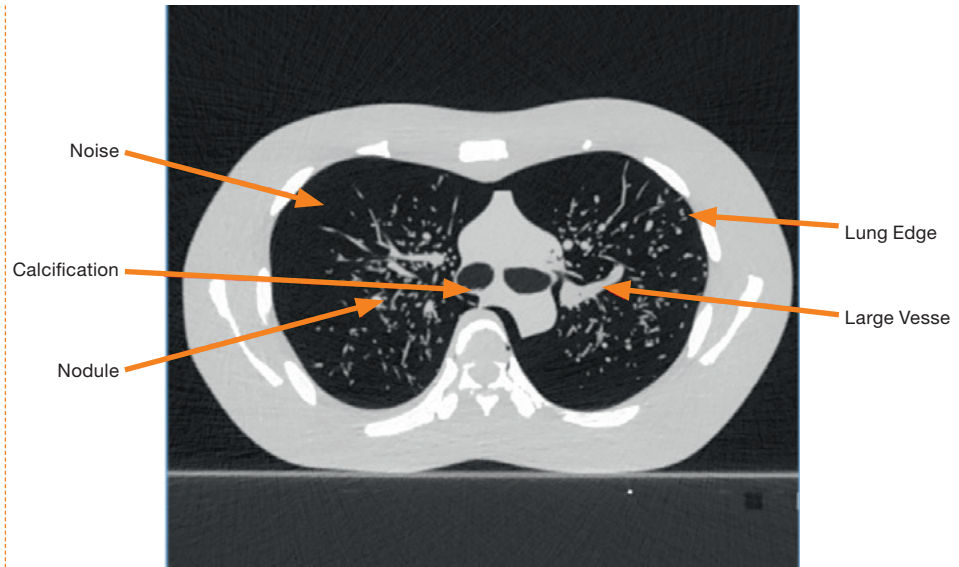**Figure1** Example of a coronal image scored by observers

Figure 2 Example of anaxial image scored by observers

For all images, the scores were totaled in order to obtain a global score for each image. For questions 1-4 the global score was given at max=16, whereas for questions 5-6 the global score for each image was given at max=5. The scores were set in order to give an overall representation of all answers and observers combined. Since the scores did not differ significantly, the overall scoring is considered valid for comparison.

Two monitors were used, one for the axial and one for the coronal views. Images were viewed using calibrated Diagnostic 24.1" EIZO monitors with 1920 x 1200 pixels and the images were loaded using a DICOM Viewing Software. All images were set to the CT lung window at a window width of 1500 and a window level of -400 similar to clinical practice(15). The observers were not allowed to manipulate the images and had to keep their distance from the monitor constant to keep the same conditions for all observers. The room lights were turned off to prevent any light reflecting onto the monitors and there was no noise in the room to distract the observers.

Objective IQ Analysis
To mathematically calculate how reconstruction techniques affect the edge definition of each anatomical structure, measurements were made using *ImageJ* software on the nodule, larger vessel

and the lung edge (16). A line was drawn from a low contrast point across the border of the structure to a high contrast point within the structure(figure 3). The middle of the line was placed on the visible outline of the structure and remained the same in each image. To analyse the pixel value a plot profile was created (figure 4). A trend line was added to the linear points in the plot profile (figure 5) from which the edge gradient was calculated using Microsoft Excel(16). The

difference between the edge gradients was converted into percentages. This procedure was replicated in all axial images.

Statistical Data Analysis

All the data was analysed using IBM SPSS Statistics Version 22 and Microsoft Excel. For the ordinal data a non-parametric test, the Wilcoxon Signed Rank test, was used to compare the data from the subjective

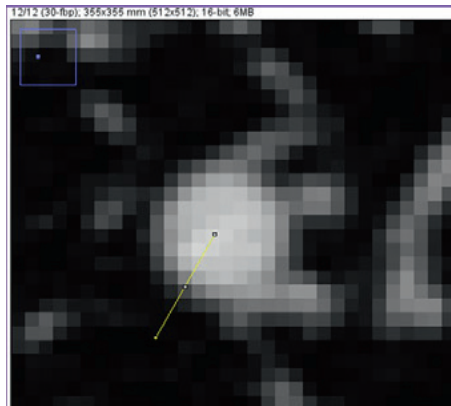**Figure 3** Line drawn from low contrast to high contrast in nodule

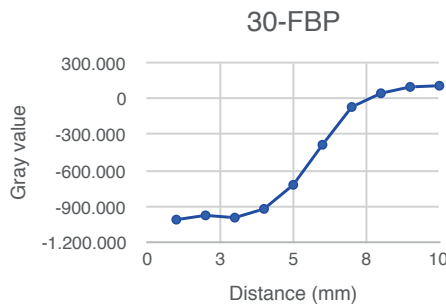**Figure 4** Graph showing the 30 mAs with FBP plot profile of the nodule

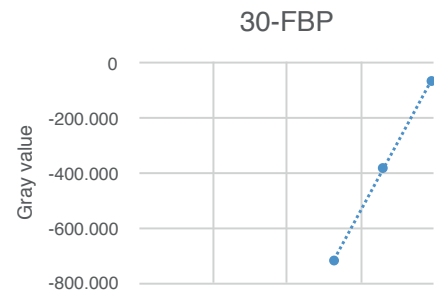**Figure 5** Graph showing the trend line from the 30 mAs with FBP nodule plot profile

Kappa value	Description
0	Same as expected by chance
< 0.40	Poor
0.40 – 0.75	Moderate
> 0.75	Excellent
1	Perfect

Table 3 Levels of Kappa values (18)

image analysis. P-values for the reconstruction techniques and mAs values were corrected with the Bonferroni Correction method. For the reconstruction techniques a p-value of <.0083 was considered significant and for the mAs values a p-value of <.0167 (17).

In order to determine the intra-observer reliability, four images were shown twice in a random order. A Kappa Test with linear weighting was performed to calculate the level of agreement, which in turn impacts the reliability of the observers (table 3).

Results
Visualisation of anatomical structures
The anatomical structures were scored using a 4-point Likert scale, with 3 being considered visible and therefore a level of acceptance for clinical practice. The values of each question were added up for all images, giving a maximum score of 16 and a level of acceptance at 12 (blue line in figures 6 and 7). However, partial identification of the anatomical structures was still possible when scored above 8 for some clinical applications.

The standard deviation shows that each reconstruction technique and mAs value causes variation in visibility, but are all still within the acceptance level. However, there was greater variation in the visualisation for the axial compared to the coronal images (figure 6 and 7).

The scores verify that some of the reconstruction techniques and mAs values compromise the partial visibility of structures, mainly at 10 mAs. For axial images reconstructed with FBP, the scores do not meet the level of acceptance in the visualisation with 10 mAs (figure 7).The results also demonstrate that the highest score was observed with 20 mAs and Safire 5 reconstruction.

The Wilcoxon Signed Rank Test showed no significant difference between the reconstruction techniques except between FBP and SAFIRE 3 ($p = 0.002$).

Figure 6 Visualisation of anatomical structures in coronal images comparing the mAs range (10-30) and 4 reconstruction techniques (FBP and Safire 1, 3, 5)

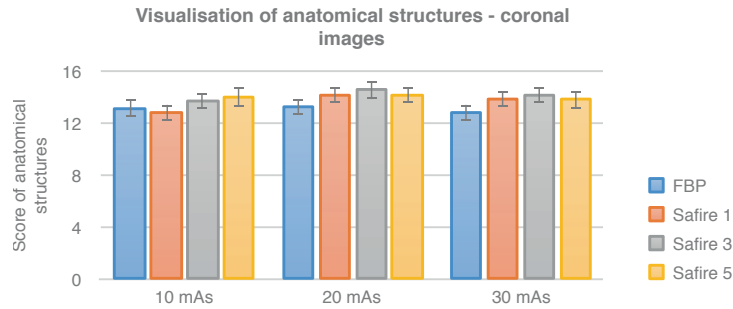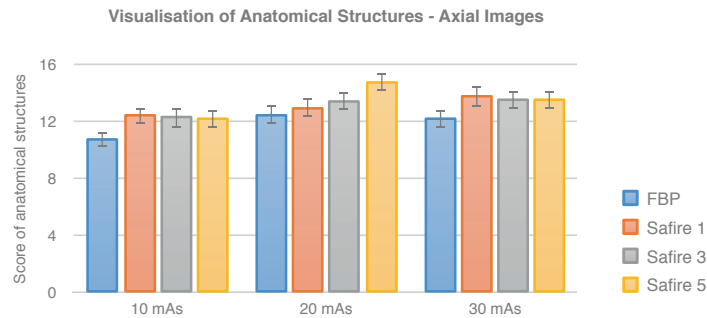Visualisation of anatomical structures - coronal images

Figure 7 Visualisation of anatomical structures in axial images comparing mAs range (10-30) and 4 reconstruction techniques (FBP and Safire 1, 3, 5)

Visualisation of Anatomical Structures - Axial Images

In addition, the test showed no significant difference between any of the mAs values with a Bonferroni correction ($p = 0.0167$).

Visualisation of image noise

FBP was compared with the SAFIRE levels used for this study and comparisons were made between these levels (figure 8). This suggests there is a reduction in image noise as mAs increases. Furthermore, it demonstrates that SAFIRE 5 has less overall image noise compared to the other reconstruction techniques for 10 and 20 mAs.

Looking at the raw data, the image noise was scored adequate, good and very good at 93.3% or higher for all mAs values per reconstruction technique. The Wilcoxon Signed Rank Test showed no significant difference between any of the reconstruction techniques except between FBP and SAFIRE 5 where there is a significant difference (FBP with SAFIRE 1, 3 and 5 respectively: $p = 0.033$; $p = 0.018$; $p = 0.001$; SAFIRE 1, 3 and 5: $p = 0.491$; $p = 0.124$; $p = 0.384$).

Figure 8 Bar chart demonstrating combined axial and coronal overall perceptual image noise score for each reconstruction technique at varying mAs values
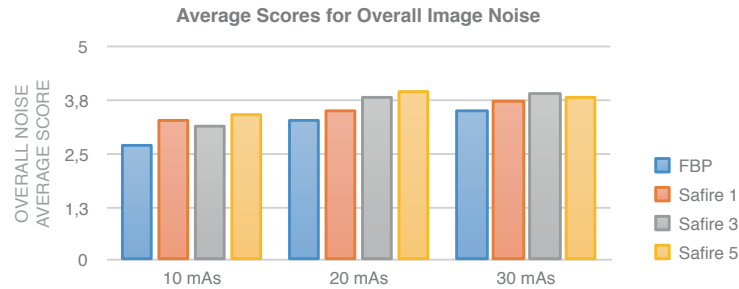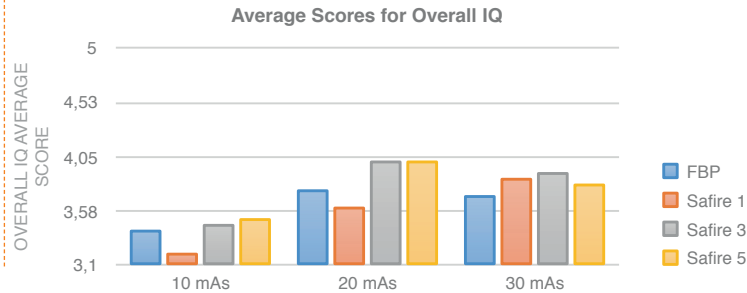
Average Scores for Overall Image Noise

OVERALL NOISE AVERAGE SCORE

FBP
Safire 1
Safire 3
Safire 5

Figure 9 Bar chart demonstrating axial and coronal IQ score combined for each reconstruction technique at varying mAs values

Average Scores for Overall IQ

OVERALL IQ AVERAGE SCORE

FBP
Safire 1
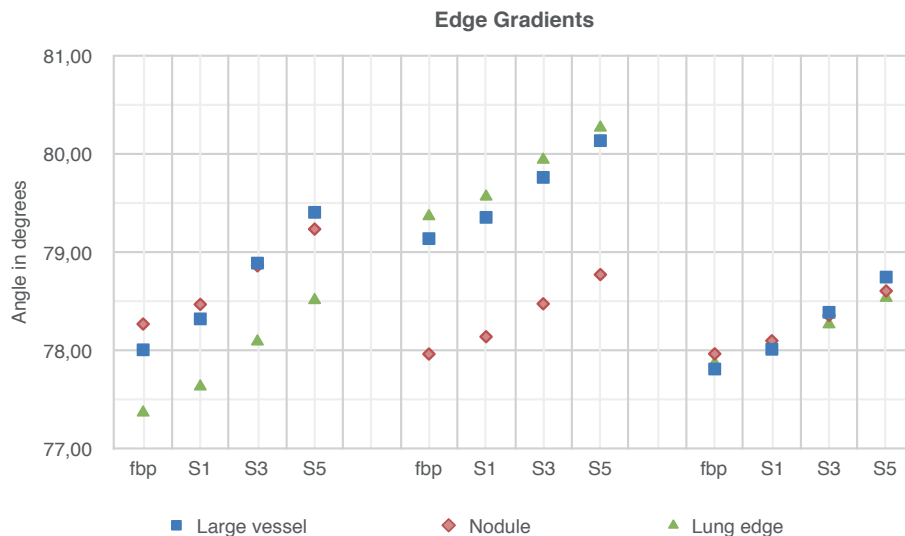Safire 3
Safire 5

Overall Image Quality

The overall IQ score is higher for 20 and 30 mAs compared with 10 mAs (figure 9). It also suggests that SAFIRE 3 produces images with higher quality than the other reconstruction techniques for 20 and 30 mAs. Just as with perceptual image noise, the observers scored the overall IQ at 93,3% or higher in the form of adequate, good and very good. The reconstruction techniques showed no significant difference between them as demonstrated by the Wilcoxon Signed Rank Test (FBP with SAFIRE 1, 3 and 5 respectively: $p = 0.405$; $p = 0.251$; $p = 0.083$; SAFIRE 1,3 and 5: $p = 0.046$; $p = 0.926$).

Objective Image Quality

The edge gradient increases when the reconstruction technique changes from FBP to SAFIRE 5 (figure 10). The sharpness of the structure is higher when the edge gradient is closer to 90°(16). This suggests that overall SAFIRE 5 at 20 mAs has a sharper outline in comparison to the other reconstruction techniques and mAs levels.

Edge Gradients

Figure 10 The calculated edge gradient against every reconstruction technique for every mAs value

The graph also shows that the biggest difference in edge gradients is between FBP and SAFIRE 5 for all mAs levels. The calculated differences between the different reconstruction techniques are minor, with a maximum increase of 1.79% (table 4).

Intra-observer reliability

The Kappa test with linear weighting suggests that observer 2 and 3 had a poor agreement level. The Kappa value for the coronal set of observer 2 could not be calculated. These observers were not excluded from the study, because of their high level of clinical experience as radiographers in CT departments. The remaining observers scored moderate for the kappa value (table 5). The kappa value of the observer 1, 4 and 5 is considered moderate.

Table 4 Difference in edge gradients between FBP and Safire 5 expressed in percentages.

mAs	Comparison of Reconstruction Techniques	Large Vessel	Nodule	Lung edge
10	FBP - SAFIRE 5	1.79%	1.23%	1.47%
20	FBP - SAFIRE 5	1.26%	1.03%	1.13%
30	FBP - SAFIRE 5	1.20%	0.82%	0.85%

	Axial	Coronal
Observer 1	0.6364	0.6924
Observer 2	0.366	N/A
Observer 3	0.1667	0.3333
Observer 4	0.5714	0.6471
Observer 5	0.7551	0.7097

Table 5   The kappa value calculated for each observer

Dose Length Product (DLP)

The DLP for the acquired images varied between 29.3, 58.6 and 87.9 mGycmfor 10, 20 and 30 mAs respectively (table 6).

Discussion

On the whole, FBP and SAFIRE 1, 3 and 5 with all mAs combinations demonstrated no significant differences in overall perceptual IQ (figure 6 and 7). For 10 mAs the observers scored different, depending on which structures they were looking at. The overall IQ was acceptable and the nodules were well defined (appendix 1). These findings are supported by other studies (19,20)bronchial polyp, solid nodule, ground glass nodule, emphysema and tree-in-bud. However, the observers could not see the calcification completely. This assumes that mAs should be considered depending on what the clinical indication is for the CT examination and also pathology protocol. Furthermore, when FBP was compared with SAFIRE, the visualisation of anatomical structures was also less defined when using FBP at 10 mAs in axial images (figure 7). This is supported by the calculated edge gradients (figure 10) and by other authors (9,21) due to the noise increase when using FBP.

This phantom based study gives an indication of potential detection of relevant structures in the clinical context for all reconstruction techniques at reduced mAs and dose. European guidelines recommend doses for CT lung below 650 mGycm. This research shows that a dose reduction of 95.5% is possible at

mAs	DLP (mGycm)	% of dose reduction against European Guidelines (650 mGycm)
10	29.3	95.5%
20	58.6	91.1%
30	87.9	86.5%

Table 6   The recorded dose for each mAs value

10 mAs (table 6). When considering the overall IQ score, a dose reduction of 91.1% can be achieved at 20mAs whilst still maintaining anatomical structure clarity. At 20 mAs, with an effective dose of 29.3 mGycm, screening for the early detection of cancer would be less harmful and spare the patient from unnecessary ionising radiation. When comparing the findings from this study with the European guidelines it is clear that it would be reasonable, as well as practicable, to lower the recommended dosage.

There were several limitations in this study, one of which was that this research was conducted on a phantom. When using a phantom the motion, breathing and heartbeat artefacts are not simulated. Also the simulated lesions are well defined and detection can be more obvious when compared to clinical exams. In addition, patients vary in size and tissue density as opposed to a phantom.

Another limitation of this study is related to the subjective IQ analysis (table 5). Observer 2 had a very low kappa value for the repeated axial images. The reliability of kappa is reduced due to few points. For observer 2 no weighted kappa could be calculated because the observed agreement was lower than the expected agreement (18,22). Subjective IQ analysis can also be influenced by the background training of the radiographers (23).

This study showed that visualisation of anatomical structures was possible even at a low mAs value of 20, and that partial visibility was made at 10 mAs. Therefore future research needs to consider values between 10 and 20mAs. Future research should include a bigger variety in clinical indications, patient size and exposure parameters (pitch, slice thickness and kVp).

Conclusion
The visual measures of IQ were largely unaffected by reconstruction techniques or mAs values. However, further work is needed for a better understanding of visual and the clinical value of reconstruction techniques at lower doses.

Acknowledgments

Bibliography

1. Commission E. EUROSTAT - Your key to European statistics [Internet]. Eurostat. 2015 [cited 2015 Aug 5]. p. Needs for health care. Available from: http://ec.europa.eu/eurostat/web/health/health-care/data/database

2. Steel P. NHS Imaging and Radiodiagnostic activity in England [Internet]. 2012 [cited 2015 Aug 12]. Available from: http://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2013/04/KH12-release-2012-13.pdf

3. Schmidt CW. CT scans: balancing health risks and medical benefits. Environ Health Perspect. 2012;120(3):118–21.

4. Team TNLSTR. Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. N Engl J Med [Internet]. 2011 Aug 4;365(5):395–409. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMoa1102873

5. Mettler F a, Huda W, Yoshizumi TT, Mahesh M. Effective doses in radiology and diagnostic nuclear medicine: a catalog. Radiology. 2008;248(1):254–63.

6. Services USD of H and H. Administration, U.S. Food and Drug [Internet]. 2015 [cited 2015 Aug 6]. Available from: http://www.fda.gov/Radiation-EmittingProducts/RadiationEmittingProductsandProcedures/MedicalImaging/MedicalX-Rays/ucm115329.htm

7. Klink T, Obmann V, Heverhagen J, Stork A, Adam G, Begemann P. Reducing CT radiation dose with iterative reconstruction algorithms: The influence of scan and reconstruction parameters on image quality and CT-DIvol. Eur J Radiol [Internet]. Elsevier Ireland Ltd; 2013;83(9):1645–54. Available from: http://dx.doi.org/10.1016/j.ejrad.2014.05.033

8. Fleischmann D, Boas FE. Computed tomography—old ideas and new technology. Eur Radiol [Internet]. 2011 Mar;21(3):510–7. Available from: http://link.springer.com/10.1007/s00330-011-2056-z

9. Willemink MJ, De Jong P a., Leiner T, De Heer LM, Nievelstein R a J, Budde RPJ, et al. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. Eur Radiol. 2013;23(6):1623–31.

10. Kalra MK, Woisetschlager M, Dahlstrom N, Singh S, Digumarthy S, Do S, et al. Sinogram-Affirmed iterative reconstruction of low-dose chest CT: Effect on image quality and radiation dose. Am J Roentgenol. 2013;201(2):235–44.

11. Kyoto Kagaku Co. Patient Simulators, Imaging Phantoms for Skill Training [Internet]. 2012 [cited 2015 Aug 18]. Available from: https://www.kyotokagaku.com/products/detail03/ph-1.html

12. Siemens. SOMATOM Definition AS [Internet]. 2015 [cited 2015 Aug 17]. Available from: http://www.healthcare.siemens.nl/computed-tomography/single-source-ct/somatom-definition-as

13. Schabel C, Fenchel M, Schmidt B, Flohr TG, Wuerslin C, Thomas C, et al. Clinical Evaluation and Potential Radiation Dose Reduction of the Novel Sinogram-affirmed Iterative Reconstruction Technique (SAFIRE) in Abdominal Computed Tomography Angiography. Acad Radiol [Internet]. Elsevier Ltd; 2013;20(2):165–72. Available from: http://dx.doi.org/10.1016/j.acra.2012.08.015

14. Commission E. European Guidelines on Quality Criteria for Computed Tomography European Guidelines on Quality Criteria [Internet]. Menzel H, Jessen K, Panzer W, Shripton P, Tosi G, editors. Europe. European Commission´s Radiation Protection Actions; 1999. 1-71 p. Available from: http://www.msct.info/CT_Quality_Criteria.htm

15. Radiantviewer.com. Radiant-Viewer [Internet]. [cited 2015 Aug 21]. Available from: http://www.radiantviewer.com/dicom-viewer-manual/change_brightness_contrast.htm

16. Manning DJ, Ethell SC, Donovan T. Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph [Internet]. The British Journal of Radiology. 2004. 231-235 p. Available from: http://www.birpublications.org/doi/abs/10.1259/bjr/28883951

17. Field A. Discovering Statistics Using IBM SPSS Statistics. 4th editio. SAGE Publications Inc.; 2013. 228-235 p.

18. Fleiss JL, Levin B, Paik MC. Statistical methods for rates and proportions. 2nd ed. New York: John Wiley; 2003.

19. Botelho MPF, Agrawal R, Gonza-lez-Guindalini FD, Hart EM, Patel SK, Töre HG, et al. Effect of radiation dose and iterative reconstruction on lung lesion conspicuity at MDCT: Does one size fit all? Eur J Radiol [Internet]. Elsevier Ireland Ltd; 2013;82(11):e726–33. Available from: http://dx.doi.org/10.1016/j.ejrad.2013.07.011

20. Christe A, Torrente JC, Lin M, Yen A, Hallett R, Roychoudhury K, et al. CT screening and follow-up of lung nodules: Effects of tube current-time setting and nodule size and density on detectability and of tube current-time setting on apparent size. Am J Roentgenol. 2011;197(3):623–30.

21. Baumueller S, Winklehner A, Karlo C, Goetti R, Flohr T, Russi EW, et al. Low-dose CT of the lung: potential value of iterative recon-structions. Eur Radiol [Internet]. 2012 Dec;22(12):2597–606. Available from: http://link.springer.com/10.1007/s00330-012-2524-0

22. Altman D. Practical Statistics for Medical Research. London: Chapman and Hall; 1991.

23. Kakinuma R, Ashizawa K, Kobayashi T, Fukushima A, Hayashi H, Kondo T, et al. Comparison of sensitivity of lung nodule detection between radiologists and technologists on low-dose CT lung cancer screening images. Br J Radiol [Internet]. 2012 Sep;85(1017):e603–8. Available from: http://www.birpublications.org/doi/abs/10.1259/bjr/75768386

Appendix 1

mAs	Score	Lung edgeVessel		Calcif	Nodule
		q1_axial	q2_axial	q3_axial	q4_axial
		Frequency			
10	Not visible	0	0	0	0
	See partially	2	3	8	1
	Visible	14	16	10	16
	Clearly defined	4	1	2	3
20	Not visible	0	0	0	0
	See partially	0	4	1	0
	Visible	8	12	13	8
	Clearly defined	12	4	6	12
30	Not visible	0	0	0	0
	See partially	0	7	1	0
	Visible	10	6	16	6
	Clearly defined	10	7	3	14

Table 5 The kappa value calculated for each observer