Research article – A comparison of Sinogram Affirmed Iterative Reconstruction and filtered back projection on image quality and dose reduction in paediatric head CT: a phantom study

Abdulfatah Ahmed[a], André Garcia[b], Astrid Bakker[c], David Tomkinson[a], Julie Salamin[d], René de Lange[c], Sergey A. Buyvidovich[e], Tina Sohrabi[e], Alexandre Dominguez[d], Cosmin Campeanu[d], Paul Plasman[c]

a) *School of Health Sciences, University of Salford, Manchester, United Kingdom*
b) *Lisbon School of Health Technology (ESTeSL), Polytechnic Institute of Lisbon, Portugal*
c) *Department of Medical Imaging and Radiation Therapy, Hanze University of Applied Sciences, Groningen, The Netherlands*
d) *Haute École de Santé Vaud – Filière TRM, University of Applied Sciences and Arts of Western Switzerland, Lausanne, Switzerland*
e) *Department of Life Sciences and Health, Radiography, Oslo and Akershus University College of Applied Sciences, Oslo, Norway*

KEYWORDS

Comparison
Filtered back projection
Sinogram-Affirmed
Iterative reconstruction
Dose reduction
Paediatric CT
Computed tomography
Image quality

ABSTRACT

Background: Computed tomography (CT) is one of the most used modalities for diagnostics in paediatric populations, which is a concern as it also delivers a high patient dose. Research has focused on developing computer algorithms that provide better image quality at lower dose. The iterative reconstruction algorithm Sinogram-Affirmed Iterative Reconstruction (SAFIRE) was introduced as a new technique that reduces noise to increase image quality.

Purpose: The aim of this study is to compare SAFIRE with the current gold standard, Filtered Back Projection (FBP), and assess whether SAFIRE alone permits a reduction in dose while maintaining image quality in paediatric head CT.

Methods: Images were collected using a paediatric head phantom using a SIEMENS SOMATOM PERSPECTIVE 128 modulated acquisition. 54 images were reconstructed using FBP and 5 different strengths of SAFIRE. Objective measures of image quality were determined by measuring SNR and CNR. Visual measures of image quality were determined by 17 observers with different radiographic experiences. Images were randomized and displayed using 2AFC; observers scored the images answering 5 questions using a Likert scale.

Results: At different dose levels, SAFIRE significantly increased SNR (up to 54%) in the acquired images compared to FBP at 80kVp (5.2-8.4), 110kVp (8.2-12.3), 130kVp (8.8-13.1). Visual image quality was higher with increasing SAFIRE strength. The highest image quality was scored with SAFIRE level 3 and higher.

Conclusion: The SAFIRE algorithm is suitable for image noise reduction in paediatric head CT. Our data demonstrates that SAFIRE enhances SNR while reducing noise with a possible reduction of dose of 68%.

INTRODUCTION

Computed Tomography (CT) is fast, precise and one of the most used modalities for diagnostic imaging[1]. It is considered the technique of choice both in adult and paediatric population, but it is also associated with high effective dose. 7 million paediatric CT scans were performed in 2007 in the USA and with this value rises almost 10% every year. Dose reduction for paediatric examinations has become a priority, since younger patients have a higher potential for an increased lifetime risk of radiation-induced malignancy[2]. However, the dose reduction should not compromise diagnostic image quality.

Different CT reconstruction algorithms have been developed over the years. Filtered back projection (FBP) is by far the most used today[3]. It is fast and applies different filters[4] and increases spatial resolution. It also increases image noise requiring a higher X-ray tube setting, resulting in a higher radiation exposure[5].

Iterative reconstruction (IR) techniques have been developed to reduce dose while maintaining or improving objective image quality by reducing noise and consequently improving Signal-to-Noise Ratio (SNR). However, these techniques require a high computing power and have been too time consuming, limiting its clinical application[6].

Table 1: Acquisition parameters for all the images

Batch	1	2	3
kV	80	110	130
mAs	50	50	50
	100	100	100
	200	200	200

Siemens has recently developed Sinogram-Affirmed Iterative Reconstruction (SAFIRE), an advanced iterative reconstruction technique that requires less computing power and uses both FBP and raw data-based iterations to remove noise and improve image quality. The corrected image is compared with the original and the process is repeated several times. SAFIRE provides 5 strength levels of noise reduction, 1 being the weakest and 5 the strongest[7]. Noise reduction and noise texture vary with the reconstruction strength selected. Strength value does not translate the number of iterations and does not affect reconstruction time[6]. Nonetheless, SAFIRE is still more time consuming than golden standard FBP[2].

This study aims to compare SAFIRE to FBP and determine whether SAFIRE permits a reduction in dose while maintaining image quality in paediatric head CT.

MATERIALS AND METHODS

CT protocol selection

For this study a simulated examination was performed using the paediatric head phantom ATOM of a five year old child, model 705[8], to acquire a series of image sets[9]. The equipment used for the simulation was a Siemens SOMATOM perspective 128 CT scanner. Prior to the study, three image sets were created using different settings for kV and mAs, indicated in Table 1.

Data reconstruction

All the image sets were reconstructed using FBP and SAFIRE strengths 1, 2, 3, 4, and 5. The kernels used for image reconstruction were 'smooth' (J30 for SAFIRE and H30 kernel for FBP). This study focused mainly on the soft tissue kernel, as it allows a better representation of soft tissue structures. Sharp kernels, used for bone tissue studies add too much image noise. All images were reconstructed using 1 and 3mm slice thickness but only 3mm thickness were selected for the final three image sets, in order to keep in-line with those used in clinical practice[10-11]. All reconstructed images were acquired from the same original dataset. Window width 50, window level 100 and field of view were kept constant for visual image quality analysis.

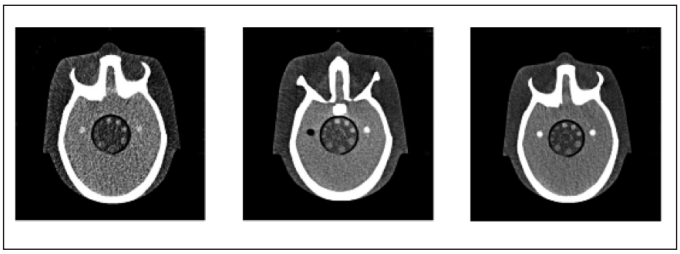


Figure 1: Reference images; FBP 80 kV 100 mAs (left), FBP 110 kV 100 mAs (middle), and FBP 130 kV 100 mAs (right).

Visual assessment of image quality

For visual analysis, a forced choice technique was used for image comparison, using two-alternative forced choice (2AFC) software[12-13]. Nine sets of parameters and six types of reconstructions totalled 54 images, which were scored by 17 observers. The 54 images were divided into three batches, each batch containing 18 images with the same kVp. In every batch, two duplicated images were included, which allowed for intra observer reliability to be assessed. Slice 18 was chosen for evaluation, through consensus decision, as it was considered that this slice represented a good visualisation of inner structures within the phantom. The SNR was calculated on all FBP images and the one with the average SNR was used as the reference image for each batch, which identified the 100mAs image, for every kVp value (Figure 1).

Eight inexperienced observers (first and second year radiography students and one high school graduate) and nine experienced observers (seven graduate radiographers and one physicist specialized in medical imaging) independently scored the images. All observers took part in a visual acuity test where a series of visibility, depth and contrast perception tests were performed to assess the visual capacity of each observer. These tests were performed to exclude observers with impaired visual acuity.

Table 2: Questionnaire used for evaluation of images

Item 1	The differentiation of contrast within the encircled structure and its immediate background (see image)?
Item 2	The visual sharpness of the bony structure displayed?
Item 3	The sharpness of the circle in the encircled structure (see image)?
Item 4	The overall image quality?
Item 5	The overall noise of the image?

Before visual image quality assessment, each individual observer received verbal instructions and a reference sheet to allow a better understanding of specific image areas evaluated in each question (appendix A). Each image was evaluated separately and compared with the reference image. The questionnaire for this evaluation contained one question for counting the circles in the central area of the phantom and five questions with a 5 point Likert scale which were used for scoring image noise, contrast and overall image quality as perceived by each individual observer (Table 2)[10,14-15]. All observers were blinded to acquisition parameters and images were displayed in a randomized fashion.

Light in the room was dimmed and constant; monitor settings (Siemens, 19 inch, resolution: 1280x1024) were kept identical for all observers, in order to further reduce observer bias.

Physical assessment of image quality

For the physical analysis SNR, uniformity and CNR measurements were performed in each image. Using the software ImageJ, three Regions of Interest (ROI's) were drawn with a 10-mm2 area for SNR calculation (Figure 2). Two other ROI's were drawn on bone and soft tissue and were used to obtain CNR values (Figure 4). ROI's were drawn in specific areas and in the same place for every image (approximately), in order to obtain better contrast measurements and to facilitate the evaluation of image quality parameters. SNR and CNR were calculated using the following standard equations:

$$CNR = \frac{Mean\ signal\ bone - mean\ signal\ tissue}{Mean\ SD}$$

$$SNR = \frac{Mean\ signal\ value\ within\ ROI}{SD\ within\ ROI}$$

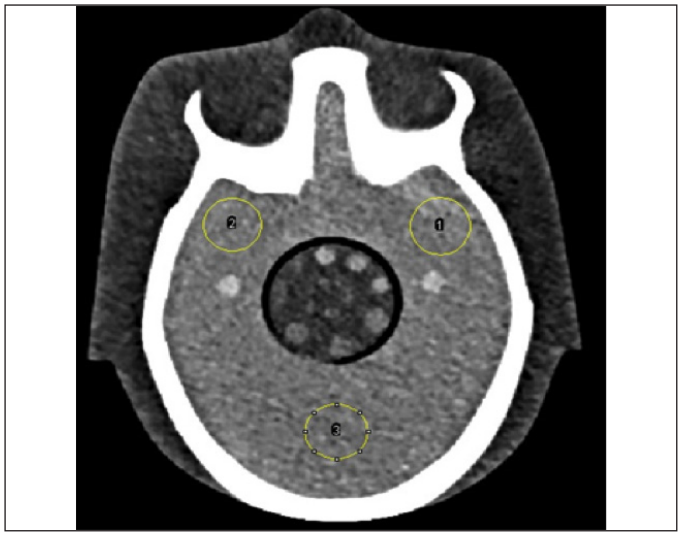


Figure 2: *ROI's used for SNR measurements.*

Figure 4: *ROI drawn to measure CNR.*

CT-EXPO dosimetry software for Monte Carlo modelling for calculating the dose[16].

Statistical Analysis

The data acquired from the visual assessment of image quality (2AFC) was imported into SPSS software (version 21.0) for statistical analysis. The data was analysed to assess a 95% confidence interval and for each IR. Descriptive statistics were used to determine the participant visual perception of the change in IR and the effect on image quality. Correlations and p-values were calculated using Microsoft Excel (2013).

RESULTS

Physical assessment of image quality

Physical image quality measures indicate the image noise was higher in the FBP compared to the SAFIRE groups (Table 1). Furthermore, a decrease in noise was observed with increasing SAFIRE strength. The decrease in noise with the increase in SAFIRE strength precipitated an increase in SNR. CNR remained constant, with minimal change in standard deviation. SAFIRE groups 3, 4, and 5 show comparable noise and SNR values.

Dose assessment

Table 3 highlights the increase in mAs within kVp values resulting in a proportional increase in DLP, effective dose (E) and the effective dose to the eyes, with a reduction at low mAs values. Figure 5 illustrates the correlation between DLP and effective dose with an r2 = 0.996, with the effective dose to the eyes displaying a correlation of r2 = 1.

Table 3: The relationship between kVp/mAs and effective dose

kVp	mAs	DLP (mGy*cm)	E(mSv)(103)	E eye (mSv)
80	50	44	0.1	3.4
80	100	88	0.2	6.9
80	200	175	0.5	13.7
110	50	103	0.3	8.1
110	100	207	0.5	16.2
110	200	413	1.1	32.3
130	50	154	0.4	12.1
130	100	309	0.8	24.2
130	200	618	1.6	48.3

Visual assessment of image quality

The division of the observers in terms of experience allowed for comparative analysis between the two groups (Table 4). Table 5 provides the results of the calculated Intra Class Correlation Coefficient (ICC) values for the experienced and non-experienced group of observers. All ICC values have a statistical significance of $p < 0.001$. The values of the intra class correlation (ICC) show weak to moderate agreement within all subject groups, where the group of experienced subjects shows a higher ICC for every criteria. The biggest difference in ICC between experienced and non-experienced observers can be observed in criteria 1, whereas criteria 4 and 5 show similar values.

Figure 6 illustrates the relationship between the mean IQ and the acquisition parameters (kVp/mAs) for FB and the SAFIRE groups with a 95% Confidence Interval (CI). An increase in the mean perceptual IQ is prominently demonstrated with increases in SAFIRE strength.

FBP and SAFIRE 1 are comparable as they display similar mean IQ with increments in parameters; however a stronger confidence interval is displayed within SAFIRE 1. SAFIRE 2 shows an increase in mean IQ at lower parameters, however this increase is more obvious with SAFIRE 3, 4 and 5.

Table 4: Data representing mean value (±SD) of image noise, CNR, SNR

		FBP	SAFIRE 1	SAFIRE 2	SAFIRE 3	SAFIRE 4	SAFIRE 5
80 KVP							
	Noise	11.1(±3.3)	10.3(±3.1)	9.6(±2.8)	8.9(±2.7)	8.2(±2.5)	7.5(±2.3)
	CNR	0.8(±0.4)	0.8(±0.5)	0.8(±0.5)	0.8(±0.5)	0.8(±0.5)	0.8(±0.6)
	SNR	5.2(±1.5)	5.7(±1.7)	6.2(±1.9)	6.9(±2.1)	7.6(±2.3)	8.4(±2.6)
110 KVP							
	Noise	6.2(±1.6)	5.7(±1.6)	5.3(±1.4)	5.0(±1.4)	4.7(±1.5)	4.2(±1.3)
	CNR	0.5(±0.4)	0.5(±0.5)	0.5(±0.5)	0.5(±0.6)	0.5(±0.7)	0.4(±0.7)
	SNR	8.2(±2.0)	8.8(±2.1)	9.4(±2.4)	10.2(±2.6)	11.2(±3.0)	12.3(±3.5)
130 KVP							
	Noise	4.9(±1.0)	4.6(±0.9)	4.2(±0.9)	3.9(±0.8)	3.6(±0.7)	3.2(±0.7)
	CNR	0.5(±0.2)	0.5(±0.2)	0.5(±0.2)	0.6(±0.2)	0.5(±0.2)	0.5(±0.3)
	SNR	8.8(±1.1)	9.4(±1.2)	10.2(±1.3)	11.0(±1.3)	12.0(±1.5)	13.1(±1.7)

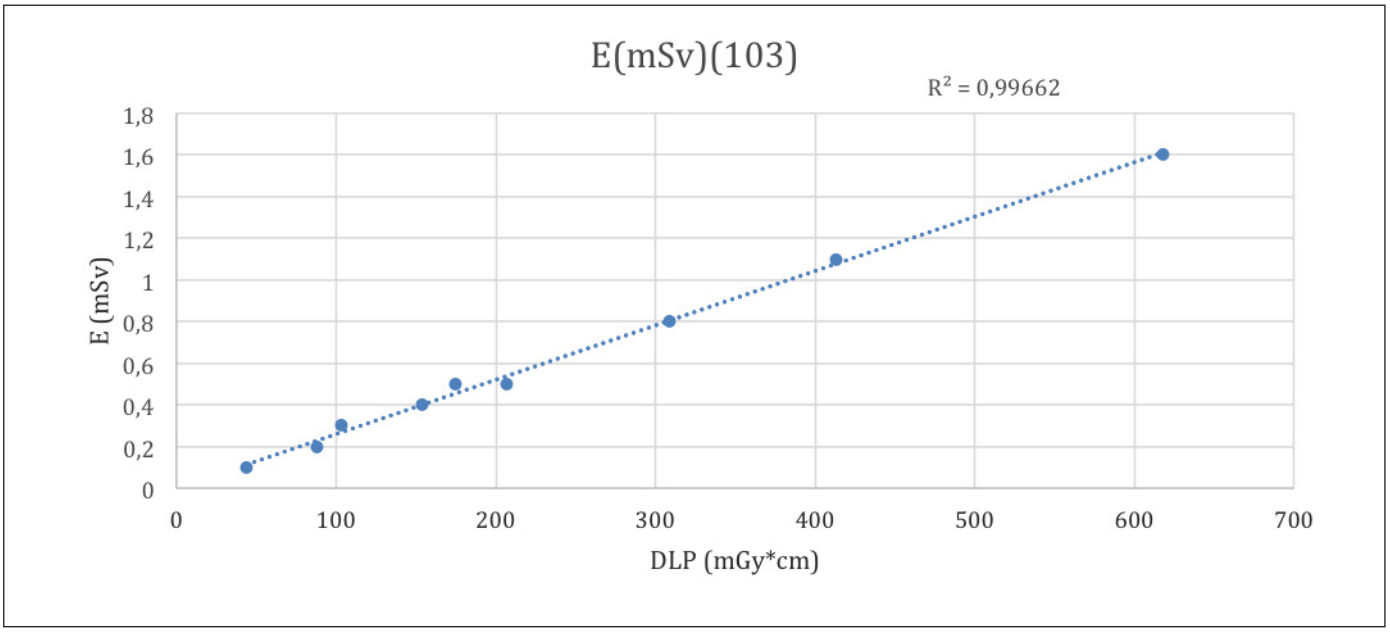


*Figure 5: Correlation between the Effective dose (103)(mSv) and DLP (mGy*cm). A strong positive correlation is shown between the effective dose and DLP with a (r = 0.99).*

Table 5: Calculated values for Intra-class correlation for experience and non-experience observer groups

	Experienced	Non-experienced
Item 1	0.536	0.211
Item 2	0.274	0.18
Item 3	0.531	0.377
Item 4	0.529	0.507
Item 5	0.677	0.624

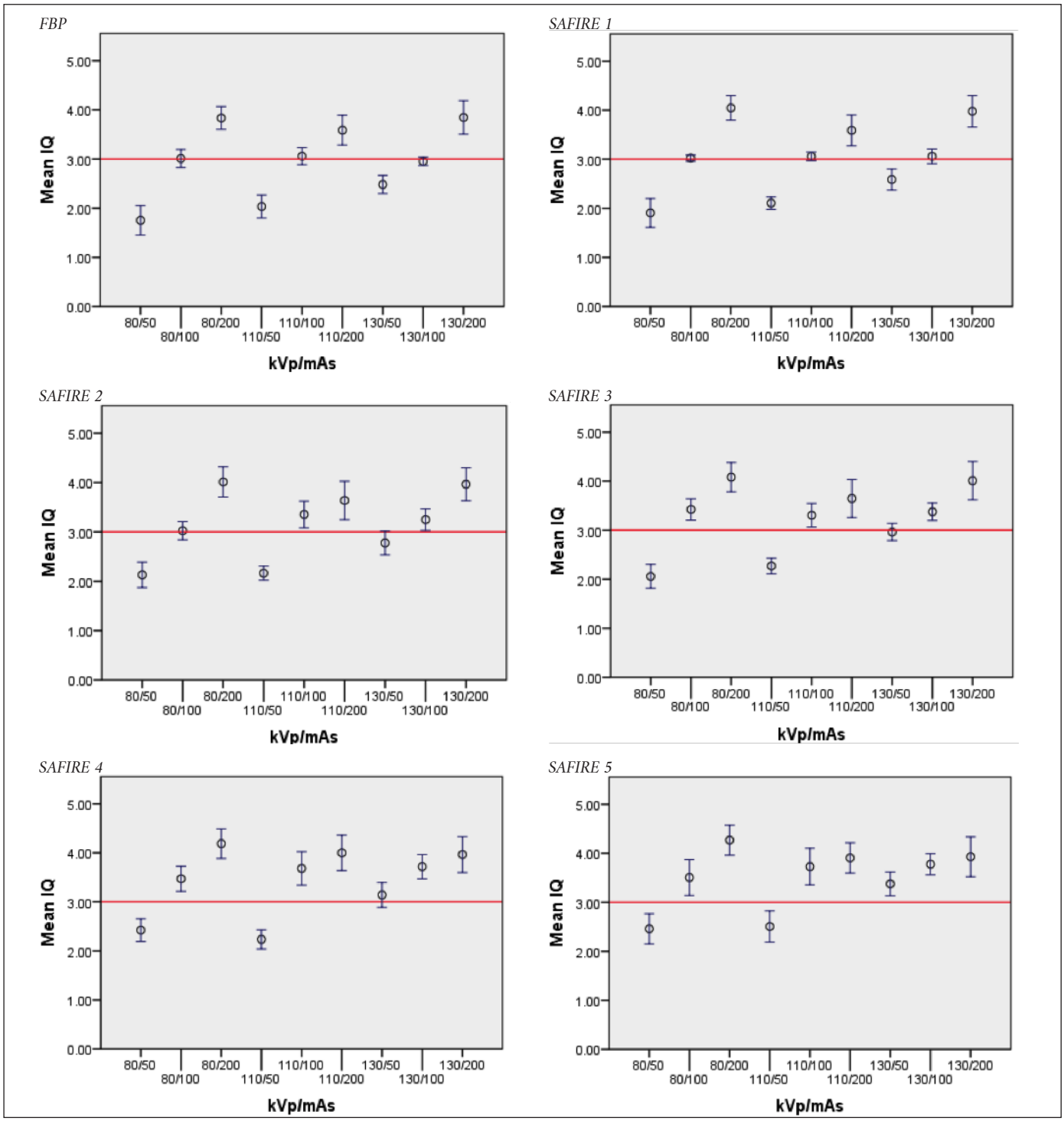


Figure 6: The mean image quality for each IR incorporating a 95% confidence interval, at each acquisition parameter. The reference image intersecting (y=3) the y axis represents an average image quality.

The Pearson Product-Moment Correlation Coefficient (r), correlation of determination (r2), and the p-value were calculated and tabulated (Table 4). This table highlights the lack of correlation between the perceptual image quality score and the physical image quality measures with $r^2 = 0.20$ and $r^2 = 0.007$ for SNR and CNR respectively. However, the p-value suggests a highly statistical significant relationship.

Mean Image quality		R2	R	P=	P
	SNR	0.2023	0.449778	2.6461E-121	p<0.0001
	CNR	0.0074	0.086023	6.8558E-121	P<0.0001

DISCUSSION

Providing acceptable image quality while reducing the radiation dose remains of paramount importance in CT examinations, more so for paediatric patients. Recent developments in iterative reconstruction technologies such as SAFIRE, shows potential in reducing the dose while improving image quality. However, the application of these techniques are limited for paediatric patients, with insignificant guidelines on the optimal SAFIRE strength for the best image quality.

This study highlights the advantages of using iterative reconstruction for a paediatric head phantom with a reduction of more than 68% in effective dose in comparison to standard Siemens SOMATON perspective 128 Computed Tomography (CT) protocol. The study also found that SAFIRE increases the SNR by an average of 22% with an average reduction in noise of 20%. When comparing the standard FBP reconstruction to SAFIRE 5, an average of 33% reduction in noise and 54% increase in SNR was calculated and the same was appreciated in the mean perceptual image quality scores, with SAFIRE 5 providing the best image quality. However, the item assessing the 'overall image quality' identified SAFIRE 4 as providing the best image quality. A factor most likely influenced by the image blurring or over smoothening reported in several studies[1-3] with an increase in SAFIRE strength recognising that a reduction in noise may not directly translate to an improvement in overall image quality1. Further research is required to identify the dose level at which diagnostic image quality can be achieved.

Observers

Experienced observers showed a higher ICC compared to inexperienced observers, which might be due to the fact that inexperienced observers have a learning effect during the task[17]. Buissink et al (2014) found that a statistically significant improvement can be observed post-training. However, all observer groups showed low to moderate ICC in all items of the IQ-criteria. The inclusion of training may have resulted in smaller difference between the groups18 and might have increased overall observer reliability.

Limitations

The simulation of paediatric head phantom ATOM five year old child, model 705[6], ensured no irradiation of human tissue[7]; however this presented a reduction in external validity. The tissue-equivalent epoxy resins present in all aspects of ATOM provided advantages in objective measures, but lacked comparable anatomical representation of a paediatric brain. This further limited the ability to adapt the European guidelines for quality criteria for computed tomography due to its reference to brain anatomy[10]. The use of a vetted criteria would have prevented the miss interpretation of the item questioning 'how many circles can you see in the image?' Nonetheless training the observers prior to the perceptual task may have improved their understanding of the question, increasing inter-observer reliability.

Images at 110kVp were acquired at different scan ranges in comparison to 80kVp and 130kVp, limiting the selection of the same slice for all images. Although the method was designed to overcome this restriction the ability to compare the batches may still exist.

CONCLUSION

In summary, the SAFIRE algorithm is suitable for image-noise reduction in paediatric head CT. Our data demonstrate that SAFIRE enhances SNR, while reducing noise, with a possible reduction in effective dose of 68%. The decrease in image quality with dose requires careful consideration of SAFIRE strength application to achieve optimal balance between image quality and noise. Our results suggest a potential for further reduction in dose and encourages an increase in external validity.

ACKNOWLEDGEMENTS

We would like to take this opportunity to acknowledge Carla Lança, for her assistance with the visual acuity tests for the observers, SIEMENS for providing us both the equipment and the reconstructed image sets. We also acknowledge Escola Superior de Tecnologia da Saúde de Lisboa for letting us use their facilities, equipment and time.

REFERENCES

1. Chang RC, Yu CC, Hsu FY, Chen TR, Hsu SM, Tyan YS. Dose assessment of the patient and the helper in emergency head computed tomography. Radiat Meas. 2011;46(12):2048-51.

2. Brenner DJ, Elliston CD, Hall EJ, Berdon WE. Estimated risks of radiation-induced fatal cancer from pediatric CT. AJR Am J Roentgenol. 2001;176(2):289-96.

3. Pickhardt PJ, Lubner MG, Kim DH, Tang J, Ruma JA, del Rio AM, et al. Abdominal CT with model-based iterative reconstruction (MBIR): initial results of a prospective trial comparing ultralow-dose with standard-dose imaging. AJR Am J Roentgenol. 2012;199(6):1266-74.

4. Korn A, Fenchel M, Bender B, Danz S, Hauser TK, Ketelsen D, et al. Iterative reconstruction in head CT: image quality of routine and low-dose protocols in comparison with standard filtered back-projection. AJNR Am J Neuroradiol. 2012;33(2):218-24.

5. Greffier J, Fernandez A, Macri F, Freitag C, Metge L, Beregi JP. Which dose for what image? Iterative reconstruction for CT scan. Diagn Interv Imaging. 2013;94(11):1117-21.

6. Han BK, Grant KL, Garberich R, Sedlmair M, Lindberg J, Lesser JR. Assessment of an iterative reconstruction algorithm (SAFIRE) on image quality in pediatric cardiac CT datasets. J Cardiovasc Comput Tomogr. 2012;6(3):200-4.

7. Grant K, Raupach R. SAFIRE: Sinogram Affirmed Iterative Reconstruction [Internet]. Siemens; 2012. Available from: https://www.healthcare.siemens.com/computed-tomography/options-upgrades/clinical-applications/safire

8. CIRS. ATOM® dosimetry verification phantoms [Internet]. Norfolk-VG: CIRSINC; 2013. Available from: http://www.cirsinc.com/file/Products/701_706/701%20706%20DS%20112613.pdf

9. Department of Health. The ionising radiation (medical exposure) regulations 2000: together with notes on good practice [Internet]. London: Department of Health; 2012. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/227075/IRMER_regulations_2000.pdf

10. Menzel HG, Schibilla H, Teunen D. European guidelines on quality criteria for computer tomography [Internet]. Brussels: European Commission; 1999. Available from: http://w3.tue.nl/fileadmin/sbd/Documenten/Leergang/BSM/European_Guidelines_Quality_Criteria_Computed_Tomography_Eur_16252.pdf

11. Haubenreisser H, Fink C, Nance JW Jr, Sedlmair M, Schmidt B, Schoenberg SO, et al. Feasibility of slice width reduction for spiral cranial computed tomography using iterative image reconstruction. Eur J Radiol. 2014;83(6):964-9.

12. Gur D, Rubin DA, Kart BH, Peterson AM, Fuhrman CR, Rockette HE, et al. Forced choice and ordinal discrete rating assessment of image quality: a comparison. J Digit Imaging. 1997;10(3):103-7.

13. Ledenius K, Svensson E, Stålhammar F, Wiklund LM, Thilander-Klang A. A method to analyse observer disagreement in visual grading studies: example of assessed image quality in paediatric cerebral multidetector CT images. Br J Radiol. 2010;83(991):604-11.

14. Yount WR. Developing scales: the Likert scale, the Thurstone scale, the Q-Sort scale, the Semantic Differential. In Yount WR, editor. Research design and statistical analysis in Christian ministry. 4th ed. Author; 2006. chap. 12.

15. Båth M, Månsson LG. Visual grading characteristics (VGC) analysis: a non-parametric rank-invariant statistical method for image quality evaluation. Br J Radiol. 2007;80(951):169-76.

16. McGreevy RL. Reverse Monte Carlo modelling. J Phys. 2001;13(46):R877.

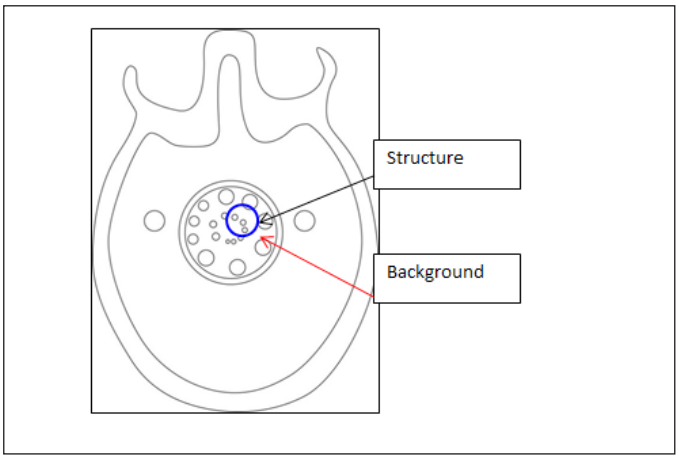
17. Brown CR, Hillman SJ, Richardson AM, Herman JL, Robb JE. Reliability and validity of the Visual Gait Assessment Scale for children with hemiplegic cerebral palsy when used by experienced and inexperienced observers. Gait Posture. 2008;27(4):648-52.

18. Buissink C, Thompson JD, Voet M, Sanderud A, Kamping LV, Savary L, et al. The influence of experience and training in a group of novice observers: a jackknife alternative free-response receiver operating characteristic analysis. Radiography. 2014;20(4):300-5.

Appendix A. Image test training

Question B

The differentiation of contrast within the encircled structure (see image) and its immediate background (see arrow)?

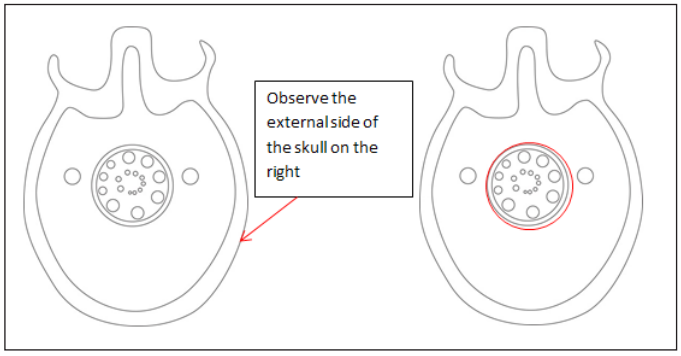


Question C

The visual sharpness of the bony structure displayed?

Question D

The sharpness of the circle in the encircled structure?



Appendix B. Questionnaire

How many circles can you count within the centre of the image?

Image Reference	1	2	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Circles																			

Appendix C

Table 5: *The average image quality score incorporating all observers with*

80 kVp								
	50 mAs	0.1	1.75 (±0.58)	1.91 (±0.57)	2.13 (±0.5)	2.06 (±0.47)	2.42 (±0.45)	2.46 (±0.6)
	100 mAs	0.2	R (±0.36)	3.02 (±0.12)	3.02 (±0.36)	3.42 (±0.42)	3.47 (±0.5)	3.51 (±0.71)
	200 mAs	0.5	3.84 (±0.45)	4.05 (±0.49)	4.01 (±0.6)	4.08 (±0.58)	4.19 (±0.59)	4.27 (±0.6)
110 kVp								
	50 mAs	0.3	2.04 (±0.45)	2.11 (±0.25)	2.16 (±0.28)	2.27 (±0.31)	2.24 (±0.38)	2.51 (±0.62)
	100 mAs	0.5	R (±0.34)	3.06 (±0.17)	3.35 (±0.53)	3.31 (±0.47)	3.68 (±0.66)	3.73 (±0.72)
	200 mAs	1.1	3.59 (±0.59)	3.59 (±0.61)	3.64 (±0.76)	3.65 (±0.75)	4 (±0.7)	3.91 (±0.6)
130 kVp								
	50 mAs	0.4	2.48 (±0.36)	2.59 (±0.42)	2.78 (±0.47)	2.96 (±0.34)	3.14 (±0.49)	3.38 (±0.47)
	100 mAs	0.8	R (±0.17)	3.06 (±0.29)	3.25 (±0.43)	3.38 (±0.35)	3.72 (±0.49)	3.78 (±0.42)
	200 mAs	1.6	3.85 (±0.66)	3.98 (±0.62)	3.96 (±0.65)	4.01 (±0.76)	3.96 (±0.71)	3.93 (±0.79)

Appendix D

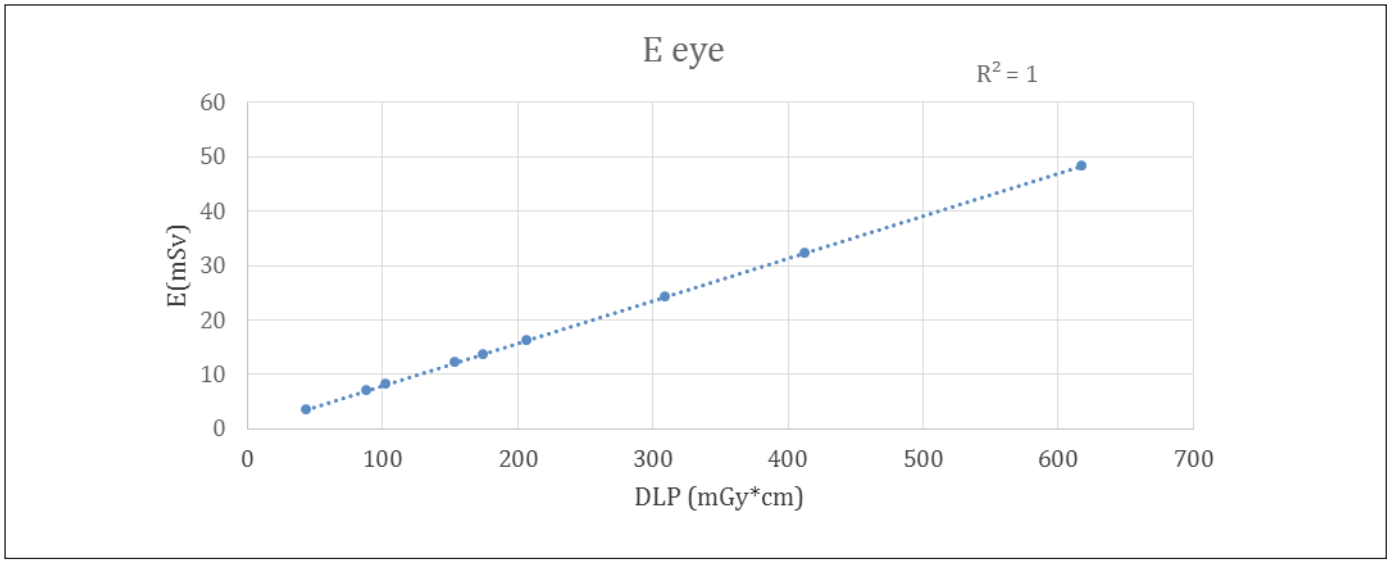


*Figure 3: Correlation between effective dose to the eyes (mSv) and the DLP (mGy*cm). A strong linear correlation is shown as the increase in DLP increases the effective dose to the eyes.*