The impact of a metal artefact reduction algorithm on treatment planning for patients undergoing radiotherapy of the pelvis

John King, Shona Whittam, David Smith *, Bashar Al-Qaisieh

Keywords: Computed tomography, Radiotherapy, Artefact reduction, Pelvis radiotherapy

ABSTRACT

Background and purpose: Metallic implants cause artefacts in computed tomography (CT) images and can introduce significant errors to structure visualisation and dosimetric calculation within the radiotherapy planning process. This study evaluated an orthopaedic metal artefact reduction algorithm and its effect on the CT number, image noise, structure delineation, and treatment dose.

Methods: Raw CT data were reconstructed using standard filtered back projection and an artefact reduction algorithm to create ‘standard’ and ‘corrected’ images. A phantom containing tissue-mimicking inserts and two titanium plugs was imaged. The average CT number was compared to baseline data acquired without metal inserts. Data from 11 pelvic external beam radiotherapy (EBRT) patients with bi- or uni-lateral hip implants were retrospectively analysed. The clinically used treatment plans were re-computed on the corrected images. A prostate-mimicking phantom containing metal ‘implants’ was imaged, and 11 observers contoured both reconstructions.

Results: The artefact reduction algorithm improved the CT number in those areas most affected by metal artefacts and decreased noise by 19 % (P = .04). Changes in dose distributions on corrected images compared to those calculated using the current clinical protocol were clinically insignificant. Volumes contoured on the corrected phantom images had larger Dice coefficients than those contoured on the standard images (P = .001), as well as a 36 % lower standard deviation in volumes.

Conclusion: This study demonstrates that the metal artefact reduction software reduces the error in CT numbers, can improve delineation accuracy, and can reduce inter-observer variability. It has the potential to streamline the planning pathway and improve treatment planning accuracy.

1. Introduction

Implanted medical devices such as dental fillings and orthopaedic implants can cause artefacts in computed tomography (CT) images. These are introduced via scatter, partial volume effects, aliasing, beam hardening and photon starvation [1,2]. The result is streaking artefacts where dark bands appear downstream of dense objects, cupping artefacts, where the periphery of objects appears falsely bright and a general reduction in image quality. Artefacts such as these can produce images with incorrect CT numbers (a relative measure of the radio density of an object). Radiotherapy planning using these images can lead to inaccuracies in the calculated dose distributions [3]. Additionally, metal artefacts can visually obscure tissues, leading to poor identification and delineation uncertainty [4–6]. As a result, this could adversely affect clinical outcomes due to potential under-dosing of the tumour volume or overdosing of organs at risk [1].

Artefacts are routinely dealt with in clinic by contouring the artefact and metal implant and manually adjusting the relative electron density values. Low density artefacts are assigned to water density, whereas the implant volumes are assigned to a density appropriate for that material. This process is time consuming and can introduce uncertainty in dose distribution.

There are various ways in which the effects of metal artefacts can be reduced, one of the most common being artefact reduction algorithms. There are multiple artefact reduction algorithms available which broadly fall into four categories (iterative, interpolation, filtering, and hybrid combinations), with many techniques specific to the CT manufacturer. To varying degrees these algorithms have been shown to...
reduce metal artefacts, improve the CT number accuracy, reduce noise in CT images, improve critical structure visualisation, aid target delineation, and improve dose calculation [1,4,5,8–12].

The primary objective of this study was to assess a commercially available reconstruction algorithm and quantify its impact on CT number accuracy, image noise, structure delineation accuracy, and calculated dose distribution for patients with orthopaedic hip implants undergoing pelvis radiotherapy.

2. Methods

2.1. Ethics statement

This study was designed to assess current care and assessed patient data retrospectively. The study protocol did not demand changes to treatment/ patient care from accepted standards for any of the patients involved, and the findings of this study were not generalizable. All patient data was anonymised before use and no data were shared outside LTHT (Leeds Teaching Hospitals Trust). In line with the health and research authority, 2017, this study was deemed a service evaluation and did not require research ethics committee review, health research authority approval, or a confidentiality advisory group application. In addition, in-house ethics approval covered all radiotherapy related patient data used in this work.

2.2. Patient selection

This study included patients with bi- or unilateral titanium hip implants (Fig. 1) that had undergone radiotherapy treatment of the pelvis. Fourteen patients were included in the study. Twelve patients had prostate treatments and two had rectum treatments. Four of these patients had bilateral hip implants.

2.3. CT number study

A tissue characterisation phantom (Model 467 electron density phantom, Gammex, Middleton, WI, USA) was imaged using a Philips Big Bore CT simulator (Philips Healthcare, Best, NL). All scans of the phantom were performed using the departmental clinical pelvis scanning protocol with the following settings: 16 × 1.5 mm collimation, 0.813 pitch, 0.75 s rotation time, 13 s scan time, 120 kV, 163 mAs, 2 mm slice thickness, 2 mm increments, 300 mm scan length, and a 600 mm field of view. The phantom was positioned at the isocentre of the scanner bore.

The tissue characterisation phantom is constructed of a disc of solid water equivalent material containing twenty holes filled with interchangeable inserts of various tissue and water equivalent materials (Fig. 2). Wax (50 % paraffin wax, 50 % bees wax) inserts were constructed in-house to replace lung-mimicking plugs and air gaps (holes containing no inserts), both of which are a contraindication for the commercially available reconstruction algorithm O-MAR (Orthopedic Metal Artefact Reduction) [7]. The phantom was imaged twice; with and without two grade 5 titanium rods replacing the inserts in positions 3 and 7, mimicking the presence of bilateral hip. Grade 5 titanium was used here as it is commonly employed for medical implants [13–17].

The raw image data were reconstructed using the standard Philips reconstruction algorithm (filtered back projection) and the O-MAR reconstruction algorithm (which incorporates an iterative projection modification process), resulting in 4 images in total. Each image was analysed using a local Matlab code (The MathWorks Inc, Natick, Massachusetts). The central slice of the CT image was aligned with a template that identifies 20 circular regions of interest (ROIs) with diameter 2.5 cm corresponding to the areas of the phantom inserts. The software extracted the average CT number and standard deviation in each ROI. A two-tailed Wilcoxon signed rank test was used to assess the statistical significance of the difference between the CT numbers in the standard and O-MAR corrected image sets [18].

The CT numbers measured on the tissue characterisation phantom containing no metal were compared to the baseline results acquired during CT simulator commissioning in 2018.

2.4. Patient contouring study

Using the raw CT data for each of the 14 patients, images were reconstructed using the standard and O-MAR algorithms. Three clinically trained observers contoured the bladder, rectum, and bowel loops on each of the ‘standard’ images, followed by the ‘corrected’ images. One observer also contoured the prostate. All contouring was conducted according to the clinical guidelines performed only on transverse slices. Observers were allowed to vary the thresholding and windowing to identify the necessary structures. Upon completion, each contour was peer reviewed to ensure its clinical suitability. A two-tailed Wilcoxon signed rank test was used to assess the statistical significance of the difference between contours conducted on the standard and corrected images [18].

2.5. Phantom contouring study

A tissue-equivalent ultrasound prostate phantom (GIRS, Norfolk, VA, USA) was used for contouring purposes. The phantom was CT imaged using the aforementioned pelvis scanning protocol, and the CT data reconstructed using the standard algorithm. The two previously described titanium inserts were positioned inside the water tank, on either side of the phantom, and aligned with the ‘prostate’ to mimic bilateral hip implants. A second CT scan of the phantom was acquired, and the data were reconstructed using the standard and O-MAR algorithms.

Twelve clinically trained observers contoured the prostate, with

---

Fig. 1. 2D transverse slices are taken from a planning CT of a patient (#4 in this study) undergoing prostate radiotherapy with bilateral hip implants. Both images are of the same slice, windowing, and levelling settings. Image (a) was reconstructed using the standard algorithm, and the image (b) using the O-MAR algorithm. Considerable streaking and photon starvation artefacts can be seen across the centre of the standard image, almost entirely obscuring the prostate.
participants contouring the ‘standard’ image first, followed by the ‘corrected’ image. All contouring was performed according to the clinical guidelines, as stated above. The Dice coefficient, precision, and sensitivity were calculated to assess each of the contours’ accuracy compared to a ‘gold standard’ contour conducted on the image containing no metal [19]. A two-tailed Wilcoxon signed rank test was used to assess the statistical significance of the difference between contours conducted on the standard and corrected images [18].

2.6. Dose distribution study

Three patients from the cohort had been CT imaged but did not continue with their treatment. Consequently, eleven patients were included in the dose distribution study, four of whom had bilateral hip implants.

The raw CT data were reconstructed using both the standard and O-MAR algorithms. The clinically used contours for the artefact area, metal implant, and PTV (planning target volume) were used to analyse both image sets. The contours and the patient’s treatment plan were delineated using the standard image set. As per the clinical protocol, the artefact volume was identified for the standard images using a CT number threshold value and forced to a relative electron density of unity. The metal implant volume was forced to a relative electron density of five. All contours, excluding the artefact volume, were copied onto the corrected images, and the electron density adjusted for the metal implant as above. The clinically used treatment plan (VMAT) was then re-calculated on each image set, and the dose distributions were calculated. A gamma analysis was carried out, and the dose-volume-histogram (DVH) statistics for the PTV were assessed to compare target dose coverage.

2.7. Treatment planning pathway audit

Following the clinical implementation of O-MAR, an independent audit was performed to identify benefits to the clinical workflow. Staff (imaging specialist radiographers, planning staff and clinicians) were asked to comment on the impact of image verification at treatment and the time saving during the planning process. The pathway for ten patients was reviewed.

3. Results

3.1. CT number study

The average CT number difference in each ROI between non-metal images and the baseline was less than 0.35 %. There was no difference in CT number between the ‘standard’ and ‘corrected’ non-metal images. For areas containing the most severe metal related artefact the O-MAR algorithm improved the CT number accuracy. For example, at position 2, the CT number rose from 5.6 % to 2.3 % compared to the baseline value (a 3.3 % percentage point increase in CT number accuracy). For the other two ROIs most affected by metal-related artefact (18 and 20), the increase in accuracy was 4.6 % points and 4.0 %, respectively. The remaining ROIs reported negligible changes in CT number accuracy. A statistically significant reduction in mean CT number standard deviation for all ROIs of 19.52 % (P = .04) was found between the ‘standard’ (78.48 [range: 42.41–260.27]) and ‘corrected’ (63.16 [range: 25.70–229.01]) images.

3.2. Patient contouring study

A statistically significant (P = .001) increase in mean prostate volume of 3.5 cm$^3$ was found for prostate delineations on the corrected images compared to the standard images. There was no statistically significant difference in volume for any other contoured organs (bladder, rectum, and bowel loop).

3.3. Phantom contouring study

A mean increase in DICE, precision and sensitivity for the corrected image vs the standard image was seen to be 0.05(P = .001), 0.08(P = .001), and 0.12(P = .0001) respectively. A statistically significant increase in volume of 7 cm$^3$ (P = .003) was found for the corrected image volume (48.5 cm$^3$) compared to the standard image volume (41.5 cm$^3$). The gold standard contour volume was 47.5 cm$^3$. Additionally, there was a reduction in standard deviation for the contoured volume of 1.9 cm$^3$ for the corrected image (3.2 cm$^3$) compared to the standard image (5.1 cm$^3$).
3.4. Dose distribution study

Using criteria of 3 mm and 3 %, the gamma pass rate was consistently greater than 99.8 % when comparing doses calculated on the ‘corrected’ images to those calculated on the ‘standard’ images. With the criteria reduced to 2 mm and 2 %, the pass rate was still over 99 % (Fig. 4). Patients 2, 4, 7, and 11 were those with bilateral hip implants.

There was no statistically significant difference in the PTV DVH statistics (D98, D50, and D2%) calculated on the standard and corrected images. In most cases, the percentage difference was less than 1 %. For one patient, the differences were more prominent than 1 % (1.4, 1.2, and 1.8 % difference for D98, D50, and D2% of the PTV respectively), however the DVH statistics were still within clinical tolerance for treatment [20].

3.5. Treatment planning pathway audit

A mean time saving of 30 min per patient was achieved during the planning process. Staff reported increased confidence in soft tissue matching due to improved image quality. In addition, regions of rectal gas were found to be more easily identifiable.

4. Discussion

A commercially available reconstruction algorithm was investigated by assessing CT number accuracy, image noise, structure delineation accuracy, and calculated dose distribution for patients with orthopaedic hip implants undergoing radiotherapy of the pelvis. The impact of implementing this algorithm on the treatment planning pathway was also assessed.

The CT number study was conducted to identify whether the O-MAR algorithm improved the accuracy of the CT number in CT images affected by metal artefacts. Inserts 18 and 20 were the most affected by streaking artefacts, followed by inserts 2 and 9, as shown in Fig. 4. At these four locations, the O-MAR algorithm significantly improved the CT number accuracy. However, the results show that in other locations, the O-MAR algorithm reduces the CT number accuracy, resulting in CT numbers which lie further away from the baseline (e.g. insert number 10) where the CT number accuracy falls from 0.8 % below the baseline to 4.3 % with O-MAR applied. This contradicts the findings of multiple published studies [21–23] where no definitive answer is presented regarding the circumstances under which the O-MAR algorithm decreases CT number accuracy. Repeat scans were conducted with the inserts in different positions within the phantom, which demonstrated similar results. One possibility is that the accuracy of the O-MAR algorithm depends on various contributing factors, such as the position of the metal, proximity to the metal/artefact, density of the area, and the location of other high-density materials to the measured area. Additionally, the negatively affected areas could result from artefacts still retained in the corrected images [24], or those produced by imperfect corrections and inaccuracies in the O-MAR segmentation process [8,25–26].

The standard deviation of the CT number provides a measure of the noise, typically correlated with image quality. It was found that the O-MAR algorithm significantly reduced the noise in images with metal artefacts. This is in agreement with multiple similar studies [1,8,22,24,25,27,28].

Analysis of the images identified that the contoured prostate volume was significantly larger on the corrected images than the standard images (mean ~8 %). Hansen et al. presented a similar result, where contoured volumes were generally larger after reducing the artefacts [29]. The patient’s bladder, rectum, and bowel loop volumes showed no significant change between the corrected and standard images. This is likely because the artefacts resulting from orthopaedic hip implants typically affected the prostate more than the other organs mentioned above due to its approximate to the hip. Therefore, the visibility of those artefacts was reduced.
Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Kwon H, Kim K, Chun Y, Woi HG, Carlson J, Park JM, et al. Evaluation of a commercial orthopaedic metal artefact reduction tool in radiation therapy of patients with head and neck cancer. The Br J Radiol 2015;105:20:4053. https://doi.org/10.1259/bjr.20140536.
[2] Ramsberg F, Dierks A, Nikolaus K, Reiser MF, Becker CR, Johnson TR. Metal artefact reduction by dual energy computed tomography using monenergetic extrapolation. Eur Radiol 2011;17:1424-9. https://doi.org/10.1007/s00330-011-2062-1.
[3] Wei J, Sandison GA, Izi WC, Ringor M, Lu X. Dosimetric impact of a CT metal artefact suppression algorithm for proton, electron and photon therapies. Phys Med Biol 2006;20:5183. https://doi.org/10.1088/0031-9155/51/20/007.
[4] Andersson KM, Dahlgren CV, Reizenstein J, Gao Y, Ahnejo A, Thunberg P. Evaluation of two commercial CT metal artefact reduction algorithms for use in proton radiotherapy treatment planning in the head and neck area. Med Phys 2018;10:4329-44. https://doi.org/10.1002/mp.13115.
[5] Andersson E, Ahnejo A, Dahlgren CV. Evaluation of a metal artefact reduction algorithm in CT studies used for proton radiotherapy treatment planning. J Appl Clin Med Phys 2014;5:112-9. https://doi.org/10.1120/jacmp.v5i15.4857.
[6] Guillouf C, Rampant P, House M. The impact of smart metal artefact reduction algorithm for use in radiotherapy treatment planning. Australas Phys Eng Sci Med 2017;385-94. https://doi.org/10.1007/s10528-017-0453-3.
[7] Metal artefact reduction for orthopedic implants (O-MAR). White Paper, Philips CT Clinical Science, Andover, Massachusetts. 2012. Available at: https://www.philips.co.uk/healthcare/education-resources/publications/hotspot/O-MAR-metal-artefact-reduction.
[8] Li H, Noel C, Chen H, Harold Li H, Low D, Moore K, et al. Clinical evaluation of a commercially available orthopaedic metal artefact reduction tool for CT simulations in radiation therapy. Med Phys 2012;12:7507-17. https://doi.org/10.1118/1.4762814.
[9] Kidoh M, Nakaura T, Nakamura S, Tokuyasu S, Osakabe H, Harada K, et al. Reduction of dental metallic artefacts in CT: value of a newly developed algorithm for metal artefact reduction (O-MAR). Clin Radiol 2014;1:11-6. https://doi.org/10.1016/j.crad.2013.08.008.
[10] Hokamp NG, Neuhaus V, Abdulhayev N, Laukamp K, Lennartz S, Motsas R, et al. Reduction artefacts caused by orthopedic hardware in the spine in spectral detector CT examinations using virtual monenergetic image reconstructions and metal-artefact-reduction algorithms. Skelet Radiol 2018;1:205. https://doi.org/10.1007/s00256-017-2776-5.
[11] Kovacs DG, Rechner LA, Appel AL, Bertheisen AK, Costa JC, Friborg J, et al. Metal artefact reduction for accurate tumour delineation in radiotherapy. Radiother Oncol 2018;3:479-86. https://doi.org/10.1016/j.radonc.2017.09.029.
[12] Chapman D, Smith S, Barnett R, Bauman G, Yartsev S. Optimization of tomotherapy treatment planning for patients with bilateral hip prostheses. Radiat Oncol 2014;1:1-6. https://doi.org/10.1186/1748-717X-9-42.
[13] Oldani C, Dominguez A. Titanium as a Biomaterial for Implants. Recent Adv Biomater Res 2018;1:1-2.
[14] Sidambe AT. Biocompatibility of advanced manufactured titanium implants—A review. Mater. 2014;12:8168-88. https://doi.org/10.3390/ma7128168.
[15] Beuzidienbout MB, Dimitrov DM, Van Staden AD, Oosthuizen GA, Dicks LM. Titanium-based hip stems with drug delivery functionality through additive manufacturing. BioMed Res Int 2015;2015. https://doi.org/10.1155/2015/104902.
[16] Hu CY, Yoon TR. Recent updates for biomaterials used in total hip arthroplasty. Biomater Res 2018;1:1-12. https://doi.org/10.1186/s40824-018-0144-8.
[17] Woolson RF. Wilcoxon signed-rank test. Wiley encyclopedia of clinical trials. 2007:1-2.
[18] Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med Imaging 2015;1:1-28. https://doi.org/10.1186/1471-2393-15-0068.
[19] Low DA, Harms WR, Murtic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. Med Phys 1998;5:656-61. https://doi.org/10.1118/1.598248.
[20] Huang JY, Kerns JR, Nute JL, Liu X, Balter PA, Stingo FC, et al. An evaluation of three commercially available metal artefact reduction methods for CT imaging. Phys Med Biol 2015;3:1047. https://doi.org/10.1088/0031-9155/60/3/1047.
[21] Andersson KM, Nowik P, Persliden J, Thunberg P, Norrman E. Metal artefact reduction in CT imaging of hip prostheses—an evaluation of commercial techniques provided by four vendors. Br J Radiol 2015;105:20140473. https://doi.org/10.1259/bjr.20140473.
[22] Hilgers G, Nixve T, Minken A. The CT number accuracy of a novel commercial metal artefact reduction algorithm for large orthopedic implants. J Appl Clin Med Phys 2014;1:274-8. https://doi.org/10.1120/jacmp.v15i1.4597.
[24] Branco D, Kry S, Taylor P, Rong J, Zhang X, Frank S, et al. Evaluation of image quality of a novel computed tomography metal artifact management technique on an anthropomorphic head and neck phantom. Phys Imaging Radiat Oncol 2021;17:111-6. https://doi.org/10.1016/j.phro.2021.01.007.

[25] Jeong S, Kim SH, Hwang EJ, Gi S, Han JK, Choi BL. Usefulness of a metal artefact reduction algorithm for orthopedic implants in abdominal CT: phantom and clinical study results. Am J Roentgenol 2015;205:307-17. https://doi.org/10.2214/AJR.14.12745.

[26] Koehler T, Brendel B, Brown KM. A new method for metal artefact reduction. In: The Second International Conference on Image Formation in X-ray Computed Tomography, June 24-27, 2012, Salt Lake City, Utah, USA; authors version; 2012.

[27] Wellenberg RH, van Osch JA, Boelhouwers HJ, Edens MA, Streekstra GJ, Ettema HB, et al. CT radiation dose reduction in patients with total hip arthroplasties using model-based iterative reconstruction and orthopaedic metal artefact reduction. Skelet Radiol 2019;48:1775-85. https://doi.org/10.1007/s00256-019-03206-z.

[28] Boomsma MF, Warringa N, Edens MA, Mueller D, Ettema HB, Verheyen CC, et al. Quantitative analysis of orthopedic metal artefact reduction in 64-slice computed tomography scans in large head metal-on-metal total hip replacement, a phantom study. Springerplus 2016;1:1-10. https://doi.org/10.1186/s40064-016-2006-y.

[29] Hansen CR, Christiansen RL, Lorenzen EL, Bertelsen AS, Åsmussen JT, Gyldenkerne N, et al. Contouring and dose calculation in head and neck cancer radiotherapy after reduction of metal artefacts in CT images. Acta Oncol 2017;6:874-8. https://doi.org/10.1080/0284186X.2017.1287427.

[30] Segedin B, Petric P. Uncertainties in target volume delineation in radiotherapy—are they relevant and what can we do about them? Radiol Oncol 2016;3:254. https://doi.org/10.1515/raon-2016-0023.

[31] McLean I. Quality assurance programme for computed tomography: diagnostic and therapy applications. IAEA Human Health Series 2012:19.

[32] Davis AT, Palmer AL, Pani S, Nibet A. Assessment of the variation in CT scanner performance (image quality and Hounsfield units) with scan parameters, for image optimisation in radiotherapy treatment planning. Physica Med 2018;45:59-64. https://doi.org/10.1016/j.ejmp.2017.11.036.