

Original Research Article

Evaluation of image quality of a novel computed tomography metal artifact management technique on an anthropomorphic head and neck phantom

Daniela Branco^{a,*}, Stephen Kry^a, Paige Taylor^a, John Rong^b, Xiaodong Zhang^a, Steven Frank^a, David Followill^a^a Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Unit 607, Houston, TX 77030, United States^b Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Unit 607, Houston, TX 77030, United States

ARTICLE INFO

Keywords:

Computed X ray tomography
Artifacts
Head and neck neoplasms
Algorithm
Gantry tilts

ABSTRACT

Background and purpose: Artefacts caused by dental amalgam implants present a common challenge in computed tomography (CT) and therefore treatment planning dose calculations. The goal was to perform a quantitative image quality analysis of our Artifact Management for Proton Planning (AMPP) algorithm which used gantry tilts for managing metal artefacts on Head and Neck (HN) CT scans and major vendors' commercial approaches.

Materials and methods: Metal artefact reduction (MAR) algorithms were evaluated using an anthropomorphic phantom with a removable jaw for the acquisition of images with and without (baseline) metal artefacts. AMPP made use of two angled CT scans to generate one artifact-reduced image set. The MAR algorithms from four vendors were applied to the images with artefacts and the analysis was performed with respective baselines. Planar HU difference maps and volumetric HU differences were analyzed.

Results: AMPP algorithm outperformed all vendors' commercial approaches in the elimination of artefacts in the oropharyngeal region, showing the lowest percent of pixels outside ± 20 HU criteria, 4%; whereas those in the MAR-corrected images ranged from 26% to 67%. In the region of interest within the affected slices, the commercial MAR algorithms showed inconsistent performance, whereas the AMPP algorithm performed consistently well throughout the phantom's posterior region.

Conclusions: A novel MAR algorithm was evaluated and compared to four commercial algorithms using an anthropomorphic phantom. Unanimously, the analysis showed the AMPP algorithm outperformed vendors' commercial approaches, showing the potential to be broadly implemented, improve visualizations in patient anatomy and provide accurate HU information.

1. Introduction

Head and neck scans represent almost 30% of all computed tomography (CT) scans done each year [1]. A substantial amount of these head and neck scans exhibit metal artifacts, particularly as produced by patients' dental work. Artifacts obscure the characterization of tissues in the oral cavity and oropharyngeal region decreasing the ability to differentiate and outline disease [2–4], potentially leading to erroneous diagnoses. This problem is also relevant in radiation oncology where CT scans are used for treatment planning; metal artifacts on CT images have been found to increase dose heterogeneity and reduce target coverage in

photon therapy [3–5]. In addition, proton dose calculations rely rigorously on the accurate representation of the HU values so that the correct relative linear stopping power is predicted. Therefore, proton treatment plans could display erroneous beam ranges and dose distributions when artifacts are present and severe.

Several solutions for metal artifact reduction (MAR) have been proposed, but are impractical, produce inaccurate CT images, and/or are not clinically available and therefore are not extensively adopted specifically in radiation therapy [5–10]. Four major commercial vendors offer MAR algorithms for clinical use. While these algorithms use proprietary information that is not shared publicly, they are basically all

Abbreviations: CT, Computed tomography; MAR, metal artifact reduction; OMAR, orthopedic metal artifact reduction; iMAR, iterative metal artifact reduction; SEMAR, single-energy metal artifact reduction; SmartMAR, Smart metal artifact reduction; AMPP, Artifact Management for Proton Planning; HU, Hounsfield Unit; OAR, Organs at Risk; kVp, Kilovoltage peak.

* Corresponding author.

E-mail address: dbranco@health.ucsd.edu (D. Branco).

<https://doi.org/10.1016/j.phro.2021.01.007>

Received 24 February 2020; Received in revised form 22 January 2021; Accepted 25 January 2021

Available online 21 February 2021

2405-6316/© 2021 The Authors. Published by Elsevier B.V. on behalf of European Society of Radiotherapy & Oncology. This is an open access article under the

CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

sinogram-based and work through the identification, deletion, and replacement of corrupted raw data. For example, one vendor offers a MAR function, which is an iterative projection modification solution that thresholds metal regions from the sinogram, removes them, and interpolates the missing data, to reconstruct the final artifact-adjusted image [11]. Similar to the commercial MAR function mentioned above, another vendor's iterative MAR algorithm removes high-contrast structures from the sinograms before interpolating the missing data [12]. A third commercial vendor's MAR algorithm segments the metal data in the projection space and replaces those data with inpainted data [13]. The fourth major vendor's MAR algorithm is also based on raw data segmentation and interpolation but incorporates different proprietary gradient correction features [14].

The effectiveness of existing MAR algorithms was found to be partial. While these algorithms reduce the presence and severity of artifacts, this reduction is far from complete. Particularly for the difficult situation of the head and neck, algorithms were found to only reduce the severity and number by 15–30% [15]. Moreover, algorithms that manipulate the sinogram data estimate values to replace missing or suspected corrupted data instead of using true measured data. This process of estimating data can introduce inaccurate substitutions [9], potentially causing additional artifacts [16]. Hence, there is a clear need to improve metal artifact management, particularly in highly heterogeneous anatomical sites, such as the head and neck.

Similar to the concept of stereoscopic imaging, our in-house developed algorithm Artifact Management for Proton Planning (AMPP) made use of two angled CT scans to generate one artifact-reduced image set [17]. With the acquisition of an angled CT scan that contains metal, the artifacts produced extend away from the areas of interest, allowing for the attainment of accurate HU information in areas that would otherwise have been affected in a typical 0° angle scan. Two CT scans at opposing angles were shown to provide enough data to reconstruct a final image free of artifacts posterior to the oral cavity that contains accurate CT numbers, not using sinogram manipulated projection data like most other commercially available MAR algorithms. The in-house algorithm was applied in the image space and therefore does not require the need to obtain and manipulate vendor proprietary raw data, making it suitable to any CT scanner that allows gantry tilts. In addition, the AMPP algorithm did not require information about the type and size of the metals in the image, making it easily applicable to any patients with dental work. One potential drawback of this technique was the possibility for anatomy changes in between scans due patient motion (swallowing, coughing, etc). These could introduce artifacts in the reconstruction and require repeated scans. The full introduction and evaluation of our in-house developed AMPP algorithm using a geometrical phantom was published by Branco et al. [17]. The purpose of the work discussed here was to further that study and perform a quantitative image quality analysis of the algorithm for managing metal artifacts on a head and neck phantom and four major vendors' commercially available approaches.

2. Materials and methods

2.1. Anthropomorphic head and neck phantom

We designed an anthropomorphic head and neck phantom composed of tissue-equivalent materials [18], with a human skull and air cavities mimicking the tissue heterogeneities in patients. This phantom was based on an Alderson phantom (The Phantom Laboratory, Salem, NY) (Fig. 1a) modified to have a jaw insert and a cylindrical insert accessible from the bottom of the neck (Fig. 1b). The cylindrical insert, which was made of Solid Water (CNMC, Nashville, TN), enabled us to introduce soft-tissue features into the phantom, including a horseshoe-shaped “tumor” in the middle of the insert, 2 parotids lateral to the tumor, and spinal cord. These features were made of Blue Water (Gammex, Middleton, WI). We based the cylinder design on human anatomy

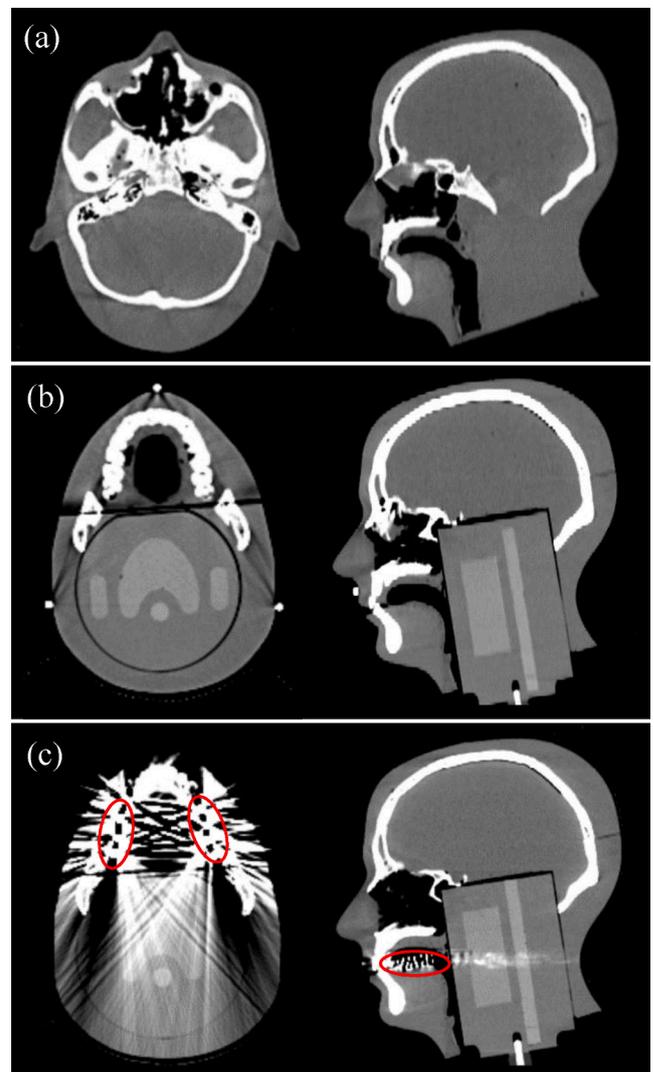


Fig. 1. (a) Axial and sagittal CT views of the Alderson phantom prior to modifications. (b) Axial and sagittal CT views of the modified phantom, which included a cylindrical insert containing a central target and 3 healthy structures, and a jaw insert. Holes were drilled in the molars in the jaw insert; these were filled with bone-equivalent materials in this panel. (c) Axial and sagittal CT views of the modified phantom with metal amalgam capsules in the tooth holes, shown by the red circles. The metal artifacts generated by the capsules are evident. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

representative of generalized oropharyngeal disease that is widely used for head and neck radiotherapy credentialing [19,20].

The jaw insert enabled us to obtain images with and without metal artifacts. We cut the lower jaw to expose the top and bottom teeth. Holes were drilled in each of the eight molars (four superior and four inferior) to hold capsules made of either bone-equivalent material to simulate a case with no fillings (Techtron HPV Bearing Grade; Gammex) or metal amalgams (Dispersalloy; Milford, DE) to simulate fillings and introduce metal artifacts. The locations, dimensions, and materials of the capsules were selected by a dental oncologist to be clinically realistic. They were placed inside the molars and were approximately 0.5 cm long. The phantom with the metal amalgam capsules in place and the artifacts generated by them were shown in Fig. 1c.

2.2. Image acquisition

The anthropomorphic phantom was scanned using Brilliance Big

Bore (Philips Healthcare System), SOMATOM Definition Edge (Siemens Healthcare), Revolution HD (GE Medical Systems), and Aquilion PRIME (Toshiba Medical) CT scanners. We used each scanner to acquire an image set with the bone-equivalent capsules (baseline scan) and an image set with the metal amalgam capsules (metal scan). Each metal scan was reconstructed using the respective vendor's MAR algorithm, Philips OMAR, Toshiba SEMAR, Siemens iMAR, GE SmartMAR (corrected scan).

We applied the in-house AMPP algorithm to image sets acquired with the Siemens SOMATOM Definition Edge scanner only, due to its performance being independent of CT scanner vendor (demonstrated in the robustness study of this manuscript – Section 2.4). The parameters of each head and neck CT protocol used for the baseline and the metal scans were shown in Supplementary Table S1.

2.3. Data analysis

The MARs algorithms were evaluated using severity of artifacts criteria over a centrally located CT slice and CT number accuracy over the introduced structures in the phantom.

Planar artifact severity. We quantified artifact severity that remained after application of different MAR algorithms by comparing the HU maps of the MAR-corrected images with those of the corresponding baseline images for a central axial slice. To ensure that the same central slice of the phantom was compared across all platforms, we selected the slices with respect to fiducials positioned on the phantom. To ensure proper registration between each baseline slice and its corresponding MAR corrected slice, we used rigid, intensity-based image registration in MATLAB. After image registration, we created CT number error maps by subtracting each MAR-corrected image from its corresponding baseline image.

We calculated the percentage of bad pixels inside the circle that defined the phantom cylinder, which was considered the region of interest in the phantom that the algorithm aimed to improve. Since AMPP was previously shown to not introduce any bias into HU values [17], all pixels with a CT number error above +20 HU or below –20 HU were considered to be bad pixels. The 20-HU threshold was based on the HU standard deviation obtained in the baseline scan. The average standard deviation calculated inside the baseline structure volumes was 8 HU; therefore, HU differences more than 2 standard deviations from the mean were considered to be too far from the baseline value and thus represented demonstrably erroneous pixels.

Structure CT number accuracy. We contoured the target and OARs on the baseline image using the Eclipse treatment planning system (version 13.6; Varian Medical Systems, Inc., Palo Alto, CA). The CT number accuracy was quantified for the volumes of interest visually affected by artifacts (WL: 40 WW:400). We copied the volumes created on the baseline scans onto the corrected metal scans to maintain consistency throughout the volumes analyzed. The mean CT numbers and standard deviations were measured inside each structure, and mean HU differences were calculated. In addition, a noise analysis was performed using beam hardening artifacts (BHA) technique from Lin et al. based on the standard deviation of the structure volumes [21]. The BHA parameter was defined as: $BHA = \sqrt{N_0^2 - N_b^2}$, where N_0 was the noise for the studied region and N_b was the noise for the background.

2.4. Robustness evaluation

We evaluated the robustness of the AMPP algorithm by applying the algorithm to images acquired using several different scanning parameters. The imaging parameters were shown in Supplementary Table S2. For example, to determine the extent to which CT x-ray tube energy affected the performance of the algorithm, we applied AMPP to images acquired at 100, 120, and 140 kVp. The AMPP-corrected scan was repeated at 120 kVp using the GE scanner to investigate the algorithm's

independence of scanner type. The AMPP algorithm was also evaluated in the context of different slice thicknesses, filter types, head tilt (HN phantom slightly tilted representing the typical head tilt patient scenario to improve tissue visualization when artifacts were present), and scanner-specific reconstruction algorithms. To obtain each baseline scan, we repeated each metal scan without the metal amalgam in place so that each AMPP-corrected scan was compared with a corresponding baseline scan obtained using the exact same imaging parameters. In each scenario, artifact severity was evaluated by assessing the percentage of bad pixels and CT number accuracy within the structures' volumes.

2.5. Registration evaluation

The results of this study hinge on the accurate registration of the image sets as described in the methods section. Good registration was evaluated visually, but in addition, we also investigated the sensitivity of the results to the image registration process. To investigate the impact of an incorrect registration, we shifted in the left–right direction the AMPP-corrected image one and two pixels compared with the baseline image and reassessed the percentage of bad pixels in the corrected image.

3. Results

3.1. Planar artifact severity

All MAR algorithms diminished the severity of the artifacts in the uncorrected images (Fig. 1c) but with varying success (Fig. 2a-d). In some instances, the MAR algorithms introduced additional artifacts in the posterior region characterized by a rippling on the posterior streaks (Fig. 2c). In contrast, the AMPP algorithm eliminated artifacts posterior to the oral cavity (Fig. 2e), showing an image nearly geometrically identical to the corresponding baseline image (Fig. 1b). Directly near the metal, none of the algorithms evaluated were able to reduce the artifacts due to the proximity to the metal implants. Hence only the posterior region of the phantom was evaluated in the analysis. This compromise was considered to be acceptable because none of the algorithms performed well in the oral cavity and one of the most common presentations on HN disease is in the oropharynx.

The quantitative differences between the corrected images and the corresponding baseline images are shown in Fig. 3. There were severe differences, displayed by the dark red and dark blue colors, between the uncorrected image and corresponding baseline images. Compared with the uncorrected image, the MAR-corrected images were improved, but they still showed visual differences to the corresponding baseline images. In comparison, the AMPP-corrected image was in good agreement with the baseline image, mostly displaying differences within ± 20 HU. The percentages of bad pixels in the circular region of interest that defined the cylinder in the phantom were 78%, 67%, 29%, 26%, 28% and 4% for the uncorrected, OMAR, SEMAR, iMAR, SmartMAR and AMPP scans respectively.

3.2. Structure CT number accuracy

The AMPP algorithm consistently outperformed the commercial algorithms, regardless of location or anatomy, and improved the HU accuracy to nearly the same as that of the uncorrected baseline scan; the maximum systematic CT number error was only 2 HU (Table 1). The standard deviations in each structure volume for the baseline, uncorrected, and corrected images were shown in Table 2. As expected, the uncorrected image had the largest standard deviation owing to the many high- and low-density streaks within the phantom. The MAR algorithms showed some improvement but did not perform as well as the AMPP algorithm, which yielded standard deviations similar to the baseline values. The BHA parameter results were presented in Supplementary Table S3. It was possible to see that the uncorrected image showed the largest noise for all structures compared to the background, while AMPP

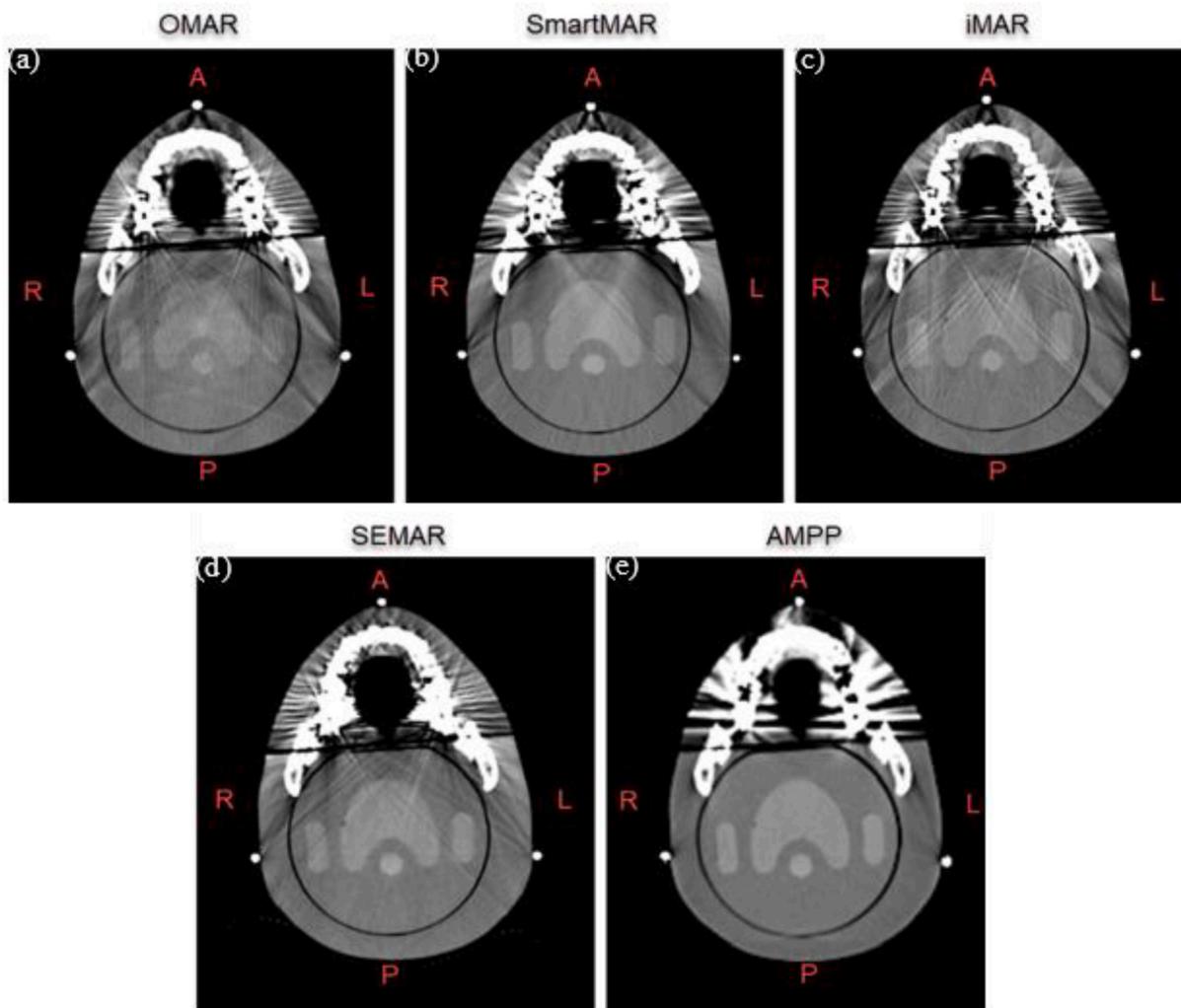


Fig. 2. (a–e) Axial CT views of the anthropomorphic phantom, corrected using the commercial MAR algorithms (a–d) or the in-house-developed AMPP algorithm (e). The same axial slice was chosen from each data set for the analysis (WL: 40 WW:400).

showed the lowest, sometimes no difference.

3.3. AMPP robustness evaluation

The percentages of bad pixels and mean CT number errors for the different imaging parameters under which the AMPP-corrected images were acquired were shown in Table 3. The AMPP algorithm performed similarly regardless of imaging parameter; each AMPP-corrected image showed similarly small HU differences and a small percentage of bad pixels as compared with the corresponding baseline images.

3.4. Registration evaluation

Whereas the bad pixel percentage of the original AMPP-corrected image was 4%, that of the AMPP-corrected image with a misalignment of one pixel was 6%, and that of the AMPP-corrected image with a misalignment of two pixels was 12%. The figure showing the error maps for the incorrect registrations was shown in Supplementary Fig. S1. It was possible to see the outline of the target and OARs as the pixel offset increases. However, the percentage of bad pixels on the visibly poorly registered AMPP-corrected images were still much lower than those on the MAR-corrected images, suggesting that the registration process was robust.

4. Discussion

In this study, a novel MAR algorithm was evaluated and compared to four currently available major commercial algorithms. An anthropomorphic head and neck phantom was designed and used for obtaining metal free scans (baseline) and artifact affected scans (with metal amalgam) to perform the analysis. Although commercial MAR algorithms generally reduced the severity of metal artifacts in the head and neck phantom, the algorithms performed slightly differently from each other, consistent with their individual proprietary distinctions. Unanimously, their performance was inferior to that of our in-house-developed AMPP algorithm.

On the basis of CT number information, all MAR algorithms reduced the percentage of bad pixels, at varying degrees of efficacy. Although the OMAR-corrected images retained several artifacts and thus still had a large percentage of bad pixels, the SEMAR, iMAR, and SmartMAR algorithms reduced the percentage of bad pixels by factors of 2–3. In contrast, the AMPP algorithm reduced the percentage by a factor of 20. In terms of the volume of the structures in the affected slices, the current commercial MAR algorithms showed inconsistent performance; whereas the SmartMAR and SEMAR algorithms performed well where the target and spinal cord were defined, the iMAR algorithm performed well where the parotids were defined. In contrast, the AMPP algorithm performed consistently well throughout the entire posterior region of the phantom, regardless of the location and intensity of the metal artifacts, even

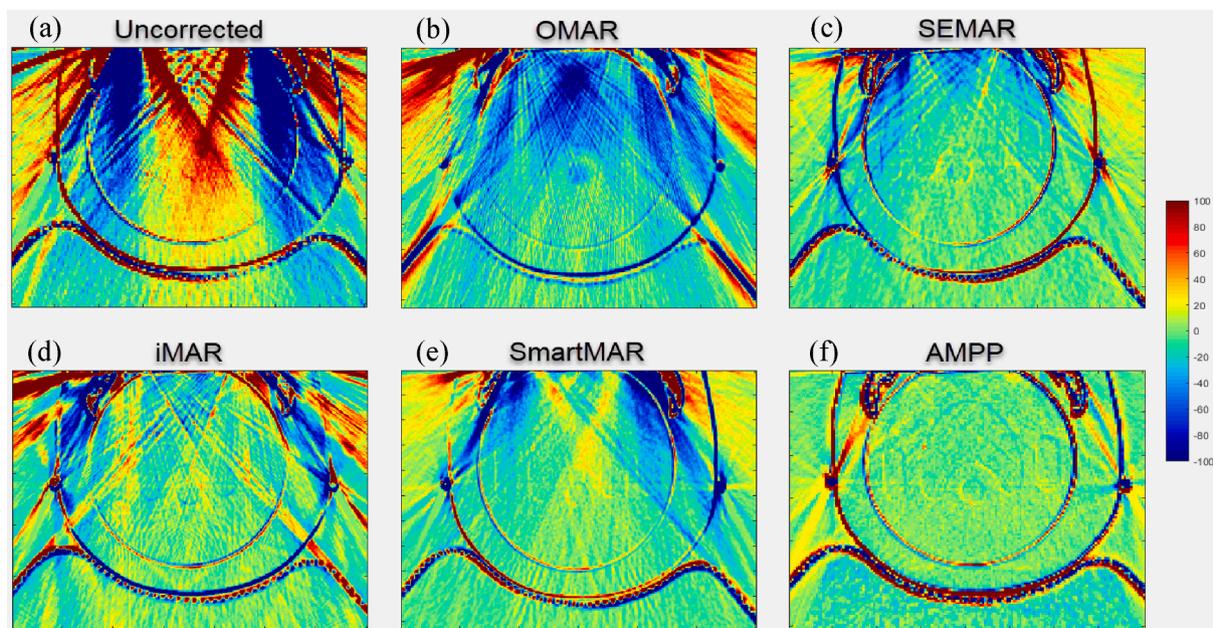


Fig. 3. HU error maps show differences between corrected images and corresponding baseline images. Severe HU differences are displayed by the dark red and dark blue colors shown by the color scale (HU).

Table 1

CT number errors (difference between each scans and corresponding baseline) within each structure volume inside the phantom by correction technique. A difference close to 0 means that the algorithm improved the CT numbers within all volumes compared to the uncorrected scan. A negative number means the baseline HU value in that structure was larger than on the metal corrected scan. The commercial MAR algorithms provided different degrees of improvement depending on structures; notably, there were structures that had systematic average CT number errors in excess of 20 HU, and all vendor algorithms had structures with average systematic errors of at least 15 HU.

Structure	CT number error within Structure Volume (HU)					
	Uncorrected	OMAR	SmartMAR	iMAR	SEMAR	AMPP
PTV	39	-18	2	-15	-3	-2
Spinal Cord	36	-13	3	-13	-1	-1
Right Parotid	-67	-34	-15	3	-20	2
Left Parotid	-63	-36	-24	-4	-19	1

eliminating the fiducials artifacts present in the baseline. This was true both in terms of eliminating systematic CT number errors as well as minimizing any spread in the HU over the structure. Our findings also showed that the AMPP algorithm performed well under different CT image acquisition parameters. The percent of bad pixels achieved with the AMPP algorithm under different parameters were all similar to each other and smaller than those achieved with the commercial MAR algorithms, ranging from 1% to 4%, whereas those in the MAR-corrected images ranged from 26% to 67%. That the AMPP algorithm was robust, as well as independent of scanner type and imaging parameters, was important for its clinical application. These results suggested that

Table 2

CT number standard deviation (SD) of each structure volume inside the phantom by correction technique.

Structure	CT number SD within Structure Volume (HU)						
	Baseline	Uncorrected	OMAR	SmartMAR	iMAR	SEMAR	AMPP
PTV	8	41	25	13	27	13	6
Spinal Cord	9	21	19	10	21	13	9
Right Parotid	11	49	19	9	16	17	7
Left Parotid	6	43	23	13	15	13	6

the AMPP algorithm, whose performance exceeded that of the commercial algorithms evaluated in the current study, could be applied to any scanner that allows for gantry tilts.

Previous evaluations of CT metals artifacts in the literature have shown contradictory results. While artifacts originated by large, geometrically simple sites such as hip prostheses are often well managed [11,16,22], the same is not true for dental artifacts. The studies that have considered more complex geometries such as the HN, have typically not been successful due to the larger number of sharp transitions between low and high attenuating materials [15,23,24]. Huang et al. also investigated CT number accuracy and severity of streaking artifacts in HN phantoms by looking at CT number differences and percentage of pixel differences that were >40 HU ('bad pixels') [15]. They found that none of the MAR algorithms analyzed were particularly successful, introducing 'out-of-plane' artifacts (due to sinogram smoothing filters) and showing a slight reduction in percentage of bad pixels and, in some instances, an increase, worsening the image. These results corroborated with the ones found in this study, marked by the limited improvement in percentage of bad pixels and introduction of additional artifacts by the algorithms analyzed here. In addition to a planar CT number accuracy analysis, we also investigated the accuracy of the CT number information in the structure volumes in the artifact affected slices (Structure CT number accuracy). We expanded the analysis to multiple slices to eliminate any potential bias individual slices may have presented. They also performed an evaluation of metal diameter accuracy that was not performed in this study as they concluded that all MAR algorithms underestimate the size of metal implants in the final reconstructed image. Kalender et al. performed a mostly qualitatively study but also found that a sinogram based algorithm that works by replacing missing

Table 3

CT number errors (difference between each scans and corresponding baseline) within structure volumes and percentages of bad pixels for different imaging parameters under which the AMPP-corrected images were acquired.

Structure	CT number error (HU)							
	Energy, 100 kVp (GE)	Energy, 120 kVp (GE)	Energy, 120 kVp (Siemens)	Energy, 140 kVp (GE)	Slice Thickness	SFOV	Head Tilt	Recon Algorithm
PTV	−1	−1	−2	−1	0	0	−2	0
Spinal Cord	0	−2	−1	0	0	0	−2	2
Right Parotid	2	−3	2	1	4	−1	−3	2
Left Parotid	2	6	1	−1	3	6	−2	2
% Bad Pixels	1.5	1.4	4.2	0.8	3.8	2.5	2	2.6

projection data was unsatisfactory in highly structured regions such as facial skull [23]. This issue is not limited to the literature, but was also supported by the vendors themselves: they often contraindicated their MAR's use when heterogeneities were present [25] and/or recommended the MAR images to be reviewed prior to being used [14].

Given the work presented here and the limited MAR algorithm performance in literature, the AMPP algorithm has shown the potential to be broadly implemented, improve visualizations in patient anatomy and provide accurate, not interpolated HU information. This can help with diagnosis as well as treatment planning in radiation oncology. In future studies, we will evaluate the proton therapy dosimetric performance of the MAR algorithms investigated in this work by establishing ideal comparisons to a metal free baseline.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.phro.2021.01.007>.

References

- [1] Schauer D, Linton O. NCRP report no 160, ionizing radiation exposure of the population of the United States, medical exposure—are we doing less with more, and is there a role for health physicists? *Health Phys* 2009;97:1–5. <https://doi.org/10.1097/01.HP.0000356672.44380.b7>.
- [2] O'Daniel J, Rosenthal D, Garden A, Barker J, Ahmad A, Ang K, et al. The effect of dental artifacts, contrast media, and experience on interobserver contouring variations in head and neck anatomy. *Am. J. Clin. Oncol.* 2007;30:191–8. <https://doi.org/10.1097/01.coc.0000256704.58956.45>.
- [3] Yusung K, Wolfgang T, Matthieu B, McNutt T, Spies L. The impact of dental metal artifacts on head and neck IMRT dose distributions. *Radiother Oncol* 2006;79:198–202. <https://doi.org/10.1016/j.radonc.2006.03.022>.
- [4] Mail N, Albarakati Y, Ahmad K, Saeedi F, Safadi N, Al-Ghamdi S, et al. The impacts of dental filling materials on RapidArc treatment planning and dose delivery: Challenges and solution. *Med Phys* 2013;40:817141–210. <https://doi.org/10.1118/1.4816307>.
- [5] Huang J, Followill D, Howell R, Liu X, Mirkovic D, Stingo F, et al. Approaches to reducing photon dose calculation errors near metal implants. *Med Phys* 2016;43:5117–30. <https://doi.org/10.1118/1.4960632>.
- [6] Zhao S, Robertson D, Wang G, Whiting B, Bae T. X-ray CT metal artifact reduction using wavelets: an application for imaging total hip prostheses. *IEEE Trans Med Imaging* 2000;19:1238–47. <https://doi.org/10.1109/42.897816>.
- [7] Man B, Nuyts J, Dupont P, Marchal G, Suetens P. An iterative maximum-likelihood polychromatic algorithm for CT. *IEEE Trans Med Imaging* 2001;20:999–1008. <https://doi.org/10.1109/42.95929>.
- [8] Yazdia M, Gingras L, Beaulieu L. An adaptive approach to metal artifact reduction in helical computed tomography for radiation therapy treatment planning: Experimental and clinical studies. *Int J Radiat Oncol Biol Phys* 2005;62:1224–31. <https://doi.org/10.1016/j.ijrobp.2005.02.052>.
- [9] Zhang X, Wang J, Xing L. Metal artifact reduction in x-ray computed tomography (CT) by constrained optimization. *Med Phys*. 2011;38:701–11. <https://doi.org/10.1118/1.3533711>.
- [10] Meyer E, Raupach R, Lell M, Schmidt B, Kachelrie M. Normalized metal artifact reduction (NMAR) in computed tomography. *Med Phys*. 2010;37:5482–93. <https://doi.org/10.1118/1.3484090>.
- [11] Li H, Noel C, Chen H, Li H, Low D, Moore K, et al. Clinical evaluation of a commercial orthopedic metal artifact reduction tool for CT simulations in radiation therapy. *Med Phys*. 2012;39:7507–17. <https://doi.org/10.1118/1.4762814>.
- [12] Axente M, Paidi A, Von E, Zeng C, Bani-Hashemi A, Krauss A, et al. Clinical evaluation of the iterative metal artifact reduction algorithm for CT simulation in radiotherapy. *Med Phys*. 2015;42:1170–83. <https://doi.org/10.1118/1.4906245>.
- [13] GE Healthcare. Smart metal artifact reduction (MAR). White paper. <https://www.gehealthcare.com/-/media/75ef28fec780481a93df7e6600cccd5a.pdf?la=en&ash=32DDB1A4A40877690EB68558CC816F6381AF5E6D>; 2003 [accessed 5 May 2020].
- [14] Toshiba Medical. Single energy metal artifact reduction. White paper. <https://us.medical.canon/download/ct-aq-one-genesis-wp-semar>; [accessed 5 May 2020].
- [15] Huang J, Kerns J, Nute J, Xinming L, Balter P, Stingo F, et al. An evaluation of three commercially available metal artifact reduction methods for CT imaging. *Phys Med Biol* 2015;60:1047–67. <https://doi.org/10.1088/0031-9155/60/3/1047>.
- [16] Das I, Paganetti H. Principles and practice of proton beam therapy. 1st ed. *Medical Physics Publishing*; 2015.
- [17] Branco D, Kry S, Taylor P, Rong J, Zhang X, Peterson C, et al. Development of a stereoscopic CT metal artifact management algorithm using gantry angle tilts for head and neck patients. *J Appl Clin Med Phys* 2020;21:120–30. <https://doi.org/10.1002/acm2.12922>.
- [18] Grant R, Summers P, Neihart J, Blatnica A, Sahoo N, Gillin M, et al. Relative stopping power measurements to aid in the design of anthropomorphic phantoms for proton radiotherapy. *J Appl Clin Med Phys* 2014;15:121–6. <https://doi.org/10.1120/jacmp.v15i2.4523>.
- [19] Molineu A, Hernandez N, Nguyen T, Ibbott G, Followill D. Credentialing results from IMRT irradiations of an anthropomorphic head and neck phantom. *Med Phys* 2013;40:0221011–221018. <https://doi.org/10.1118/1.4773309>.
- [20] Branco D, Taylor P, Zhang X, Li H, Guindani M, Followill D. An anthropomorphic head and neck quality assurance phantom for credentialing of intensity-modulated proton therapy. *Int J Particle Ther* 2017;3:40–7. <https://doi.org/10.14338/IJPT-17-00005.1>.
- [21] Lin X, Miao F, Li J, Dong P, Shen Y, Chen K. High-definition CT gemstone spectral imaging of the brain. *Comput Assist Tomogr* 2011;35:294–7. <https://doi.org/10.1097/rct.0b013e3182058d5c>.
- [22] Hilgers G, Nuver T, Minken A. The CT number accuracy of a novel commercial metal artifact reduction algorithm for large orthopedic implants. *J Appl Clin Med Phys* 2014;1:274–8. <https://doi.org/10.1120/jacmp.v15i1.4597>.
- [23] Kalender W, Hebel R, Ebersberger J. Reduction of CT artifacts caused by metallic implants. *Radiol* 1987;164:576–7. <https://doi.org/10.1148/radiology.164.2.3602406>.
- [24] Branco D, Kry S, Taylor P, Zhang X, Rong J, Frank S, et al. Dosimetric impact of commercial CT metal artifact reduction algorithms and a novel in-house algorithm for proton therapy of head and neck cancer. *Med Phys* 2020. <https://doi.org/10.1002/mp.14591>.
- [25] Hey S, Hoogenraad D, Elanchezian V, van Meel M, Admiraal-Behloul F. Metal Artifact Reduction for Orthopedic Implants (O-MAR). White Paper. <https://philips-productcontent.blob.core.windows.net/assets/20180315/747912f2871a472da8eba8a400e4a95d.pdf>; 2016 [accessed 5 May 2020].