

Improved Image Quality in Head and Neck CT Using a 3D Iterative Approach to Reduce Metal Artifact

W. Wuest, M.S. May, M. Brand, N. Bayerl, A. Krauss, M. Uder, and  M. Lell

ABSTRACT

BACKGROUND AND PURPOSE: Metal artifacts from dental fillings and other devices degrade image quality and may compromise the detection and evaluation of lesions in the oral cavity and oropharynx by CT. The aim of this study was to evaluate the effect of iterative metal artifact reduction on CT of the oral cavity and oropharynx.

MATERIALS AND METHODS: Data from 50 consecutive patients with metal artifacts from dental hardware were reconstructed with standard filtered back-projection, linear interpolation metal artifact reduction (LIMAR), and iterative metal artifact reduction. The image quality of sections that contained metal was analyzed for the severity of artifacts and diagnostic value.

RESULTS: A total of 455 sections (mean \pm standard deviation, 9.1 ± 4.1 sections per patient) contained metal and were evaluated with each reconstruction method. Sections without metal were not affected by the algorithms and demonstrated image quality identical to each other. Of these sections, 38% were considered nondiagnostic with filtered back-projection, 31% with LIMAR, and only 7% with iterative metal artifact reduction. Thirty-three percent of the sections had poor image quality with filtered back-projection, 46% with LIMAR, and 10% with iterative metal artifact reduction. Thirteen percent of the sections with filtered back-projection, 17% with LIMAR, and 22% with iterative metal artifact reduction were of moderate image quality, 16% of the sections with filtered back-projection, 5% with LIMAR, and 30% with iterative metal artifact reduction were of good image quality, and 1% of the sections with LIMAR and 31% with iterative metal artifact reduction were of excellent image quality.

CONCLUSIONS: Iterative metal artifact reduction yields the highest image quality in comparison with filtered back-projection and linear interpolation metal artifact reduction in patients with metal hardware in the head and neck area.

ABBREVIATIONS: FBP = filtered back-projection; HU = Hounsfield unit; MAR = metal artifact reduction; IMAR = iterative MAR; LIMAR = linear interpolation MAR; NMAR = normalized MAR

Imaging plays a crucial role in the staging of oral cancers and is essential for determining tumor resectability, choosing suitable anatomic reconstruction, and planning radiation therapy. The imaging method of choice for evaluating the oral cavity and oropharynx is MR imaging because it provides higher soft-tissue contrast and is less susceptible to artifacts caused by dental hardware. Yet, the limited availability and higher costs of MR imaging, as well as individual patient conditions (breathing or swallowing disorders, claustrophobia, electronic implants such as pacemakers or ferromagnetic foreign bodies), make CT an important alternative option for many patients. Thus, CT is used frequently to

stage or follow-up patients because of its wide availability, relatively low cost, and very short scan time. In patients with dental fillings or implants, however, image quality can be degraded by photon starvation and beam hardening.¹ Due to these artifacts, tumors may be only partially visible or completely obscured, making it challenging to define tumor extent. Moreover, streak artifacts may obscure ipsilateral or contralateral lymph node metastases, which can potentially change the therapeutic approach.

The use of high-resolution kernels and extended CT-value ranges² improves image quality; evaluating the surrounding soft tissue, however, remains challenging or even impossible in many cases and can lead to missed findings. For metal artifact reduction (MAR),^{3,4} sinogram in-painting methods have been proposed. Areas affected by metal artifacts are regarded as missing data and are filled in by different interpolation techniques, such as linear interpolation metal artifact reduction (LIMAR). Because LIMAR is associated with algorithm-induced artifacts, normalized MAR (NMAR) was developed, and it has demonstrated the potential to

Received November 27, 2014; accepted after revision February 9, 2015.

From the Radiological Institute (W.W., M.S.M., M.B., N.B., M.U., M.L.), Friedrich-Alexander-University-Erlangen-Nuremberg, Erlangen, Germany; and Medical Imaging (A.K.), Siemens Healthcare, Erlangen, Germany.

Please address correspondence to Wolfgang Wuest, MD, Radiological Institute, Maximiliansplatz 1, 91054 Erlangen, Germany; e-mail: wolfgang.wuest@uk-erlangen.de

<http://dx.doi.org/10.3174/ajnr.A4386>

improve image quality in patients with artifacts from dental hardware and to improve the diagnostic accuracy of head and neck and of pelvic CT^{5,6} while minimizing algorithm-induced artifacts.

An extension of the MAR methods (ie, LIMAR and NMAR) is a frequency-split technique that also recovers noise texture and anatomic details in close proximity to metal. In a previous study of pelvic CT, this technique delineated adjacent bone and tissue next to metal implants more accurately than NMAR.⁶

The aim of this study was to evaluate a novel 3D iterative approach using normalized and frequency split metal artifact reduction in clinical routine head and neck imaging. The resulting image quality was compared with that of filtered back-projection (FBP) reconstructions and LIMAR.

MATERIALS AND METHODS

Study Population

From January to December 2013, consecutive patients scheduled for neck CT were screened for study participation. Each patient signed informed consent; the study protocol was approved by the local institutional review board. Raw datasets from 50 patients who met the inclusion criteria (no contraindication to contrast-enhanced CT and no motion artifacts) and whose testing resulted in impaired image quality caused by metallic dental hardware were enrolled. The study population consisted of 23 female and 27 male patients with a mean (\pm standard deviation) age of 61 \pm 15.1 years (range, 24–86 years). Each examination was performed on a single-source CT system (Definition AS+; Siemens, Erlangen, Germany) with the following parameters: 0.5-second gantry rotation time, 128- \times 0.6-mm section collimation using a z-flying focal spot, and 160 reference milliamperes-second tube current with automatic exposure control at a tube voltage of 120 kV. The contrast agent (350 mg of iodine/mL [Iomeron; Bracco, Milan, Italy]) was injected at a flow rate of 3 mL/s (volume, 90 mL) followed by a saline bolus (3 mL/s [volume, 30 mL]). A scan delay of 80 seconds was used for each patient. The raw data were transferred to an external workstation equipped with prototype LIMAR software, and 3 datasets (from FBP, LIMAR, and IMAR) were reconstructed with identical (anatomically adapted) fields of view, 2.5-mm section thicknesses, and standard soft-tissue (B35f) and bone (B70f) reconstruction kernels.

IMAR

IMAR combines 2 previously introduced MAR algorithms, NMAR⁷ and frequency-split MAR,⁸ in an iterative update scheme. NMAR replaces those parts of the sinogram that are affected by metal through normalized interpolation. The aim of NMAR is to avoid the introduction of new artifacts tangentially to high-contrast objects, which is often observed with other sinogram in-painting methods. This is achieved by removing high-contrast structures from the sinogram before interpolation and reinserting them afterward. A prior image is calculated from the initial image by assigning soft-tissue pixels (identified by thresholding) to 0 Hounsfield units (HU). The prior image is forward-projected, and the initial sinogram is divided pixel-wise with the prior sinogram. Linear interpolation is performed on the relatively flat normalized sinogram followed by denormalization with the prior sinogram. NMAR images are finally obtained by recon-

struction of the corrected sinogram and reinsertion of the metal pixels from the uncorrected images. Frequency-split MAR combines the low spatial frequencies of a metal artifact-corrected image with the high spatial frequencies of the corresponding initial image. Low- and high-frequency images are obtained by Gaussian filtering. The aim of frequency-split MAR is to preserve both the natural image impression and the edge information of the uncorrected image, which is often affected by pure sinogram in-painting methods, especially in the vicinity of the metal implants. The drawback of the frequency-split operation is the reinsertion of high-frequency streak artifacts into the corrected images. IMAR repeatedly performs the normalized sinogram interpolation and frequency-split operations by using the result of each iteration as input for the next iteration, which effectively reduces the remaining artifacts of the prior image and consequently improves the quality of NMAR in each iteration. The performance of IMAR depends on the choice of several user-selectable model parameters, such as the number of iterations, HU thresholds for metal segmentation and for prior image calculation, and the filter parameter of the frequency-split operation. Those parameters are vendor specific. However, the user can select from a list of parameter configurations that are optimized for several metal implant types, such as dental fillings, hip prostheses, spine implants, and cardiac pacemakers. All reconstructions in this study were performed with the dental-fillings parameter configuration.

Image Analysis

Images obtained by using FBP, LIMAR, and IMAR were displayed side by side on a dual-monitor 3D postprocessing platform (syngo.via; Siemens) in random order for each acquisition after all identifying information had been removed. The images were reviewed in the soft-tissue window (window level, 50 HU; window width, 400 HU) and in the bone window (window level, 300 HU; window width, 2,500 HU).

To assess image quality, both subjective and objective parameters were evaluated. Subjective image quality of the FBP, LIMAR, and IMAR reconstructions was assessed by using a 5-point Likert scale (1, indicates severe artifacts, largely not diagnostic; 2, poor image quality, partly nondiagnostic; 3, moderate image quality, limited diagnostic confidence; 4, good image quality, sufficient for diagnosis; 5, excellent image quality, no artifacts). The structure with the least favorable diagnostic quality defined the rank for each category.

To obtain objective parameters of image quality, regions of interest were placed in the soft tissue of the tongue, cheeks, and muscles of the neck bilaterally. The standard deviation was measured for all the ROIs and regarded as an indicator of the presence of artifacts.

Statistical Analysis

Values are given as means \pm their standard deviation. One-way analysis of variance and nonparametric Friedman-ANOVA were performed for subjective and objective, respectively, image quality scores and values after the Kolmogorov-Smirnov test for normal distribution. Subsequent Bonferroni and Tamhane T2 post hoc tests, depending on variances in the Levene statistic, were performed for 1-way ANOVA. Pairwise post hoc tests, as pro-

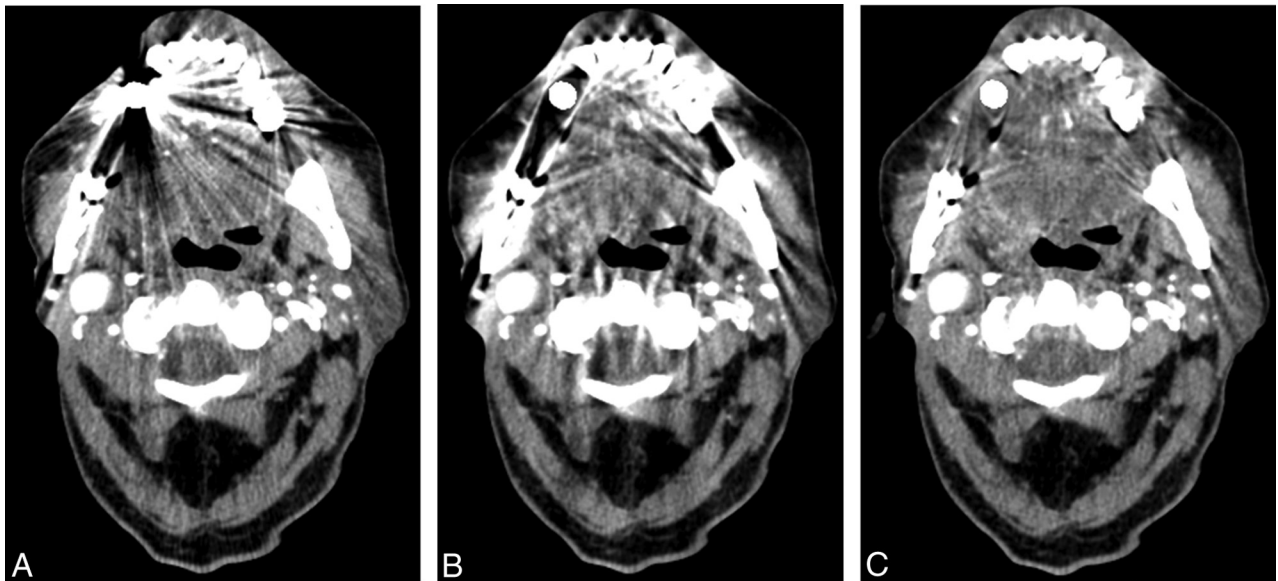


FIG 1. The effects of metal artifacts from dental hardware are highest with FBP (A), lower with LIMAR (B), and lowest with IMAR (C); a better visualization of the oral cavity was achieved with IMAR (C).

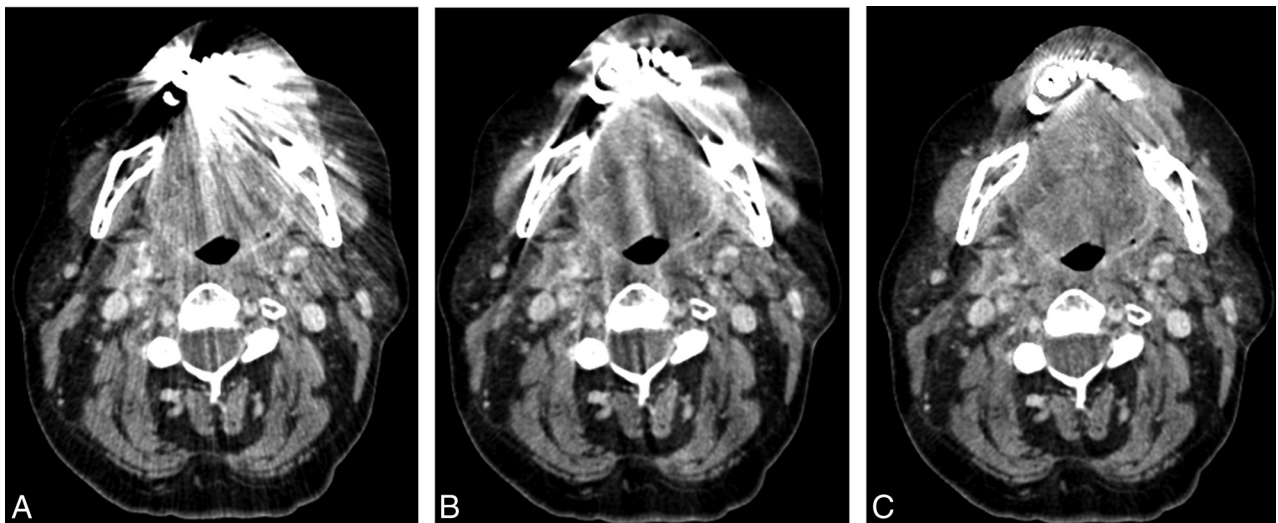


FIG 2. The effects of metal artifacts from dental hardware are highest with FBP (A), lower with LIMAR (B), and lowest with IMAR (C); a better visualization of the oral cavity was achieved with IMAR (C).

posed by Conover,⁹ were performed for the Friedman tests. A significance level of .05 was assumed. Statistical analysis was performed by using the software package SPSS Statistics version 19 (IBM, Armonk, New York).

RESULTS

Filtered back-projection, LIMAR, and IMAR reconstructions were performed successfully for each patient. A total of 455 sections (9.1 ± 4.1 sections per patient) contained metal artifacts and were evaluated with each reconstruction method. Sections without metal artifacts were not affected by the algorithms and had identical image quality.

IMAR significantly increased the subjective overall image quality compared with LIMAR and FBP ($P < .001$), and there was no significant difference between image quality after FBP and that after LIMAR ($P = .822$). The ratings were 2.1 ± 1.1 (FBP), 2 ± 0.9 (LIMAR), and 3.7 ± 1.2 (IMAR).

Of the sections, 38% were considered nondiagnostic with FBP,

31% with LIMAR, and only 7% with IMAR. Thirty-three percent of the sections had poor image quality with FBP, 46% with LIMAR, and 10% with IMAR. Moderate image quality was rated for 13% of the sections with FBP, 17% with LIMAR, and 22% with IMAR, good image quality was rated for 16% of the sections with FBP, 5% with LIMAR, and 30% with IMAR, and excellent image quality was rated for 1% of the sections with LIMAR and 31% with IMAR (Figs 1–3). These results are summarized in the Table.

The mean number of sections with severe artifacts was 3.5 ± 2.6 with FBP, 2.8 ± 2.2 with LIMAR, and 0.6 ± 1.1 with IMAR. With LIMAR, the mean number of sections with excellent image quality was 0.1 ± 0.3 , and with IMAR it was 2.8 ± 1.7 .

The mean standard deviation in the soft tissue of the tongue, the right cheek, and the left cheek were significantly higher with FBP than with LIMAR or IMAR ($P < .001$), and there was a significant difference between LIMAR and IMAR ($P < .001$):

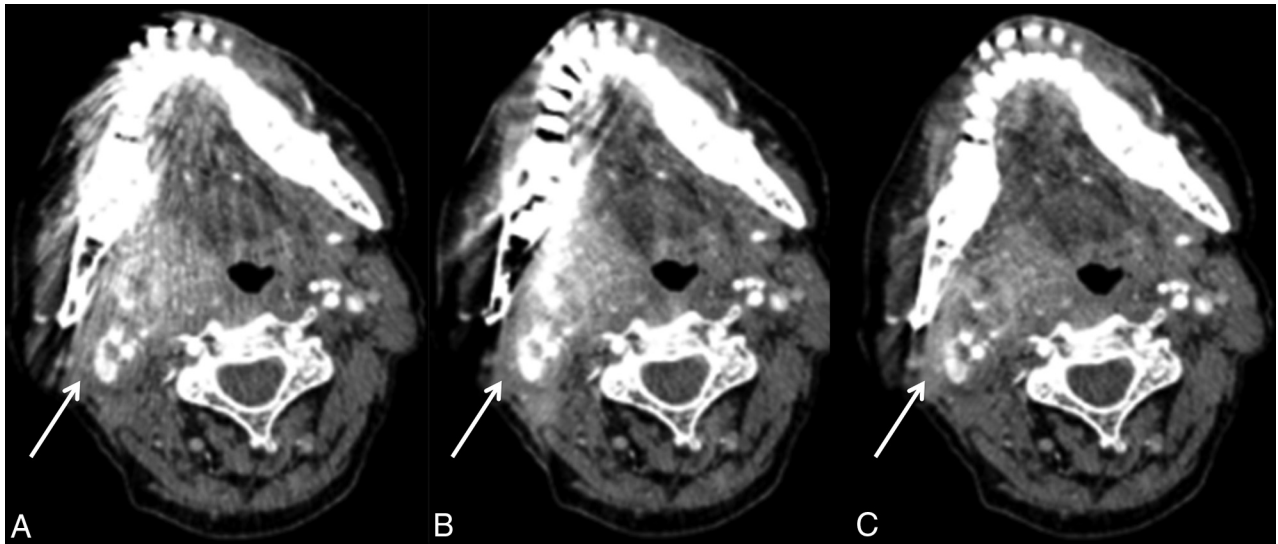


FIG 3. Imaging of oropharyngeal cancer on the right side with FBP (A), LIMAR (B), and IMAR (C). The effects of metal artifacts from dental hardware are highest with FBP (A), lower with LIMAR (B), and lowest with IMAR (C); full tumor extension is discernable only with IMAR (C, arrow).

Artifacts with FBP, LIMAR, and IMAR according to rating^a

Rating ^b	FBP	LIMAR	IMAR
1	174	140	31
2	150	211	47
3	61	78	100
4	71	22	135
5	0	4	142

^a Values shown are the number of sections with artifacts.

^b A rating of 1 indicates severe artifacts, largely not diagnostic; 2, poor image quality, partly nondiagnostic; 3, moderate image quality, limited diagnostic confidence; 4, good image quality, sufficient for diagnosis; and 5, excellent image quality, no artifacts.

162 ± 202 HU for FBP, 42 ± 19 HU for LIMAR, and 21 ± 6 HU for IMAR for the tongue; 73 ± 67 HU for FBP, 22 ± 10 HU for LIMAR, and 15 ± 5 HU for IMAR for the right cheek; and 63 ± 50 HU for FBP, 25 ± 14 HU for LIMAR, and 16 ± 5 HU for IMAR for the left cheek.

No significant difference was found between FBP and LIMAR in the muscles on either side of the neck ($P = .1$), but the IMAR mean standard deviation values were significantly lower than those of LIMAR and FBP ($P < .001$): 13 ± 4 for FBP, 12 ± 4 for LIMAR, and 11 ± 3 for IMAR on the left side and 14 ± 4 for FBP, 13 ± 4 for LIMAR, and 11 ± 3 for IMAR on the right side.

Cortical delineation of the alveolar process of the maxilla and mandible at the level of metal hardware improved with IMAR; however, IMAR induced some new artifacts next to metal hardware, which affected cortical delineation in 29 (58%) of 50 patients (Fig 4). Also, IMAR induced new artifacts in more remote areas, such as the spinal cord, in 9 (18%) of 50 patients (Fig 5).

DISCUSSION

Artifacts based on metallic implants and dental restorations are a frequently encountered obstacle in head and neck imaging, and advanced MAR algorithms might be a solution for this problem.^{5,7} Dental hardware affects not only CT imaging but also the attenuation correction in positron-emission tomography, dose calculation, and target definition for intensity-modulated radiation therapy.¹⁰

In our study, IMAR yielded objective and subjective image quality that was higher than that with FBP or LIMAR, and more than four-fifths of the sections that were not of diagnostic quality with FBP were evaluable with IMAR. Significantly more images were of diagnostic image quality with IMAR than with both FBP and LIMAR, which results in improvements in tumor detection and/or exclusion.

Tumor staging for squamous cell carcinoma of the oral cavity is based on size and extension into adjacent structures. The assessment of tumor infiltration depth is especially clinically challenging, and cross-sectional imaging is performed to gain that information. Yet, assessments of soft tissue in the oral cavity are often limited with CT and, to a lesser degree, MR imaging by metal artifacts. The first step for improving image quality and metal-streak artifacts is to remove all metal hardware from the scan range; however, that is not possible in many cases. In clinical routine, additional scans angulated to the mandible are often performed to increase diagnostic confidence for lymph node assessment and evaluation of the posterior neck. Parts of the oral cavity may still remain incompletely evaluated, however, and this approach increases radiation exposure and prolongs examination time. Application of an extended CT scale, thin-section collimation, a small FOV, dedicated reconstruction kernels,² and an increase of the tube voltage and current are options for reducing these artifacts; however, increasing the tube voltage and current increase patient radiation exposure, and none of these options have been dramatically successful. More elaborate strategies include monoenergetic processing of dual-energy CT data, which works nicely for surgical plates and implants,^{11,12} but its effect is limited with dental hardware. Sinogram in-painting methods^{8,13,14} and iterative,^{15,16} statistical,^{17,18} and filtering methods^{13,19} have been suggested, but for various reasons, they have not made their way into clinical practice. NMAR is an in-painting-based MAR method that is designed to reduce metal artifacts and to prevent the introduction of new artifacts by replacing raw data from the metal trace more reliably.⁷ Previously, the potential of NMAR to reduce artifacts from dental hardware was evaluated

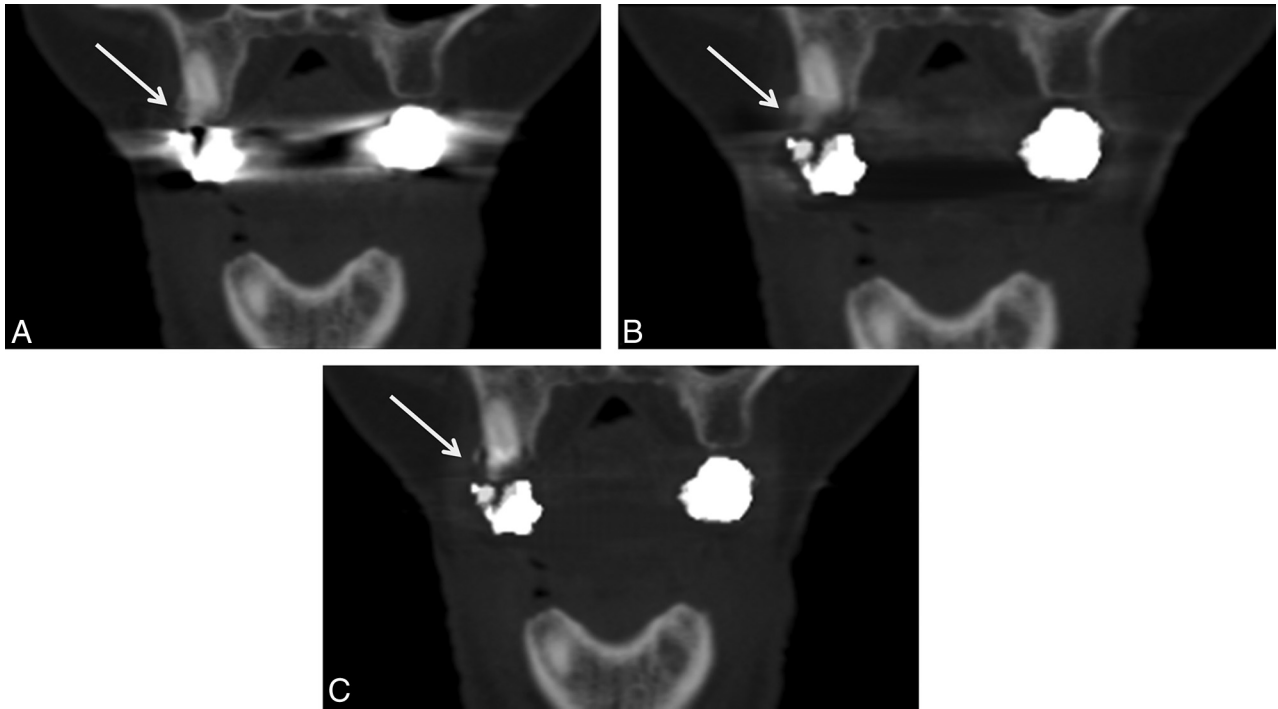


FIG 4. The effects of metal artifacts from dental hardware are highest with FBP (A), lower with LIMAR (B), and lowest with IMAR (C). Partial loss of cortical structures can be seen in B but is more pronounced in C (arrows).

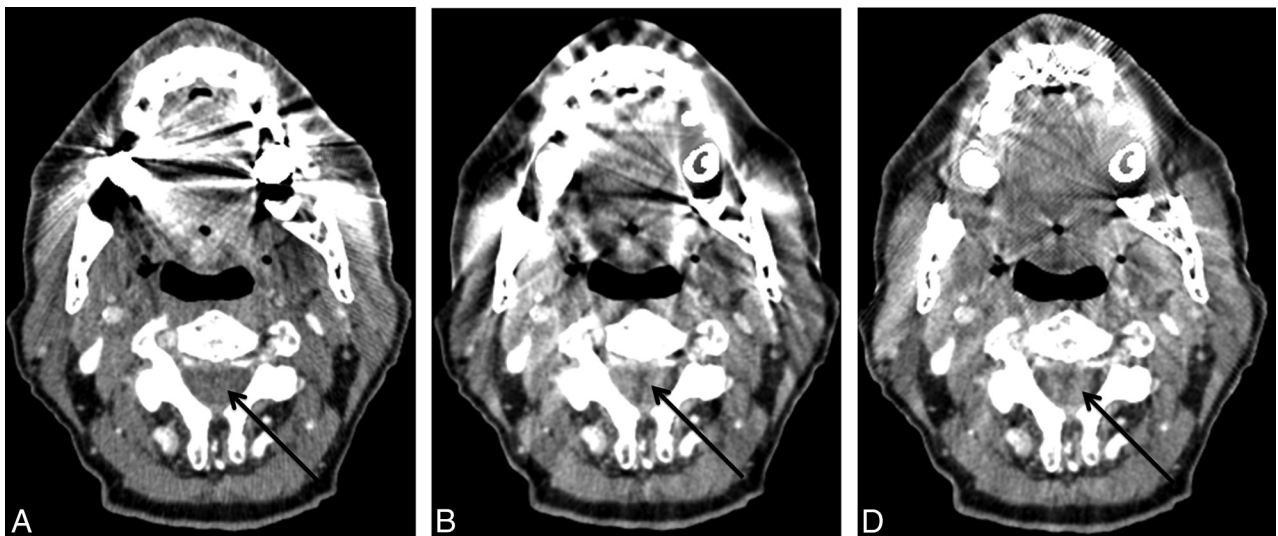


FIG 5. The effects of metal artifacts from dental hardware are highest with FBP (A), lower with LIMAR (B), and lowest with IMAR (C); however, LIMAR and IMAR induced new artifacts in the spinal cord (arrows).

in the head and neck region. The number of nondiagnostic sections with FBP was reduced by 50% with NMAR, which improved image quality and diagnostic accuracy. However, a drawback of LIMAR image reconstruction is that the tissues next to metal (eg, the bone trabeculae²⁰ or adjacent soft tissue) may be blurred and therefore not fully assessable. The frequency-split technique was introduced to address this problem of local blurring. With IMAR, anatomic information from the original images is recovered by high-pass filtering during the frequency-split iteration combined with multiple iterations in NMAR. This algorithm was evaluated in patients with hip prostheses. Image quality and the accuracy of pelvic abnormality assessments were compared in FBP,

LIMAR, and IMAR. IMAR reduced metal artifacts significantly and improved number measurements with CT and the confidence in depicting pelvic abnormalities.²¹

In our study, we found a significant improvement in soft-tissue delineation in the oral cavity and oropharynx, but we also found a degradation of bone delineation; artificial defects in the IMAR datasets of the bone abutting the metallic implants appeared in a number of cases (58%). Because of the reduction of streak artifacts, however, a better delineation of cortical bone in more remote areas in the sections containing metal hardware was achieved. Because the surgical approach is substantially influenced by tumor infiltration of the mandible or maxilla, which

leads to more extensive reconstruction methods to preserve function, the correct evaluation of bony structures is of high importance. Because of the imperfect bone delineation with IMAR, both FBP and IMAR images need to be reconstructed and evaluated to improve the overall diagnostic value in certain cases, which could be a significant limitation at the present time. Thus, detecting osseous involvement in tooth-bearing areas remains difficult with cross-sectional imaging (both CT and MR imaging). Additional limitations are that we did not investigate the clinical impact of our findings on treatment planning and prognosis, and only the metal artifact algorithm of one vendor could be evaluated, so no direct comparisons with other algorithms are possible.

CONCLUSIONS

In our patient population and with our specific CT scanner, IMAR yielded the highest image quality in comparison with FBP and LIMAR in patients with metal hardware in the head and neck area. We found significant improvement in the evaluation of soft tissue that was nondiagnostic with FBP and LIMAR.

Disclosures: Wolfgang Wuest—UNRELATED: Payment for Lectures (including service on speakers bureaus); Siemens AG, Healthcare Sector. Matthias Stefan May—UNRELATED: Payment for Lectures (including service on speakers bureaus); Siemens AG, Healthcare Sector; Payment for Development of Educational Presentations: Siemens AG, Healthcare Sector. Michael Brand—UNRELATED: Payment for Lectures (including service on speakers bureaus); Siemens Healthcare. Michael Lell—UNRELATED: Consultancy: Bracco and Guerbet; Grants/Grants Pending: Bayer,* Siemens,* and German Government (BMBF)*; Payment for Lectures (including service on speakers bureaus): Bayer and Siemens. Andreas Krauss—UNRELATED: Employment: Siemens AG; Patents (planned, pending or issued): Siemens AG; Stock/Stock Options: Siemens AG. *Money paid to the institution.

REFERENCES

- Barrett JF, Keat N. **Artifacts in CT: recognition and avoidance.** *RadioGraphics* 2004;24:1679–91 CrossRef Medline
- Lee MJ, Kim S, Lee SA, et al. **Overcoming artifacts from metallic orthopedic implants at high-field-strength MR imaging and multi-detector CT.** *RadioGraphics* 2007;27:791–803 CrossRef Medline
- Kalender WA, Hebel R, Ebersberger J. **Reduction of CT artifacts caused by metallic implants.** *Radiology* 1987;164:576–77 CrossRef Medline
- Glover GH, Pelc NJ. **An algorithm for the reduction of metal clip artifacts in CT reconstructions.** *Med Phys* 1981;8:799–807 CrossRef Medline
- Lell MM, Meyer E, Kuefner MA, et al. **Normalized metal artifact reduction in head and neck computed tomography.** *Invest Radiol* 2012;47:415–21 CrossRef Medline
- Lell MM, Meyer E, Schmid M, et al. **Frequency split metal artefact reduction in pelvic computed tomography.** *Eur Radiol* 2013;23:2137–45 CrossRef Medline
- Meyer E, Raupach R, Lell M, et al. **Normalized metal artifact reduction (NMAR) in computed tomography.** *Med Phys* 2010;37:5482–93 CrossRef Medline
- Meyer E, Raupach R, Lell M, et al. **Frequency split metal artifact reduction (FSMAR) in computed tomography.** *Med Phys* 2012;39:1904–16 CrossRef Medline
- Conover WJ. *Practical Nonparametric Statistics.* New York: Wiley; 1980
- Kennedy JA, Israel O, Frenkel A, et al. **The reduction of artifacts due to metal hip implants in CT-attenuation corrected PET images from hybrid PET/CT scanners.** *Med Biol Eng Comput* 2007;45:553–62 CrossRef Medline
- Bamberg F, Dierks A, Nikolaou K, et al. **Metal artifact reduction by dual energy computed tomography using monoenergetic extrapolation.** *Eur Radiol* 2011;21:1424–29 CrossRef Medline
- Zhou C, Zhao YE, Luo S, et al. **Monoenergetic imaging of dual-energy CT reduces artifacts from implanted metal orthopedic devices in patients with fractures.** *Acad Radiol* 2011;18:1252–57 CrossRef Medline
- Abdoli M, Ay MR, Ahmadian A, et al. **Reduction of dental filling metallic artifacts in CT-based attenuation correction of PET data using weighted virtual sinograms optimized by a genetic algorithm.** *Med Phys* 2010;37:6166–77 CrossRef Medline
- Prell D, Kyriakou Y, Kachelrie M, et al. **Reducing metal artifacts in computed tomography caused by hip endoprostheses using a physics-based approach.** *Invest Radiol* 2010;45:747–54 CrossRef Medline
- De Man B, Nuyts J, Dupont P, et al. **An iterative maximum-likelihood polychromatic algorithm for CT.** *IEEE Trans Med Imaging* 2001;20:999–1008 CrossRef Medline
- Zhang X, Wang J, Xing L. **Metal artifact reduction in x-ray computed tomography (CT) by constrained optimization.** *Med Phys* 2011;38:701–11 CrossRef Medline
- Aootaphao S, Pintavirooj C, Sotthivirat S. **Penalized-likelihood reconstruction for metal artifact reduction in cone-beam CT.** *Conf Proc IEEE Eng Med Biol Soc* 2008;2008:2733–36 CrossRef Medline
- Buzug T, Oehler M. **Statistical image reconstruction for inconsistent CT projection data.** *Methods Inf Med* 2007;46:261–69 Medline
- Naranjo V, Lloréns R, Alcañiz M, et al. **Metal artifact reduction in dental CT images using polar mathematical morphology.** *Comput Methods Programs Biomed* 2011;102:64–74 CrossRef Medline
- Liu PT, Pavlicek WP, Peter MB, et al. **Metal artifact reduction image reconstruction algorithm for CT of implanted metal orthopedic devices: a work in progress.** *Skeletal Radiol* 2009;38:797–802 CrossRef Medline
- Morsbach F, Bickelhaupt S, Wanner GA, et al. **Reduction of metal artifacts from hip prostheses on CT images of the pelvis: value of iterative reconstructions.** *Radiology* 2013;268:237–44 CrossRef Medline