Clinical relevance of metal artefact reduction in computed tomography (iMAR) in the pelvic and head and neck region: Multi-institutional contouring study of gross tumour volumes and organs at risk on clinical cases

Marius Hagen, Matthias Kretschmer, Florian Würschmidt, Tobias Gauer, Christian Giro, Elias Karsten and Jörn Lorenzen

1 Asklepios Medical School, Hamburg, Germany
2 Radiological Alliance, Hamburg, Germany
3 Department of Radiotherapy and Radio-Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
4 Department of Radiotherapy, University Medical Center Schleswig-Holstein, Kiel, Germany

Abstract

Introduction: Artefacts caused by dental implants and hip replacements may impede target volume definition and dose calculation accuracy. The iterative metal artefact reduction (iMAR) algorithm can provide a solution for this problem. The present study compares delineation of gross tumour volumes (GTVs) and organs at risk (OARs) in the pelvic and the head and neck (H & N) regions using computed tomography (CT) with and without iMAR, and thus the practical applicability of iMAR for routine clinical use.

Methods: The native planning CT and CT-iMAR data of two typical clinical cases with image-distorting artefacts were used for multi-institutional contouring and analysis using the Dice similarity coefficient (DSC). GTV/OAR contours were compared with an intraobserver approach and compared to predefined reference structures.

Results: Mean volume for GTV前列腺 in the intraobserver approach decreased from $87 \pm 44$ cm$^3$ (native CT) to $75 \pm 22$ cm$^3$ (CT-iMAR) ($P = 0.168$). Compared to the reference, DSC values for GTV前列腺 increased from $0.68 \pm 0.15$ to $0.78 \pm 0.07$ (CT vs. iMAR) ($P < 0.05$). In the H & N region, the reference for GTV舌 (34 cm$^3$) was underestimated on both data sets. No significant improvement in DSC values ($0.83 \pm 0.06$ (native CT) versus $0.86 \pm 0.06$ (CT-iMAR)) was observed.

Conclusion: The use of iMAR improves the anatomical delineation at the transition of prostate and bladder in cases of bilateral hip replacement. In the H & N region, anatomical residual structures and experience were apparently sufficient for precise contouring.

Key words: contouring study; Dice similarity coefficient; iMAR; iterative metal artefact reduction; target volume definition.

Introduction

Radiation therapy is one of the major pillars in cancer treatment worldwide. The clinically established gold standard of three-dimensional conformal external radiation therapy is computed tomography (CT). High density and high atomic metal implants cause dark and bright streaking artefacts and thus disrupt the quality and resolution of CT imaging. Such artefacts are the result of beam hardening, photon starvation and scatter, leading to a loss of image information. Consequently, anatomical identifiability is impaired and the CT numbers can be modified. Electron densities result from the CT numbers and represent a crucial measure for further treatment planning. Inaccuracies in CT numbers result in incorrect electron densities, and such image reconstruction errors...
may cause errors in the process of target volume definition, with negative impacts such as tumour recurrence (in case of underestimation of target volumes) or increased radiation toxicity (in case of volume overestimation) with unnecessary inclusion of healthy tissues. Implants in the hip and head and neck areas can produce massive artefacts, which are often in the tumour area or in its immediate vicinity. As a consequence, contouring of target volumes and organs at risk (OARs) is more difficult, leading to a more elaborate contouring process of target volumes and OARs. In addition, the prevalence of orthopaedic prostheses and dental fillings is increasing in elderly patients. According to current projections, the demand for hip arthroplasty in the Organisation for Economic Co-operation and Development (OECD) countries will increase from 1.8 million per year in 2015 to 2.8 million in 2050. Additionally, according to statistical estimated data, 450,000 new cases of prostate cancer were predicted for 2018 in Europe, and radiation therapy is one of the main treatment methods for prostate cancer; an increasing frequency of artefact-affected CT data due to hip arthroplasty can be expected in the coming years. To delineate such artefact-affected tissues, radiation oncologists have to rely on experience. In the literature, numerous different approaches to metal artefact reduction have been described in recent years. However, some of these approaches are only theoretical models. Although some have been proven to reduce artefacts, not all have found their way into clinical routine. The effectiveness of the iterative metal artefact reduction (IMAR) algorithm was first demonstrated by means of technical phantoms. Further studies, using anthropomorphic phantoms, also demonstrated the efficacy of IMAR. Patient-based studies with IMAR are rarely described in the literature, and only a few retrospective studies with IMAR on real patients have been published so far, albeit no contouring studies for the pelvis or head and neck (H & N) region have been performed. We used the IMAR algorithm for our study. Two previously introduced MAR algorithms, the normalized metal artefact reduction (NMAR) and the frequency split metal artefact reduction (FSMAR), in an iterative update scheme, are combined by the IMAR algorithm. The aim of this study was to analyse the clinical significance of IMAR through a multi-institutional contouring study on representative clinical cases in the pelvic and H & N regions.

**Methods**

The study involved native planning CT and CT-iMAR data of two typical clinical cases in our radiotherapy department. The planning CT was performed in the supine position using a Somatom AS 20 (Siemens, Erlangen, Germany) with a 2-mm slice thickness. For each patient, corresponding CT-iMAR data were generated by software-based postprocessing. All procedures were performed in accordance with the institutional ethics requirements.

**Patient characteristics**

1. A 76-year-old patient with elevated prostate-specific antigen (PSA) recurrence of a low-risk prostate gland adenocarcinoma and a bilateral hip replacement (Fig. 1).
2. An 88-year-old patient with recurrence of a squamous cell carcinoma of the tongue and cervical lymph
node metastases and extensive dental implants (Fig. 2).

**Study design**

Reference volume ($V_{ref}$) parameters were generated by two highly experienced radiation oncologists from the Radiological Alliance Hamburg (FW for the pelvic region and CG for the H & N region) in Eclipse Treatment Planning System version 13.6, with one doctor for each of the two cases. To generate the reference structures (gross tumour volumes (GTVs) and organs at risk (OARs)), both image data sets (native CT and CT-iMAR) of each case were simultaneous accessible for their contouring process. The final version of the $V_{ref}$ was then contoured only on the CT-iMAR data set and used as a structure set for further analysis. Structure sets for the $V_{ref}$ generated on the CT-iMAR were transferred to the native CT data set, which resulted in uniform $V_{ref}$ on CT-iMAR and native CT for both cases. The clinical experience in oncological CT imaging amounts to more than 15 years. Each of them was a specialist in the particular region. Here, ‘GTV$_{Prostate}$’ (without seminal vesicles) and ‘GTV$_{Tongue}$’ were defined as the primary tumour mass; both organs were contoured in their entire extension. The right parotid gland for the H & N region and the urinary bladder and rectum for the pelvic region were contoured as OARs. All these contours were determined as the $V_{ref}$. Based on specific instructions, radiation oncologists and radiation oncologists in training from different participant institutions were requested to delineate target volumes and OARs using their segmenting tools and clinical experience. A minimum number of years in clinical practice were not required. However, experience in target definition for such cases was mandatory. The contouring process of the observers was performed within the Eclipse Treatment Planning System version 10 and 13.6 (Varian Medical Systems, Palo Alto, CA, USA), depending on availability. Observers were blinded to all other contouring plans to guarantee independence of results. A guideline of contouring tasks was designed to minimise memory bias;

![Fig. 2. Axial and sagittal images from the patient’s planning CT for case 2: (a) and (c) without IMAR and (b) and (d) after correction with IMAR (comparable windowing).](image-url)
the contouring process was initiated on the native CT data set that presented loss of information and where GTVs were partly located in areas with total signal extinction or very high Hounsfield unit (HU) values. Subsequently, the same contouring was performed on the CT-iMAR data set that supposedly presents more image information. Finally, for both cases, contours were obtained for all regions of interest (ROIs) on both data sets (native CT and CT-iMAR) for each participant. Figures 3 and 4 show examples on sagittal/coronal planes of the $V_{\text{ref}}$ and all contours delineated by each observer on both data sets. Prior to delineation, observers received limited clinical information on the two cases to define GTVs and OARs.

**Fig. 3.** Sagittal views of the pelvic region with all contoured GTVs (red) and OARs (green and yellow): (a) and (b) without iMAR and (c) and (d) after correction with iMAR. $V_{\text{ref}}$ denotes the previously defined reference for the GTV$_{\text{Prostate}}$ (turquoise), OAR$_{\text{Rectum}}$ (purple) and OAR$_{\text{Bladder}}$ (purple) (comparable windowing).

**Fig. 4.** Sagittal (a, c) and coronal (b, d) views of the H & N region with all contoured GTVs (red) and OARs (yellow): (a) and (b) without iMAR and (c) and (d) after correction with iMAR. $V_{\text{ref}}$ denotes the previously defined reference for the GTV$_{\text{Tongue}}$ (green) and the OAR$_{\text{Parotid gland}}$ (purple) (comparable windowing).
Data analysis

All structures were analysed using Eclipse Treatment Planning System version 13.6. Cranio-caudal cutting of all ROIs and $V_{\text{ref}}$ to a main artefact area allowed us to evaluate just the effect of iMAR. Structures were already cut in the presented figures. For a quantitative detection of the main artefact area, we used a ‘Segment High Density Artefacts’ tool in Eclipse to identify artefact with high density (threshold 2.83 g/cm3). The first and last slice on the CT scan that showed such segments were defined as the borders of the main artefact area. Thereafter, we evaluated the Dice similarity coefficient (DSC) among the reference structures and each observer volume on the native CT scan and CT-iMAR data set to study the iMAR effect and to quantify the results ($V_{\text{ref}}$–observer$_n$). In addition, we evaluated the effect of iMAR within each observer (observer$_{n}$–observer$_{ref}$) [intraobserver approach]. The DSC is defined as follows:

$$\text{DSC} = 2 \times A \cap B / A + B$$

It provides values between 0 and 1, where 0 represents no intersection and 1 reflects clear overlap of structures A and B. In the literature, a DSC > 0.7 is commonly reported to indicate an excellent match.24 The DSC parameter not only considers absolute volumes but also local positions of two (A and B) compared volumes. Single absolute volumes of two structures are independent of the location. Volumetric and DSC data were collected using the DICOM-statistic tool in Eclipse which is embedded in Eclipse Contouring; volume was measured in cubic centimetres. Descriptive statistics were used; minimum, maximum and mean values, as well as standard deviation (SD), were calculated for each structure’s data (on native CT and CT-iMAR) and for the DSC analysis. Average DSC and average volumetric values for the targets (GTVs and OARs) were analysed with paired t-tests after checking for normal distribution (Kolmogorov-Smirnov test). All analyses and boxplots were performed with a $P$-value set to $<0.05$ in Excel (Microsoft, Redmond, USA). In addition, we asked the observers to evaluate the assistance and benefit of iMAR in categories (1 = excellent, 2 = good; 3 = adequate; 4 = sufficient; 5 = poor; 6 = unsatisfactory). Work experience in radiotherapy (in years) was also inquired.

Results

Table 1 and 2 show the volumetric parameters of the GTVs and OARs in the pelvic region and in the H & N region. Distributions of absolute volumes and DSC parameters are illustrated in Figures 5 and 6 and represented as boxplots. Overall, delineated prostate volumes ranged from 28.3 cm$^3$ to 195.6 cm$^3$ using the native CT data set and from 37.4 cm$^3$ to 117.5 cm$^3$ with the CT-iMAR data set. Average GTV$_{\text{prostate}}$ was 87 ± 44 cm$^3$ on the native CT data set and 75 ± 22 cm$^3$ on the CT-iMAR data set. Thus, the volume decreased by 12 cm$^3$, and the SD halved. However, there were no significantly different results in volumetric parameters due to iMAR imaging contouring of the prostate as target volume. No significant differences were found between native CT and CT-iMAR imaging for contouring the rectum as an OAR ($P = 0.714$). On both imaging data sets (native CT and CT-iMAR), mean volume and SD were nearly similar. Significant evidence ($P = 0.013$) could be provided for

### Table 1. Evaluated GTV and OAR parameters for the pelvic region

| Pelvic region | GTV$_{\text{prostate}}$ | OAR$_{\text{rectum}}$ | OAR$_{\text{bladder}}$ |
|---------------|------------------------|-----------------------|------------------------|
| Observer      | $V_{\text{CT}}$ [cm$^3$] | $V_{\text{iMAR}}$ [cm$^3$] | $V_{\text{CT}}$ [cm$^3$] | $V_{\text{iMAR}}$ [cm$^3$] | $V_{\text{CT}}$ [cm$^3$] | $V_{\text{iMAR}}$ [cm$^3$] |
| 1             | 89.0                   | 61.7                  | 103.2                  | 113.0                  | 212.5                  | 237.7                  |
| 2             | 49.3                   | 56.2                  | 68.1$_{\text{min}}$   | 79.0$_{\text{min}}$   | 219.3                  | 220.4                  |
| 3             | 97.9                   | 84.6                  | 93.8                   | 86.2                   | 223.3                  | 226.2                  |
| 4             | 101.3                  | 85.3                  | 144.1                  | 114.1                  | 241.9$_{\text{max}}$   | 223.3                  |
| 5             | 141.0                  | 117.5$_{\text{max}}$ | 81.2                   | 93.4                   | 222.2                  | 234.2                  |
| 6             | 28.3$_{\text{min}}$   | 46.0                  | 106.8                  | 85.6                   | 194.8                  | 237.9$_{\text{max}}$   |
| 7             | 195.6$_{\text{max}}$  | 97.5                  | 117.9$_{\text{max}}$  | 114.5                  | 170.8$_{\text{min}}$   | 215.6$_{\text{min}}$   |
| 8             | 29.4                   | 37.4$_{\text{max}}$  | 86.2                   | 86.3                   | 199.9                  | 205.8                  |
| 9             | 75.7                   | 88.9                  | 84.6                   | 87.2                   | 215.0                  | 220.9                  |
| 10            | 69.9                   | 68.0                  | 102.5                  | 98.9                   | 223.9                  | 235.7                  |
| 11            | 87.5                   | 87.6                  | 106.2                  | 109.1                  | 215.5                  | 231.1                  |
| 12            | 77.7                   | 64.7                  | 115.3                  | 119.2$_{\text{max}}$  | 210.7                  | 224.2                  |
| 13            | 83.6                   | 75.7                  | 103.9                  | 110.5                  | 218.4                  | 232.5                  |
| $V_{\text{ref}}$ | 76.0                   | 76.0                  | 83.9                   | 83.9                   | 236.3                  | 236.3                  |
| Mean ± SD    | 87 ± 44                | 75 ± 22               | 101 ± 19               | 100 ± 14               | 213 ± 17               | 227 ± 10               |
| $P$           | 0.168                  | 0.714                 | 0.013                  |

Min/Max, minimal and maximal volume; SD, standard deviation; $V_{\text{CT}}/V_{\text{iMAR}}$, volume on CT/iMAR data set; $V_{\text{ref}}$, reference structure on iMAR data set.
delineation of the bladder in the two different imaging data sets. Mean volumes of the bladder on native CT and CT-iMAR were 213 ± 17 cm³ and 227 ± 10 cm³, respectively, ranging from 170.9 cm³ to 241.9 cm³ (on native CT) and from 215.6 cm³ to 237.9 cm³ (on CT-iMAR). Data for the H & N region were not significantly different. The reference structure GTV Tongue (34 cm³) was slightly underestimated both on the native CT data set (28 ± 6 cm³) and on the CT-iMAR data set (30 ± 7 cm³).

The statistical evaluation of the intraobserver DSC calculations and the DSCs between observer and the reference for each ROI is presented in Table 3. There was no significant improvement in DSC values (0.83 ± 0.06 (native CT) compared to 0.86 ± 0.06 (CT-iMAR)) for the GTV Tongue. Contoured volumes of the parotid gland corresponded to the reference (18 cm³). All resulting mean DSC values in the intraobserver modality for each organ were higher (>0.7) than the mentioned limit. Compared with the reference, the mean DSC values for the prostate increased significantly from 0.68 ± 0.15 to 0.78 ± 0.07 (P = 0.01). No significant change in mean DSC values for the rectum could be identified (P = 0.696). In both imaging data sets, the mean DSC for the rectum was > 0.7. For the bladder, we already found high mean DSC values on native CT (0.91 ± 0.04). However, the small improvement of mean DSC values to 0.94 ± 0.01 due to iMAR was significant (P = 0.008).

The average professional work experience of the participating observers was 11.6 ± 9.2 years for the pelvic region and 13.5 ± 10.0 years for the H & N region. Median values for assistance due to iMAR were 3.0 (pelvic region) and 2.8 (H & N region).

**Discussion**

The goal of this study was to evaluate the importance of the iMAR algorithm in routine clinical use, especially for tissue demarcation in the contouring process of extremely metal artefact-affected CT data images. Several studies have previously shown an influence of iMAR on HU reestablishment and dose calculation accuracy by removing metal artefact. However, so far, no contouring study can be found that evaluates the practical applicability of iMAR in routine clinical use for prostate or H & N cases. In our study, we found for the pelvic
region that physicians were able to delineate the prostate \((P = 0.010)\) and the bladder \((P = 0.008)\) significantly more precisely using iMAR image data. Moreover, the distribution of the values was lower with iMAR. The visibility of the cranial part of the prostate, its apex and the dorso-basal area of the bladder at the transition to the prostate gland was especially improved by iMAR, which is also reflected in the contoured volume of both organs. Due to iMAR, the average volume of the bladder was larger, while that of the prostate was smaller compared to the reference. The iMAR had no influence on the distinctness of the rectum in our study. The participants were probably able to compensate for artefact-related misinformation in the area of the rectum. Overall, without using iMAR, visibility in the area between rectum and bladder was very limited by bright and dark streak artefacts; thus, important adjoining anatomical structures and geometrical borders were not discernible. With iMAR, this area was filled with soft tissue. Though borders between tissues seem slightly blurred on the iMAR reconstructed data, physicians were able to delineate GTVs and OARs successfully because image quality was sufficiently improved.

In the H & N region, about one-third of the volume of the parotid gland was erased. Contrary to the assumption that this was a difficult case, it appeared that clinical experience and non-artefact-affected residual anatomical structures on the CT data set were apparently sufficient for precise volume definition. Accordingly, there was no significant improvement in the target volume definition in the CT-iMAR data set compared to the reference structure for the GTV\(_{\text{Tongue}}\) \((P = 0.238)\) and OARs \((P = 0.508)\).

For our study, radiation oncologists and radiation oncologists in training from university hospitals and from an experienced practice were chosen as participants to...
rule out a bias in the predetermined contouring progress. This also correlates with the moderate iMAR score of 3.0 and 2.8 for the pelvic and H & N region. The group of participants with different training statuses was deliberately chosen to test a wide range of clinically active physicians in a radiotherapy department to gain additional information, such as the importance of iMAR in individual training status. The wide range of findings in contoured volumes did not correlate with individual training status. We found that even some junior doctors showed remarkably high and reliable values. This strengthened the argument that iMAR is an important additional tool regardless of work experience. Average benefit scores of 2–3 were consistent with the findings that iMAR reconstruction of tissues was well (but not perfectly) improved, as if there were no implants. Though iMAR reliably reconstructs soft tissues, the boundaries between soft tissues were slightly indistinct but adequate for delineation of GTVs and OARs.

The iMAR technique reduces dose errors significantly. Bär et al. investigated the influence of artefact from dental implants on the dose distribution. Comparing intensity-modulated radiation therapy (IMRT) plans from corrected and uncorrected images, dose differences in the range of ±5% were discovered in target volume and OARs, depending on the treatment site. Also, März et al. showed a significantly improved dose calculation accuracy for IMRT and volumetric modulated arc therapy (VMAT) plans with correction of artefact by iMAR. iMAR-corrected CTs are also beneficial in the image-guided radiation therapy (IGRT) process, delivering improved reference data for cone beam computed tomography (CBCT) and two-dimensional (2D) images. To achieve a maximum of tumour control with a minimum of toxicity, the highest possible precision of radiotherapy is required to further improve treatment results. Recent research has shown that there is a trend towards hypofractionated radiotherapy (e.g. in prostate carcinomas and in some H & N tumours). The phase III prostate advances in comparative evidence (PACE) trial even investigates extreme hypofractionation in prostate cases, using doses of >7 Gy per fraction. Therefore, high image quality will be of increasing importance as safety margins can be expected to be tight, especially in extreme hypofractionation schedules.

The DSC values for the GTV prostate did not show perfect results, and GTV site anomalies also affected planning target volumes (PTVs). In turn, dose distribution will change, making tumour control more difficult, especially in areas of common recurrences. Consequently, we can assume that iMAR can have a positive effect on accurate dose distributions and, therefore, on tumour control. The seminal vesicles are located in the transitional zone and are often the site of recurrences. Likewise, artefact reduction occurred in the apex, and tissue contrast improved. By increasing contouring precision, the probability of recurrences might be reduced. Axente et al. were able to show that using a CT-iMAR data set for contouring can increase the confidence of the participants. They also demonstrated that iMAR is an indispensable method for the reconstruction of the local anatomy in prostate cases with bilateral hip implants; our findings are in line with their results. Conversely, in

| Table 2. Evaluated GTV and OAR parameters for the H & N region. |
|---------------------------------|-----------------|-----------------|
| Head and neck                   | GTV | OAR Parotid gland |
| Observer                        | V<sub>CT</sub> | V<sub>iMAR</sub> | V<sub>CT</sub> | V<sub>iMAR</sub>| |
| 1                               | 24.1 | 32.1 | 17.9 | 18.1 |
| 2                               | 25.6 | 21.9 | 12.3 | 13.7 |
| 3                               | 31.6 | 27.6 | 12.8 | 13.8 |
| 4                               | 26.3 | 21.5 | 15.3 | 17.0 |
| 5                               | 5.3  | 35.3 | 23.8 | 18.6 |
| 6                               | 35.5 | 31.7 | 21.0 | 18.5 |
| 7                               | 20.6 | 31.2 | 24.9 | 26.0 |
| 8                               | 26.4 | 21.8 | 18.5 | 18.5 |
| 9                               | 46.8 | 26.4 | 21.8 | 21.8 |
| 10                              | 35.0 | 22.2 | 15.6 | 15.3 |
| Mean ± SD                       | 34 ± 18 | 21.0 | 18.0 | 18.0 |

The DSC values for the GTV prostate did not show perfect results, and GTV site anomalies also affected planning target volumes (PTVs). In turn, dose distribution will change, making tumour control more difficult, especially in areas of common recurrences. Consequently, we can assume that iMAR can have a positive effect on accurate dose distributions and, therefore, on tumour control. The seminal vesicles are located in the transitional zone and are often the site of recurrences. Likewise, artefact reduction occurred in the apex, and tissue contrast improved. By increasing contouring precision, the probability of recurrences might be reduced. Axente et al. were able to show that using a CT-iMAR data set for contouring can increase the confidence of the participants. They also demonstrated that iMAR is an indispensable method for the reconstruction of the local anatomy in prostate cases with bilateral hip implants; our findings are in line with their results. Conversely, in

| Table 3. Results for dice similarity coefficient |
|---------------------------------|-----------------|
| Pelvic region                   | Obs<sub>iMAR</sub> – Obs<sub>CT</sub> | Obs<sub>iMAR</sub> – V<sub>ref</sub> | Obs<sub>CT</sub> – V<sub>ref</sub> |
| GTV prostate                    | 0.75 ± 0.13     | 0.68 ± 0.15     | 0.78 ± 0.07     | 0.010   |
| OAR rectum                     | 0.88 ± 0.04     | 0.73 ± 0.05     | 0.74 ± 0.06     | 0.696   |
| OAR bladder                    | 0.93 ± 0.03     | 0.91 ± 0.04     | 0.94 ± 0.01     | 0.008   |
| Mean ± SD                      | 11.6 ± 9.2 [n = 13] |
| P                               | 3.0 [n = 13]    |
| GTV tongue                      | 0.85 ± 0.07     | 0.83 ± 0.06     | 0.86 ± 0.06     | 0.238   |
| OAR parotid gland               | 0.79 ± 0.09     | 0.74 ± 0.05     | 0.75 ± 0.05     | 0.508   |
| Mean ± SD                      | 13.5 ± 10.0 [n = 12] |
| P                               | 2.8 [n = 12]    |

iMAR score, median value; Obs<sub>CT</sub>, observer on CT data set; Obs<sub>iMAR</sub>, observer on iMAR data set; V<sub>ref</sub>, reference structure on iMAR data set.
cases of H & N cancer, Axente et al. showed that iMAR does not influence the contouring confidence of physicians. To improve contouring accuracy, co-registration of CT with magnetic resonance imaging (MRI) is often used in the clinical routine. Successful definition of target volumes in cases with bilateral hip replacements in pelvic radiotherapy could be demonstrated using co-registration of CT and MRI.33,34 Further studies using iMAR and MRI might demonstrate an increase in the precision of treatment planning. Based on our findings, it can be surmised that iMAR might show similar effects for some gynaecological tumours. For our study, two cases with extreme metal artefact were selected. However, these two cases were not representative of all patients’ cases having metal implants or inserts, as delineation and dose calculation will vary depending on each patient’s anatomy and implants especially in H & N cases. Nonetheless, using iMAR at our institution is the standard of care in cases of detected metal artefact during creation of planning CT scans.

In conclusion, prostate cancer patients with bilateral hip implants, iMAR is helpful regarding target volume definition because of quantitatively improved image reconstruction, especially in the transition of the prostate and bladder and the apex of the prostate, thus reducing uncertainties in target definition and further processes. For the H & N region, experience and anatomical residual structures are sufficient for precise contouring of target volumes and OARs. Due to its ease of use, the authors recommend using iMAR, if available, for delineation in cases of disruptive metal artefacts regardless of anatomical regions but particularly for prostate or head and neck cases.

Acknowledgements
The authors want to thank all participating doctors for providing their data sets. The data sets used and/or analysed during the current study are available from the corresponding author via reasonable request. Agreement on the use of patient image data is available.

References
1. Barrett JF, Keat N. Artifacts in CT: recognition and avoidance. Radiographics 2004; 24: 1679–91.
2. Bär E, Schwahoffer A, Kuchenbecker S et al. Improving radiotherapy planning in patients with metallic implants using the iterative metal artifact reduction (iMAR) algorithm. Biomed Phys Eng Exp 2015; 1: 025206.
3. Giantsoudi D, De Man B, Verburg J et al. Metal artifacts in computed tomography for radiation therapy planning: dosimetric effects and impact of metal artifact reduction. Phys Med Biol 2017; 62: R49–80.
4. Weiss J, Schabel C, Bongers M et al. Impact of iterative metal artifact reduction on diagnostic image quality in patients with dental hardware. Acta Radiol 2017; 58: 279–85.
5. Pabinger C, Lothaller H, Portner N et al. Projections of hip arthroplasty in OECD countries up to 2050. Hip Int 2018; 1120700018757940.
6. Ferlay J, Colombet M, Soerjomataram I et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries and 25 major cancers in 2018. Eur J Cancer 2018; 103: 356–87.
7. O’Daniel JC, Rosenthal DI, Garden AS 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.
8. Kalender WA, Hebel R, Ebersberger J. Reduction of CT artifacts caused by metallic implants. Radiology 1987; 164: 576–7.
9. Meyer E, Raupach R, Leli M et al. Normalized metal artifact reduction (NMAR) in computed tomography. Med Phys 2010; 37: 5482–93.
10. 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.
11. Bamberg F, Dierks A, Nikolau K et al. Metal artifact reduction by dual energy computed tomography using monoenergetic extrapolation. Eur Radiol 2011; 21: 1424–9.
12. Meyer E, Raupach R, Leli M et al. Frequency split metal artifact reduction (FSMAR) in computed tomography. Med Phys 2012; 39: 1904–16.
13. Morsbach F, Wurming M, Kunz DM et al. Metal artefact reduction from dental hardware in carotid CT angiography using iterative reconstructions. Eur Radiol 2013; 23: 2687–94.
14. 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.
15. Higashigaito K, Angst F, Runge VM et al. Metal artifact reduction in pelvic computed tomography with hip prostheses: comparison of virtual monoenergetic extrapolations from dual-energy computed tomography and an iterative metal artifact reduction algorithm in a phantom study. Invest Radiol 2015; 50: 828–34.
16. Maerz M, Mittermair P, Krauss A et al. Iterative metal artifact reduction improves dose calculation accuracy: Phantom study with dental implants. Strahlenther Onkol 2016; 192: 403–13.
17. Kovacs DG, Rechner LA, Appelt AL et al. Metal artefact reduction for accurate tumour delineation in radiotherapy. Radiother Oncol 2018; 126: 479–86.
18. Bolstad K, Flatabo S, Aadnevik D et al. Metal artifact reduction in CT, a phantom study: subjective and objective evaluation of four commercial metal artifact reduction algorithms when used on three different orthopedic metal implants. Acta Radiol 2018; 284185117751278.
19. Axente M, Paidi A, Von Eyben R et al. Clinical evaluation of the iterative metal artifact reduction
algorithm for CT simulation in radiotherapy. *Med Phys* 2015; 42: 1170–83.

20. Diehn FE, Michalak GJ, DeLone DR et al. CT dental artifact: comparison of an iterative metal artifact reduction technique with weighted filtered back-projection. *Acta Radiol Open* 2017; 6: 2058460117743279.

21. Aissa J, Thomas C, Sawicki LM et al. Iterative metal artefact reduction in CT: can dedicated algorithms improve image quality after spinal instrumentation? *Clin Radiol* 2017;72(5): e288.e7-e12.

22. Aissa J, Boos J, Sawicki LM et al. Iterative metal artefact reduction (iMAR) in postsurgical chest CT: comparison of three iMAR-algorithms. *Br J Radiol* 2017; 90: 20160778.

23. Kachelrieß M. Iterative Metal Artifact Reduction (iMAR): Technical Principles and Clinical Results in Radiation Therapy 2015 [Cited 25 March 2018.]. Available from URL: https://www.healthcare.siemens.de/medical-imaging/medical-imaging-options-upgrades/dg2/lp-imar-full/features

24. Gunther JR, Liauw SL, Choi S et al. A prostate Fossa contouring instructional module: implementation and evaluation. *J Am Coll Radiol* 2016; 13: 835–41.e1.

25. Dearnaley D, Syndikus I, Mossop H et al. Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHIP trial. *Lancet Oncol* 2016; 17: 1047–60.

26. Incrocci L, Wortel RC, Alemayehu WG et al. Hypofractionated versus conventionally fractionated radiotherapy for patients with localised prostate cancer (HYPRO): final efficacy results from a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncol* 2016; 17: 1061–9.

27. Lee WR, Dignam JJ, Amin MB et al. Randomized phase III noninferiority study comparing two radiotherapy fractionation schedules in patients with low-risk prostate cancer. *J Clin Oncol* 2016; 34: 2325–32.

28. Catton CN, Lukka H, Gu CS et al. Randomized trial of a hypofractionated radiation regimen for the treatment of localized prostate cancer. *J Clin Oncol* 2017; 35: 1884–90.

29. Teckie S. High-dose hypofractionated radiotherapy is effective and safe for tumors in the head-and-neck. *Oral Oncol* 2016; 60: 74–80.

30. Dixon LM, Douglas CM, Shaukat SI et al. Conventional fractionation should not be the standard of care for T2 glottic cancer. *Radiat Oncol* 2017; 12.

31. Pathmanathan AU, van As NJ, Kerkmeijer LGW et al. Magnetic resonance imaging-guided adaptive radiation therapy: A “Game Changer” for prostate treatment? *Int J Radiat Oncol Biol Phys* 2018; 100: 361–73.

32. Takeda T, Tin AL, Corradi RB et al. Topography of prostate cancer recurrence after radiation therapy: a detailed mapping study of salvage radical prostatectomy specimens. *Eur Urol* 2018; 73: 488–90.

33. Charnley N, Morgan A, Thomas E et al. The use of CT-MR image registration to define target volumes in pelvic radiotherapy in the presence of bilateral hip replacements. *Br J Radiol* 2005; 78: 634–6.

34. Rosewall T, Kong V, Vesprini D et al. Prostate delineation using CT and MRI for radiotherapy patients with bilateral hip prostheses. *Radiother Oncol* 2009; 90: 325–30.