

ARTICLE INFO

Submitted: 19/01/2020

Accepted: 17/09/2020

Online: 24/12/2020

Quantification of Metal Artefacts from Orthodontic Brackets in CT Images: A Modified Method

Mahmud Mohammed^a, Norma Ab. Rahman^{b*}, Ahmad Hadif Zaidin Samsudin^c, Johari Yap Abdullah^d

^a*School of Dental Sciences, Universiti Sains Malaysia, 16150 Kota Bharu, Kelantan, Malaysia*

^b*Orthodontic Unit, School of Dental Sciences, Universiti Sains Malaysia, 16150 Kota Bharu, Kelantan, Malaysia*

^c*Department of Radiology, School of Medical Sciences, Universiti Sains Malaysia, 16150 Kota Bharu, Kelantan, Malaysia*

^d*Craniofacial Imaging Laboratory, School of Dental Sciences, Universiti Sains Malaysia, 16150 Kota Bharu, Kelantan, Malaysia*

*Corresponding author: drnorma@usm.my

To cite this article: Mohammed M, Ab. Rahman N, Samsudin AHZ, Abdullah JY (2020). Quantification of metal artefacts from orthodontic brackets in CT images: A modified method. *Arch Orofac Sci*, 15(2): 109–117. <https://doi.org/10.21315/aos2020.15.2.428>

To link to this article: <https://doi.org/10.21315/aos2020.15.2.428>

ABSTRACT

Any metallic object within the CT scanning field can produce metal artefacts, which will degrade the diagnostic image quality. Previous methods described for quantifying this kind of artefacts were complicated and difficult to reproduce. The purpose of this study was to introduce a new rapid method for quantifying the artefacts produced in craniofacial CT images. This is an in-vitro experimental study. Four different compositions of orthodontic brackets were bonded consecutively in the tooth surfaces of a cadaveric skull head. All scans were performed by a single operator using the same CT machine followed by a standard scanning protocol. Artefact intensity for all data sets was quantified by following a modified method with a freely available open-source software ImageJ. All datasets were duplicated where metal artefacts were quantified according to the previous conventional method. Statistical analysis included independent samples *t*-test for validation and intraclass correlation coefficient (ICC) with a 95% confidence interval for both intra- and inter-examiner reliability. The modified method of measuring artefact score exhibited excellent intra (0.997–0.941) and inter-rater (0.996–0.905) reliability. In addition, no significant difference ($p = 0.072$) of mean artefact score was noted between the groups measured by the modified method and the conventional method. This modified method for measuring the artefact intensity is valid and reliable.

Keywords: *Artefacts; computer-assisted image analysis; diagnostic imaging; multislice CT; radiographic image interpretation*

INTRODUCTION

In CT scan imaging, metal artefacts are common occurrences, which resulted from the presence of metal objects such as dental restorations, orthodontic brackets, implants,

surgical plates, and pins in the scanning field. The metal objects are highly attenuated by the X-ray beam resulting in incorrect high attenuation values of objects behind the metal and produce artefacts within the images (Klinke *et al.*, 2012). Metallic objects

with higher atomic number, such as stainless steel or platinum produce more artefacts in comparison with low atomic number metals, such as titanium and nickel (Prell *et al.*, 2010; Boas and Fleischmann, 2011).

Several studies have been conducted to reduce the metal artefacts (Boas and Fleischmann, 2011; Mouton *et al.*, 2013). Previously, it was found that metal artefacts can be reduced by using newer reconstruction techniques or approaches such as metal artefact reduction for Orthopaedic Implants (O-MAR) (Kidoh *et al.*, 2014).

However, there is no appropriate method to measure the size of metal artefacts because they will show ill-defined edge and usually will be distributed irregularly in the form of dark and bright streaks within the image. In one study, the metal artefact was quantified using the naked eye and scoring of the artefact was done by two or more expert radiologists (Gunzinger *et al.*, 2014). This kind of artefact quantification is more effective for in-vivo study where the radiologist scores the artefact according to the soft tissue visibility whereas, for in-vitro phantom based study, it is difficult to score the artefacts because there is no soft tissue structure in the phantoms. Apart from that, several studies were conducted to quantify the metal artefacts in CT images using different types of techniques (Baek *et al.*, 2008; Chindasombatjaroen *et al.*, 2011; Pauwels *et al.*, 2013; Hirschinger *et al.*, 2015; Hokamp *et al.*, 2020). However, these methods had some limitations, which includes the complex procedure, not applicable in all type of artefact quantification, complex description of methods and inadequate information regarding the method.

The purpose of this study was to quantify the artefacts produced in craniofacial CT images with a simple modified method of artefact quantification. The outcome of this study

can provide an idea about the number of artefacts produced in CT images and result in easier quantification of the metal artefact in medical images for future study.

MATERIALS AND METHODS

This is an in-vitro experimental study where single unit conventional CT scan machine (TOSHIBA CGGT-032A, Japan) and human cadaveric skull including teeth present in both jaws were used for all the scans. Before conducting this study, ethical approval was obtained from the Human Research Ethics Committee of Universiti Sains Malaysia (USM/JEPPEM/17070352). This research was conducted following relevant guidelines and regulations. Scanning of the cadaveric skull without any orthodontic brackets was done for image registration. For calibration, each time before scanning, the cadaveric skull head was placed in a constant position by the help of the laser light of the CT machine and a custom-made box. A total of four types of brackets were bonded consecutively to the cadaveric skull by following the manufacture bonding and debonding procedure. A standard scanning and reconstruction protocol was used for all the scanning (matrix size 512512 field of view 200200, slice thickness 1 mm, peak tube voltage 120 kV, tube current 225 mAs and pitch 0.75s).

Sample Size Calculation

Sample size calculation was done by using G power software, version-3 with the power of 80%, alpha error probability 0.05 and a large effect size 0.39 a total of 64 samples were calculated. The effect size was calculated from the mean and standard deviation (SD) of the previous study (Klinke *et al.*, 2012). Therefore, according to the sample size calculation we have selected 64 artefact affected axial slices.

Image Analysis

a) Image registration

Signal loss and signal amplification within all datasets were visualised by using the freely available open-source software (ImageJ version 1.52e Wayne Rasband, National Institute of Health, USA). After visualising all the datasets, only metal artefact affected axial slices were accounted by a radiologist for image analysis. From each scan, 16 axial

slices were selected for artefact evaluation. After that, all 16 axial slices were manually registered with the corresponding control slices with ImageJ (Figs. 1 and 2). The same series of sagittal slices were also used for image registration. Total 64 axial slices were selected for evaluating artefact. As the artefacts within the images were distributed irregularly, the placement of region of interest (ROI) cannot be performed in the sagittal slices.

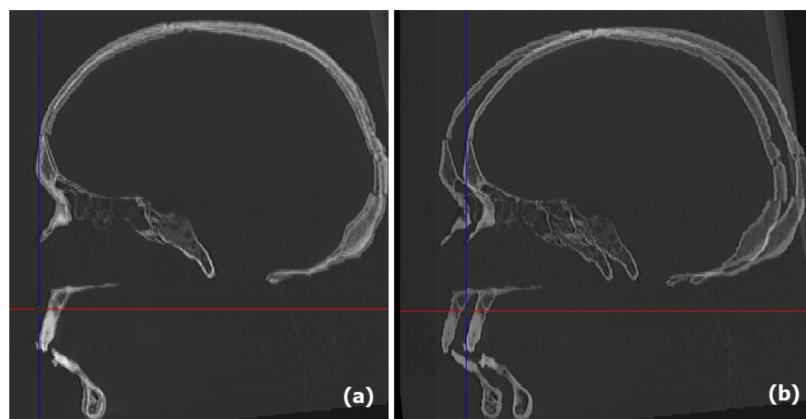


Fig. 1 Manual alignment and synchronisation of the stack with control stack (a) before synchronisation of the stack; (b) after synchronisation of the stack.

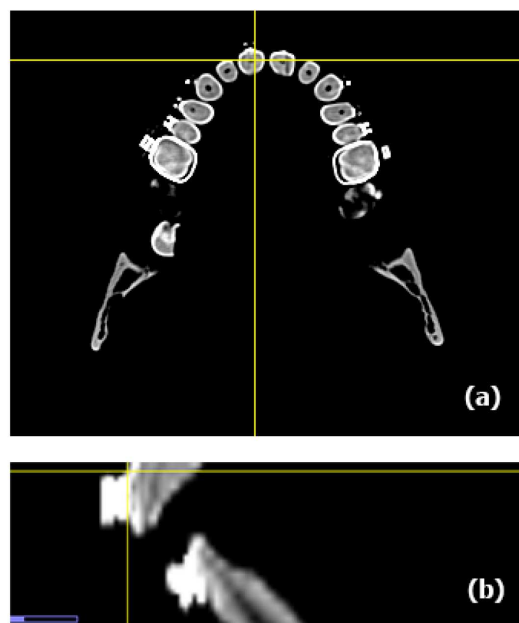


Fig. 2 Registration of axial and sagittal slice (a) axial synchronised slice of stainless-steel bracket; (b) sagittal synchronised slice of stainless-steel bracket.

b) Placement of the ROI

A fixed ROI (364254 pixels) was placed in the axial view of all synchronised axial slices (Fig. 3). This ROI covered the entire area of 364254 pixels including the teeth, bone and orthodontic appliance itself. The radiologist determined the location of the ROIs. The location of the ROIs was constant for all of the datasets, which were determined by the anatomical landmark. Six different anatomical landmarks were selected for placing the ROIs, which made ROIs constant in all axial slices. There are total six points in the ROI (Fig. 3), from there the top middle point ($x, y = 252, 22$) and the lower middle point ($x, y = 252, 258$) need to be placed

between the upper central incisor. The lower right point and lower left point of the ROI were placed next to the lower border of the right ($x, y = 141, 225$) and left condyle ($x, y = 361, 222$) of the mandible. The right and left middle point need to be placed parallel to the distal canal of right ($x, y = 96, 140$) and left ($x, y = 408, 140$) second molars.

After that, all datasets were duplicated, the metal and bones were removed manually from that duplicated datasets and calculated the artefact intensity according to the conventional method (Fig. 4b). The original datasets were used for measuring the artefact intensity according to the modified method.

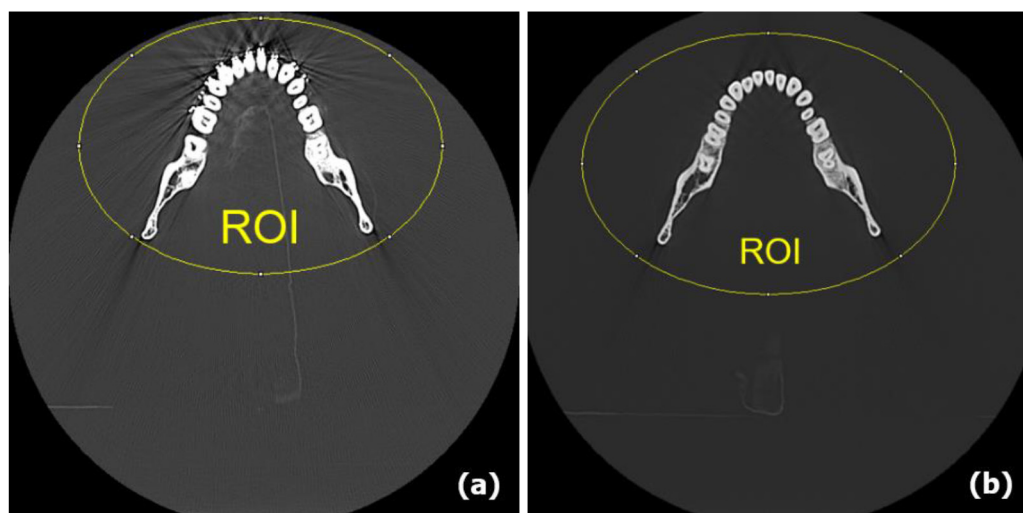


Fig. 3 Placement of ROI within metal artefact affected slice (a) stainless steel orthodontic bracket; (b) ROI within the same series of control slice.

c) Calculation of artefact intensity with the modified method

Artefacts intensity within the original CT images were measured by using ImageJ. Removal of the teeth, bone and orthodontic brackets from the ROIs were done by manual thresholding (Fig. 4). As it was seen before, there was no appropriate standardised segmentation method for removing the metal and bone from the ROIs. For this reason, the maximum and minimum threshold

for all images were selected manually for excluding the bone, teeth and metal which was constant for all images. The maximum threshold for all of the data was +100 HU to exclude the bone, teeth and appliance from ROIs and the minimum lower limit threshold for all the data was -3072 HU. For selecting the maximum and minimum threshold, grey value for metal, bone and teeth was calculated to remove them from the images. Grey values which lied within the red threshold areas of ROIs were only accounted

for image analysis (Fig. 4a). This maximum thresholding value depends on the maximum grey value of the metal and bones.

After doing the procedures mentioned above, the ImageJ automatically measured the SD within the fixed ROIs of all axial slices. After calculating the SD, the actual SD within all stacks were divided by this maximum (max) SD and multiplying with 100 $[(\frac{SD}{max\ SD}) \times 100]$ and this will represent

the actual SD in a percentage of its max SD. A higher percentage indicates a more pronounced artefact score. The percentage of the max SD estimates the overall extent of darkening and brightening streak artefact within the ROIs. For the modified method, the max SD was constant for all the datasets as we have removed the bone and metal from the images with the help of thresholding tool. In the modified method, there was no need for calculating the max SD for all images.

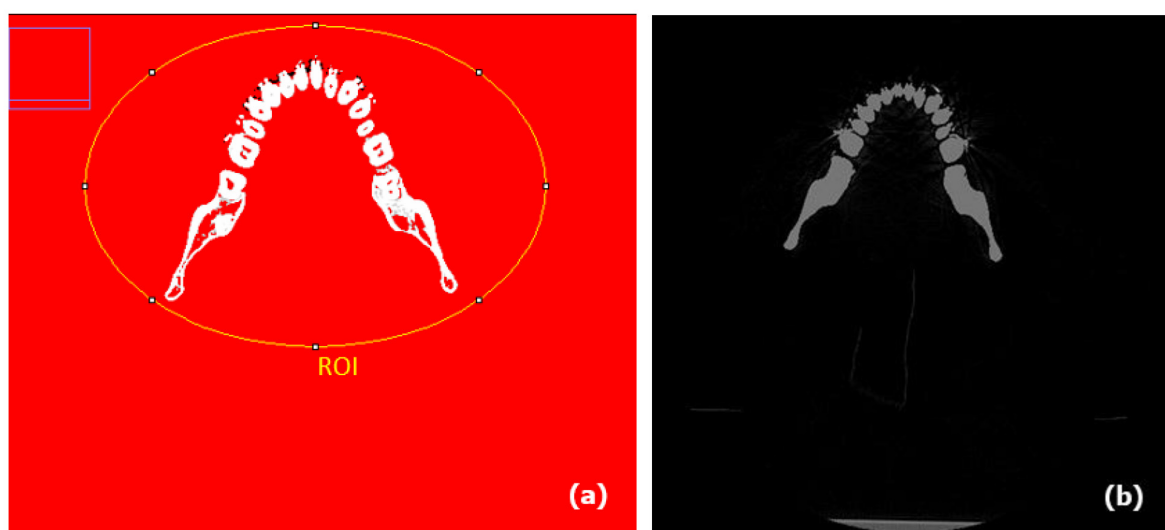


Fig. 4 Removal of bone, teeth and orthodontic brackets within the ROI (a) by thresholding, the red areas representing the threshold area between the maximum and minimum threshold and the white areas representing the area which is more than +100 threshold; (b) by the conventional method.

- d) Calculation of artefact intensity with the conventional method

$$\max SD = \sqrt{\frac{(\max\ grey\ values - \text{mean})^2 + (\min\ grey\ values - \text{mean})^2}{(n-1)}}$$

The mean, SD, maximum and minimum grey values for duplicated datasets were also calculated for measuring the artefact intensity according to the conventional method (Pauwels *et al.*, 2013). The theoretical max SD was calculated via the maximum and minimum grey values for each axial slice. In the conventional method, the bones, teeth and metal were removed manually from all the axial slices. For that reason, the max SD need to be calculated in all images according to the equation given by (Pauwels *et al.*, 2013). The equation for calculating the max SD within the datasets is:

Repeated Measurement for Data Analysis

One observer conducted each measurement twice at an interval of two weeks to consider the variability (Arshad *et al.*, 2017). In this current study, there was a two-week interval between the first and second observation to eliminate memory bias. Therefore, for each data, the measurements were carefully repeated four times by two observers to achieve an acceptable consistency. The values of all four measurements were used for the validity analysis.

Statistical Analysis

Statistical analyses were performed by using the SPSS software version 24.0. The level of significance was considered, $p < 0.05$. The normality of data distribution was confirmed by the descriptive histogram. Intra and inter-rater reliability of artefact measurement methods were tested by the intraclass correlation coefficient (ICC) test with a 95% confidence interval. Two-way mixed-effect, absolute agreement, average measure ICC (2,1) models were indicated for the intra-rater and inter-rater reliability test (Koo and Li, 2016). Independent sample t -test was applied to detect the mean difference of artefact score between the modified method and the conventional method. For intra and inter-examiner reliability of this modified method, the ICC tests were done. The level of significance was considered at $p < 0.05$.

RESULT

Reliability Study

The purpose of this study was to quantify the artefacts produced in craniofacial CT

images with a simple modified method of artefact quantification. In this study, artefacts within all 64 CT images were calculated by the conventional and modified methods. According to the result, the modified method for measuring artefact intensity exhibited a similar output like the conventional method. There was no significant difference ($p = 0.072$) of mean artefact score noted between the images measured by the modified and conventional method (Table 1). The mean artefact score measured by the modified method and the conventional method was respectively 6.811 ± 0.493 and 6.769 ± 0.499 .

Validation Study

The artefact scores exhibited excellent intra and inter-rater reliability because the ICC values of intra-rater reliability test were 0.939 and the ICC values for inter-rater reliability was 0.923 (Tables 2 and 3).

DISCUSSION

In this experimental study, the artefacts produced in craniofacial CT images were

Table 1 Artefact score between the modified method and conventional method

Variable	Mean $\pm$ SD		T-statistic	p-value
	Modified method (n = 64)	Conventional method (n = 64)		
Artefact score	6.811 $\pm$ 0.493	6.769 $\pm$ 0.499	0.375	0.709

Note: The mean difference is significant at the 0.05 level; n = number of samples.

Table 2 Intra-rater reliability of artefact scores

Variable	Measure	ICC	95% Confidence interval		F-test with true value 0			
			Lower bound	Upper bound	Value	df1	df2	p-value
Artefact score	Single	0.939	0.889	0.967	31.756	39	39	< 0.001

Note: Two-way mixed-effect, absolute agreement, average measure ICC (2,1) models were indicated for the intra-rater reliability test.

Table 3 Inter-rater reliability of artefact scores

Variable	Measure	ICC	95% Confidence interval		F-test with true value 0			
			Lower bound	Upper bound	Value	df1	df2	p-value
Artefact score	Single	0.923	0.860	0.958	25.497	39	39	< 0.001

Note: Two-way mixed-effect, absolute agreement, average measure ICC (2,1) models were indicated for the inter-rater reliability test.

quantified with a simple modified method. In a previous study, Sanders *et al.* (2007) used the mean grey values instead of the SD of grey values for measuring the artefact intensity in CT images. Sometimes, the mean grey values of CT image can be a negative value, which can mislead the quantification. In such condition, the SD of the grey values becomes the appropriate unit to be used instead of the mean grey values in quantifying the artefact intensity.

In another study, metal artefacts from different types of dental material were quantified in cone beam CT (CBCT) and multi-detector CT images. According to that method, black and white part of the artefact were quantified independently (Chindasombatjaroen *et al.*, 2011). There are some limitations with this method as the phantom itself generated some artefacts making it difficult to distinguish the artefacts. Additionally, overlapping of the black and white parts of the artefacts can mislead the quantification (Gunzinger *et al.*, 2014). Pauwels *et al.* (2013) proposed a new method for quantifying the artefact in CBCT images through the variation of the SD of the grey values. In that study, there was no clear indication regarding the removal of bone and metal component from the image. The researchers removed the bone and metal from the images manually and the whole procedure was quite complex and time-consuming.

In this study, artefact intensity was evaluated using the ImageJ by comparing the SD of grey values of the images, with higher values representing more pronounced artefacts and vice versa. SD being a function of variance, would reflect artefact intensities because larger variance within the ROIs indicated more artefact-induced noise.

On the other hand, a lower SD represents a homogenous image (Pauwels *et al.*, 2013; Queiroz *et al.*, 2018). Furthermore, the present study converted the SD into the percentage of its max SD and termed as artefact score, which made the result

comparable with the other studies as all the CT datasets do not use the conventional 12-bit scale (4,096 grey values). In addition, bones and metals were removed from all the CT images by manual thresholding, the modified method introduced by the present study, which made the method simple and less time-consuming.

One of the limitations of measuring artefact score (SD percentage) is the placement of ROI within the images. For this reason, in all cases while measuring the artefact score, the ROIs had to be placed in a standard position around the bony landmark of images. To confirm the consistency of the modified method measuring the artefact score, the intra and inter-rater reliability tests were required. In one particular study, the researchers had examined the intra and inter-rater reliability of their measured artefact score (Pauwels *et al.*, 2013). In that study, all measurements were performed by two examiners to estimate the inter-rater reliability. Each examiner performed the measurement two times within the same data to check the intra-examiner reliability. In the current study, all datasets were selected for assessing the intra and inter-rater reliability. According to Koo and Li (2016), with a 95% confidence interval, the ICC value should be more than 0.9 to ensure excellent and intra and inter-rater reliability. In this present study, the ICC values for inter-rater reliability test was 0.923 and intra-rater reliability test was 0.939, which means that the modified method for measuring the artefact intensity is consistent and reliable.

Previously, Hirschinger *et al.* (2015) quantified artefacts in CT images from different types of orthodontic brackets according to the method described by Pauwels *et al.* (2013) with few modifications. Hirschinger *et al.* (2015) used the line profile tool instead of a ROI tool to measure the SD of grey values. The line profile tool can only quantify the variation of the grey values along the line. Hirschinger *et al.* (2015) bonded a single orthodontic bracket on the buccal surface of a molar tooth and quantified the

artefact from it and because of this, the line profile tool was the choice of measuring tool. However, the line profile tool is used for calculating the pixels, which lied within a straight line. This means it can calculate the pixels, which lie within a given line. On the other hand, the ROI tool has various shapes, such as a square or a circular shape. When pixels are distributed irregularly within the image, the ROI tools are more effective than the line profile tool. For the same reason, in this current study, the ROI tool is more appropriate instead of a line profile tool.

The modified method is less time consuming and simple to perform, which made the artefact quantification easier in CT images. In this study, the simplified quantification method used ImageJ to reduce the time and cost when measuring the artefact intensity in the CT images. Previously, there was no standardised method to eliminate the metal and bone from CT images, which made the artefact quantification time consuming and became a complex procedure (Pauwels *et al.*, 2013). In this study, metal and bone from all CT images were removed by the thresholding tool of the ImageJ which helps removing the bone and metal from the image without any complex procedure, and make the artefact quantification more accurate and less time-consuming.

CONCLUSION

This modified method for quantifying the artefact intensity is valid and reliable. This method makes the quantification of the artefact or noise easier in the medical images without any complex mathematical analysis or algorithm. This will help to reduce the time when measuring the artefacts and overcome the misdiagnosis in the CT images. Additionally, all the image analysis procedures were done by using a freely available open-source software that could be reproduced without any unnecessary costing.

REFERENCES

- Arshad AI, Alam MK, Khamis MF (2017). Assessment of complete unilateral cleft lip and palate treatment outcome using EUROCRAN index and associated factors. *Int J Pediatr Otorhinolaryngol*, **100**: 91–95. <https://doi.org/10.1016/j.ijporl.2017.06.025>
- Back CH, Chung MK, Son YI, Choi JY, Kim HJ, Yim YJ *et al.* (2008). Tumor volume assessment by 18F-FDG PET/CT in patients with oral cavity cancer with dental artifacts on CT or MR images. *J Nucl Med*, **49**(9): 1422–1428. <https://doi.org/10.2967/jnumed.108.051649>
- Boas FE, Fleischmann D (2011). Evaluation of two iterative techniques for reducing metal artifacts in computed tomography. *Radiology*, **259**(3): 894–902. <https://doi.org/10.1148/radiol.11101782>
- Chindasombatjaroen J, Kakimoto N, Murakami S, Maeda Y, Furukawa S (2011). Quantitative analysis of metallic artifacts caused by dental metals: Comparison of cone-beam and multi-detector row CT scanners. *Oral Radiology*, **27**(2): 114–120. <https://doi.org/10.1007/s11282-011-0071-z>
- Gunzinger JM, Delso G, Boss A, Porto M, Davison H, von Schulthess GK *et al.* (2014). Metal artifact reduction in patients with dental implants using multispectral three-dimensional data acquisition for hybrid PET/MRI. *EJNMMI Phys*, **1**: 102. <https://doi.org/10.1186/s40658-014-0102-z>
- Hirschinger V, Hanke S, Hirschfelder U, Hofmann E (2015). Artifacts in orthodontic bracket systems in cone-beam computed tomography and multislice computed tomography. *J Orofac Orthop*, **76**(2): 152–163. <https://doi.org/10.1007/s00056-014-0278-9>

- Hokamp NG, Eck B, Siedek F, dos Santos DP, Holz JA, Maintz D *et al.* (2020). Quantification of metal artifacts in computed tomography: Methodological considerations. *Quant Imaging Med Surg*, **10**(5): 1033–1044. <https://doi.org/10.21037/qims.2020.04.03>
- Kidoh M, Nakaura T, Nakamura S, Tokuyasu S, Osakabe H, Harada K *et al.* (2014). Reduction of dental metallic artefacts in CT: Value of a newly developed algorithm for metal artefact reduction (O-MAR). *Clin Radiol*, **69**(1): e11–e16. <https://doi.org/10.1016/j.crad.2013.08.008>
- Klinke T, Daboul A, Maron J, Gredes T, Puls R, Jaghsi A *et al.* (2012). Artifacts in magnetic resonance imaging and computed tomography caused by dental materials. *PLoS One*, **7**(2): e31766. <https://doi.org/10.1371/journal.pone.0031766>
- Koo TK, Li MY (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*, **15**(2): 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Mouton A, Megherbi N, Van Slambrouck K, Nuyts J, Breckon TP (2013). An experimental survey of metal artefact reduction in computed tomography. *J Xray Sci Technol*, **21**(2): 193–226. <https://doi.org/10.3233/XST-130372>
- Pauwels R, Stamatakis H, Bosmans H, Bogaerts R, Jacobs R, Horner K *et al.* (2013). Quantification of metal artifacts on cone beam computed tomography images. *Clin Oral Implants Res*, **24**(A100): 94–99. <https://doi.org/10.1111/j.1600-0501.2011.02382.x>
- Prell D, Kyriakou Y, Kachelrie M, Kalender WA (2010). Reducing metal artifacts in computed tomography caused by hip endoprostheses using a physics-based approach. *Invest Radiol*, **45**(11): 747–754. <https://doi.org/10.1097/RLI.0b013e3181e94384>
- Queiroz PM, Oliveira ML, Groppo FC, Haiter-Neto F, Freitas DQ (2018). Evaluation of metal artefact reduction in cone-beam computed tomography images of different dental materials. *Clin Oral Investig*, **22**(1): 419–423. <https://doi.org/10.1007/s00784-017-2128-9>
- Sanders MA, Hoyjberg C, Chu CB, Leggitt VL, Kim JS (2007). Common orthodontic appliances cause artifacts that degrade the diagnostic quality of CBCT images. *J Calif Dent Assoc*, **35**(12): 850–857.