Frontiers in Dentistry

Relationship between Voxel Value of Cone Beam CT and Hounsfield Unit of CT: Systematic Review and Meta-Analysis

Daryush Goodarzipour[1], Ahmad Reza Shamshiri[2], Zahra Khodadadi[3]*

1. Department of Radiology, School of Dentistry, Tehran University of Medical Sciences, Tehran, Iran
2. Research Center for Caries Prevention, Dentistry Research Institute, Department of Community Oral Health, School of Dentistry, Tehran University of Medical Sciences, Tehran, Iran.
3. Department of Oral and Maxillofacial Radiology, School of Dentistry, Shahed University of Medical Sciences, Tehran, Iran

Article Info

Article type:
Systematic Review

Article History:
Received: 20 Jan 2025
Accepted: 25 Apr 2025
Published: 28 Jan 2026

**** Corresponding author:***
Department of Oral and Maxillofacial Radiology, School of Dentistry, Tehran University of Medical Sciences, Tehran, Iran.

Email: z.khodadadi@shahed.ac.ir

A B S T R A C T

This systematic review aims to investigate the relationship between voxel value obtained from Cone Beam Computed Tomography (CBCT) in studies compared to Hounsfield of Multidetector Computed Tomography (MDCT) in homogeneous and heterogeneous samples. A literature search was carried out in the databases PubMed, Scopus, Embase, and Web of Science searching for relevant literature until February 2022 (updated at July 2023). A risk of bias assessment of the studies was performed using a modified checklist based on the Cochrane Collaboration's tool and the Journal of Biomedical Informatics. The software version 20.104 of MedCalc was used to conduct the meta-analysis of correlation coefficients. Out of 4750 articles in the initial search, 13 met the eligibility criteria. Out of the articles, eight studies were included in the meta-analysis. Both heterogeneous and homogenous samples showed a strong correlation between the voxel value of CBCT and Hounsfield Unit (HU), with high heterogeneity (r=0.900 and 0.998 respectively and I2>70%). Two other meta-analyses were conducted for kVp and voxel size, which showed a high correlation. The 95% confidence interval was used to present the estimated pooled correlation. The strong correlation of voxel value and HU indicates the possible potential of CBCT in radiographic bone density measurement. However, further research is needed to obtain an accurate conversion equation for translating voxel values of CBCT to HU. Calibration of voxel values within each scan using a reference object and consideration of both linear and non-linear regression could improve accuracy.

Keywords: Cone-Beam Computed Tomography, Hounsfield Unit, Voxel Value, Gray Scale

➢ **Cite this article as:** Goodarzipour D, Shamshiri AR, Khodadadi Z. Relationship between Voxel Value of Cone Beam CT and Hounsfield Unit of CT: Systematic Review and Meta-Analysis. *Front Dent*. 2026:23:05. http://doi.org/10.18502/fid.v23i5.20888

INTRODUCTION

Cone Beam Computed Tomography (CBCT) is widely used in dentistry as a complementary imaging technique to conventional methods. CBCT scan is an important imaging technique for implant treatment (evaluation of bone quantity before the surgery and examination of the adjacent vital structures such as the inferior alveolar canal and maxillary sinus), assessment of the maxillofacial lesions, trauma, and osteomyelitis. In addition, the voxel value of CBCT has been shown to aid in determining the type of bone for implant insertion, and prediction of implant stability. [1, 2]

The standard reference for radiographic measurement of Bone Material Density (BMD) is the Multidetector Computed Tomography (MDCT). BMD is used to evaluate the success of implant treatment and anchorage tools in orthodontics. [3-5] Hounsfield Unit (HU) of MDCT is a quantitative measurement of BMD and is proportional to the amount of X-ray

attenuation and is assigned to each pixel to display an image. Materials with higher attenuation coefficient (higher density) have higher HUs. The standard formula for calculating the HU is as follows [6]:

$$HU_{\text{material}} = 1000 \times \frac{(\mu_{\text{material}} - \mu_{\text{water}})}{\mu_{\text{water}}}$$

In CBCT, the attenuation rate is indicated by gray scale or voxel value. Although CBCT software manufacturers refer to the gray scale as HU, it is essential to note that these HU measurements are not necessarily accurate [1, 2].

CBCT, in addition to lower dose and cost and less image acquisition time, has better spatial resolution, contrast range, and gray scales than MDCT. However, due to the artifact associated with beam hardening, pyramid shape of X-ray, the lack of a standard method for the CBCT scanner, and the exposure parameters, it is unclear whether the voxel value can represent HU [7, 8].

MDCT units typically operate at 120 kV, while dental CBCT units operate within a nominal kV range of 70 to 120 kV. These different operating voltages result in varying effective energies between the two types of scanners, with CBCT being associated with higher levels of scatter radiation and artifacts compared to MDCT. This can make estimating bone density using CBCT more challenging. Despite these limitations, some studies have identified a linear relationship between the Hounsfield Units (HU) obtained from MDCT and the voxel values derived from CBCT. Moreover, authors have suggested that it may be possible to estimate bone density using the voxel values of CBCT by applying a standard conversion formula that relates gray values (GV) to HU [1, 7-12]. On the other hand, some other studies are concerned about the possibility of converting gray value to HU [6, 13, 14]. There are fundamental differences between MDCT and CBCT, which complicate the use of quantitative voxel value (e.g., size of field size (FOV), relatively high scatter radiation and limited reconstruction algorithms) [15].

There is no consensus about the use of the voxel value of CBCT for BMD assessment. This systematic review aims to investigate the relationship between voxel value obtained from CBCT in studies compared to the HU of MDCT in two groups of homogeneous and heterogeneous samples.

MATERIALS AND METHODS

This systematic review was done according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) checklist.

Study design:

This review involves in vitro and in vivo studies according to the following PICO: (P) Population: homogenous and heterogeneous samples, (I) Intervention: GV of CBCT scan, (C) Comparison: HU of MDCT, (O) Outcome: correlation of the GV and HU.

The homogenous samples include almost uniform-density materials, such as iodine solution. The heterogeneous group consists of samples with various densities throughout the material and subsequently various densities in the ROI.

Inclusion and exclusion criteria:

All the studies conducted on heterogeneous and homogeneous samples focusing on comparing the voxel value and HU are included. The included studies must use a gold standard to determine the HU: phantom or material with known HU or density, or MDCT scan. Additionally, they must report the correlation coefficient (r-value) or the coefficient of determination (R2-value) of GV of CBCT and HU.

Exclusion criteria are review studies, letter to editor, personal opinion, book chapter, conference paper, and studies with no gold standard or with a different aim.

Search strategy and Study selection:

PubMed, Scopus, Embase, and Web of Science databases were searched for relevant literature until February 25, 2022. The search was subsequently updated as of July 11, 2023. Also, a google scholar search was done for gray literature. The search strategy is mentioned in Supplementary 1. No date or language limit was applied.

The selection process of the studies was performed independently by two authors (DG, ZK): first, by title and abstract, followed by a review of the full text. Disagreements were

resolved by a third opinion (AS).

Data collection process and data items:

All the included studies were evaluated according to the inclusion and exclusion criteria, and the key features were extracted: author(s), publication year, record number(s) (The sample size was considered the number of records used for r or R2 calculation, which is not the same as the patient or phantom's number.), sample type (homogeneous or heterogeneous), CBCT machine brand, exposure parameters, r or R2-value(s) (determined by Pearson correlation test). If available, mean GV and HU were also noted.

Risk of bias assessment:

A modified checklist was designed based on the Cochrane Collaboration's tool for assessing the risk of bias (RoB) in randomized trials and the Journal of Biomedical Informatics (JBI) to evaluate the risk of bias in the included studies. Since both checklists have applicable and non-applicable items for the purpose of the present study, the most relevant items were chosen. The modified checklist consists of five sections: performance of the study (blinding), reporting of outcomes, similarity of measurement method between the index test and standard reference, reliability of the measurement, and reproducibility of procedure (see Supplementary 2). The studies with 2 or 3 unclear scores in any section considered as unclear RoB and 4 or 5 unclear scores considered as high RoB. High score in any section indicated a high RoB.

Two authors (ZK and DG) cooperatively assessed Rob of the included studies. Any disagreements were resolved through the involvement of a third opinion (AS).

Statistical method:

The current study aimed to conduct a systematic review and meta-analysis of studies reporting r or R2-values and record's number. Rather than using sample size, reported number of the records was extracted from tables, figures, or text for the correlation meta-analysis. In studies reported R2-values, the square root was computed and employed as the R-value for analysis purposes.

The present study involved the classification of data into two distinct groups, namely heterogeneous and homogeneous samples. Moreover, subgroup analyses were used to determine the potential sources of heterogeneity. By combining the Pearson correlation coefficients of the included studies, a forest plot was produced. Pearson correlation coefficients were classified as weak (0.1 to 0.3), moderate (0.4 to 0.7), and strong (0.8 to 1) based on their interpretation. This study utilized Cochran's Q test to determine the presence of statistical heterogeneity, with a P-value<0.10 indicating statistically significant heterogeneity. Additionally, the magnitude of statistical heterogeneity between studies was assessed through the use of I2, where values of 25%, 50%, and 75% were considered low, moderate, and high heterogeneity levels, respectively [16]. The software version 20.104 of MedCalc was used to conduct the meta-analysis of correlation coefficients. The 95% confidence interval (CI) was used to present the estimated pooled correlation.

RESULTS

Study selection:

The initial search on the database identified 4750 papers. After removing the duplication, 2511 citations remained. Following the title and abstract evaluation, 58 were selected for full-text assessment. One paper [17] was included through cross-reference of the included papers, and another paper [18] included by the search update. Finally, 13 papers [7,11,17-27] met the eligibility criteria and were included in the systematic review. A flowchart summarizing this process is shown in Figure 1.

Study characteristics:

The included studies were conducted from 2010 to 2022, evaluating the correlation between CBCT voxel value and HU. They were evaluated in two groups of homogenous and heterogeneous samples, which were determined by the density distribution throughout the sample. The homogenous group consists of four papers [17,21,25,26] reporting 53 different scan protocols with the corresponding correlation and one study [22] reporting one R2-value for four scan protocols (Table 1). The heterogeneous group contains eight studies [7,11,18-20,23,24,27] and 21 different scan protocols (Table 2).

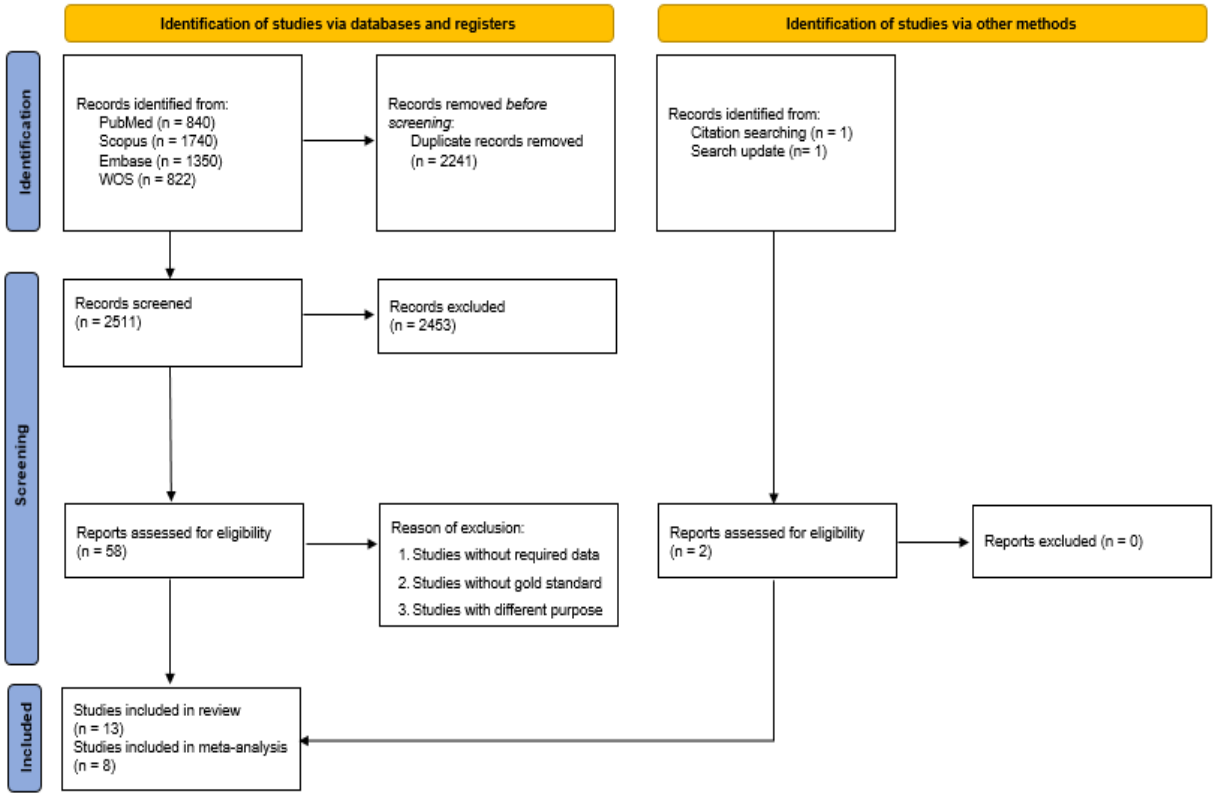

Fig 1. Preferred Reporting Items for Systematic Reviews (PRISMA) flow diagram for study selection

Risk of bias assessment:

Most of the included studies showed a low risk of bias for all evaluated scopes (selective reporting, similarity of the measurement methods, reference standard, reproducibility, and reliability of method), except the blindness, which was not mentioned in the studies (Table 3). The records number of Haghanifar et al.'s study [25] and Razi et al.'s study [20] and exposure factors of Khavid et al. [23] and Nomura et al.'s study[22] was not mentioned in the paper.

Synthesis of results

The included studies were grouped based on their samples, which were categorized as homogeneous or heterogeneous. In the homogeneous group, the maximum number of records was reported by Pauwels et al. (2013) [21] , while the minimum was reported by Nomura et al. [22] (145 and 12, respectively). In the heterogeneous group, Cassetta et al. [27] reported the maximum number of records, whereas Kamaruddin et al. reported the minimum (300 and 32, respectively). Regarding the r-values, Pauwels et al. (2013) [21] reported the minimum value (0.7014 for 3D Accuitomo XYZ, small FOV), whereas r-values ≥ 0.9990 were reported by Pauwels et al. (2011) [17] and Chindasombatjaroen et al [26] . In the heterogeneous group, Rostetter et al. [19] reported the minimum r-value (r-value = 0.469), while the maximum was reported by Cassetta et al. [27] and Khavid et al. [23] (r-value ≥ 0.97). Unfortunately, we were unable to determine the number of records in Haghanifar et al., Gaur et al. and Razi et al.'s paper [18,20,25] and the exposure parameters for Khavid et al. and Nomura et al.'s papers [22,23] either from the reported results or by corresponding with the authors. Therefore, meta-analysis was conducted with eight studies [11,17,19,21,24,26-28] .

The lowest correlation in the included studies was related to evaluation of left zygomatic bone of 19 patients (heterogeneous group) which was considerably different from its right counterpart (r-value= 0.877 (right side) compared to 0.469 (left side)).

Table 1. Summary of articles with homogeneous samples

Author	CBCT Device	KVp	mAs	Voxel size (µm)	FOV (mm)	Reference	sample	Number of records[1]	Correlation (r- value)
Chindasombatjaroen (2011) [16]	Alphard Vega 3030	80	102	390	200×200	MDCT (LightSpeed QX/i)	Iopamidol (at 8 different concentrations)	8	0.9998
		80	119	390	200×200			8	0.9998
		80	153	390	200×200			8	0.9998
		80	170	390	200×200			8	0.9998
		80	204	390	200×200			8	0.9998
		100	102	390	200×200			8	0.9998
		100	119	390	200×200			8	0.9998
		100	153	390	200×200			8	0.9998
		100	170	390	200×200			8	0.9998
		100	204	390	200×200			8	0.9999
Nomura (2010) [17]	3D Accuitomo					MDCT (Somatom)	Iodine with different concentrations	12	0.9909
Pauwels (2013) [18]	3D Accuitomo 170	90	155	250	17×12	MDCT (Somatom)	PMMA[2], air, aluminum, HA[3] 50 mg/cm[3], HA100, and HA 200	5	0.9984
		90	90	250	17×12			5	0.9982
		90	155	80	6×6			5	0.9967
		90	90	80	6×6			5	0.9969
	3D Accuitomo XYZ	80	72	125	4×3			5	0.7014
	GALILEOS Comfort	85	28	290	15×15			5	0.9965
	i-CATH NG	120	36.5	300	23×16			5	0.9951
		120	18.5	300	23×16			5	0.9948
		120	36.5	250	16×13			5	0.9972
		120	20	250	16×13			5	0.9885
		120	18.5	400	16×13			5	0.998
		120	10	400	16×13			5	0.9905
	Kodak 9000 3D	70	110	76	5×3.6			5	0.7997
	Kodak 9500	90	110	300	20×18			5	0.9959
		90	110	200	14.5×8.3			5	0.9991
	NewTom Vgi	110	180	240	12×8			5	0.9983
		110	40	240	12×8			5	0.9982
	PaX-Uni3D	85	120	200	5×5			5	0.9595
	Picasso Trio	85	115.2	200	12×7			5	0.9957

Table 1 cont'd

		85	72	300	12×7			5	0.9936
	ProMax 3D	84	168	160	8×8			5	0.9942
		84	168	320	8×8			5	0.9941
	SCANORAH 3D	85	57	200	10×7.5			5	0.9967
		85	30.4	200	10×7.5			5	0.9996
		85	30.4	300	10×7.5			5	0.9982
	SkyView	90	97.5	340	17×17			5	0.9958
		90	52	340	17×17			5	0.9957
		90	58.5	340	17×17			5	0.9947
	Veraviewepocs 3D	70	51	125	8×8			5	0.7777
Pauwels (2011) [19]	GALILEOS Comfort	85	28	300	150×150	MDCT (GE Prospeed, SOMATOM)	PMMA, air, aluminum, HA 50 mg/cm^3, HA100, and HA 200	5	0.998
	i-CAT Classic High-dose	120	35	200	160×80			5	0.9973
	i-CAT Classic Low-dose	120	10	400	160×80			5	0.987
	ILUMA Elite	120	76	200	210×140			5	0.9946
	ProMax 3D High-dose	84	168	160	80×80			5	0.9996
	ProMax 3D Low-dose	84	20	320	80×80			5	0.9993
	SCANORA 3D High-dose	85	36	130	60×60			5	0.9988
	SCANORA 3D Low-dose	85	24	200	60×60			5	0.9993
	SkyView High-dose	90	96	340	173×173			5	0.999
	SkyView Low-dose	90	52	340	173×173			5	0.9986
	Veraviewepocs 3D	70	51	130	80×80			5	0.6864
Haghanifar (2017) [20]	SORDEX	89	75.6	NA	6×7	MDCT (Somatom)	Iodine with different concentrations	NA	0.9969
	NEWTOM	110	6.912	NA	18×16			NA	0.9787

[1]The number of records used by the author to calculate r-value; [2]Polymethyl methacrylate; [3]Hydroxyapatite

Table 2. Summary of articles with heterogeneous samples

Ref.[1]	CBCT Device	KVp	Voxel size (µm)	FOV (mm)	Reference	Sample	Records[2]	VV[3]	HU	Diff.[4]	R[5]
[27]	Soredex SCANORA	90	250	13 × 14.5	MSCT (SOMATOM)	Cortical-cancellous bone (dry mandible)	100	1,053.31	744.35	308.96	0.977
						Cancellous bone (dry mandible)	100	816.62	572.45	244.17	0.931
						Cortical bone (dry mandible)	100	1505.26	1354	151.26	0.978
[24]	Planmeca	90	NA	NA	MDCT (Aquilion PRIME)	Head phantom	32	1098.162	1265.5	-167.34	0.911
[23]	NewTom GIANO	NA	NA	NA	MDCT (SOMATOM)	Bone blocks of cow ribs	52	1084.5	805.9	278.60	0.978
[7]	Newtom 5G	110	150	120×80	MDCT (Philips)	Cadaver (human mandible)	80	773.28	348.25	425.03	0.968
[11]	Accuitomo 170	90	80	40×40	MSCT (Philips)	Cadaver (human mandible)	20	377.49	222.85	154.64	0.89
[19]	KaVo 3D eXam	120	400	NA	MSCT (SOMATOM)	Dens axis (Patient)	19	309.9	538.2	-228.30	0.812
						Right maxilla (Patient)	19	318.4	468.7	-150.30	0.830
						Left maxilla (Patient)	19	262.4	420	-157.60	0.848
						Right mandible (Patient)	19	192	217.7	-25.70	0.648
						Left mandible (Patient)	19	237.6	221.2	16.40	0.734
						Right sphenoid (Patient)	19	-6.3	206	-212.30	0.842
						Left sphenoid (Patient)	19	22	233.6	-211.60	0.818
						Right mastoid (Patient)	19	-7.9	204.2	-212.10	0.8779
						Left mastoid (Patient)	19	-89.9	6.8	-96.70	0.888

Table 2 cont'd

						Right zygomatic bone (Patient)	19	488	705.6	-217.60	0.877
						Left zygomatic bone (Patient)	19	242.6	556	-313.40	0.469
[20]	Newtom VGi	110	300	NA	MSCT (SOMATOM)	Soft tissue (patient)	NA	NA	NA	NA	0.92
						Hard tissue (patient)	NA	NA	NA	NA	0.86
						General (patient)	NA	NA	NA	NA	0.95
[18]	NewTom	90	NA	110×80	MSCT (SOMATOM)	Goat head	NA	355.14	257.25	97.89	0.496

[1] Reference number; [2]The number of records used by the author to calculate r-value; [3]Voxel value; [4] Difference; [5] Correlation (r-value)

Table 3. Risk of bias assessment of the included studies

Author	Blinding	Reporting	Measurement	Reliability	Reproducibility
Chindasombatjaroen (2011) (26)	U[1]	L[2]	L	L	L
Nomura (2010) (22)	U	L	L	L	L
Pauwels (2013) (21)	U	L	L	L	L
Pauwels (2011) (17)	U	L	L	L	L
Cassetta (2013) (27)	U	L	L	L	L
Kamaruddin (2016) (24)	U	L	L	L	L
Khavid (2021) (23)	U	L	L	L	L
Parsa (2012) (7)	U	L	L	L	L
Parsa (2013) (11)	U	L	L	L	L
Rostetter (2015) (19)	U	L	L	L	L
Razi (2019) (20)	U	L	L	L	L
Haghanifar (2017) (25)	U	L	U	U	L
Gaur (2022) (18)	U	L	U	L	L

[1]Unclear; [2]Low

Although the robust measurement method of the study, which involved merging DICOM datasets and using the mean of 10 dots as the final radiodensity of the anatomical area [25], this r-value seemed an outlier data and no explanation was found in the paper. Thus, it was excluded from the present meta-analysis.

Meta-analysis

The included studies that report kVp, mA, number of records and R2/r-value were used to conduct the meta-analysis. The homogeneous group exhibited a wide spectrum of tube voltage ranging between 70-120 kVp.

Consequently, this group was further partitioned into two subgroups based on tube voltage; low voltage (70-89kVp) and high voltage (90-120kVp), for more robust meta-analytical investigations. Due to the high heterogeneity of the studies, a random effect model was used. The result showed a high correlation in each group: 0.998 (95% CI: 0.996 to 0.999) for homogeneous samples, 0.997 (95% CI: 0.992 to 0.999) for homogeneous samples with 70-89kVp, 0.999 (95% CI: 0.997 to 0.999) for homogeneous samples with 90-120kVp, 0.900 (95% CI: 0.843 to 0.937) for heterogeneous samples, and 0.993 (95% CI: 0.989 to 0.996) for all studies. All groups were found to be heterogeneous, with significant Q-test result (P<0.10) and inconsistency (I2) values ranging from 41.48%to 86.95%. With the exception of the high voltage homogeneous sample group with an I2 value of 41.48%, the inconsistency test results for all groups were high (I2>64%) (Table 4).

Additionally, two other meta-analyses were conducted for kVp and voxel size. The included studies were grouped based on the scan voltage: 70-89, and 90-120kVp, and due to the high heterogeneity of the data, random effect meta-analyses were used. The correlation coefficient was 0.997 (95% CI: 0.992 to 0.999), and 0.988 (95% CI: 0.979 to 0.993), respectively. Also, correlation meta-analyses of voxel size were 0.988 (95% CI: 0.979 to 0.993) for 75-250µm voxel size and 0.995 (95% CI: 0.987 to 0.998) for 250<µm voxel size (Table 4).

Table 4. Correlation meta-analysis of voxel value of CBCT and HU

Subgroup variable	Number of records	Correlation coefficient (random effect)	95% CI	Significance level (p-Value)	Inconsistency (I^2)
Sample type					
Heterogeneous samples	622	0.900	0.843 to 0.937	<0.001	86.95%
Homogeneous samples	280	0.998	0.996 to 0.999	<0.001	64.72%
Homogeneous samples (70-89kVp)	130	0.997	0.992 to 0.999	<0.001	76.35%
Homogeneous samples (90-120kVp)	150	0.999	0.997 to 0.999	<0.001	41.48%
Voltage (kVp)					
70-89	130	0.997	0.992 to 0.999	<0.001	76.35%
90-120	772	0.988	0.979 to 0.993	<0.001	89.70%
Total	902	0.993	0.988 to 0.996	<0.001	89.85%
Voxel size					
75-250µm	515	0.988	0.979 to 0.993	<0.001	75.87%
250<µm	355	0.995	0.987 to 0.998	<0.001	93.05%
Total	870	0.993	0.989 to 0.996	<0.001	89.85%

Forest plot for correlations of GV and HU in homogeneous samples with 70-89kVp and 90-120kVp is shown in Figure 2 and 3, respectively. Forest plot for correlations of GV and HU in heterogeneous samples is shown in Figure 4. The plot size is proportional to the precision of the estimated effect sizes (r-values), and the bars indicate the corresponding 95 % CIs. The diamond is placed on the summary correlation coefficient of the observational studies.

The overall summary estimate suggests a moderately strong positive correlation between GV and HU. However, the wide confidence interval around this estimate indicates that there is substantial uncertainty about the true underlying effect size. This may be attributed to the high heterogeneity in the studies and small sample size.

The subgroup analyses conducted suggest that sample type, low tube voltage and tube current may be sources of heterogeneity, as evidenced by the variation in Q and I2 values across different groups, which could explain some of the observed variability in effect sizes across studies. Tube voltage at 70kVp was found to be particularly problematic, resulting in low correlation in some exposure protocols with homogeneous samples (r-value < 0.8) [17,21]. When groups with identical tube voltage (90-120kVp) were compared for their heterogeneity, including homogeneous and heterogeneous samples as well as the total, it was found that the homogeneous samples exhibited the lowest inconsistency. (Table 4) Voxel size was also investigated; however, the findings remained inconclusive due to the presence of both protocol scans with lower correlation for voxel sizes less than 130 µm and those with a 400 µm voxel size. Certainly, there are other sources of heterogeneity that cannot be investigated thoroughly in this meta-analysis, such as differences in measurement methods, other parameters of exposure, variations in study conditions, and discrepancies between CBCT units from similar and different commercial brands.

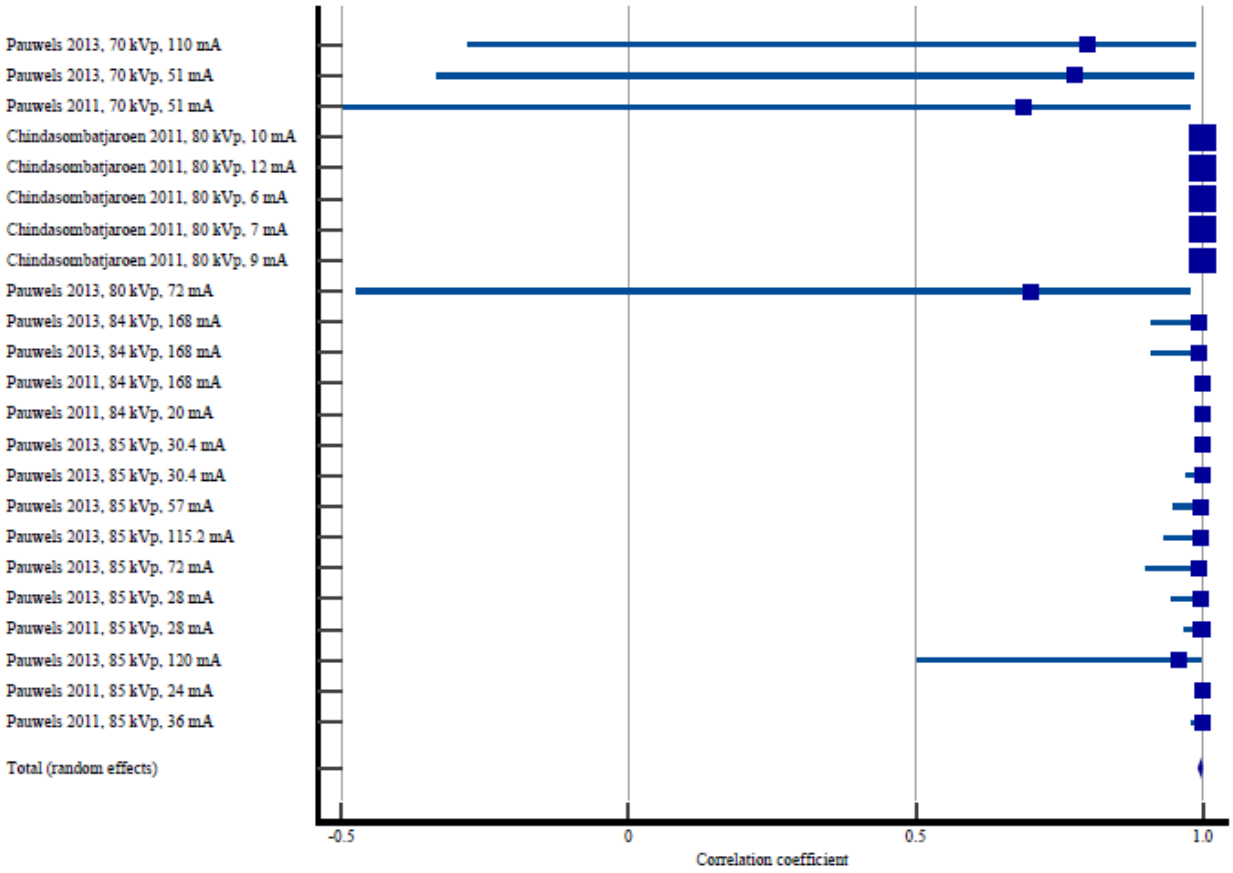

Fig 2. Forest plot of homogeneous sample with 70-89kVp

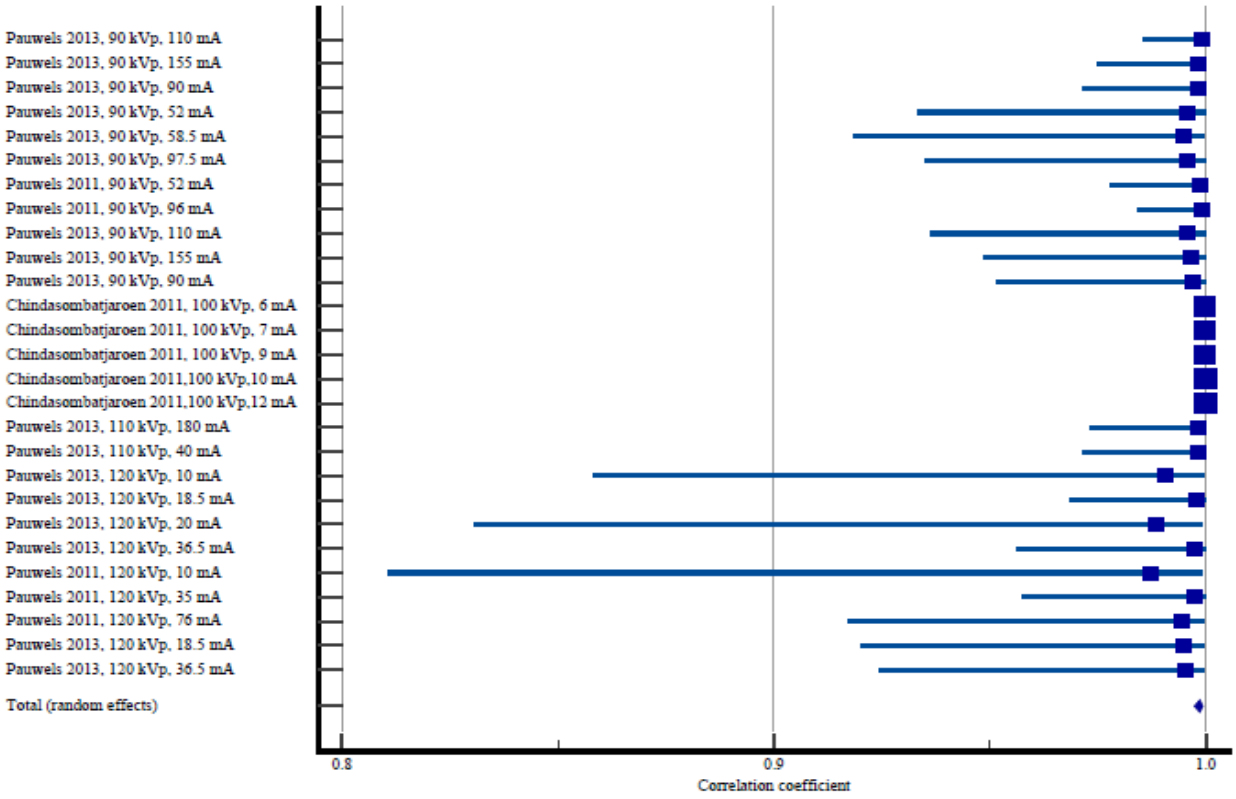

Fig 3. Forest plot of homogeneous sample with 90-120kVp

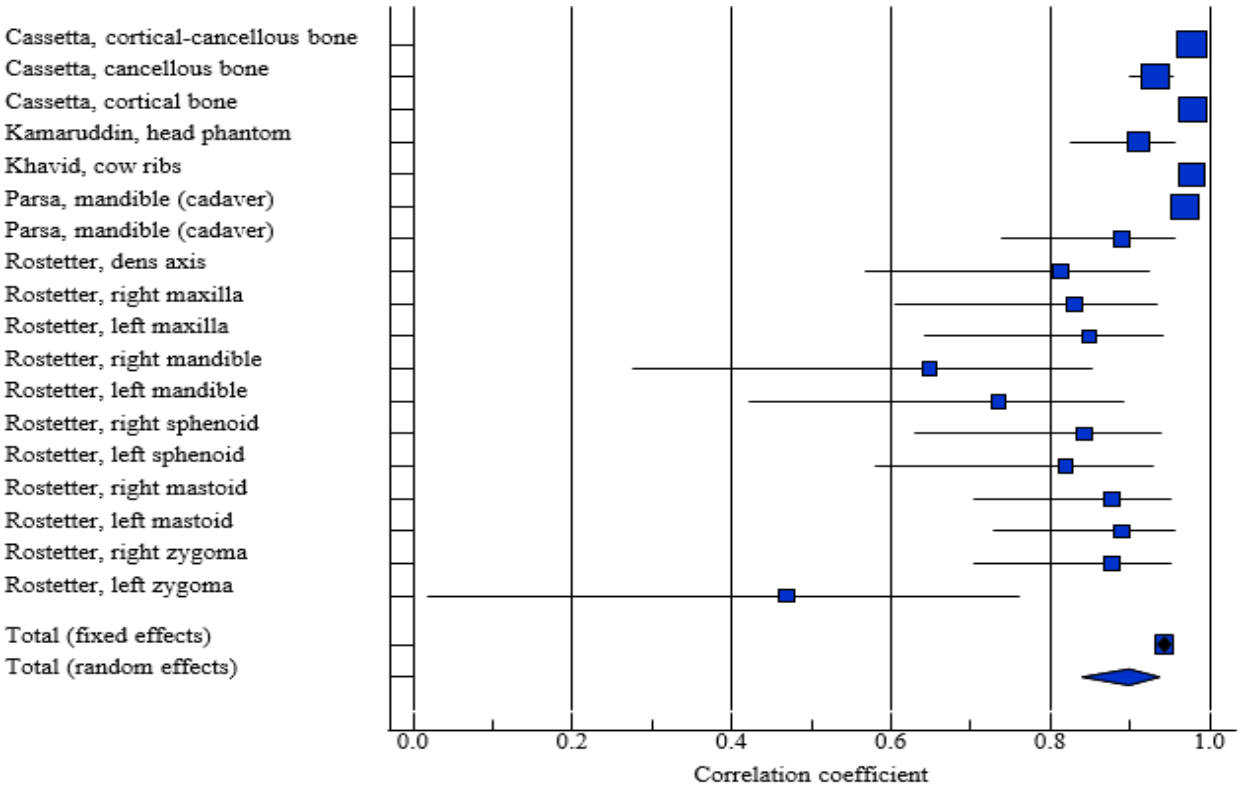

Fig 4. Forest plot of heterogeneous sample

DISCUSSION

CBCT is a widely used modality in dentistry, especially for preoperative implant planning. However, the reliability of the GV of CBCT has been questioned due to increased noise level, inherent scattering artifacts of the scanning technology, limitation of currently applied reconstruction algorithms that leads to lower contrast resolution than MDCT [6,7]. This systematic review and meta-analysis investigated the available evidence on the correlation of GV of CBCT and HU of MDCT as the reference standard, which demonstrated a moderately strong correlation (r=0.993). In addition to the overall meta-analysis, data were analyzed in subgroups of heterogeneity/ homogeneity sample and some exposure parameters.

Tube voltage is one of the most important exposure parameters that affect GV. In a literature review, El-Tabarany et al. classified this parameter as a significant factor [29]. In the present review, kVp of the studies were evaluated in two groups (70-89, 90-120kVp), and meta-analysis of the data revealed an almost similar correlation in both groups (r-value= 0.997, and 0.988, respectively). Also, it was found that the tube voltage served as a source of heterogeneity in the dataset of this study and was found to be associated with lower correlation in some exposure protocols (specifically, at 70kVp) [21]. In their study, Chindasombatjaroen et al. [26] found that increasing the tube voltage in MDCT led to a decrease in HU value, while tube current did not have a significant effect on density measurement. On the other hand, both tube current and tube voltage affect GV in CBCT scans, with an increase in either resulting in a decrease in GV. Additionally, Vasconcelos et al. demonstrated that higher kVp and the Metal Artifact Reduction (MAR) algorithm can reduce metal artifacts, resulting in less deviated GV [30]. Similarly, CBCT unites with high filtration and high kVp may decrease the risk of beam hardening artifacts, ultimately leading to more accurate measurements of attenuation coefficients and therefore GV [31].

Voxel size is another element that may influence the CBCT gray value, contributing to spatial resolution. Smaller voxel size enables CBCT images to exhibit fine details but at the cost of higher image noise, which indiscriminately affects the mean voxel value [11]. Voxel size is usually attached to the other exposure parameters, such as FOV and kVp. Pauwels et al. [21], evaluating the impact of different voxel sizes of 13 CBCT devices (ranging from 0.125 to 0.4mm) noted that this factor did not affect the accuracy of CBCT voxel value. Due to the variability of scan protocols (kVp, mA, FOV, etc.) and methodologies across the enrolled studies, this paper could not draw conclusive findings on voxel size.

Regarding the region of interest (ROI), fusing the MDCT and CBCT can reduce the possible human errors in the measurement of voxel value. Among the twelve studies included in this review, five employed observer-independent software to ensure accurate matching of CBCT and MDCT data, including four studies with heterogeneous samples [7,19,24,27] and one study with homogeneous samples [26]. Two other heterogeneous sample studies [11,20] used reference points to merge CBCT and MDCT images. When evaluating anatomical structures of patients who may have different head positions during scanning, determining the ROI is particularly crucial due to increased susceptibility to metal and motion artifacts. The studies with heterogeneous samples included evaluations of cadaver [7,11], head phantom [24], cow bone block [23], patient [19,20], and dry mandible [27]. However, these sample types have limitations that may deviated from the clinical situation and should be interpreted with caution; for instance, dry bone samples lack soft tissue and bone moisture, and the homogeneity of different samples in head phantoms while the physiological tissue are heterogeneous [7,20]. Homogeneous sample group evaluated different concentrations of media. In two studies [17,21], in addition to three different concentrations of hydroxyapatite, aluminum and air were evaluated as object with very high and very low density, respectively, which were associated with higher correlation of CBCT voxel value and HU. On the other hand, voxel

value of object with medium density showed larger error value and lower correlation. In another study [22], the samples included aluminum, but its correlation was not evaluated separately. In the homogeneous sample group, two studies using the same phantom, filled the sample container with polymethyl methacrylate (PMMA) and another study [22] used a water phantom to simulate a human head.

To the best of our knowledge, this is the first systematic review conducting a meta-analysis of the correlation of CBCT voxel value and MDCT HU in two groups of homogenous and heterogeneous samples. The heterogeneous samples revealed a lower correlation between CBCT voxel value and HU with higher heterogeneity of the studies (r= 0.900, I2=86.95%). Furthermore, when comparing heterogeneous samples with high voltage homogeneous samples (where the tube voltage for both groups was identical and between 90-120 kVp), it was observed that the second group exhibited less heterogeneity in the studies (I2 = 41.48% compared to 86.95%). These findings suggest that type of sample could be a potential source of heterogeneity in the studies, which has an impact on the correlation of GV and HU. It can be attributed to the presence of different densities along the FOV, such as soft tissue and adjacent anatomical structures, which makes the sample more similar to the clinical situation. Also, different densities such as bony lesions or anatomic structures adjacent to implants cause random effect of the noise on the measured gray value [21]. Furthermore, it must be considered that reconstruction algorithms assign gray values based on the amount and distribution of densities in the FOV to achieve an optimized contrast. To put it in another way, the presence of higher densities shifts the histogram, leading to a decrease the gray values in the reconstructed image [21]. For this problem, using a reference object in the FOV with known density could make it possible to calibrate the histogram, which results in a more accurate gray value [21,32]. Additionally, Komori et al. investigating ossicular chain and inner ear soft tissue, proposed to use fixed gray value thresholds for each purpose exclusively to standardized volume-rendering of CBCT images [33].

There have been two previous systematic reviews conducted on this topic. A recent study evaluating these two variables showed a positive correlation (r=0.669). The present study searched more databases and used a larger sample from the included papers to perform a meta-analysis. Our result is consistent with Selvaraj et al. [34] who found a good correlation between CBCT voxel value and HU, although we found a higher correlation (r=0.993), which may be attributed to a larger sample size and also, involving homogeneous samples in this study which showed a stronger correlation. In addition, Eguren et al. systematically reviewed the possibility of GV conversion to HU. They defined three steps: equipment calibration, correlation, prediction equation model, and standard formula. This study mentioned a lack of CBCT standardization and calibration as an obstacle to voxel value conversion to HU [5]. Most of the studies included in this review revealed a strong correlation and linear relationship between HU and GV across different densities and in both homogeneous and heterogeneous samples. However, Rostetter et al. [19] evaluating different anatomical structures of patients, reported a linear relationship, but noted that the strength of this relationship depends on the anatomical position in the head and its corresponding density. In addition, Nomura et al. [22] found that both linear and non-linear regression models demonstrated high r-value, but non-linear regression provided a better fit to the plots.

In the terms of comparing the mean values of HU and GV, most of the studies included in this review reported higher HU values than GV [7,19,22.24.26], while three studies [18,23,25,27] reported lower HU values in comparison to GVs. However, only two of these studies which belong to the first group (HU values > GV), had the same exposure factors in MDCT and CBCT scans. Rostetter et al. [19] used the same tube voltage for both CBCT and MDCT scans, while Chindasombatjaroen et al. [26] compared scan protocols with the same tube current

and tube voltage. It could be attributed to multiple factors, including elevated scatter radiation levels, noise, and artifacts of CBCT scan technology. When reconstructing images, scatter radiation can lead to an underestimation of absorption, which is dependent on the object and proportional to the amount of scatter present. To minimize radiation exposure, CBCT machines use lower tube currents than medical CT, leading to a lower signal-to-noise ratio and increased noise levels, resulting in more inconsistency and larger standard deviations in GVs. Furthermore, as CBCT acquires a larger volume compared to highly collimated fan beam medical CT, artifacts have a more pronounced effect on image quality [35].

This systematic review aimed to evaluate the correlation of GV and HU, and for this purpose, articles reporting the correlation coefficient (r-value) and coefficient of determination (R2-value) were included. However, it is essential to distinguish between them. The correlation coefficient (r-value) indicates the strength and the direction of a linear relationship between two variables ranging between -1 and 1. The coefficient of determination (R2-value) assesses how well a model predicts an outcome ranging between 0 and 1. In other words, less scattered data point higher R2 value [6] . Although the r or R2 value of the literature regarding the correlation of GV and HU is typically high (Tables 1 and 2), it should be kept in mind that this is not necessarily proof of the accuracy of GV and must be interpreted with caution. Despite the high correlation found in this meta-analysis, significant deviation (25.7-313.40 numerical value in the included studies, Table 2) in GV makes it unreliable for clinical use. Only at R2 values close to 1 (>0.999), these deviations are small enough to consider clinical use [6]. Eguren et al.'s systematic review [5] found a high correlation between actual and expected GV but with significant variability which was 500 GV numerical value with a deviation of 20%.

Bone density assessment is crucial in implant placement and the ultimate success of the procedure. A classification for jaw bone density first proposed by Misch is based on the HU and defined in five groups: D1: >1250, D2: 850–1250, D3: 350–850, D4: 350-0, and D5<0 HU. [36] Considering these ranges, in the case of a slight deviation of voxel value below the proposed ranges, it would be possible to use CBCT for implant planning. Pauwels et al. [21] have investigated 13 different CBCT machines and stated that some exposure protocols demonstrate reasonable accuracy for bone density estimation. However, the authors did not recommend differential diagnosis of lesions by quantitative assessment of CBCT voxel value. It is in line with Etoz et al. [37], that found no relationship between the voxel value of CBCT and pathological diagnosis of apical cyst and granuloma.

The main limitation of this study was the heterogeneity of the methods, samples, and data. Additionally, all the information was not mentioned in some studies, such as the interval between CBCT and MDCT, the conditions under which viewers assessed bone density, and whether these modifications could affect radiologic bone density. Furthermore, CBCT devices are not identical, and there are noticeable differences in exposure protocols, software, and hardware.

CONCLUSION

In conclusion, the strong correlation between voxel values and HU indicates the potential for CBCT in measuring radiographic bone density. However, further research is needed to obtain an accurate conversion equation for translating voxel values of CBCT to HU. Calibration of voxel values within each scan using a reference object and consideration of both linear and non-linear regression could improve the accuracy of voxel value measurement. Despite existing variations, some trends can be observed, although the translation of GV to HU remains debatable and requires more consistent scientific evidence.

ACKNOWLEDGEMENT

This study was supported by Tehran University of Medical Sciences, School of Dentistry.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Razi T, Niknami M, Alavi Ghazani F. Relationship between Hounsfield Unit in CT Scan and Gray Scale in CBCT. J Dent Res Dent Clin Dent Prospects. 2014 Spring;8(2):107-10.
2. Mallya S, Lam E. White and Pharoah's Oral radiology: principles and interpretation. St. Louis, United states, Elsevier, 2019: 659.
3. Shapurian T, Damoulis PD, Reiser GM, Griffin TJ, Rand WM. Quantitative evaluation of bone density using the Hounsfield index. Int J Oral Maxillofac Implants. 2006 Mar-Apr;21(2):290-7.
4. Shahlaie M, Gantes B, Schulz E, Riggs M, Crigger M. Bone density assessments of dental implant sites: 1. Quantitative computed tomography. Int J Oral Maxillofac Implants. 2003 Mar-Apr;18(2):224-31.
5. Eguren M, Holguin A, Diaz K, Vidalon J, Linan C, Pacheco-Pereira C, et al. Can gray values be converted to Hounsfield units? A systematic review. Dentomaxillofac Radiol. 2022 Jan 1;51(1):20210140.
6. Pauwels R, Jacobs R, Singer SR, Mupparapu M. CBCT-based bone quality assessment: are Hounsfield units applicable? Dentomaxillofac Radiol. 2015;44(1):20140238.
7. Parsa A, Ibrahim N, Hassan B, Motroni A, van der Stelt P, Wismeijer D. Reliability of voxel gray values in cone beam computed tomography for preoperative implant planning assessment. Int J Oral Maxillofac Implants. 2012 Nov-Dec;27(6):1438-42.
8. Valiyaparambil JV, Yamany I, Ortiz D, Shafer DM, Pendrys D, Freilich M, et al. Bone quality evaluation: comparison of cone beam computed tomography and subjective surgical assessment. Int J Oral Maxillofac Implants. 2012 Sep-Oct;27(5):1271-7.
9. Mah P, Reeves TE, McDavid WD. Deriving Hounsfield units using grey levels in cone beam computed tomography. Dentomaxillofac Radiol. 2010 Sep;39(6):323-35.
10. Aranyarachkul P, Caruso J, Gantes B, Schulz E, Riggs M, Dus I, et al. Bone density assessments of dental implant sites: 2. Quantitative cone-beam computerized tomography. Int J Oral Maxillofac Implants. 2005 May-Jun;20(3):416-24.
11. Parsa A, Ibrahim N, Hassan B, van der Stelt P, Wismeijer D. Bone quality evaluation at dental implant site using multislice CT, micro-CT, and cone beam CT. Clin Oral Implants Res. 2015;26(1):e1-7.
12. Dahiya K, Kumar N, Bajaj P, Sharma A, Sikka R, Dahiya S. Qualitative Assessment of Reliability of Cone-beam Computed Tomography in evaluating Bone Density at Posterior Mandibular Implant Site. J Contemp Dent Pract. 2018 Apr 1;19(4):426-430.
13. Scarfe WC, Angelopoulos C. Maxillofacial cone beam computed tomography: principles, techniques and clinical applications.Switzerland, Springer, 2018.
14. Silva IM, Freitas DQ, Ambrosano GM, Bóscolo FN, Almeida SM. Bone density: comparative evaluation of Hounsfield units in multislice and cone-beam computed tomography. Braz Oral Res. 2012 Nov-Dec;26(6):550-6.
15. González-García R, Monje F. The reliability of cone-beam computed tomography to assess bone density at dental implant recipient sites: a histomorphometric analysis by micro-CT. Clin Oral Implants Res. 2013 Aug;24(8):871-9.
16. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to meta-analysis. West Sussex, United Kingdom, John Wiley & Sons Ltd. Chichester, 2009: 421.
17. Pauwels R, Stamatakis H, Manousaridis G, Walker A, Michielsen K, Bosmans H, et al. Development and applicability of a quality control phantom for dental cone-beam CT. J Appl Clin Med Phys. 2011 Nov 15;12(4):3478.
18. Gaur A, Dhillon M, Puri N, Sethi Ahuja U, Rathore A. Questionable accuracy of CBCT in determining bone density: A comparative CBCT-CT in vitro study. Dent Med Probl. 2022 Jul-Sep;59(3):413-419.
19. Rostetter C, Metzler P, Schenkel JS, Seifert B, Luebbers HT. Comparison of in vivo cone-beam and multidetector computed tomographic scans by three-dimensional merging software. Br J Oral Maxillofac Surg. 2015 Dec;53(10):1021-6.
20. Razi T, Emamverdizadeh P, Nilavar N, Razi S. Comparison of the Hounsfield unit in CT scan with the gray level in cone-beam CT. J Dent Res Dent Clin Dent Prospects. 2019 Summer;13(3):177-182.
21. Pauwels R, Nackaerts O, Bellaiche N, Stamatakis H, Tsiklakis K, Walker A, et al. Variability of dental cone beam CT grey values for density estimations. Br J Radiol. 2013 Jan;86(1021):20120135.
22. Nomura Y, Watanabe H, Honda E, Kurabayashi T. Reliability of voxel values from cone-beam computed tomography for dental use in evaluating bone mineral density. Clin Oral Implants Res. 2010 May;21(5):558-62.
23. Khavid A, Sametzadeh M, Godiny M, Moarrefpour MM. Comparison of the hounsfield unit values obtained from cone-beam computed tomography (CBCT) and multidetector computed tomography (MDCT) images for different bone densities. J Contemp Med Sci. 2021 Mar;7(2):92-5.
24. Kamaruddin N, Rajion ZA, Yusof A, Aziz ME. Relationship between Hounsfield unit in CT scan and gray scale in CBCT. *AIP Conf Proc.* 19 December 2016; 1791 (1): 020005.
25. Haghanifar S, Yousefi S, Moudi E, Abesi F,

Bijani A, Moghadamnia AA, et al. Accuracy of densitometry of two cone beam computed tomography equipment in comparison with computed tomography. Electron Physician. 2017 May 25;9(5):4384-4390.
26. Chindasombatjaroen J, Kakimoto N, Shimamoto H, Murakami S, Furukawa S. Correlation between pixel values in a cone-beam computed tomographic scanner and the computed tomographic values in a multidetector row computed tomographic scanner. J Comput Assist Tomogr. 2011 Sep-Oct;35(5):662-5.
27. Cassetta M, Stefanelli LV, Pacifici A, Pacifici L, Barbato E. How accurate is CBCT in measuring bone density? A comparative CBCT-CT in vitro study. Clin Implant Dent Relat Res. 2014 Aug;16(4):471-8.
28. Parsa A, Ibrahim N, Hassan B, Motroni A, van der Stelt P, Wismeijer D. Influence of cone beam CT scanning parameters on grey value measurements at an implant site. Dentomaxillofac Radiol. 2013;42(3):79884780.
29. El-Tabarany KS, Aboulfotouh MM, Ashmawy MS. VALIDITY OF CONE BEAM COMPUTED TOMOGRAPHY VOXEL DENSITY VALUES: A REVIEW. Int J Adv. Res. 2018 Sep; 6(9):103-121.
30. Vasconcelos TV, Leandro Nascimento EH, Bechara BB, Freitas DQ, Noujeim M. Influence of Cone Beam Computed Tomography Settings on Implant Artifact Production: Zirconia and Titanium. Int J Oral Maxillofac Implants. 2019 Sep/Oct;34(5):1114-1120.
31. Molteni R. Prospects and challenges of rendering tissue density in Hounsfield units for cone beam computed tomography. Oral Surg Oral Med Oral Pathol Oral Radiol. 2013 Jul;116(1):105-19.
32. Reeves TE, Mah P, McDavid WD. Deriving Hounsfield units using grey levels in cone beam CT: a clinical application. Dentomaxillofac Radiol. 2012 Sep;41(6):500-8.
33. Komori M, Miuchi S, Hyodo J, Kobayashi T, Hyodo M. The gray scale value of ear tissues undergoing volume-rendering high-resolution cone-beam computed tomography. Auris Nasus Larynx. 2018 Oct;45(5):971-979.
34. Selvaraj A, Jain RK, Nagi R, Balasubramaniam A. Correlation between gray values of cone-beam computed tomograms and Hounsfield units of computed tomograms: A systematic review and meta-analysis. Imaging Sci Dent. 2022 Jun;52(2):133-140.
35. Schulze R, Heil U, Gross D, Bruellmann DD, Dranischnikow E, Schwanecke U, et al. Artefacts in CBCT: a review. Dentomaxillofac Radiol. 2011 Jul;40(5):265-73.
36. Resnik R. Misch's contemporary implant dentistry. 4th ed. St. Louis, United states: Elsevier; 2020 2021. 1241 p.
37. Etöz M, Amuk M, Avcı F, Yabacı A. Investigation of the effectiveness of CBCT and gray scale values in the differential diagnosis of apical cysts and granulomas. Oral Radiol. 2021 Jan;37(1):109-117.