UNIVERSIDAD PERUANA CAYETANO HEREDIA

Facultad de ESTOMATOLOGÍA

¿LOS VALORES DE GRIS PUEDEN SER CONVERTIDOS EN UNIDADES HOUNSFIELD? UNA REVISIÓN SISTEMÁTICA

CAN GRAY VALUES BE CONVERTED TO HOUNSFIELD UNITS? A SYSTEMATIC REVIEW

TESIS PARA OPTAR POR EL TÍTULO DE ESPECIALISTA EN ORTODONCIA Y ORTOPEDIA MAXILAR

AUTOR
ANDERSON OSWALDO HOLGUIN RICCER

ASESOR

JOSE ANTONIO VIDALON CASTILLA

CO- ASESOR

CARLOS YURI LIÑAN DURAN

LIMA - PERÚ

2022

JURADO

Presidente: MG. ESP. ORLANDO TUESTA DA CRUZ

Vocal: MG. ESP. JORGE CARLOS MELGAR GUTIERREZ

Secretario: Mg. Esp. DIANA CECILIA BECERRA NUÑEZ

Fecha de Sustentación: 03 de Agosto 2022

Calificación: Aprobado – Nota: 88

ASESORES DE TESIS

ASESOR

Mg. Esp. CD. Jose Antonio Vidalon Castilla

Departamento Académico de Ortodoncia y Ortopedia Maxilar

ORCID: 0000-0002-8168-0006

CO-ASESOR

Carlos Yuri Liñan Duran

Departamento Académico de Ortodoncia y Ortopedia Maxilar

ORCID: 0000-0003-2669-842X

DEDICATORIA

Esta investigación va dedicada a Dios, por ser mi guía y mi inspiración en la vida; a mis padres y mi hermana; por su apoyo incondicional en este hermoso camino a convertirme en ortodoncista. Quisiera dejar una dedicación especial al grupo de investigación de ortodoncia de la Universidad Peruana Cayetano Heredia, que se conformó por residentes y docentes de la Facultad de Estomatología Roberto Beltrán en el año 2020, grupo con quienes realizamos esta investigación y otras más con grandes frutos en los tiempos de la pandemia por covid-19

AGRADECIMIENTOS

Agradezco a Dios y a la vida, por haberme dado la oportunidad de conocer este mundo tan maravilloso como es la ortodoncia. Así como a mis padres por los valores y buenas costumbres impartidas. Agradezco a toda la gente que fue parte de este proyecto, ya que sin ellos no hubiese sido posible haberlo realizado.

FUENTES DE FINANCIAMIENTO

No aplica

DECLARACIÓN DE CONFLICTO DE INTERÉS

Los autores declaran no tener conflictos de interés.

RESULTADO DEL INFORME DE SIMILITUD

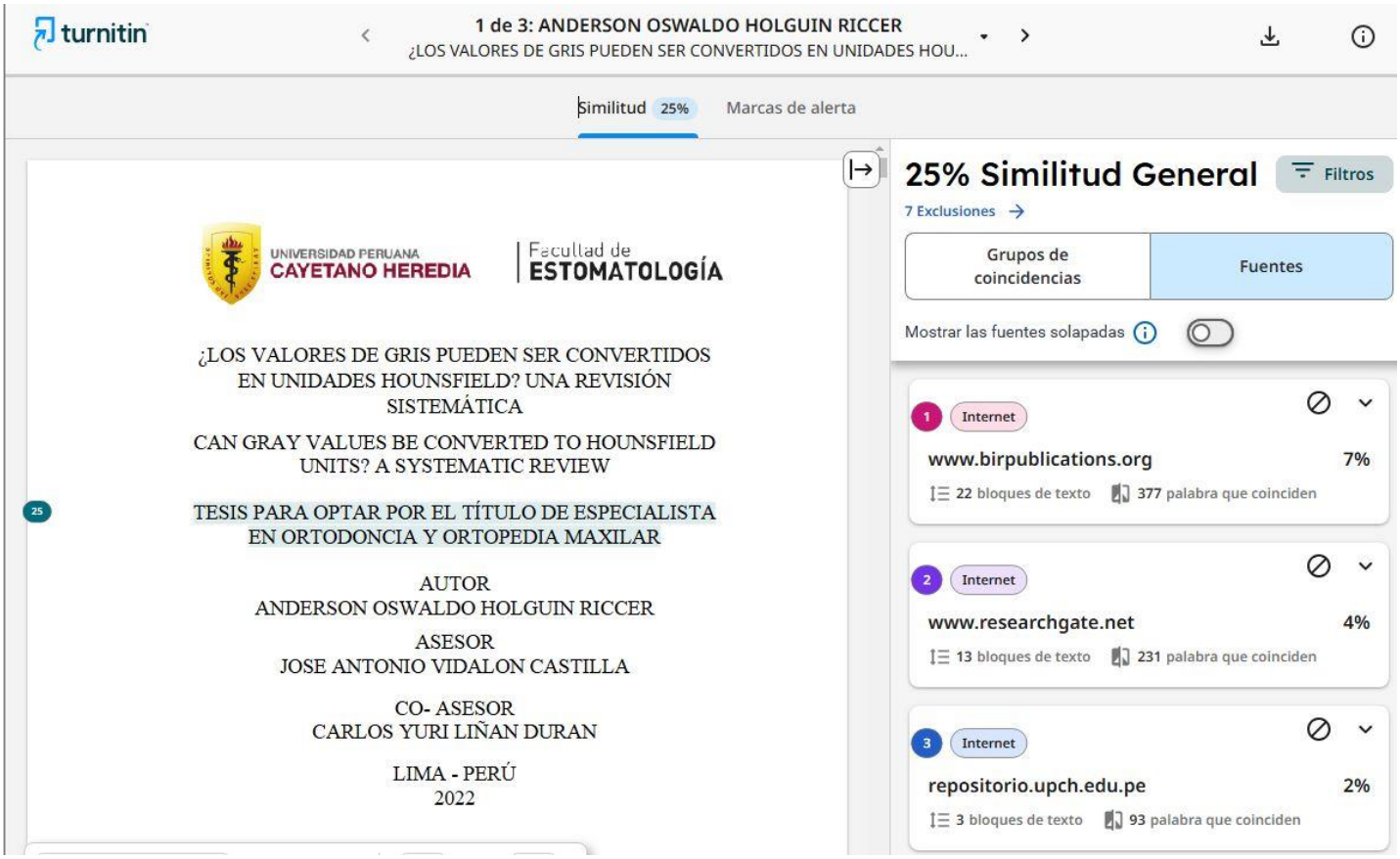

TABLA DE CONTENIDOS

RESUMEN

Objetivos: El propósito de esta revisión sistemática fue responder a la pregunta de focalizada: "¿Podrían los valores de gris (GV) de CBCT (tomografía computarizada de haz cónico) convertirse en unidades Hounsfield (HU) de tomografía computarizada multidetector (TCMD)?" **Métodos:** Los estudios incluidos intentan responder la pregunta de investigación de acuerdo con la estrategia PICO. Los estudios se recopilaron mediante la búsqueda de diferentes bases de datos electrónicas y literatura gris parcial hasta enero de 2021 sin restricciones de idioma o tiempo. La evaluación metodológica de los estudios se realizó utilizando la herramienta de evaluación de la salud bucal (OHAT) para los estudios in vitro y la evaluación de la calidad de los estudios de precisión diagnóstica (QUADAS-2) para los estudios in vivo. Se aplicó el instrumento Grading of Recommendations Assessment, Development and Evaluation (sistema GRADE) para evaluar el nivel de evidencia en los estudios. **Resultados:** Se obtuvieron 2710 artículos en la Fase 1 y quedaron 623 citas después de eliminar los duplicados. En esta revisión sólo se incluyeron tres estudios mediante un proceso de selección de dos fases y después de aplicar los criterios de elegibilidad. Todos los estudios fueron metodológicamente aceptables, aunque en términos generales con bajo riesgo de sesgo. Algunos estudios incluidos tuvieron estimaciones de evidencia y fuerzas de recomendación bastante bajas y limitadas; evidenciando la necesidad de estudios clínicos con capacidad diagnóstica que sustenten su uso. **Conclusiones:** Esta revisión sistemática demostró que las VG de CBCT no se pueden convertir en HU debido a la falta de estudios clínicos con capacidad diagnóstica que sustenten su uso. Sin embargo, se evidencia que se necesitan tres pasos de conversión

(calibración del equipo, modelos de ecuaciones de predicción y una fórmula estándar (conversión de GV a HU)) para obtener pseudo valores de Hounsfield en lugar de solo obtenerlos de una regresión o directamente del software.

Palabras clave Densidad ósea; Tomografía computarizada de haz cónico; Unidad de Hounsfield; Tomografía computarizada multidetector; diagnóstico por imagen; revisión sistemática.

ABSTRACT

Objectives: The purpose of this systematic review was to answer the focus question: “Could the gray values (GVs) from CBCT (cone beam computed tomography) be converted to Hounsfield units (HUs) in multidetector computed tomography (MDCT)?” **Methods:** The included studies try to answer the research question according to the PICO strategy. Studies were gathered by searching several electronic databases and partial grey literature up to January 2021 without language or time restrictions. The methodological assessment of the studies was performed using The Oral Health Assessment Tool (OHAT) for in vitro studies and the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) for in vivo studies. The Grading of Recommendations Assessment, Development and Evaluation (GRADE system) instrument was applied to assess the level of evidence across the studies. **Results:** 2710 articles were obtained in Phase 1, and 623 citations remained after removing duplicates. Only three studies were included in this review using a two-phase selection process and after applying the eligibility criteria. All studies were methodologically acceptable, although in general terms with low risks of bias. There are some included studies with quite low and limited evidence estimations and recommendation forces; evidencing the need for clinical studies with diagnostic capacity to support its use. **Conclusions**: This systematic review demonstrated that the GVs from CBCT cannot be converted to HUs due to the lack of clinical studies with diagnostic capacity to support its use. However, it is evidenced that three conversion steps (equipment calibration, prediction equation models, and a standard formula (converting GVs to HUs)) are needed to obtain

pseudo Hounsfield values instead of only obtaining them from a regression or directly from the software.

Keywords: Bone density; Cone-beam computed tomography; Hounsfield unit; Multidetector Computed Tomography; diagnostic imaging; systematic review.

Introduction

In dentistry, there are several reasons to justify the demand for valid tools to assess bone mineral density (BMD). It is important for preoperative planning of implants and temporary anchorage device (TADs) placements,[1-3] to assess the height, width, and the distance to other anatomical structures such as the sinus region or the mandibular canal. Quality and quantity of the receptor bone is required due to its influence on primary stability and success of orthodontic temporary anchorage devices.[3,4] Furthermore, identifying the bone degenerative processes of the jaws, temporomandibular joints (TMJ),[5] and systemic conditions help the clinicians in the diagnosis and treatment planning of patients.

The gold standard for evaluating radiological BMD is the Multislice Computed Tomography (MSCT) using Hounsfield Units (HU), which is the coefficient of ray attenuation for bones (absolute density).[6,7] Nowadays, cone-beam computed tomography (CBCT) is increasingly compared to MSCT in dentistry to assess mineralized tissues because it provides adequate image quality. Besides, CBCT is associated with lower radiation exposure doses, lower costs, faster scanning times, good spatial resolutions, gray density ranges and contrasts, as well as a good pixel/noise ratio compared to MSCT.[8] However, due to the associated artifacts, the lack of standardization of CBCT scanners and the acquisition parameters, it is uncertain that gray values (GV) could be correlated to HU.[9]

The HU are defined as linear transformations of measured X-ray attenuation coefficients of a material with reference to water. Some studies have mentioned the possibility to obtain HU through CBCT because the relationship between GV and linear attenuation coefficient persists in CBCT and HU can be calculated using a conversion equation, for example a regression equation and standard conversion formula.[10-17] This way, the standard conversion formula used to calculate HU for any material was (HU material = $\frac{\mu^{material} - \mu water}{\mu water} \times 1000$).

Nevertheless, other studies do not support the ability of converting GV to HU due to the CBCT acquisition mode, stating that the BMD obtained through CBCT does not correlate to MSCT.[9,18-21] There are crucial differences between MSCT and CBCT, which complicates the use of quantitative GV that are inherently associated with this technique (i.e. the limited field size, relatively high amount of scattered radiation and limitations of currently applied reconstruction algorithms).[22] There is contradictory information about whether CBCT can be used for BMD similar to HU in MSCT.[9] Thus, obtaining HU from gray levels or voxel values is controversial using CBCT and is insufficiently studied.[9,21]

Therefore, the objective of this systematic review was to answer a focused question: "Could the GV from CBCT be converted to the HU in MSCT?"

Material and Methods

Study design

This study involved *in vitro* (phantoms or dry bones) and *in vivo* (patients) studies that evaluate GV from CBCT and HU from MSCT according to different conversion formulas.

The included studies should help answer the research question according to the PIRD strategy as follows: **P**opulation: CBCT and MSCT images; **I**ndex text: GV from CBCT; **R**eference text: HU from MSCT; **D**iagnosis of Interest: equation formulas which compare GV with HU.

Eligibility criteria

Studies in which the primary objective was to assess reliability of CBCT voxel GV measurements using HU derived from MSCT were included. It was essential that the eligible studies demonstrated the statistical correlation and linear regression for the conversion formula. Also, the MSCT should be taken at the same region of interest (ROI) as the CBCT scans. No language or time restrictions were applied.

The following studies were excluded: Studies that used other imaging technique devices (micro-CT, Dual-energy X-ray absorptiometry, ultrasonography and magnetic resonance); reviews, personal opinions, conference abstracts and letters; experiments in animal models; and studies without reference to a gold standard. An additional eligibility criterion was added in phase 2 in which the studies that did not present a conversion formula were not eligible.

Search strategy

Electronic search strategies were applied in PubMed, Cochrane, EMBASE, LILACS and Ovid MEDLINE databases up to May 2020. No date or language restrictions were applied. (Appendix A). Also, journals related to the topic, grey literature (clinical trials.gov and google scholar) and the references of included studies were screened to identify any missed publications. Search results were collected, and duplicate references were removed. The

search was updated in all databases until 10 January 2021, no additional studies were finding for inclusion in this review.

Study selection

A two-phase selection process was developed according to inclusion criteria. In phase 1, two authors (AHR and MEL) selected articles by title and abstract independently. In phase 2, the same authors reviewed the full text of all potential articles to apply the eligibility criteria. Disagreements were resolved by consensus with a third author (JVC). The final selection was always based on the full text of the publication and discussion between evaluators.

Data collection process

One author (AHR) extracted the required data from the included articles and a second author (MEL) reviewed all the retrieved information. The key features were crosschecked, with a one-week interval, by the same authors. Again, disagreements among them were solved by consensus with the third author (JVC). Using a standardized template, the descriptive characteristics of included studies and their key features such as author, year and country; sample characteristics (scanned structure, index text-CBCT, reference standard-MSCT); ROI; the intervention characteristics (method of conversion of GV to HU) and the specific results related to our research question were recorded.

Risk of bias (RoB) in individual studies

The RoB of selected studies was evaluated using The Oral Assessment tool (OHAT) for *in vitro* studies.[23] The Quality Assessment tool for Diagnostic Accuracy (QUADAS-2)

evaluated the *in vivo* studies.[24] Previously, an author with experience in RoB (KD) tools trained the evaluators (AHR, MEL) in two rounds. In both rounds, randomized articles were independently analyzed and a favorable level of agreement value of 0.9 was reached. The authors (AHL and MEL) applied the tools for the rest of the studies independently and, if any disagreement was found, items were discussed between co-authors (P-PC and KD).

The OHAT evaluates randomization, allocation concealment, experimental condition, blinding, incomplete data, exposure characterization, outcome assessment, reporting and other biases related to the methodological structure. OHAT scores were definitely low RoB (++), probably low RoB (+), probably high RoB (-/NR), definitely high RoB (--). QUADAS-2 assesses patient selection, index test, reference standard, flow and timing, as well as its concerns regarding applicability. Guiding questions within every domain were scored as yes, no or unclear and topic conclusions and applicability were registered as low, high or unclear RoB.

Summary measures

The capacity of CBCT scans to identify BMD according to conversion formulas was considered as the primary outcome. Any type of outcome measurement was considered in this review (categorical and continuous variables). No meta-analysis was performed due to heterogeneity of the data.

Level of evidence

The grading of recommendation, assessment, development, and evaluation (GRADE system) instrument evaluates the quality of evidence.[25] Two authors (KD, CP) rated the quality of the evidence as well as the strength of the recommendation according the

following aspects: study design, RoB, consistency, directness, precision, publication bias and other aspects reported by studies included in this systematic review.[26] The quality of the evidence was characterized as high, moderate, low, or very low. The GRADE was assessed using tools from their website http://gradepro.org.

Results

Study selection

Figure 1 shows a flow diagram detailing the process of identification, inclusion and exclusion of the studies. In phase 1, a total of 2711 articles were obtained by title and abstract. After removing duplicates, 623 different citations remained. Following a detailed evaluation of abstracts, 483 articles were selected for phase 2. A manual search was performed on Google Scholar and Gray literature. One study was identified from the reference lists and was included in this review. Finally, after full-text reading, 34 studies were excluded due to multiple reasons specified in Appendix B. Therefore, three studies were finally included, as per the flow diagram.

Study characteristics

The three included studies evaluated if the GV can be compared to HU according to statistical correlation, linear regression and conversion formulas.[14,17,27] In total, this systematic review assessed 77 scans, *in vitro* studies were done using phantoms[14,27] and *in vivo* using patient scans.[17] *In vitro* studies described different types of phantoms, ROI and method of conversion and *in vivo* studies included 61 patients with CBCT scans and attenuation coefficients from the National Institute of Standards and Technology (NIST) (Table I).[14,17,27] In an attempt to gather missing information and to retrieve the formula

methodologies, we tried to contact the corresponding authors of the included articles but were not successful.

Results of individual studies and Synthesis of results

In-vitro studies

Mah *et al*[14] converted CBCT grey values into HU using a calculated linear attenuation coefficient. They were derived from a linear regression model to obtain the best linear fit "The effective energy 63 keV" a result of plotting attenuation coefficients of 8 materials provided by NIST and 2 MSCT scanners against the GV of these materials obtained from 11 CBCT scanners. The linear relationship between GV and CT numbers of each of the materials exists (R^2=0.9999). This linear regression model was used to obtain attenuation coefficients and calculate HU. The calculated attenuation coefficients were transformed into HU according to a or the? standard formula.

Magill *et al* [27] used half value layer (HVL) for each X-ray tube to determine the average of beam energy to each scanner (46.7 to 51.7 keV) and obtain the best R^2value. A linear regression was fit, and the resulting values of linear attenuation coefficients were converted to CT numbers using the standard conversion formula. This study shows the consistency of the water value (0 to 1 HU) for all three CBCT scanners.

In-vivo study

Reeves *et al*[17] used the same conversion methodology (GV to HU) for an *in vitro* study.[14] They found the difference between the calculated and actual HU to be less than 3%, whereas the relationship between GV and HU was defined as linear. The Asahi Alphard

3030 (Belmont Takara, Kyoto, Japan) CBCT scan showed calculated HU for outer bone equivalent (OBE) material at 1465.9 HU and inner bone equivalent (IBE) material at 254.7 HU using 70 keV; while the Planmeca ProMaxTM 3D (Planmeca, Helsinki, Finland) CBCT scan showed calculated HU for OBE material at 124.2 HU and IBE material at 251.9 HU at the same keV. To get those results, linear regression studies were developed for different keV (Table II). At 70 keV, the linear regression equation was $y = 1.7198\,x + 317.45$, showing a strong correlation. Then the standard conversion formula was applied.

As a summary of results, Table II shows the specific formula applied from each selected study.

RoB assessment

For the OHAT, the *in vitro* studies[14,27] showed direct evidence of low RoB for all evaluated scopes. The selection of study participants and confounding-modifying variables were not applicable for the *in vitro* studies (Table III OHAT). Quadas-2 assessed the *in vivo* studies[17] Regarding participant selection, *in vivo* study did not clearly report whether randomization was consecutive, or case-control design was not used, or whether appropriate exclusions were done. The index and the standard test results showed a low risk for introducing bias and applicability in this domain. Time intervals for both tests were adequate, as well as processing of the results. Study flow charts were reproduced for better understanding of the process and all studies have a low RoB for index tests, reference standards, flow-timing domains, and applicability concerns (Figure 2)

Grading the "body of evidence"

Regarding the GRADE tool, the sensitivity and specificity columns were eliminated since these data were not provided (Table IV). GRADE evaluation was assessed for all studies according to guidelines for test accuracy.[28,29]However, the studies showed a low RoB and direct evidence from the data. Heterogeneity and imprecise results were obtained since the CBCT and MSCT equipment had independent configurations that difficult them from being compared. Added to the fact that only one study was done with humans and the others were done with individualized phantoms, influenced an overall very low estimation of the evidence with a limited recommendation force.

Summary of evidence

Several studies proposed a variety of methods for assessing the BMD by Dual-energy X-ray absorptiometry[30], quantitative CT[31], micro-CT [31] and MSCT.[32] These were used to evaluate the radiological density of tissues in which HU is used as the gold standard. Some studies have mentioned the possibility of obtaining HU through CBCT, due to a possible linear relationship between GV and the attenuation coefficients of CBCT scanners. This leads to the possibility of converting GV of CBCT to CT numbers and then in HU with a standard conversion formula.[14-17] However, other studies concluded that the HU derived from CBCT is not identical.[9,18-22] Thus, the use of CBCT to obtain HU is still controversial. The increasing application of CBCT for BMD assessment impelled researchers to evaluate the new system in relation to obtaining HU by CBCT. As the literature is scarce and controversial on this subject to date, there is one systematic review concerning the capability of CBCT to identify low-BMD patients, suggesting the potential of CBCT for this purpose.[33] However, there is a lack of consensus among the investigations about the capability of CBCT to identify BMD, and a more detailed investigation is required. The

objective of this systematic review was to answer a focused question: "Could the GV from CBCT be converted to HUs in MSCT?" and to guide future research in this area.

Variability of CBCT scanners

One of the disadvantages of most CBCT machines is a lack of machine standardization. Like snowflakes and fingerprints, no two CBCT models are the same, demonstrating essential differences in terms of exposure, hardware and reconstruction.[20] The Mah *et al*[14] *in vitro* study used 12 different CBCT and 2 MSCT scanners. The Magill *et al*[27] *in vitro* study used 3 different CBCT scanners and CT numbers from American College of Radiology and Reeves et al[17] *in vivo* study used 2 CBCT scanners and CT numbers from NIST. Lagravère *et al*,[34] found that the HU values obtained from CBCTs differ between scanner models, settings and types of software used. Therefore, the GV corrections are typically only applicable for each device model.

Tube voltage

Tube voltage, mA and time of exposure are indirectly proportional to the amount of X-Ray photons and energy distribution. Nemtoi *et al*,[35] found the optimal tube voltage for CBCT imaging of the hard tissues is between 70 and 120 kV. Different studies showed that the x-tube voltage affects the GV from CBCT, thus, various kVp and mA values were used in the selected studies.[14,27] They found that the highest available voltage (i.e., 110 kV) resulted in the highest image quality in terms of general impression, sharpness, noise, and artefacts. On the other hand, it is well known that x-ray tubes and motion detectors can cause blurring so using pulsed x-ray beams are used to reduce this effect. Reduction of the mA to a

minimally acceptable value followed by reduction of the exposure time affects the reconstructed voxel size for several CBCT models in a pre-set manner. Fast-scan protocols showed equal or slightly better image quality compared to the standard-scan mode. It is important to emphasize that the ALARA (as low as reasonably achievable) principle should always be applied; thus, the protocol must be tailored to each case. The number of studies assessing the impact of voxel size variation on the diagnostic outcome in CBCT imaging in dentistry is small.[36]

Size of the Field of View (FOV)

Rodrigues *et al*,[37] found that the GV determined in CBCT images are significantly influenced by the FOV size. The determination of a small size FOV (< 50 mm) or a large FOV (>16 cm) depends on the structure to be evaluated. The GV in CBCT and MSCT with the same FOV size were significantly different for all materials and CBCT showed lower values in most of the comparisons. Also, CBCT had a significant correlation for plaster in all of the FOV sizes and for oil substances in the 16x8mm and 16x13mm FOV, except for water that did not show significant correlations. Katsumata *et al*,[11] observed the highest density variability in the smallest FOV scans whereas large FOV scans had more consistent density values. This was attributed to using an FOV greater than the patient diameter to prevent the truncation of data in the axial slice. The phantom size is known to produce cupping artifacts with a drop in CT number, due to the inability of the reconstruction algorithm to properly account for objects outside the FOV.[38] Mah *et al*,[14] used a small or large water container to place the phantom in to provide some level of soft tissue. The acrylic/water volume around the object influenced the GV in that the

greater the mass, the smaller the GV. However, the elimination of exomass (the structures located outside the FOV) by using a greater FOV resulted in a small variation of GV when the total mass was altered. Nackaerts *et al,*[21] found that just as the size of FOV is important, the placement of the ROI in the center of the FOV is also important because it minimizes the variability of results in the measurements.

Previous calibration of CBCT machine and effective energy

To calculate HU, it is necessary to calibrate the scanner against the X-ray absorption of air and water. Also, to determine adequate kilovoltage energy it is important to calibrate the CBCT. For this reason, linear regressions were performed in all studies to find the best linear fit as ''effective energy'' or "average beam". Before the *in vivo* study[17], an *in vitro* study[14] was done using a phantom and several other materials with well know attenuation values to find the best effective energy. The two previous studies used the same technique to find the keV. Magill *et al*[27] used HVL for each x-ray tube to determine the average beam energy instead of the best R^2 value. The HVL reduces the intensity of an x-ray beam entering that material by one-half. The accuracy of the water value is very important for obtaining HU values. When looking precisely at the Mah *et al*[14] and Reeves *et al*[17] tables, we note that the accuracy of water values was inconsistent and widely variable. The Magill *et al*[27]study shows more consistent water values because the conversion method was able to produce the water value of 0 HU within 1 SD for three CBCT scanners, and the linear attenuation coefficients from the determined average effective energy were specific to each CBCT x-ray tube.

Conversion and regression alone are not enough

This systematic review excluded 30 articles which compared GV with HU. The reason was that only three studies (two *in vitro* studies[14,27] and one *in vivo* study[17]) showed correlation, linear regression and conversion formulas. Most articles were excluded because they only reported correlation/regression without conversion formulas. Pauwels *et al,*[20,22] found that correlation and regression are not enough to convert GV to HU. Also, it is important to understand that correlation coefficients (R) and coefficients of determination (R^2) are different. The literature shows high values of R and R^2 above 0.95, which could make us think CBCT GV have the potential to be used as HU. However, simulated scatter plots indicate that although high R^2 values exist, a deviation of 20% and a numerical value deviation of 500 GV appeared. This demonstrates the large variability between actual and expected GV even for high R-values.

Mah *et al,*[14] and Reeves *et al,*[17] demonstrated linear relationships between GV and attenuation coefficients at some ‘‘effective’’ energy. Attenuation coefficients were obtained from the linear regression equation and CT numbers in HU were derived using the standard equation. Magill *et al*[27] used HVL to determine the average beam energy. At the extrapolated average energy, linear attenuation coefficients were derived for converting mean pixel values in GV to linear attenuation coefficient values. The resulting linear attenuation coefficient values were converted to CT numbers (HU) using the standard formula. General results have demonstrated that the GV taken from CBCT can be used to derive HU in a clinical environment. However, there is no consistent scientific evidence to support the routine use of CBCTs in BMD evaluations and more studies are needed to demonstrate the diagnosis capacity.

Discussion

This systematic review indicates in general terms a low RoB, as there are some limitations in the included studies. Significant methodological differences were identified such as different CBCT and MSCT scanners, *in vitro* vs *in vivo* studies and conversion methods). The principal limitation was the absence of the sensitivity and specificity, values data classically reported in diagnostic capability studies. However, for the purposes of this review, RoB and GRADE tools were applied because they were considered to be the most suitable for their evaluation. The second limitation was that the *in vitro* studies were done using phantoms, and they did not have clinical parameters such as soft tissue. On the other hand, the reliability of obtaining GV between the same subjects on the same day in *in vivo* studies has not been demonstrated due to obvious ethical implications involving radiation safety concerns.[39]

Another limitation is the lack of CBCT standardization which makes it difficult to compare the HU relative values obtained in different CBCT scanners making it impossible to extrapolate the data. The results obtained from these conversion methods are unique to each CBCT machine. Besides, all included studies used the mass attenuation coefficients from Form X on the NIST website, and the website describes an error of interpolation that conditions the mass attenuation values calculating the attenuation coefficients that are subject to a degree of uncertainty because of this factor.[14] These methods of conversion require training and more knowledge on radiological and statistical concepts to understand the process of CBCT machine calibrations and conversions which are limiting factors when considering the clinicians routine.

Clinical application

One of the biggest benefits to obtain BMD in dentistry is the implant selection.[40] An accurate assessment of bone quality is essential in implant surgery since implant failure

rates tend to be higher when placed in poor quality bone.[41] This review found some limitations on the conversion of GV to HU; nevertheless, the calculated HU in CBCT data could be useful to approximate the understanding about BMD which can be an outlet to use the GV taken directly from CBCT. Another clinical application is related to miniscrew placement. These relatively new devices and insertion techniques in most cases just take into consideration the anatomical characteristics and subjective BMD but do not take BMD as an objective factor for installation planning. A calculated HU can aid in the determination of the volume and density of cortical bone available and necessary for the primary stability for miniscrews.[42] Nowadays, the assessment of airways to treat obstructive sleep apnea syndrome is where orthodontic research is heading.[43-46] A calculated HU could help to determine airway boundaries for a more accurate airway volume and how they change with treatments such as RME, MARPE or SARME.

Usually, maxillofacial surgeons use GV for differential diagnosis despite HU being the tool chosen for this purpose. Studies have found that GV for differentiating lesions should be taken with care because they could show scattered values.[47] Also, calculated HU could help patients with osteonecrosis of the jaws. The inability to identify borders of the osteonecrosis lesions makes surgeons perform block resection to ensure the complete removal of necrotic tissue.[48] The capacity to identify and demarcate borders of osteonecrosis lesions by analysis of BMD could change the surgical procedure.

These are some of the reasons that justify the value of obtaining HU from CBCT. Although, converted GV cannot be taken as definitive HU values, Pauwels *et. al*.[22] proposes it may be more appropriate to call these values pseudo-Hounsfield. Much work remains to be done, and future research should be directed to conduct clinical studies that allow for better use of the gray scale and be able to provide more solid conclusions.

Conclusion

This systematic review has demonstrated that the GV from CBCT cannot be converted to HU due to the lack of clinical studies of diagnostic capacity to support its use. However, it is evidenced that three conversion steps (equipment calibration, regression formulas and standard formula) are needed to obtain pseudo-Hounsfield values instead of only getting them from a regression or directly from the software.

Conflict of Interests

The authors declare no potential conflict of interest with respect to the authorship and/or publication of this article. There was no funding for this research project.

Author ContributionsMEL, contributed to conception design, data acquisition, analysis, interpretation, drafted and critically analyzed the manuscript. AHR also contributed to conception design, data acquisition, analysis, interpretation, drafted and critically analyzed the manuscript. VJ and LDC helped with the research and corrected the writing. DK corrected the writing and made methodological quality assessment of the studies drafted and critically analyzed the manuscript. PPC contributed to analysis, interpretation and critically analyzed the manuscript. LM, contributed to analysis, interpretation and critically analyzed the manuscript. All authors read and approved the final manuscript.

References

1. 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; **21**: 290–7.
2. 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; **20**: 416-24.
3. Nucera R, Lo Giudice A, Bellocchio AM, Spinuzza P, Caprioglio A, Perillo L, et al. Bone and cortical bone thickness of mandibular buccal shelf for mini-screw insertion in adults. Angle Orthod 2017; **87**: 745-51.
4. Turkyilmaz I, Tözüm TF, Tumer C. Bone density assessments of oral implant sites using computerized tomography. J Oral Rehabil 2007; **34**: 267-72.
5. dos Anjos Pontual ML, Freire JS, Barbosa JM, Frazão MA, dos Anjos Pontual A. Evaluation of bone changes in the temporomandibular joint using cone beam CT. Dentomaxillofac Radiol 2012; **41**: 24-9.
6. Shahlaie, M., Gantes, B., Schulz, E., Riggs, M. & Crigger, M. Bone density assessments of dental implant sites: 1. Quantitative computed tomography. The International Journal of Oral & Maxillofacial Implants. 2003; 18: 224–231.
7. 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; **21**: 290–7.
8. Arisan V, Karabuda ZC, Avsever H, Özdemir T. Conventional multi-slice computed tomography (CT) and cone-beam CT (CBCT) for computer- assisted implant placement. Part I: relationship of radiographic gray density and implant stability. Clin Implant Dent Relat Res 2013; **15**: 893-906.

9. Silva IM, Freitas DQ, Ambrosano GM, Bóscolo FN, Almeida SM. Bone density: comparative evaluation of Hounsfield units inmultislice and cone-beam computed tomography. Braz Oral Res 2012; **26**: 550-6.
10. Lee S, Gantes B, Riggs M, Crigger M. Bone density assessments of dental implants sites: 3. Bone quality evaluation during osteotomy and implant placement. Int J Maxillofac Implants 2007; **22**: 208-12.
11. Katsumata A, Hirukawa A, Okumura S, Naitoh M, Fujishita M, Ariji E, et al. Relationship between density variability and imaging volume size in cone-beam computerized tomographic scanning of the maxillofacial region: an *in vitro* study. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009; **107**: 420-5.
12. Naitoh M, Aimiya H, Hirukawa A, Ariji E. Morphometric analysis of mandibular trabecular bone using cone beam computed tomography: an in vitro study. Int J Oral Maxillofac Implants 2010; **25**: 1093-8.
13. 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; **24**: 871-9.
14. Mah P**,** Reeves TE, McDavid WD. Deriving Hounsfield units using grey levels in cone beam computed tomography. Dentomaxillofac Radiol 2010; **39**: 323–35.
15. Lagravère MO, Fang Y, Carey J, Toogood RW, Packota GV, Major PW. Density conversion factor determined using a cone-beam computed tomography unit NewTom QR-DVT 9000. Dentomaxillofac Radiol 2006; **35** :407–9.
16. 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; **27**: 1438-42.
17. Reeves TE, Mah P, McDavid WD. Deriving Hounsfield units using grey levels in cone beam CT: a clinical application. Dentomaxillofac Radiol 2012; **41**: 500-8.
18. Cassetta M, Stefanelli LV, Di Carlo S, Pompa G, Barbato E. The accuracy of CBCT in

measuring jaws bone density. Eur Rev Med Pharmacol Sci 2012; **16**: 1425-9.

19. 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; **35**: 662-5.
20. Pauwels R, Jacobs R, Singer SR, Mupparapu M. CBCT- based bone quality assesment: are Hounsfield units applicable? Dentomaxillofac Radiol 2015; **44**: 20140238.
21. Nackaerts O, Maes F, Yan H, Couto Souza P, Pauwels R, Jacobs R. Analysis of intensity variability in multislice and cone beam computed tomography. Clin Oral Implants Res 2011; **22**: 873-9.
22. 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; **86**: 20120135.
23. NTP (National Toxicology Program). 2013b. OHAT Implementation of Systematic Review. Available: http://ntp.niehs.nih.gov/go/38673 [accessed 10 June 2020]
24. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al.QUADAS-2 Group QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; **155**: 529–36. doi: https://doi.org/10.7326/0003- 4819-155-8-201110180-00009
25. Schünemann H, Brozek J, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendation. Version 3.2 [updated March 2009]. The GRADE Working Group; 2009. Available from: https://gdt.gradepro.org/app/handbook/handbook.html [Accessed 6 September 2020].
26. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011; **64**: 401-6.
27. Magill D, Beckmann N, Felice MA, Yoo T, Luo M, Mupparapu M. Investigation of dental cone- beam CT pixel data and a modified method for conversion to Hounsfield

unit (HU). Dentomaxillofac Radiol 2018; **47**: 20170321.

28. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 1. Study design, risk of bias, and indirectness in rating the certainty across a body of evidence for test accuracy. J Clin Epidemiol 2020; **122**: 129-41.
29. Schünemann HJ, Mustafa RA, Brozek J, Steingart KR, Leeflang M, Murad MH, et al. GRADE guidelines: 21 part 2. Test accuracy: inconsistency, imprecision, publication bias, and other domains for rating the certainty of evidence and presenting it in evidence profiles and summary of findings tables. J Clin Epidemiol 2020; **122**: 142-52.
30. Gulsahi A, Paksoy CS, Yazicioglu N, Arpak N, Kucuk NO, Terzioglu H. Assessment of bone density differences between conventional and bone-condensing techniques using dual energy x-ray absorptiometry and radiography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2007; **104**: 692-8.
31. Burghardt AJ, Link TM, Majumdar S. High-resolution computed tomography for clinical imaging of bone microarchitecture. Clin Orthop Relat Res 2011; **469**: 2179–93.
32. Turkyilmaz I, Ozan O, Yilmaz B, Ersoy AE. Determination of bone quality of 372 implant recipient sites using Hounsfield unit from computerized tomography: a clinical study. Clin Implant Dent Relat Res 2008; **10**: 238–44.
33. Guerra ENS, Almeida FT, Bezerra FV, Figueiredo PTDS, Silva MAG, De Luca Canto G, et al. Capability of CBCT to identify patients with low bone mineral density: a systematic review. Dentomaxillofac Radiol 2017; **46**: 20160475.

34. Lagravère MO, Carey J, Ben-Zvi M, Packota GV, Major PW. Effect of object location on the density measurement and Hounsfield conversion in a NewTom 3G cone beam computed tomography unit. Dentomaxillofac Radiol 2008; **37**: 305–8.

35. Nemtoi A, Czink C, Haba D, Gahleitner A. Cone beam CT: a current overview of devices. Dentomaxillofac Radiol 2013; **42**: 20120443. doi: 10.1259/dmfr.20120443.

36. Spin-Neto R, Gotfredsen E, Wenzel A. Impact of voxel size variation on CBCT-based diagnostic outcome in dentistry: a systematic review. J Digit Imaging 2013; **26**: 813-20. doi:10.1007/s10278-012-9562-7

37. Rodrigues AF, Campos MJ, Chaoubah A, Fraga MR, Farinazzo Vitral RW. Use of gray values in CBCT and MSCT images for determination of density. Implant Dent 2015; **24**: 155-9. doi:10.1097/id.0000000000000179.

38. Seet KY, Barghi A, Yartsev S, Van Kyk J. The effects of field-of-view and patient size on CT numbers from cone-beam computed tomography. Phys Med Biol 2009; **54**: 6251-62.

39. Isfeld D, Lagravere M, Leon-Salazar V, Flores-Mir C. Novel methodologies and technologies to assess mid-palatal suture maturation: a systematic review. Head Face Med 2017; **13**: 13. doi: 10.1186/s13005-017-0144-2.

40. Arisan V, Karabuda ZC, Ozdemir T. Accuracy of two stereolithographic guide systems for computer-aided implant placement: a computed tomography-based clinical comparative study. J Periodontol 2010; **81**: 43-51.

41. 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; 19: 426-30.

42. Lee RJ, Moon W, Hong C. Effects of monocortical and bicortical mini-implant anchorage on bone-borne palatal expansion using finite element analysis. Am J Orthod Dentofacial Orthop 2017; **151**: 887-97.

43. Tsui WK, Yang Y, McGrath C, Leung YY. Improvement in quality of life after skeletal advancement surgery in patients with moderate-to-severe obstructive sleep apnea: a longitudinal study. Int J Oral Maxillofac Surg 2020; **49**: 333-41.
44. Vale F, Albergaria M, Carrilho E, Francisco I, Guimarães A, Caramelo F, Maló L. Efficacy of Rapid Maxillary Expansion in the Treatment of Obstructive Sleep Apnea Syndrome: A Systematic Review with Meta-analysis. J Evid Based Dent Pract 2017; **17**: 159-68.
45. Li Q, Tang H, Liu X, Luo Q, Jiang Z, Martin D, et al. Comparison of dimensions and volume of upper airway before and after mini-implant assisted rapid maxillary expansion. Angle Orthod 2020; doi:10.2319/080919-522.1. Epub ahead of print.
46. Buck LM, Dalci O, Darendeliler MA, Papadopoulou AK. The Effect of Surgically-Assisted Rapid Maxillary Expansion on Upper Airway Volume: A Systematic Review. J Oral Maxillofac Surg 2016; **74**: 1025-43.
47. Ata-Ali J, Diago-Vilalta JV, Melo M, Bagán L, Soldini MC, Di-Nardo C, et al. What is the frequency of anatomical variations and pathological findings in maxillary sinuses among patients subjected to maxillofacial cone beam computed tomography? A systematic review. Med Oral Patol Oral Cir Bucal 2017; **22**: e400-e409.
48. Gaêta-Araujo H, Vanderhaeghen O, Vasconcelos KF, Coucke W, Coropciuc R, Politis C, et al. Osteomyelitis, osteoradionecrosis, or medication-related osteonecrosis of the jaws? Can CBCT enhance radiographic diagnosis? Oral Dis 2020 Jul; 5. doi: 10.1111/odi.13534. Epub ahead of print.

Figure 1. Flow Diagram of the Literature

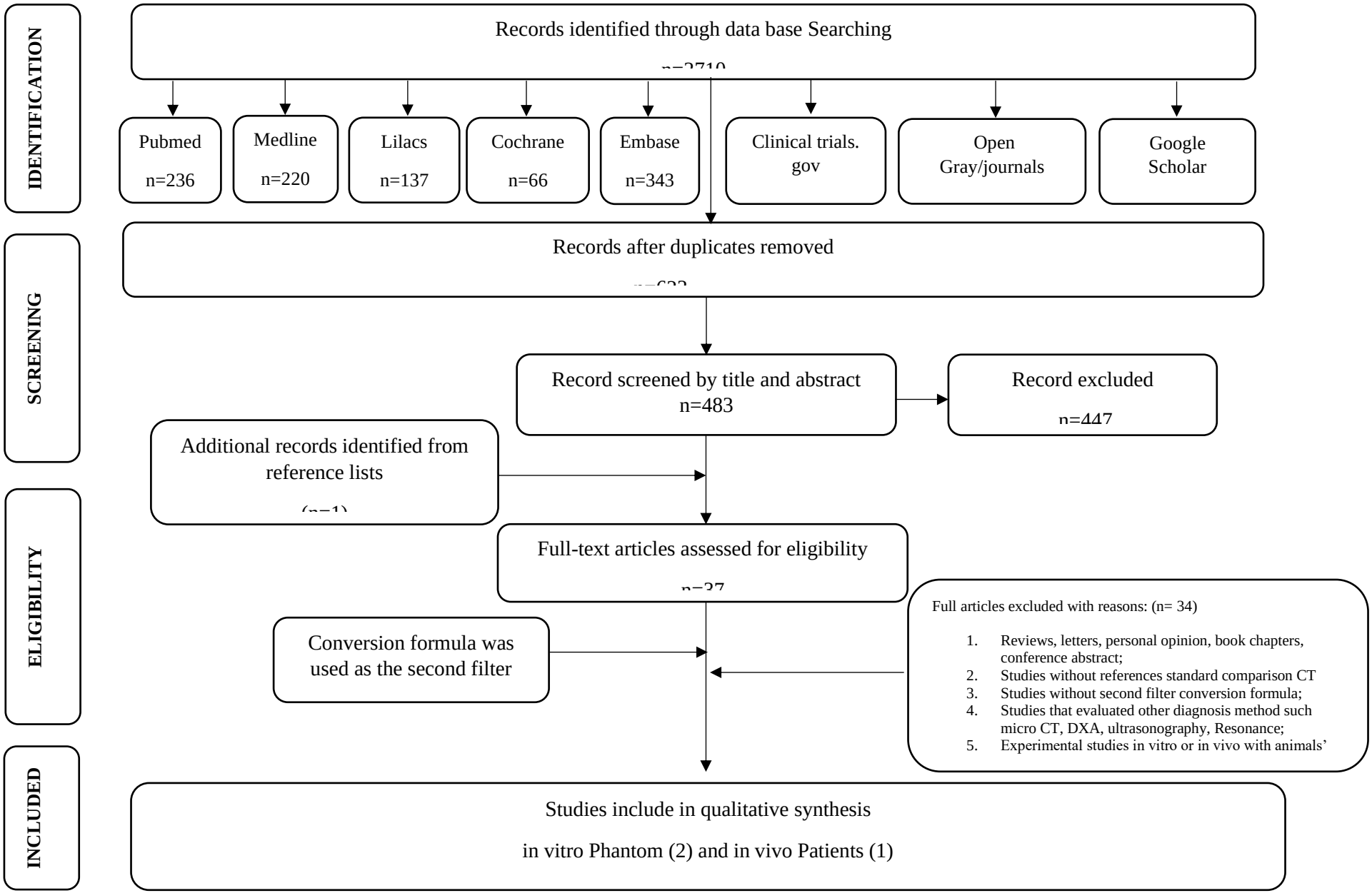

Figure 2. QUADAS- 2 summary.

Study	RISK OF BIAS				APPLICABILITY CONCERNS		
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Reeves, 2012	? Unclear Risk	☺Low Risk	☺Low Risk	☺Low Risk	☺Low Risk	☺Low Risk	☺Low Risk

☺Low Risk ☹High Risk ? Unclear Risk

Table I. *In vitro* and *in vivo* studies

Study Characteristics					***Results***		
Author, Year and Country	***Sample***	***Index test***	***References Standard- Gold Standard***	***ROI***	***Method of convertion GV to HU***	***Results***	***Main Conclusion***
Mah et al., 2010 USA	*In vitro* 3D dental phantom with 8 different materials	11 CBCT scanners - 11 settingss: Asahi Alphard 3030; Hitachi CB MercuRay; I-CAT Classic; I-CAT Next Generation; Iluma ; Morita Accuitomo FPD; Morita Veraview Epochs; NewTom VG; Planmeca ProMax 3D; Galileos and Scanora 3D.	Aquilon 64 slice CT: 120 kV, 300 mA. Briliance 64 CT: 120 kv, 300 mA	Central region within each of 8 materials in the 3D phantom.	1. Statistical correlation 2. Lineal/quadratic regression 3. Conversion formula	R2= 0.999 Highest correlation at effective bean energy	HU can be derived from the GV from CBCT scanners using linear attenuation coefficients as an intermediate step.
Magill et al., 2018 USA	*In vitro* Cylindrical standardized phantom - five different materials: air, polyethylene, acrylic, water and bone-equivalents.	3 CBCT Scanners: CS9300 - Matrix: 557 x 557, Pixel Size (µm): 300, Slice Thickness (mm): 0.3, 90 kVp; 3D Accuitomo - Matrix: 512 x 512, Pixel Size (µm): 250, Slice Thickness (mm): 1.0, 90 kVp; ProMax 3D Mid - Matrix: 500 x 500, Pixel Size (µm): 400 , Slice Thickness (mm): 0.4, 90 kVp.	Phantom specifications for materials in the CT number accuracy module published by the ACR at 120 kVp.	For each image set, ROIs were drawn in the five phantom materials in consecutive axial slices.	1. Statistical correlation 2. Lineal regression 3. Conversion formula	All material HU values for the modified technique fell within 2.4 σ, for the manufacturer-reported HU values ranged from 2.6 to 13.5 σ	The study was able to formulate a valid conversion technique that can provide HU readings for three CBCT units.
Reeves et al., 2012 USA	*In vivo* n= 61 patient scans	Asahi Alphard 3030 (31 scans): 80kv, 5mA, 17 s, 3 exposure mode (200x178 mm FOV, 0.39mm VS / 154X154 mm FOV, 0.3mm VS / 102X102 mm FOV, 0.2mm VS). ProMaxTM 3D (30 scans) - 5 settings: 80kV- 84 kV / 8 mA - 14mA / 18 s / 0.32 mm VS - 0.16 mm VS / 80x80mm FOV.	Linear attenuation coefficients derived from NIST tables of X-ray mass attenuation coefficients and mass energy absorption coefficients for the elemental components in each material.	Average grey levels within a square, 10x10 pixel, for 5 five materials. ROI with the highest grey levels was chosen.	1. Statistical correlation 2. Lineal/quadratic regression 3. Conversion formula	ProMax 3D 60keV: r2=0.9982 ProMax 3D 70keV: r2=0.9987 ProMax 3D 80keV: r2=0.997	A method for deriving Hounsfield units from grey levels in CBCT.

PMMA, polymethyl methacrylate; HA, hydroxyapatite; CBCT, cone- beam computed tomography; MSCT, multislice computed tomography; kV, tube voltage; mAs, milliampere seconds; s, scan time; FOV, field of view; ROI, region of interest; HU, Hounsfield unit; GV, grey values; VV voxel values; kVp, peak kilovoltage; keV, effective energy kilovoltage; r, correlation coefficient; R^2,coefficient of determination; ACR, American College of Radiology; σ, standard deviation,

Table II. Resume of regression and conversion formulas

Author, Year	*Effective energy*	*Conversion Formulas*
Mah, et al, 2010	63 keV	$y = 0.0001734x + 0.2844124$; $HU\ material = \frac{(\mu\ \text{material} - \mu\ \text{water})}{\mu\ \text{water}} x\ 1000$
Magill, et al, 2018	46.7 kVp 48.4 kVp 51.7 kVp	$HVL = \frac{In\ (2)}{\mu \text{Aluminium}}$; $HU\ material = \frac{(\mu\ \text{material} - \mu\ \text{water})}{\mu\ \text{water}} x\ 1000$
Reeves, et al, 2012	70 keV	$y = 1.7198\,x + 317.45$; $HU\ material = \frac{(\mu\ \text{material} - \mu\ \text{water})}{\mu\ \text{water}} x\ 1000$

Table III. *In vitro* studies risk of bias - OHAT tool

	Randomization	Allocation concealment	Experimental condition			Blinding during study	Incomplete data	Exposure characterization	Outcome assessment	Reporting	Other Bias
Study, year	1. Was the exposure level adequately randomized?	2. Was allocation to study groups adequately concealed?	3. Did selection of study participants result in appropriate comparison groups?	4. Did the study design or analysis account for important confounding and modifying variables?	5. Were experimental conditions identical across study groups?	6. Were the researchers blinded to the study?	7. Were outcome data complete without attrition or exclusion from analysis?	8. Can we have confidence in the exposure characterization?	9. Can we be confident in the outcome assessment?	10. Have all measured results been reported? ("methods" and "results" section of the paper)	11. Were there no other potential threats to internal validity
Mah et al, 2010	++	++	na	na	++	++	++	++	++	++	+
Magill et al, 2018	++	++	na	na	++	++	++	++	++	++	++

Low risk of bias (++), probably low risk of bias (+), probably high risk of bias (-/NR), definitely high risk of bias (--).