ABSTRACT

Objective: To compare the accuracy of craniometric measurements made on lateral cephalograms and on cone beam computed tomography (CBCT) images.

Materials and Methods: Ten fiducial markers were placed on known craniometric landmarks of 25 dry skulls with stable occlusions. CBCT scans and conventional lateral headfilms subsequently were taken of each skull. Direct craniometric measurements were compared with CBCT measurements and with cephalometric measurements using repeated measures analysis of variance (ANOVA). All measurements were repeated within a 1-month interval, and intraclass correlations were calculated.

Results: No statistically significant difference was noted between CBCT measurements and direct craniometric measurements (mean difference, 0.1 mm). All cephalometric measurements were significantly different statistically from direct craniometric measurements (mean difference, 5 mm). Significant variations among measurements were noted. Some measurements were larger on the lateral cephalogram and some were smaller, but a pattern could be observed: midsagittal measurements were enlarged uniformly, and Co-Gn was changed only slightly; Co-A was always smaller.

Conclusion: CBCT craniometric measurements are accurate to a subvoxel size and potentially can be used as a quantitative orthodontic diagnostic tool. Two-dimensional cephalometric norms cannot be readily used for three-dimensional measurements because of differences in measurement accuracy between the two exams. (Angle Orthod. 2011;81:28–37.)

KEY WORDS: CBCT; Accuracy; Cephalometry; 3D, Quantitative; Diagnosis

INTRODUCTION

Craniometric measurements have been used to aid orthodontic diagnosis over the past century. Direct craniometric and anthropometric measurements were used before the discovery of the x-ray and the introduction of the cephalometric method. The cephalostat, a modified version of the craniostat used for measuring skulls, was introduced by Broadbent in 1931. Since then, cephalometry has become a commonly used diagnostic tool in orthodontics. Many studies have looked at the reliability of lateral cephalograms and have found them to be reproducible. Few studies, however, have attempted to assess the accuracy of cephalometric measurements as applied three-dimensionally because of known intrinsic limitations of these images, such as distortion and magnification.

Recently, a new technology is catching the attention of the orthodontic community—cone beam computed tomography (CBCT). According to a systematic review of this technology, before 2007 only 14 articles had
been published that were related to CBCT and orthodontics; this number is closer to 300 today. CBCT is an evolution of the original computed tomography (CT) proposed by Hounsfield and Comark. CBCT scans allow the orthodontist to assess the patient’s hard and soft tissue in three dimensions (3D). The accuracy and reliability of such images have been tested and were found to be adequate for implant planning, periodontal disease quantification, and assessment of tumor/lesion volumes.

CBCT application as a craniofacial diagnostic tool often has been underutilized, with the orthodontist gathering 3D data and then synthesizing conventional two-dimensional (2D) films with which he or she is more familiar (eg, lateral headfilms, panoramic radiographs). These reconstructed images are accurate and reliable when compared with conventional radiographs and simulate the way lateral cephalometric or panoramic films are magnified and distorted. This so-called “bridge” from 3D to 2D images has helped orthodontists use the advantages of CT scans without having to add a lateral cephalometric exposure for craniofacial diagnosis.

Today, existing software allows us to take full advantage of CT scans in performing 3D measurements and developing 3D craniofacial analyses. These 3D measurements, made on CT images, can be more accurate and reproducible and have the potential to aid in the craniofacial diagnosis of facial asymmetries, functional shifts, and canted occlusal planes.

Few studies have looked at the accuracy of 3D craniometric CBCT measurements and the potential use of CBCT as a quantitative orthodontic diagnostic tool. The aim of this study is to assess the accuracy and reliability of craniometric measurements made on CBCT scans and lateral cephalograms using dry skulls and fiducial markers.

MATERIALS AND METHODS

The Ethics Committee of the “Museu do Homem de Sambaqui” approved this ex vivo, self-controlled, unblinded experiment. Twenty-five skulls were selected from an overall sample of 500 specimens. To fit the inclusion criteria, the skulls and mandibles had to be well preserved and have a stable occlusion. The skulls were number coded to allow identification of each specimen.

Ten fiducial markers (plastic beads, each with a hole 0.5 mm in diameter) were placed on commonly used craniometric landmarks (Table 1). The skull and the mandible were articulated in presumed centric occlusion to achieve maximum tooth contact.

Twelve direct craniometric measurements (Table 2) were made using a digital caliper (27-500-90, GAC, Bohemia, NY) and were recorded on Excel 2007 (Microsoft, Redmond, Wash) spreadsheets specifically designed for this study.

CT scans were made using the iCAT Next Generation (Imaging Sciences International, Hatfield, Pa) CBCT unit. A standardized protocol of the iCAT for the extended (17 × 23 cm) field of view (FOV) with 0.3 mm slice thickness, 26.9 seconds acquisition time was used (Table 3). The raw images were exported using the iCAT native software (iCAT Vision) into DICOM 3 multifiles. Each file was identified with the specific skull descriptor.

The DICOM images were loaded into SimPlant Ortho 2.0 (Materialise Dental, Lueven, Belgium) software. A custom analysis was created using the dedicated “3D Cephalometric” software module. The custom 3D analysis (COMPASS 3D) was saved to be used with all CT scans. Twelve points were created and identified on the multiplanar reconstructed (MPR) images that were generated by the software (ie, axial, sagittal, and coronal slices). Image enhancing tools and maximum zoom were used to make sure each point was marked precisely at the center of the plastic beads, in all three planes of space (Figure 1).

Based on these 12 points, reference lines were determined and displayed as “3D Cephalogram” (Figure 2). Ten linear measurements between points were determined and calculated by the software. The results were output automatically by the software into a “.CSV” file and were copied to the customized spreadsheets.

The same skulls subsequently were placed on a specially developed stand and were positioned on the cephalostat of the x-ray unit, Orthophos 3C (Siemens, Erlagen, Germany), to facilitate obtaining lateral cephalograms of each specimen. The x-ray unit was calibrated optimally at 60 kV, 66 mA, and 0.16 seconds to achieve the best-quality images while still allowing the markers to be identified.

The films were traced on acetate paper in a darkened room with the aid of a custom black mask to enhance visualization, and the contours of the beads were outlined. Bilateral structures were averaged (the distance between the center of the two bilateral beads was divided in two to determine where the cephalometric landmark would be located) and measured on the lateral cephalogram. Eight (8) linear distances between the cephalometric points were measured using a digital caliper (27-500-90, GAC, Bohemia, NY), with measurements made to the nearest 0.01 mm. The results were input into a custom spreadsheet.

Method Error

To quantify the method error, all measurements were repeated within a 1-month interval, and intraclass
Table 1. Landmarks Selected for the Study

<table>
<thead>
<tr>
<th>Landmark Name</th>
<th>Anatomic Region</th>
<th>Lateral View</th>
<th>Axial View</th>
<th>Coronal View</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glabella (Gl)</td>
<td>Contour of the bony forehead</td>
<td>Anterior-most point at the supraciliary level</td>
<td>Middle-anterior-most point</td>
<td>Middle-anterior-most point</td>
</tr>
<tr>
<td>Point A (A)</td>
<td>Premaxilla</td>
<td>Posterior-most point on the curve of the maxilla between the anterior nasal spine and the supradentale</td>
<td>Middle-anterior-most point on the tip of the premaxilla</td>
<td>Middle point between upper central incisor apices</td>
</tr>
<tr>
<td>Gnathion (Gn)</td>
<td>Contour of the bony chin</td>
<td>Anterior-inferior-most point</td>
<td>Middle-anterior-most point</td>
<td>Middle-inferior-most point</td>
</tr>
<tr>
<td>Menton (Me)</td>
<td>Lower border of the mandible</td>
<td>Inferior-most point</td>
<td>Middle-most point</td>
<td>Inferior-most point</td>
</tr>
<tr>
<td>Posterior nasal spine (PNS)</td>
<td>Median-posterior, sharp bony process of the maxilla</td>
<td>Point on the posterior tip of the maxilla</td>
<td>Middle-posterior-most point</td>
<td>Middle-posterior-most point between the pterygoid processes of the sphenoid bone</td>
</tr>
</tbody>
</table>
Table 1. Continued

<table>
<thead>
<tr>
<th>Landmark Name</th>
<th>Anatomic Region</th>
<th>Lateral View</th>
<th>Axial View</th>
<th>Coronal View</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior nasal spine (ANS)</td>
<td>Median, sharp bony process of the maxilla</td>
<td>Point on the anterior tip of the maxilla</td>
<td>Anterior-most point</td>
<td>Middle-anterior–most point</td>
</tr>
<tr>
<td>Right mandibular gonion (Go-R)</td>
<td>Angle of the right mandibular body</td>
<td>Lateral-most point along the angle</td>
<td>Posterior-most point</td>
<td>Inferior-most point</td>
</tr>
<tr>
<td>Left mandibular gonion (Go-L)</td>
<td>Angle of the left mandibular body</td>
<td>Lateral-most point along the angle</td>
<td>Lateral-posterior–most point</td>
<td>Lateral-inferior–most point</td>
</tr>
<tr>
<td>Right condyliion (Co-R)</td>
<td>Right condyle</td>
<td>Superior-posterior–most point</td>
<td>Middle point between lateral poles of the condylar head</td>
<td>Middle-superior–most point</td>
</tr>
<tr>
<td>Left condyliion (Co-L)</td>
<td>Left condyle</td>
<td>Superior-posterior–most point</td>
<td>Middle point between lateral poles of the condylar head</td>
<td>Middle-superior–most point</td>
</tr>
</tbody>
</table>
correlations (ICCs) were calculated for each measurement method. The CBCT ICC was excellent ($r = 0.99$) (0.01 mm mean difference between measurements; 0.04 mm standard deviation). The direct craniometric measurement ICC also was excellent ($r = 0.98$) (0.1 mm difference between measurements; 0.5 mm standard deviation). The ICC for the lateral cephalogram was very high ($r = 0.98$) (0.1 mm difference between measurements; 0.5 standard deviation).

**RESULTS**

No statistically significant difference was noted between the CBCT measurements and the direct craniometric measurements on analysis of variance (ANOVA) ($P > .05$). The mean difference was 0.01 mm, and all measurements were no more than 0.3 mm different from the craniometric gold standard (Table 4).

The difference between all craniometric and cephalometric measurements was statistically significant (Tukey test, $P < .05$). Great variation was noted among different measurements. Some measurements were larger (Figure 3) on the lateral cephalogram, and some were smaller (Figure 4). A pattern was observed among groups of measurements (midsagittal, mandibular, and maxillary). Measurements between midsagittal points were magnified uniformly. The distance between Condylion and Gnathion seemed to be least affected by the cephalometric magnification and distortion. On the other hand, the distance between Condylion and Point A was smaller on the lateral cephalogram than that determined by direct measurement of the skull (Table 4).
better results are expected if we are to use CBCT scans as a refined research tool. More recently, Baumgaertel et al. compared 10 direct dental measurements made with a digital caliper on 30 dry skulls vs those made on 3D rendered volumes from CBCT scans. They found that the CBCT measurements underestimated direct measurements, and that this difference became statistically significant only when compound measurements were used. Compound measurements require the same landmark to be located multiple times, which can add another source of error. This can be avoided if a dedicated 3D Cephalometric module is used, where landmarks are located only once and the software computes all distances. These dedicated 3D Cephalometric modules are only now becoming commercially available, and each must be validated individually.

The Periago, Hilgers, and Baumgartel studies did not use fiducial markers; all measurements were performed on the 3D rendered volumes, which adds another source of error and limitation to their methods. According to Grauer and coworkers, generating measurements in 3D volumetric images rather than simultaneously in sets of 2D MPR images introduces...
error because of the difficulty involved in locating landmarks in 3D space and the inaccuracies of the user entered threshold used for the construction of 3D virtual surface models (ie, segmentation). Rendered 3D objects may be adequate for qualitative assessment (ie, superimposition). However, accurate quantitative assessment should be performed on the MPR images.

Hilgers and coworkers\(^{21}\) in 2005 compared direct measurements of the temporomandibular joint region vs those made on the MPR images of a CBCT scan (iCAT) with 0.4 mm slice thickness of 25 dry skulls, and found that CBCT measurements were accurate and reproducible.

More recently, Berco and coworkers\(^{17}\) used a single skull, where fiducial radiopaque markers (stainless steel balls 0.5 mm diameter) were used to identify the landmarks to be measured. Landmarks were identified on the iCAT MPR images with a 0.4 mm slice thickness. In this single skull study, investigators were able to demonstrate much greater accuracy (0.2 mm mean difference) than was reported previously; however, because of the study’s extremely small sample size, these results should be validated further.

Clearly, measurement accuracy is affected by the method used. Not only is the quality of the images (the voxel size) and the instrument used to perform those measurements (the precision of the caliper and software) important, but the size, material, and shape of the fiducial marker can influence the study results.\(^{22}\)

The current study used a refined method that favored the reproducibility of direct, cephalometric, and CBCT measurements. Plastic beads instead of previously used steel balls were used because a pilot study by our group determined that it was difficult to reproduce direct measurements over convex surfaces (such as steel balls). Also the plastic material used offered sufficient contrast to be identified on CBCT images, thus reducing the effects of artifacts (ie, scatter images).

Lagrave\(^{15}\) and colleagues conducted an experiment that demonstrated the extremely high reproducibility of CBCT measurements on a prototype mandible using titanium markers with a hollow funnel-like shape. Therefore, the current study used plastic beads with holes in the center, so the caliper could be attached firmly, increasing measurement reproducibility.

The scan resolution in our study also was greater than all those published previously because a new CBCT unit with higher resolution capability (0.3 mm slice thickness) was available. The caliper used in our study offered measurements made to the nearest 0.01 mm, and the dedicated “3D Cephalometric” module of the SimPlant software used has been validated for various purposes by many studies in the industrial and biomedical field.\(^{18,19}\)

In contrast to most previous investigations, the ICC of all measuring methods used in our study was very high,\(^{16,23,24}\) This high level of correlation was due to our study design (using fiducial markers), which virtually

### Table 4. Descriptive Statistics, Pair-wise Group Comparisons for Each Measurement, Using Repeated Measures ANOVA (\(P > .05\)) With Bonferroni Correction

<table>
<thead>
<tr>
<th>Measurement</th>
<th>SKULL(^a) Mean SD</th>
<th>CBCT(^b) Mean SD</th>
<th>CEPH(^b) Mean SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gl-Me</td>
<td>123.9^A 8.9</td>
<td>130.2^A 8.9</td>
<td>142.7^B 9.7</td>
</tr>
<tr>
<td>Gl-ANS</td>
<td>65.1^A 3.7</td>
<td>65.5^A 3.6</td>
<td>71.6^A 3.9</td>
</tr>
<tr>
<td>ANS-Me</td>
<td>65.9^A 6.9</td>
<td>65.7^A 7.0</td>
<td>72.2^A 7.7</td>
</tr>
<tr>
<td>ANS-PNS</td>
<td>52.4^A 3.9</td>
<td>52.4^A 3.7</td>
<td>56.5^A 4.1</td>
</tr>
<tr>
<td>Co-Gn (right)</td>
<td>128.1^A 5.1</td>
<td>128.0^A 5.1</td>
<td>128.0^B 5.7</td>
</tr>
<tr>
<td>Co-Gn (left)</td>
<td>128.3^A 5.3</td>
<td>128.2^A 5.2</td>
<td>128.0^B 5.7</td>
</tr>
<tr>
<td>Co-A (right)</td>
<td>101.6^A 3.9</td>
<td>101.4^A 4.0</td>
<td>95.5^A 4.1</td>
</tr>
<tr>
<td>Co-A (left)</td>
<td>101.8^A 3.8</td>
<td>101.5^A 3.7</td>
<td>95.5^A 4.1</td>
</tr>
<tr>
<td>Co-Go (right)</td>
<td>61.5^A 5.0</td>
<td>61.6^A 5.0</td>
<td>64.9^A 5.7</td>
</tr>
<tr>
<td>Co-Go (left)</td>
<td>59.5^A 7.5</td>
<td>59.5^A 7.7</td>
<td>64.9^A 5.7</td>
</tr>
<tr>
<td>Go-Me (right)</td>
<td>88.8^A 4.6</td>
<td>88.8^A 4.6</td>
<td>82.3^A 4.8</td>
</tr>
<tr>
<td>Go-Me (left)</td>
<td>90.7^A 4.1</td>
<td>90.7^A 4.2</td>
<td>82.3^A 4.8</td>
</tr>
</tbody>
</table>

\(^a\) Mean measurements and standard deviations (SD) on the dry skulls (SKULL), CBCT scans (CBCT), and lateral cephs (CEPH).

\(^b\) A indicates that differences between the skull and the images derived from the CBCT are not statistically significant; B, that differences between the skull and the images derived from the CBCT are statistically significant.

![Figure 3. Mean difference between three groups in millimeters.](image-url)
eliminated other sources of error such as landmark identification. Had we not used fiducial markers, the reliability of all measurements, particularly those on the lateral cephalogram, undoubtedly would have been a lot less accurate and less reproducible.

High accuracy and reproducibility of cephalometric measurements are particularly important because we also compared the CBCT measurements vs those made on the lateral cephalograms to assess whether 2D cephalometric norms could be used as normative values.

Figure 4. Mean difference between three groups in millimeters.

Figure 5. X-ray distortion and magnification.
for 3D measurements. Overall, these differences were not only statistically but also clinically significant. Wide variation was noted among different cephalometric measurements, but a pattern could be identified: mid-sagittal measurements were enlarged uniformly, while mandibular length (Co-Gn) remained relatively unchanged and midfacial length (Co-A) always was smaller.

This variability among measures was a result of distortion caused by the nonparallel x-ray beam after hitting a 3D object, projecting its shadow on a 2D head film. The shadow of the object sometimes is smaller than the object itself (distortion; Figure 3); on the other hand, the nonparallel x-rays sometimes can enlarge the shadow of the object (magnification; Figure 4). This complex interaction between distortion and magnification (Figure 5) is what causes cephalometric measurements to be so inaccurate, except when the distortion is compensated for by the magnification, as, for example, when Co-Gn is measured. Furthermore, the left mandibular length (Co[L]-Gn) was more accurate on the lateral cephalogram than the right mandibular length (Co[R]-Gn). This difference occurs at least in part because the left side is closer to the film than is the right side, leading to less distortion on the side closest to the film. Understanding how the image is distorted at each measurement may help us develop ways to correct these distortions and derive normal values for 3D measurements based on already existing cephalometric norms.

CONCLUSIONS

- CBCT craniofacial measurements obtained by a dedicated “3D Cephalometric module” are extremely accurate and can be used for craniofacial analysis.
- Lateral cephalograms have intrinsic limitations that result in distorted images, enlarged in some areas and reduced in others. A pattern seems to exist, and mathematical algorithms may correct these distortions in the future, so that 3D normal values for 3D quantitative assessment and diagnosis can be derived from previously known 2D norms, without exposing untreated patients to radiation.

ACKNOWLEDGMENTS

We wish to thank the “Museu do Homem de Sambaqui” for allowing us to use their sample of more than 500 skulls (some up to 5000 years old) excavated in the early 1960s by Father João Alfredo Horr. We also thank Dr Vinicius Machado for lending us the “Orthophos 3C” x-ray unit, and KaVo Dental Corporation of Brazil for installing the cephalometric unit and the “iCAT Next Generation” CBCT unit at the museum.

REFERENCES


