Cephalometric measurements from 3D reconstructed images compared with conventional 2D images

Natalia Zamora; Jose M. Llamas; Rosa Cibrián; Jose L. Gandia; Vanessa Paredes

ABSTRACT

Objective: To assess whether the values of different measurements taken on three-dimensional (3D) reconstructions from cone-beam computed tomography (CBCT) are comparable with those taken on two-dimensional (2D) images from conventional lateral cephalometric radiographs (LCRs) and to examine if there are differences between the different types of CBCT software when taking those measurements.

Material and Methods: Eight patients were selected who had both an LRC and a CBCT. The 3D reconstructions of each patient in the CBCT were evaluated using two different software packages, NemoCeph 3D and InVivo5. An observer took 10 angular and 3 linear measurements on each of the three types of record on two different occasions.

Results: Intraobserver reliability was high except for the mandibular plane and facial cone (from the LCR), the Na-Ans distance (using NemoCeph 3D), and facial cone and the Ans-Me distance (using InVivo5). No statistically significant differences were found for the angular and linear measurements between the LCRs and the CBCTs for any measurement, and the correlation levels were high for all measurements.

Conclusion: No statistically significant differences were found between the angular and linear measurements taken with the LCR and those taken with the CBCT. Neither were there any statistically significant differences between the angular or linear measurements using the two CBCT software packages. (Angle Orthod. 2011;81:856–864.)

KEY WORDS: 2D; 3D; CBCT; LCR; Cephalometric; Measurement

INTRODUCTION

Cephalometric analysis is one of the key tools in undertaking an accurate diagnosis in orthodontics, even though it presents a number of limitations given that it reduces to two dimensions a three-dimensional (3D) object by projecting all structures onto a single plate, thereby creating difficulties when it comes to understanding and undertaking an adequate analysis. Moreover, cephalometric analysis has other technical limitations as the images obtained can be distorted because of mistakes associated with the X-ray apparatus or errors in the positioning of the patient’s head.

Conventional medical computed tomography (CT) was developed to analyze 3D structures on the three spatial planes and so provide a more realistic image, although the high cost of 3D reconstruction and the high dosage of radiation involved limited its use. Cone-beam computed tomography (CBCT)1,2 was then developed to reduce the radiation dosage, achieve greater precision on the three spatial planes, and reduce the costs associated with CT.3

The accuracy4–16 and reproducibility11,14 of cephalometric landmarks have been widely studied for CT4,17,18 and CBCT5,10,12,13,19 by evaluating the location of those landmarks20 or by comparing the linear and angular measurements15,21–23 taken on lateral cephalometric radiographs (LCR) with two-dimensional (2D) projections (orthogonal, in perspective, maximum-intensity projection [MIP], projection similar to x-ray [RaySum]), obtained from slices of a CBCT scan. The results for landmark location were similar for both methods,20–22
greater reproducibility being found in the projections from CBCT than from LCR, although there were no clinically significant results.\textsuperscript{23} It has also been observed that the RaySum projection is more reproducible than the MIP and that either projection could replace LCR, thus making LCRs unnecessary when a CBCT has previously been undertaken on a patient.\textsuperscript{15,21--23}

However, for frontal projections, other authors found statistically significant differences, as they saw that the positioning of the patient’s head\textsuperscript{24} influenced the measurement.

Other studies concluded that in order to adjust the location of cephalometric landmarks, slices were more accurate than the 3D view provided, because 3D projection did not represent real surfaces.\textsuperscript{14} Nevertheless, these results differed from those of other studies that found linear measurements undertaken directly on 3D images to be more accurate.\textsuperscript{25}
As there are few studies that compare the cephalometric values obtained from LCR with those from 3D projection of CBCT, and as we wanted to determine the compatibility of different CBCTs in common cephalometric analyses, the aims of our study were, first, to check whether the angular and linear measurements of 3D reconstructions obtained from a CBCT are equal to those of 2D images obtained from the LCR, and second, to check whether there are differences between two CBCT software packages (Beta NemoCeph 3D, Software Nemotec, SL, Madrid, Spain and InVivo5 Anatomage, San Jose, Ca) when taking those measurements.

MATERIALS AND METHODS

The sample was distributed as follows: four women and four men between the ages of 8 and 30 years old. Mean (±SD) age = 15.55 ± 6.25 years old.

Eight patients were selected. Those patients had undergone both an LCR and a CBCT before treatment and within 3 months of each other. The CBCTs had...
been taken because some of these patients were going to undergo orthognathic surgery and the others presented with inclusions of maxillary canines.

The LCRs were taken using a cephalostat Orthophos Plus DS (Sirona Dental Systems GmbH, Germany) with a radiation time of 16 seconds for an exposure time of 0.4 seconds, the tube voltage being 73 kV and the intensity 15 mAs, whereas the CBCTs were taken using i-CAT equipment (Imaging Sciences International, Hatfield, Pa) that generates a total of 326 slices, with an image matrix size of 400 x 400 in medium-quality and high-resolution mode. The field of view (FOV) used was the portrait mode, which gathers data in extended FOV mode and covers the full skull of 170 mm in height x 230 mm in diameter with a scanning time of 8.9 seconds. Voxel size was 0.4 x 0.4 x 0.4 mm, the tube voltage being 120 kV and the intensity 23.87 mAs. The size of the data files created is 35 megabytes.

To carry out this study, 17 cephalometric landmarks were located. These are described from the three spatial planes in Table 1a. Then, 10 angular and 3 linear measurements were taken (Table 1b).

The LCRs were digitalized using a conventional scanner-type HP Scan Jet II Cx/T (Hewlett Packard Development Company LP, Palo Alto, Ca), and the location of the landmarks was measured using the

Figure 2. Location of the cephalometric landmarks in the InVivo® software package.

Figure 3. Comparison between the three methods. Means of all values provided.
Nemotec Dental Studio NX program (Software Nemotec SL, Madrid, Spain).

The raw data and slices obtained from the CBCTs were imported to two different software packages where 3D reconstruction was undertaken. Location of the cephalometric landmarks for each of the two CBCT software packages was undertaken by creating different types of projections (multiplanar reconstruction, RaySum, MIP) on the sagittal, coronal, and axial views and to the volumetric reconstruction (3D).

First, to fine-tune the location of points using the Beta NemoCeph 3D software (Software Nemotec SL, Madrid, Spain), two windows were created for the sagittal view and another two for the coronal or frontal view, which, moreover, can be enlarged to full screen using the zoom. By doing so, we can simultaneously observe the right and left sides or have two projections of the same side open for the sagittal view, or, for the coronal view, have two different projections open at the same time (Figure 1). To locate each landmark, the slice of the most appropriate plane was selected before fitting it onto the other planes for greater accuracy.

Second, to locate landmarks using the InVivo5 software (Anatomage, San Jose, Calif), the program allows one to view either a part or the whole skull by navigating across the slices so as to choose the most appropriate one for locating each point (Figure 2).

All measurements were undertaken by a single observer previously trained and qualified in the location of cephalometric landmarks and with six years of experience/background in orthodontics. The measurements were taken on two different occasions with an interval of one week between the first and second measurements so as to measure the intraobserver error.

All variables and measurements were introduced onto a version 12.0 Excel spreadsheet (Microsoft Corp, Redmond, Wa) and then analyzed using version 17.0 of the statistics package SPSS for Windows (IBM Corp, Somers, NY). To calculate the reproducibility of each of the systems used, a t-test was undertaken for related samples. To determine reliability of the different types of systems, a repeated measurement analysis of variance was undertaken using a variance factor, by means of Scheffé’s adjustment for multiple comparisons, through the means obtained from the first and second observations for that purpose. Differences were considered significant at \( P < .05 \).

RESULTS

The results obtained from this study are shown in Tables 2 and 3 and in Figure 3.

Table 2 shows the intraobserver reproducibility between the first and second measurement for each of the methods. In the measurements taken on the LCR, a statistically significant difference can be observed in the angle of the mandibular plane, where the maximum difference between the two measurements (first and second) was determined as being 5.78° ([5.78 \( \pm \) 3.48] * 100 = 11.25), which means an error of 11.25%, as well as in the facial cone, with a maximum difference of 3.60° ([3.60 \( \div \) 65.43] * 100 = 5.5), which means an error of 5.5%. In the measurements taken with the Beta NemoCeph 3D, a difference was found for the Na-Ans distance, with a maximum difference between measurements of 2.5 mm ([2.5 \( \div \) 51.18] * 100 = 5), which means 5% of error in that measurement. Lastly, in the measurements taken with InVivo5, a statistically significant difference was found for the facial cone (maximum difference: 3.48°; [3.48 \( \div \) 66.15] * 100 = 5.3), which means 5.3% of error, and for the Ans-Me distance (maximum difference: 5.78; [5.78 \( \div \) 65.46] * 100 = 8.8), which means an error of 8.8%.

### Table 2. Intraobserver Reliability Between the First and Second Measurement

<table>
<thead>
<tr>
<th>Measurement</th>
<th>LCR</th>
<th>CBCT (NemoCeph 3D)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean±SD 1st–2nd Measurement</td>
<td>Mean±SD 1st–2nd Measurement</td>
</tr>
<tr>
<td>SNA</td>
<td>83.36 ± 5.24</td>
<td>83.08 ± 4.22</td>
</tr>
<tr>
<td>SNB</td>
<td>80.68 ± 6.91</td>
<td>80.29 ± 6.07</td>
</tr>
<tr>
<td>NS-Palatal plane</td>
<td>7.53 ± 6.25</td>
<td>7.02 ± 6.20</td>
</tr>
<tr>
<td>NS-Mandibular plane</td>
<td>31.98 ± 6.62</td>
<td>32.57 ± 6.35</td>
</tr>
<tr>
<td>Interincisal angle</td>
<td>133.10 ± 9.24</td>
<td>134.51 ± 1.34</td>
</tr>
<tr>
<td>Upper Incisor-NS</td>
<td>100.80 ± 8.27</td>
<td>100.8 ± 8.95</td>
</tr>
<tr>
<td>Tweed angle</td>
<td>92.51 ± 8.93</td>
<td>91.92 ± 8.02</td>
</tr>
<tr>
<td>Facial cone</td>
<td>65.43 ± 6.34</td>
<td>64.81 ± 5.69</td>
</tr>
<tr>
<td>Cranial deflexion</td>
<td>29.84 ± 2.35</td>
<td>30 ± 2.30</td>
</tr>
<tr>
<td>Facial depth</td>
<td>91.36 ± 5.15</td>
<td>90.73 ± 5.32</td>
</tr>
<tr>
<td>Na-Me</td>
<td>114.70 ± 8.16</td>
<td>115.21 ± 0.56</td>
</tr>
<tr>
<td>Na-Ans</td>
<td>51.18 ± 3.13</td>
<td>50.36 ± 3.69</td>
</tr>
<tr>
<td>Ans-Me</td>
<td>66.36 ± 6.79</td>
<td>67.08 ± 6.70</td>
</tr>
</tbody>
</table>

* Statistically significant difference at \( P \leq .05 \).
DISCUSSION

Each of the 13 variables used in this study were measured on the LCR and on the 3D reconstruction of each CBCT using the two software packages (Beta NemoCeph 3D and InVivo5) in order to compare their reproducibility and the statistically significant differences between each of the methods for the measurements analyzed.

To determine the reproducibility of the measurements and the intraobserver error, a single observer repeated all the measurements on a second occasion. We consider that it is sufficient to do so twice, as additional measurements do not provide any relevant additional information. Moreover, the possible systematic errors made by an observer will always be the same for the three methods. If the measurements of more observers are combined, they may lead to errors when it comes to evaluating the real variability of the measurements.

In general, reproducibility between the first and the second measurement, both for the linear and angular measurements, was high for all three methods, with correlation indices of intraclass correlation > 0.86 in all cases except for cranial deflection and the Na-Ans distance on the LCR and InVivo5 and for facial depth with the InVivo5 software package.

The only statistically significant differences found were between the first and second measurement for the SNa-mandibular plane angle and for the facial cone measured on the LCR, for the Na-Ena distance measured using the NemoCeph 3D software package, and for the facial cone and Ans-Me distance using the InVivo5 software package.

For these measurements, a systematic error rather than a random one was found between the first and second measurements providing greater values for one measurement than the other. For this reason those measurements appear with statistically significant differences. However, in all cases the relative error determined from the maximum discrepancy has not exceeded the method error, as all the measurements taken would fall within method variability (obtained by dividing the standard deviation [SD] between the mean of each value).

On analyzing the correlations of each of the variables for the three methods, no angular or linear measurement was found with a statistically significant difference using the three techniques, although we have observed that facial depth presents the worst correlation between the three variables (between LCR and InVivo and between Beta NemoCeph 3D and InVivo5), which implies that there is a greater difference between the values of this measurement in the three methods.

In contrast to our results, Van Vlijmen et al. found statistically significant differences in several measurements that were measured on the LCR and on the CBCT; for example, the SNa-mandibular plane angle and the SNB angle. However, as in our study, the difference between most of the measurements was less than the SD, confirming that that difference was not clinically significant, although reproducibility of the measurements on the LCR was higher than in the 3D reconstructions. In the cases in which a 2D line and 3D plane were used for undertaking a measurement, they observed that there may be a clinically relevant difference, so 3D tracings would not be recommendable for undertaking longitudinal studies where pre-treatment records had been recorded in 2D. However, measurements undertaken on a single patient who has 3D records would be of use for comparing pre- and posttreatment changes or changes due to growth. In another study by the same authors, which compared 2D and 3D frontal views, clinically relevant differences were found for both types of measurement. In a third study, the same authors compared measurements between two types of CBCT (Iluma Imtech, Ardmore, OK and iCAT, Imaging Sciences International, Inc, Hatfield, Pa), concluding that the measurement reproducibility of 3D reconstructions in Iluma was higher than in the i-CAT, finding, in addition, statistically significant differences between several angular measurements, although only two of them presented clinically relevant differences.

Yitschaky et al. compared angular, linear, and ratio measurements between the LCR and CT and did not
find statistically significant differences, except in some measurements, such as the Witts appraisal, or in others that included the Sella point (because of the differences in their location between both records). As our results, they did not find statistically significant differences for either the Na-Me or Ans-Me distances, concluding that it is possible to use most of the values previously established in 2D for measuring 3D records, as the differences between both systems were small.

Other authors have also compared angular measurements between LCR and CT.\textsuperscript{30} They found only statistically significant differences in 2 of the 14 measurement that they carried out, upper incisor-Na and upper incisor-NS. However, in our study, we did not find such difference for the upper incisor-NS angle. They found, as we did, differences that were not statistically significant for SNA, SNB, NS-palatal plane, and Tweed angle. They concluded that 3D angular cephalometric analysis was a fairly reliable method, like the traditional 2D cephalometric analysis, and that it was also more suitable for the diagnosis of cases with complex orthodontic anomalies. Currently, it could also be a suitable alternative method to 2D cephalometry with the decrease in radiation exposure and costs in the future.

To carry out this study, records of patients who had already undergone both an LCR and a CBCT for various reasons were used. Other studies have used dry skulls to undertake the measurements, as patient irradiation is not justified unless strictly necessary. However, and despite the fact that certain studies state that soft tissues distort the measurements,\textsuperscript{26} it is necessary to undertake them with all the tissues in place.

\begin{table}
\centering
\begin{tabular}{ |l|l|l|l|l| }
\hline
Measurement & (I) & (J) & Difference of the Mean (I–J) & Error & Significance* \\
\hline
SNA & LCR & NemoCeph 3D & 0.27500 & 2.24546 & .993  \\
 & & InVivo5 & 1.22375 & 2.24546 & .863  \\
 & NemoCeph 3D & InVivo5 & 0.94875 & 2.24546 & .915  \\
SNB & LCR & NemoCeph 3D & 0.38125 & 3.14926 & .993  \\
 & & InVivo5 & 1.18000 & 3.14926 & .932  \\
 & NemoCeph 3D & InVivo5 & 0.79875 & 3.14926 & .968  \\
NS-Palatal plane & LCR & NemoCeph 3D & 0.51250 & 3.30267 & .988  \\
 & & InVivo5 & −0.54875 & 3.30267 & .986  \\
 & NemoCeph 3D & InVivo5 & −1.06125 & 3.30267 & .950  \\
NS-Mandibular plane & LCR & NemoCeph 3D & −0.59375 & 3.12667 & .982  \\
 & & InVivo5 & −1.73000 & 3.12667 & .859  \\
 & NemoCeph 3D & InVivo5 & −1.13625 & 3.12667 & .936  \\
Interincisal angle & LCR & NemoCeph 3D & −1.46250 & 5.19933 & .961  \\
 & & InVivo5 & 1.27750 & 5.19933 & .970  \\
 & NemoCeph 3D & InVivo5 & 2.74000 & 5.19933 & .871  \\
Upper incisor-NS & LCR & NemoCeph 3D & 0.06875 & 4.31088 & 1.000  \\
 & & InVivo5 & −0.90000 & 4.31088 & .978  \\
 & NemoCeph 3D & InVivo5 & −0.96875 & 4.31088 & .975  \\
Tweed angle & LCR & NemoCeph 3D & 0.58750 & 4.30435 & .991  \\
 & & InVivo5 & 0.14125 & 4.30435 & .999  \\
 & NemoCeph 3D & InVivo5 & −0.44625 & 4.30435 & .995  \\
Facial cone & LCR & NemoCeph 3D & 0.62500 & 3.03655 & .979  \\
 & & InVivo5 & −0.72250 & 3.03655 & .972  \\
 & NemoCeph 3D & InVivo5 & −1.34750 & 3.03655 & .907  \\
Cranial deflexion & LCR & NemoCeph 3D & −0.16250 & 1.02769 & .988  \\
 & & InVivo5 & −0.33375 & 1.02769 & .949  \\
 & NemoCeph 3D & InVivo5 & −0.17125 & 1.02769 & .986  \\
Facial depth & LCR & NemoCeph 3D & 0.62500 & 2.51701 & .970  \\
 & & InVivo5 & 2.34125 & 2.51701 & .654  \\
 & NemoCeph 3D & InVivo5 & 1.71625 & 2.51701 & .795  \\
Na-Me & LCR & NemoCeph 3D & −0.48125 & 5.06312 & .995  \\
 & & InVivo5 & −0.98125 & 5.06312 & .981  \\
 & NemoCeph 3D & InVivo5 & −0.50000 & 5.06312 & .995  \\
Na-Ans & LCR & NemoCeph 3D & 0.81250 & 1.84284 & .908  \\
 & & InVivo5 & −0.77000 & 1.84284 & .917  \\
 & NemoCeph 3D & InVivo5 & −1.58250 & 1.84284 & .696  \\
Ans-Me & LCR & NemoCeph 3D & −0.71875 & 3.40104 & .978  \\
 & & InVivo5 & 0.89250 & 3.40104 & .966  \\
 & NemoCeph 3D & InVivo5 & 1.61125 & 3.40104 & .894  \\
\hline
\end{tabular}
\caption{Multiple Comparisons Using the Scheffe Method}

* Significance between three systems: LCR, NemoCeph 3D (CBCT), and InVivo5 (CBCT).
included in order to be able to check more accurately and to get as close as possible to the daily clinical reality as far as measuring patients is concerned. The drawback of carrying out a study of these characteristics is that irradiating actual patients for this purpose is not justified and therefore the sample in our study was restricted to only eight patients.

CONCLUSIONS

- No statistically significant differences were found for angular and linear measurements between the measurements taken with the LCR and those of the CBCT. The correlation indices of all measurements are very high.

- There were also no statistically significant differences for any of the angular and linear measurements between the two software packages of the CBCT (NemoStudio and InVivo5). Likewise, the correlation indices found between the different measurements were very high.

REFERENCES


models of human skulls derived from two different cone beam CT scanners. *Clin Oral Invest*. 2010; Jul 17 (Epub ahead of print).


