# A study on the reproducibility of cephalometric landmarks when undertaking a three-dimensional (3D) cephalometric analysis

**Natalia Zamora [1], José-María Llamas [2], Rosa Cibrián [3], José-Luis Gandia [4], Vanessa Paredes [5]**

[1] PhD, MSc, Orthodontist, Assistant Professor, Department of Orthodontics, Faculty of Medicine and Dentistry, University of Valencia, Valencia, Spain
[2] PhD, Orthodontist, Associate Professor and Director of Postgraduate Orthodontics Masters Course, Department of Orthodontics, Faculty of Medicine and Dentistry, University of Seville, Seville, Spain
[3] PhD, Associate Professor, Department of Physiology, Faculty of Medicine and Dentistry, University Valencia, Valencia, Spain
[4] PhD, Orthodontist, Associate Professor and Department Chair of Postgraduate Orthodontics Masters Course, Department of Orthodontics, Faculty of Medicine and Dentistry, University of Valencia, Valencia, Spain
[5] PhD, Orthodontist and PhD Assistant Professor, Department of Orthodontics, Faculty of Medicine and Dentistry, University of Valencia, Valencia, Spain

*Correspondence:*
*Clínica Odontológica*
*Departamento Ortodoncia 5° piso*
*c/ Gasco Oliag n°l. 46010*
*Valencia. Spain*
*nataz84@hotmail.com*

Received: 26/05/2011
Accepted: 20/07/2011

## Abstract

Objectives: Cone Beam Computerized Tomography (CBCT) allows the possibility of modifying some of the diagnostic tools used in orthodontics, such as cephalometry. The first step must be to study the characteristics of these devices in terms of accuracy and reliability of the most commonly used landmarks. The aims were 1- To assess intra and inter-observer reliability in the location of anatomical landmarks belonging to hard tissues of the skull in images taken with a CBCT device, 2- To determine which of those landmarks are more vs. less reliable and 3- To introduce planes of reference so as to create cephalometric analyses appropriated to the 3D reality. Study design: Fifteen patients who had a CBCT (i-CAT®) as a diagnostic register were selected. To assess the reproducibility on landmark location and the differences in the measurements of two observers at different times, 41 landmarks were defined on the three spatial axes (X,Y,Z) and located. 3.690 measurements were taken and, as each determination has 3 coordinates, 11.070 data were processed with SPSS® statistical package. To discover the reproducibility of the method on landmark location, an ANOVA was undertaken using two variation factors: time (t1, t2 and t3) and observer (Ob1 and Ob2) for each axis (X, Y and Z) and landmark. The order of the CBCT scans submitted to the observers (Ob1, Ob2) at t1, t2, and t3, were different and randomly allocated. Multiple compari-

sons were undertaken using the Bonferroni test. The intra- and inter-examiner ICC´s were calculated. Results: Intra- and inter-examiner reliability was high, both being ICC ≥ 0.99, with the best frequency on axis Z. Conclusions: The most reliable landmarks were: Nasion, Sella, Basion, left Porion, point A, anterior nasal spine, Pogonion, Gnathion, Menton, frontozygomatic sutures, first lower molars and upper and lower incisors. Those with less reliability were the supraorbitals, right zygion and posterior nasal spine.

*Key words:* *Cone Beam Computed Tomography, cephalometry, landmark, orthodontics, reliability.*

## Introduction

Modern Cone Beam Computerized Tomography (CBCT) systems applied to oral and maxillofacial regions provide a helpful diagnostic tool in orthodontics (1). This technology eliminated the errors of the conventional radiographs (magnification, image distortion, superimposition of anatomical structures) and the drawbacks that medical computed tomography (CT) presented (high economic costs, high radiation doses) (2). Several authors have compared these radiation doses between all kinds of devices (3-5), all of them concluding that conventional systems (panoramic radiography and lateral radiography) continue to emit the lowest radiation doses, followed by CBCT and lastly conventional CT. Nevertheless, it has been observed that several orthodontic patients need not only panoramic radiography and lateral radiography, but also other additional radiographs; such as a posteroanterior radiography to assess asymmetries, periapical series in periodontal problems, occlusal radiographs or even a magnetic resonance in order to assess the temporomandibular articulation. The sum of the effective doses of all these additional radiographs exceeded the effective dose of the CBCT. Therefore, in these cases, the use of the CBCT would be recommended instead of undertaking all those above mentioned radiographs separately (5,6). CBCT also allow the possibility of modifying some of the diagnostic tools used in orthodontics, such as cephalometry. The first step to undertake cephalometric studies on CBCTs, must be to study the characteristics and limitations of these devices in terms of accuracy and reliability of the most commonly used landmarks. Despite the fact that many studies have evaluated the accuracy and reliability both in CTs (7-10) and CBCTs (11-23) many of them evaluate linear distances. This methodology presents the limitation that in case the measurement is seen to be lacking in accuracy, we cannot determine which of the two points forming the line is the inaccurate one (7). Training and familiarization with the location of cephalometric landmarks is essential because landmark identification errors are considered a major source of cephalometric errors (15). Moreover, landmarks must be located from whichever viewpoint as exact points and with a unique anatomical location (8) and must be defined for each of the 3D spatial planes as coordinates (20). Once the easier points to locate are known, we can introduce planes of reference for a 3D cephalometric analysis.

The aims of this study were; 1- To assess intra and inter-observer reliability in the location of anatomical landmarks belonging to hard tissues of the skull in images taken with a CBCT device, 2- To determine which of those landmarks are more vs. less reliable and 3- To introduce planes of reference so as to create cephalometric analyses appropriated to the 3D reality.

## Material and Methods

A study approved by the ethical committee of the Clinical University Hospital of the University of Valencia was undertaken. Fifteen (n=15) patients were selected of between 8 and 27 years of age (mean age: 15.27 ± 5.34 (SD) years old), 73.4% females and 26.6% males. The CBCTs had been taken because some of these patients were scheduled for orthognathic surgery, while the other ones presented with impacted maxillary canines.

Scans of these patients were undertaken using the CBCT i-CAT® (Imaging Sciences International, Hatfield, Pa) equipment. This apparatus generates a total of 306 slices with an image matrix size of 575x575. It was set at medium quality and high resolution. The portrait mode field of view (FOV) was employed, which gathers data in extended FOV mode and includes the entire head (170 mm x 230 mm) with a scanning time of 8.9 seconds. The voxel size was 0.4 x 0.4 x 0.4 mm. Tube voltage is 120 kVp and its intensity 23.87 mAs. The gross data and the slices obtained were imported to Beta NemoStudio® software (Software Nemotec SL, Madrid, Spain) where the post-processing of the medical images, their conversion into DICOM (Digital Imaging and Communications in Medicine) format, and 3D reconstructions were undertaken.

41 landmarks belonging to the hard tissues of the skull were defined. The landmarks chosen are commonly used in orthodontics for locating craniofacial structures, and they represent landmarks that clinicians are familiarized to recognise and locate without difficulty (Table 1). The software automatically determined the origin of the coordinates (0,0,0) at the right anterior lower corner of the cube that contained the 3D image and 3 axes were defined: X parallel to the right-left direction, Y in

**Table 1.** Definition of the three spatial planes of the 41 points used in this study. Anteroposterior point(APP), Mid point (MP), Posterior point(PP), Lowest Point(LP), Upper Point(UP), Anterlower Point (ALP), Anteriorupper.Point(AUP), Posteriorlower Point(PLP), Highest Point(HP), Inner Point (IP).

| Point Name | Anatomical definition | Sagittal or lateral view | Coronal or frontal view | Axial view |
|---|---|---|---|---|
| **Sella turcica (S)** | APP MP pituitary fossa sphenoid bone | MP APP width | MP lateral width fossa, determined antero-posteriorly by the other two slices(2) | MP APP and lateral width fossa |
| **Nasion (Na)** | Most AP frontonasal suture | Most AP | MP | Most AP+MP anterior contour |
| **Basion (Ba)** | Most AP foramen magnum | Most PP+LP | MP foramen, determined antero-posteriorly by the 2 | Most AP foramen |
| **Crista Galli (Cg)** | Most UP crista galli from ethmoid bone | UP | MP | MP determined supero-inferiorly by the 2 |
| **Right and left Porion (PoR and PoL)** | UP and MP external right roof auditory meatus | UP+MP | UP | MP determined supero-inferiorly by the 2 |
| **Point A (A)** | Most PP maxillar curvature, between anterior nasal spine and supradental point | Most PP | MP determined antero-posteriorly by the 2 | AP+MP |
| **Point B (B)** | Most PP anterior surface mandibular symphysis | Most PP | MP determined antero-posteriorly by the 2 | AP+MP |
| **Pogonion (Pg)** | Most AP mandibular symphysis | Most AP | MP | AP+MP |
| **Gnathion (Gn)** | Most ALP mandibular symphysis | MA+LP | MP+LP | AP, LP+MP |
| **Menton (Me)** | LP mandibular symphysis | LP | LP | LP+MP |
| **Anterior Nasal Spine (Ans)** | Most AP maxillary process nasal floor region | Most AP | AP+MP | AP+MP |
| **Posterior Nasal Spine (Pns)** | Most PP and mid-point palantine bone contour | Most PP | PP+MP | PP+MP |
| **Right and left Condylion (CdR and CdL)** | UP point head right condyle | UP+most PP | Most UP+MP | Most PP |
| **Right and left Gonion (GoR and GoL)** | Most PP edge branch. Bisection tangents posterior edge branch and lower body | Most PP | Most PP+MP | Most PP determined supero-inferiorly by the 2 |
| **Right and left Maxillar (MxR and MxL)** | Zygomatic-alveolar crest. Maximum concavity maxillar contour around molars and lower contour maxillo-cygmatic process | MP 1 upper molar | Most IP concavity | IP determined supero-inferiorly by the 2 |
| **Right and left Orbital (OrR and OrL)** | Most AUP infraorbital margin orbital | Most AP | UP+MP | Most AP |
| **Right and left Supraorbital (SOrR and SOrL)** | Most ASP supraorbital margin orbital | Most AP | LP+MP | Most AP |
| **Right and left Bucal (BcR and BcL)** | External surface zygomatic arch, where arch becomes straight and inwards | Most AP | MP | Most AP+MP |
| **Anterior point of the right and left branch (RR and RL)** | Deepest point anterior edge branch | Most PP+MP | MP | MP determined supero-inferiorly by the 2 |
| **Incisal edge upper right central incisor (UIR)** | LP incisal edge central right incisor | LP | Most MP mesiodistal width | Most AP+MP |
| **Incisal edge of the lower right central incisor (LIR)** | UP incisal edge central right incisor | UP | Most MP mesiodistal width | Most AP+MP |
| **1st upper right and left molar (A16, A26)** | Most PP+MP distal surface molar | Most PP+MP | MP determined antero-posteriorly by the 2 | Most PP+MP |
| **1st lower right and left molar (B46, B36)** | Most PP+MP distal surface molar | Most PP+MP | MP determined antero-posteriorly by the 2 | Most PP+MP |
| **Right and left Tuberosity (TbR, TbL)** | Most PLP distal contour maxillar tuberosity | Most PP+LP | LP+MP | Most PP+MP |
| **Right and left Frontozygomatic Suture (FzR and FzL)** | Most ALP side frontozygomatic suture | Most AP | MP | MP determined supero-inferiorly by the 2 |
| **Left and Right Zygion (ZyL and ZyR)** | Most LP external rim zygomatic arch | Most AIP | Most LP | Most AP+MP determined supero-inferiorly by the 2 |
| **Right and left Antegonion (AgR and AgL)** | HP concavity lower edge mandibular branch at its union with the body of the mandible | Most UP | UP+MP | MP determined supero-inferiorly by the 2 |

posterior-anterior direction and Z in the upper-lower direction. Once the coordinate system had been standardized, spatial positions of each landmark were represented as numerical values (in mm) on each axis. In order to fine-tune the location of points, the software permitted the visualization of two windows for the sagittal view and another two for the coronal or frontal view, which, moreover, could be enlarged to full screen using the zoom. By doing so, we were able to 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 (MPR= multiplanar reconstruction, Raysum= X-ray projection, MIP=maximum intensity projection) (Fig. 1). To locate each landmark, the slice of the most

To discover the reproducibility of the method on landmark location, an ANOVA was undertaken using two variation factors: time (t1, t2 and t3) and observer (Ob1 and Ob2) for each axis (X, Y and Z) and landmark. The order of the CBCT scans submitted to the observers (Ob1, Ob2) at t1, t2, and t3, were different and randomly allocated. Multiple comparisons were undertaken using the Bonferroni test. The intra- and inter-examiner ICC´s were calculated. In our study, we decided to use the ICC since the measurements were numeric and not categorical, in which case we would have used the Cohen's kappa test. Besides, the ICC had been used by other authors reviewed in our work and this allows us to compare the results in a very similar way. Reproducibility
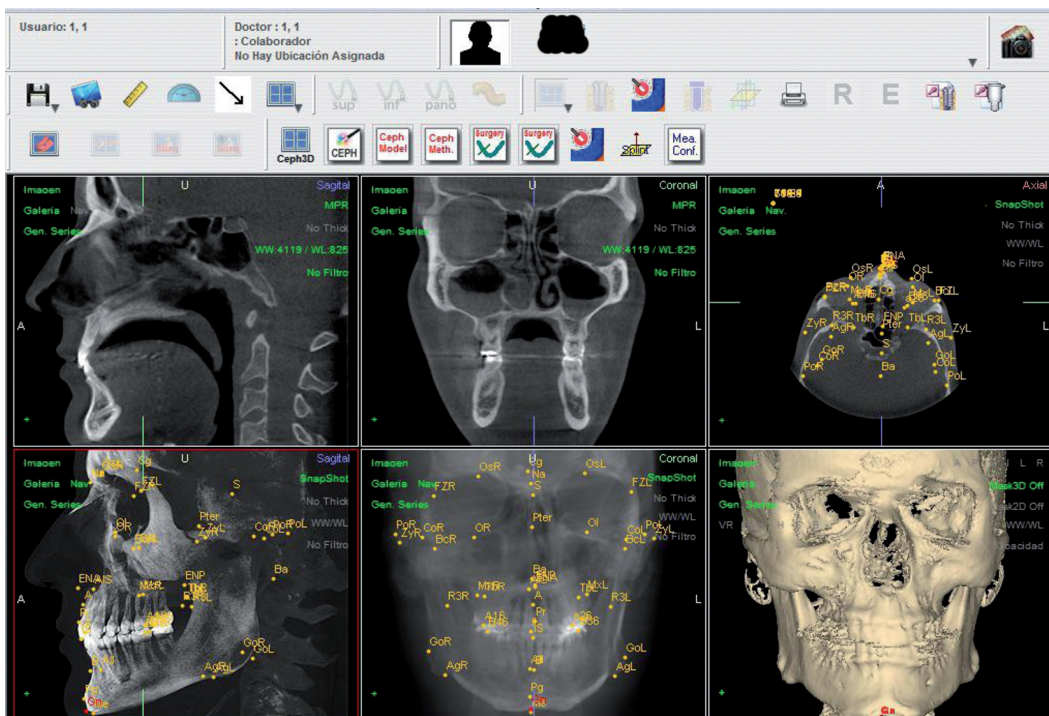


**Fig. 1.** Sagittal, coronal, axial and 3D reconstruction windows.

appropriate plane was selected before fitting it onto the other planes for greater accuracy.

To evaluate reproducibility and intra- and inter-observer error, two (2) previously trained and qualified observers in the location of cephalometric landmarks, both with six years of experience/background in orthodontics, repeated the measurements on three occasions at intervals of one week. 3.690 measurements were taken and, as each determination has 3 coordinates, 11.070 data were processed with SPSS® statistical package 17.0 for Windows (IBM Corporation, Sommers, NY).

errors were found for each landmark. A regression line was determined for comparing the values between Ob1 and Ob2, estimating the slope, ordinate at origin and the Pearson correlation coefficient. So that the measurements of both observers could be considered equivalent, the confidence interval of the slope had to contain the 1 and the confidence interval of the ordinate at origin, the 0.The difference between ratios was determined estimating the confidence interval of that difference. Differences were considered significant for $p < 0.05$ and confidence intervals were determined at 95%.

## Results

We analysed the 1.845 determinations made by each observer (Ob) considering the value of each axis independently, so as to find out whether the reproducibility of locating the different landmarks was associated with one or other axis of coordinates.

*-Intra and inter-observer time variability*

Prior to calculating the reliability of each landmark, an evaluation of the raw data was performed. The global mean and the standard deviation of each observer (Ob1, 2) was calculated in Time (1, 2, 3) in the X,Y,Z planes. A statistically significant difference was found for Ob 1 for axis Y in t1 compared with t2 and t3 (p=0.006 between t1-t2; and p=0.008 between t1-t3). There was no statistically significant difference for Ob1 between t2-t3 and for Ob2 on any of the axes.

The ICC determined was ≥ 0.99 for all the axes in intra-observer measurements and ≥ 0.99 in inter-observer measurements, with the highest values corresponding to axis Z (ICC > 0.996).

*-Reliability and measuring method errors*

The value of the coordinate of each landmark has been represented in a single graphic. The number of values represented amounts to 5.535. We have represented t-1, t-2 and t-3 of Ob 1(abscissae) versus t-1, t-2 and t-3 of Ob 2 (ordinates) (Fig. 2).

The points on the bisection were taken into account for determining the method sensitivity error. The points outside the bisection corresponded to the landmarks with the highest errors between observers.

For each landmark 6 determinations were made. The observations that diverged from the bisection could be a consequence either for being one determination that diverges from the other five, or because, even though the intra-observer measurements were similar, there were differences between observers. The determination which deviated more than 10 mm from the mean and those where the mean of the values from the three determinations undertaken by Ob 1 diverged more than 10 mm from the mean of those of Ob 2, for the same point in the same patient, were considered as wrong determinations.

With these criteria there were 61 wrong determinations (Fig. 2). This means an error of 0.55% CI95% [0.54% - 0.56%]. Ob1 presented 0.45% CI95% [0.44% - 0.66%] and Ob 2, 0.65% CI95% [0.64% - 0.66%]. They did not have the same distribution on the three axes. 13 (0.11%) correspond at X, 38 (0.34%) at Y and 10 (0.09%) at Z (Table 2). In evaluating the data, the decision was made to remove from the data set any identified landmarks that were technical errors. We decided to do it in that way because large outliers were due to technical errors in the use of
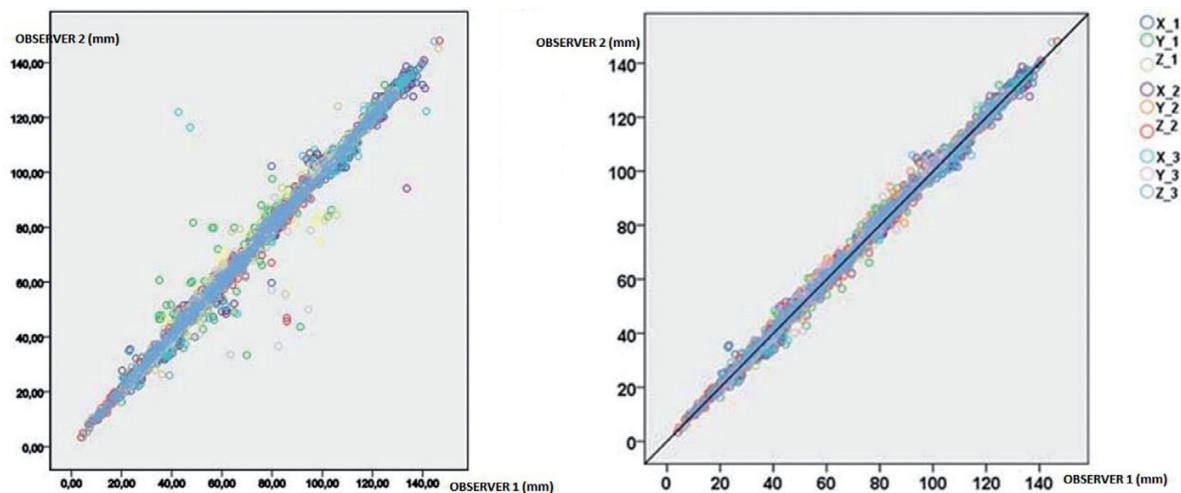


**Fig. 2.** Single graphic with Values (mm) of the 1845 determinations with standard deviation (SD) of less than 0.5 mm and maximum of 2 mm, undertaken by Observer 1 versus the corresponding 1845 determinations undertaken by Observer 2, for each axis.

**Table 2.** Landmarks with high error. Each anatomical point has been measured 90 times, as the 3 measurements have been undertaken by each observer on 15 patients. Dark and light grey represent the landmarks with the highest error.

| Landmark | Nº errors | % of error |
|----------|-----------|------------|
| A26 | 1 | 1.1% |
| RL | 1 | 1.1% |
| PoR | 1 | 1.1% |
| SorR | 9 | 10.0% |
| A16 | 1 | 1.1% |
| GoL | 2 | 2.2% |
| ZyR | 7 | 7.8% |
| ZyL | 1 | 1.1% |
| OrL | 1 | 1.1% |
| B | 2 | 2.2% |
| Cg | 2 | 2.2% |
| AgL | 3 | 3.3% |
| TbL | 1 | 1.1% |
| MxL | 1 | 1.1% |
| Pns | 6 | 6.7% |
| GoR | 3 | 3.3% |
| TbR | 1 | 1.1% |
| OrR | 3 | 3.3% |
| RR | 2 | 2.2% |
| SOrL | 11 | 12.2% |
| MxR | 2 | 2.2% |

the software and not due to the misunderstanding of the landmark definition or inability to locate the landmark in the 3-D image. With those criteria, the 61 wrong determinations were eliminated, in order to determine the method sensitivity error. The adjustment line of these points has a Pearson correlation coefficient of 0.998 with a slope of 0.999 CI95% [0.997, 1.001] and an ordinate at origin of 0.007 CI95% [-0.124, 0.137], showing that the measurements of both observers are perfectly comparable.

Standard deviation (SD) of each coordinate (X, Y, Z) was analysed for each landmark measured on each patient. According to these results, accuracy was quite similar on all 3 axes (Table 3).

The average SD of all the landmarks was 1.0 mm, which corresponds to an average relative error of 1.3%. 74% of the measurements on X axis, 76.5% on Y axis and 69.7% on Z axis were below 1.5% of error. The 95% CIs showed only a statistically significant difference between the percentage of Y axis [74.1%, 78.8%] and that of X axis [67.1%, 72.2%].

The average SD that corresponded to each anatomical landmark was also determined and classified by areas of interest (Table 4).

**Table 3.** Maximum standard deviations (SD) and percentage (%) of determinations with standard deviations (SD) of less than 0.5 mm and 2 mm, for each observer (Ob1, Ob2) and each coordinate (X, Y, Z).

| Axis | Percentage (%) of determinations with SD<=0.5mm | | Percentage (%) of determinations with SD<=2mm | | Maximum SD | |
|------|------|------|------|------|------|------|
| | Ob1 | Ob2 | Ob1 | Ob2 | Ob1 | Ob2 |
| X | 50.2% | 46.1% | 95.3% | 93.3% | 4.5mm | 6.3mm |
| Y | 59.8% | 57.7% | 93.3% | 95.4% | 5.4mm | 5.7mm |
| Z | 58.5% | 52.4% | 95.3% | 94.1% | 4.1mm | 7.1mm |

**Table 4.** Average standard deviation (SD) in milimiters (mm) corresponding to each anatomic landmark. Those represented in shading had SD on all axes (X, Y, Z) below the average value (equal to 1.0 mm) and present less measurement error.

|  | Landmark | SD_X (mm) | SD_Y (mm) | SD_Z (mm) |
|---|---|---|---|---|
| **Cranial region** | **Na** | 0.36 | 0.32 | 0.49 |
|  | **Ba** | 0.79 | 0.51 | 0.49 |
|  | **S** | 0.89 | 0.41 | 0.61 |
|  | **Cg** | 0.36 | 1.24 | 0.89 |
|  | **PoR** | 0.76 | 0.38 | 0.28 |
|  | **PoL** | 0.9 | 0.46 | 0.37 |
| **Orbital-zygomatic region** | **ZyR** | 0.83 | 1.72 | 0.82 |
|  | **ZyL** | 0.7 | 1.38 | 0.78 |
|  | **FzR** | 0.28 | 0.41 | 0.37 |
|  | **FzL** | 0.32 | 0.57 | 0.38 |
|  | **BcR** | 1.08 | 1.07 | 0.83 |
|  | **BcL** | 1.03 | 1.25 | 0.83 |
|  | **OrR** | 1.35 | 0.65 | 0.28 |
|  | **OrL** | 1.15 | 0.59 | 0.29 |
|  | **SOrR** | 1.34 | 0.59 | 0.43 |
|  | **SOrL** | 1.16 | 0.74 | 0.45 |
| **Maxilar region** | **Ans** | 0.96 | 0.66 | 0.3 |
|  | **Pns** | 1.21 | 1.17 | 0.5 |
|  | **A** | 0.86 | 0.41 | 0.93 |
|  | **TbR** | 1.27 | 0.82 | 1.52 |
|  | **TbL** | 1.48 | 0.85 | 1.73 |
|  | **MxR** | 0.54 | 1.04 | 1.07 |
|  | **MxL** | 0.59 | 0.88 | 1.22 |
|  | **A16** | 0.63 | 0.57 | 0.81 |
|  | **A26** | 0.47 | 0.39 | 0.68 |
|  | **UIR** | 0.86 | 0.31 | 0.28 |
| **Mandibular region** | **Pg** | 0.16 | 0.23 | 0.67 |
|  | **Gn** | 0.15 | 0.44 | 0.45 |
|  | **Me** | 0.5 | 0.63 | 0.24 |
|  | **B** | 0.78 | 0.23 | 0.79 |
|  | **GoR** | 0.36 | 0.85 | 1.02 |
|  | **GoL** | 0.48 | 0.97 | 1.07 |
|  | **AgR** | 0.67 | 1.55 | 0.57 |
|  | **AgL** | 0.73 | 1.28 | 0.57 |
|  | **RR** | 0.2 | 0.13 | 0.69 |
|  | **RL** | 0.33 | 0.15 | 0.85 |
|  | **B36** | 0.55 | 0.52 | 0.83 |
|  | **B46** | 0.54 | 0.37 | 0.94 |
|  | **LIR** | 0.75 | 0.31 | 0.31 |
|  | **CdR** | 1.09 | 0.67 | 0.84 |
|  | **CdL** | 1.32 | 0.59 | 1.04 |

## Discussion

In the present study the reproducibility in the location of the different landmarks commonly used for orthodontic diagnosis when making a cephalometric tracing has been determined. Of all the studies reviewed, our work is the one with more CBCTs and cephalometric points used. In order to carry out this study, records of patients who had already undergone a CBCT for various reasons were used. Other studies use 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, it is necessary to undertake them with all the tissues included, in order to be able to check more accurately and to simulate the clinical situation(15,23). 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 fifteen patients who had required CBCTs for surgical reasons or for the presence of impacted teeth.

In this study each of the 41 landmarks was also described and defined for each of the coordinates of the space (X, Y, Z). As recommended by Oliveira et al. (20), this definition in the three spatial planes facilitates the true landmark position. The problem of some studies is that they do not give a detailed definition of the landmarks in each of the spatial planes as we do (15,16). All the landmarks of this study belonged to anatomical structures. In CT and CBCT, superimpositions of structures are not generated and, therefore, the points based on them have little interest in this type of records (6).

In the present study, two observers determined in three occasions the location of 41 landmarks and 11.070 data were processed unlike the study of Schlicher et al.(16) where nine examiners made the measurements but only in one occasion (5472 data), analyzing only half of our measurements. In the overall intra-observer time variability a statistically significant difference was only found for Ob 1 for axis Y in t1. This result can be interpreted both as a learning process that takes place in the location of the landmarks in this axis in the last two measurements, or as a fortuitous situation associated with the measuring process. As in our study, Park et al.(9) and Kumar et al.(24) did not find inter-examiner statistically significant differences between the 19 landmarks or the 14 landmarks measured, respectively. Schlicher et al.(16) did not find either significant improvement in the identification of the 32 landmarks measured. The results coincide with ours but with a smaller number of points to locate (19, 14 and 32 respectively compared with the 41 points of our work). Although intraobserver landmark identification errors are generally lower than interobserver errors in Lagravère´s study (15) and ours not completely coincide since they showed that intra-

and interobserver ICCs were similar. The high intra-observer and inter-observer ICCs found indicate high degree of reliability in locating landmarks, following the observers' prior learning process. Our results coincide with those of Oliveira et al. (20) who found an ICC ≥0.9 for 85% of the intra-observer measurements and for 65.5% of the inter-observer measurements. They obtained the best ICC, as we did, on the Z axis (93.3%). In their study only 1% of the intra-observer and 3% of the inter-observer measurements had an ICC< 0.45. Our results agree partially with those of Lagravère et al.(22) which found an ICC > 0.97 for the intra-observer and an ICC> 0.92 for the inter-observer measurements; with the results of another study taken by the same authors (15) with an ICC>0.99 for the intra-observer and inter-observer measurements; and with those of Park et al.(9) who concluded that the reliability of all their landmarks was high, showing a good location on the three axes. As Schlicher et al. (16) calculated the average coordinates of each landmark (calculated for the two observers in the three occasions) but a gold standard was not included in the study.

In our study, of the 41 landmarks measured, half of them, 20 did not present errors in their determination in any of the 90 measurements taken (Na, S, Ba, PoL, A, Ans, Pg, Gn, Me, FzR, FzL, B36, B46 UIR, LIR, BcR, BcL, CdR, CdL, AgR). This suggests that these landmarks could be used safely when establishing their position. However, of the 21 remaining landmarks, SOrL, SOrR, ZyR and Pns, with more than 6 errors in their determination, would correspond with landmarks that should not be used as a basis for devising cephalometries or for which greater training would be necessary before using them in an analysis. As Schlicher et al. (16) stated, familiarity and anatomy could be responsible for the poorer performance in locating these landmarks. Those landmarks that presented the lower standard deviation (SD) on all the axes below the average value (1.0 mm) were Na, S, Ba, PoR, PoL, A, Ans, RR, RL, Pg, B, Gn, Me, FzR, FzL, A16, A26, B36, B46, UIR, LIR. If we consider the points which not only have not presented any error in their determination but have the least standard deviation (SD) value, then Na, S, Ba, PoL, A, Ans, FzR, FzL, Pg, Me, Gn, B36, B46, UIR, LIR appear to be the most reliable ones. These can be considered as reliable landmarks for being used in 3D cephalometric analyses.

If we analyse them per areas of interest, the cranial region showed the greatest reliability. Clinicians are accustomed to identify these landmarks in conventional cephalometry, especially Na, S and Ba, so this may explain their high reliability. High reliability of the frontozygomatic sutures (FzR, FzL) in the orbital-zygomatic region was also found. The landmarks with a greater margin of error were the supraorbitals (SOrR, SOrL).

This may be due to the difficulty of locating them on those CBCTs that do not include the complete cranium. In the maxillary region, the landmarks of greatest reliability were the first upper molars (A16, A26), the upper incisor (UIR), the anterior nasal spine (Ans) and point A. Landmarks such as the maxillaris (MxR, MxL) or the retromolar tuberosities (TbR, TbL) were those with the greatest margin of error. The difficulty in locating these landmarks may be caused by the lack of practice at identifying them, because they are not usually used in conventional lateral cephalometrics. In the mandibular region, Gnathion (Gn), Menton (Me), Pogonion (Pg) and point B obtained the highest reliability. However, the condyles (CdR, CdL), gonions (GoR, GoL) or antegonials (AgR, AgL) had lower reliability.

As in our study, Lagravère et al. (15) found, in general, mean differences between intra-observer measurements of less than 1mm. In their study, those landmarks that were between 1-2 mm of error were: OrL, S, Ba, Ans, Pns in the X plane, GoR, GoL, PoL and Pns in the Y plane and point B and LIR in the Z plane. On the other hand, the mean differences between intra-observer measurements were, in general higher than 1mm, this results being different to those found in our study. They found that the less reliable landmarks (with > 2mm of error) were: OrR, OrL, PoR, PoL, CdR, CdL in the X plane, GoR, GoL and Ans in the Y plane; and GoR, GoL and LIR in the Z plane. In our study we also found low reliability for Pns, OrR, OrL, CdR, CdL in the X plane, Pns in the Y plane and GoR, GoL in the Z plane (although in our results the SD of these landmarks is always < 2mm). Our results did not totally coincide with those of Oliveira et al. (20) In their study, the two landmarks that presented low reliability were: Y coordinates of the right and left mandibular ramus and Z coordinate of the right and left condylion. We found low reliability in the right condylion at the X coordinate. However, our results coincide with those of Muramatsu et al. (10) who evaluated the reliability of 19 landmarks in images taken with CT, observing that the Basion had the lowest confidence area ellipse on all planes, which indicated great reliability. Other authors (16) found S point to be the most reliable landmark, as we did, and PoR and OrR the most unreliable ones.

In general, in our study we found that landmark location using CBCT has less than 1.5% of error. This indicates high reliability in the location of all the landmarks and, therefore, is of great interest for clinical application in 3D cephalometric analyses. Nevertheless, we must not forget that introducing a certain cephalometric analysis cannot be based only on its reproducibility, but also on its clinical significance and other factors.

Current 3D records are very useful, especially to accurately locate bilateral and transversal points that are better observed from the coronal plane. In this way it is possible to study and analyse, in greater depth and reliability, dentofacial asymmetries, problems on the transversal level of the maxillary bone structures, as well as the size of the structures that are involved in the skull.

When analysing the results, greater reliability is to be expected on the sagittal plane as clinicians are trained in locating landmarks in conventional 2D cranial lateral teleradiography. However, with a properly calibration and training, observers can locate, with high reliability, the landmarks on each of the three planes of the space. In general, landmark location in 3D requires greater time than in 2D; firstly because prior training is necessary to familiarize with the different slices presented, and secondly because a first location on one of the planes and then a plotting on the other two is required for accuracy (12). As this location differs from the 2D conventional radiography location and as many of our patients are asked to have a CBCT as a diagnostic tool, it has been established that with correct instruction and with adequate learning, clinicians should locate properly these landmarks directly on CBCTs, obtaining much more diagnostic information about the different relationships between the different craniofacial structures of the patient.

One of the new things presented in this paper we have not seen in others is that based on the reliability of the landmarks studied, we have also defined three planes of reference. In line with this criterion of greater reliability, various anatomical landmarks could be chosen as a basis for defining these three planes of reference on the three spatial axes (Fig. 3):

• Mid-sagittal plane (XZ): anterior-posterior vertical plane that divides the body in two halves (right and left portions), defined by the points Na, S, Ba, points of high reliability and of easy location.

• Horizontal or transverse plane (XY): Horizontal plane perpendicular to the median sagittal plane that divides the body into upper and lower halves. We can obtain it and make it pass through each point that we choose, for example through Na.

• Coronal or frontal plane (YZ): vertical plane perpendicular to the two previous ones that goes from one side of the body to another dividing it into two parts (anterior and posterior). It cuts through the mid-sagittal plane in the middle.
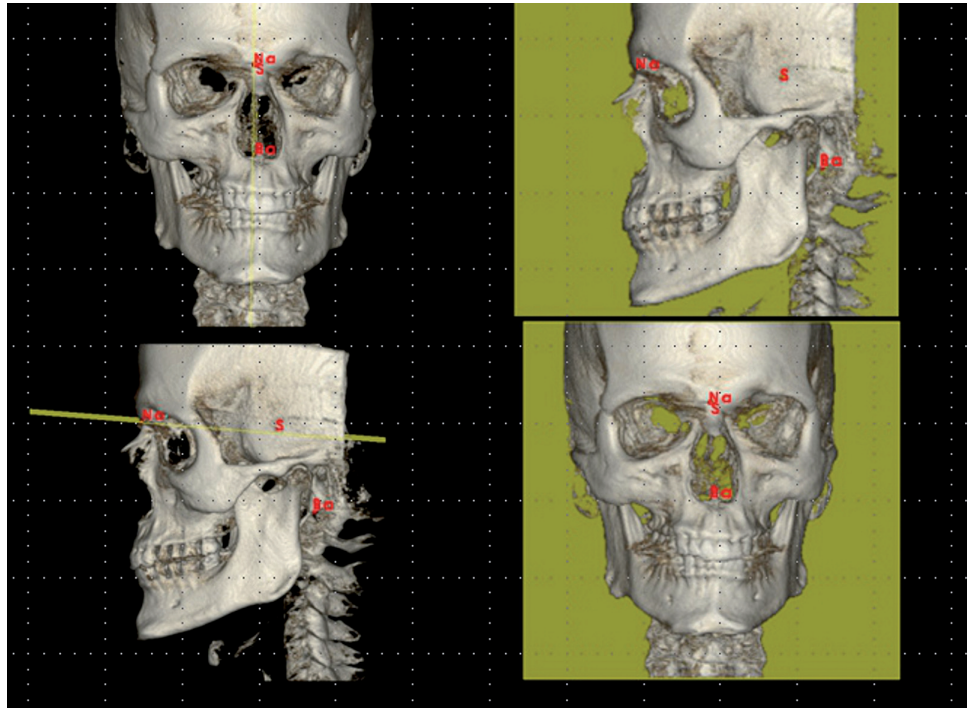
**Fig. 3.** Planes of reference: Mid-sagittal plane (XZ): defined by the points Na, S, Ba; Horizontal or transversal plane (XY); Coronal or frontal plane (YZ).

## References

1. Mozzo P, Procacci C, Tacconi A, Martini PT, Andreis IA. A new volumetric CT machine for dental imaging based on the cone-beam technique: preliminary results. Eur Radiol. 1998;8:1558-64.

2. Quintero JC, Trosien A, Hatcher D, Kapila S. Craniofacial imaging in orthodontics: Historical perspective, current status, and future developements. Angle Orthod 1999;69:491-506.

3. Silva MA, Wolf U, Heinicke F, Bumann A, Visser H, Hirsch E. Cone-beam computed tomography for routine orthodontic treatment planning: a radiation dose evaluation. Am J Orthod Dentofacial Orthop. 2008;133:640.e1-5.

4. Chau AC, Fung K. Comparison of radiation dose for implant imaging using conventional spiral tomography, computed tomography, and cone-beam computed tomography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2009;107:559-65.

5. De Vos W, Casselman J, Swennen GR. Cone-beam computerized tomography (CBCT) imaging of the oral and maxillofacial region: a systematic review of the literature. Int J Oral Maxillofac Surg. 2009;38:609-25.

6. Cattaneo PM, Melsen B. The use of cone-beam computed tomography in an orthodontic department in between research and daily clinic. World J Orthod. 2008;9:269-82.

7. Lou L, Lagravère MO, Compton S, Major PW, Flores-Mir C. Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2007;104:402-11.

8. Katsumata A, Fujishita M, Maeda M, Ariji Y, Ariji E, Langlais RP. 3D-CT evaluation of facial asymmetry. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2005;99:212-20.

9. Park SH, Yu HS, Kim KD, Lee KJ, Baik HS. A proposal for a new analysis of craniofacial morphology by 3-dimensional computed tomography. Am J Orthod Dentofacial Orthop. 2006;129:600.e23-34.

10. Muramatsu A, Nawa H, Kimura M, Yoshida K, Maeda M, Katsumata A et al. Reproducibility of maxillofacial anatomic landmarks on 3-dimensional computed tomographic images determined with

the 95% confidence ellipse metod. Angle Orthod. 2008;78:396-402.

11. Pinsky HM, Dyda S, Pinsky RW, Misch KA, Sarment DP. Accuracy of three-dimensional measurements using cone-beam CT. Dentomaxillofac Radio.l 2006;35:410-6.

12. Mischkowski RA, Pulsfort R, Ritter L, Neugebauer J, Brochhagen HG, Keeve E et al. Geometric accuracy of a newly developed cone-beam device for maxillofacial imaging. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2007;104:551-9.

13. Moerenhout BA, Gelaude F, Swennen GR, Casselman JW, Van Der Sloten J, Mommaerts MY. Accuracy and repeatability of cone-beam computed tomography (CBCT) measurements used in the determination of facial indices in the laboratory setup. J Craniomaxillofac Surg. 2009;37:18-23.

14. Lagravère MO, Carey J, Toogood RW, Major PW. Three-dimensional accuracy of measurements made with software on cone-beam computed tomography images. Am J Orthod Dentofacial Orthop. 2008;134:112-6.

15. Lagravère MO, Low C, Flores-Mir C, Chung R, Carey JP, Heo G et al. Intraexaminer and interexaminer reliabilities of landmark identification on digitized lateral cephalograms and formatted 3-dimensional cone-beam computerized tomography images. Am J Orthod Dentofacial Orthop. 2010;137:598-604.

16. Schlicher W, Nielsen I, Huang JC, Maki K, Hatcher DC, Miller AJ. Consistency and precision of landmark identification in three-dimensional cone beam computed tomography scans. Eur J Orthod. 2012;34:263-75.

17. Delamare EL, Liedke GS, Vizzotto MB, da Silveira HL, Ribeiro JL, Silveira HE. Influence of a programme of professional calibration in the variability of landmark identification using cone beam computed tomography-synthesized and conventional radiographic cephalograms. Dentomaxillofac Radiol. 2010;39:414-23.

18. Stratemann SA, Huang JC, Maki K, Miller AJ, Hatcher DC. Comparison of cone beam computed tomography imaging with physical measures. Dentomaxillofac Radiol. 2008;37:80-93.

19. Periago DR, Scarfe WC, Moshiri M, Scheetz JP, Silveira AM, Farman AG. Linear accuracy and reliability of cone beam CT de-

rived 3-dimensional images constructed using an orthodontic volumetric rendering program. Angle Orthod. 2008;78:387-95.

20. De Oliveira AE, Cevidanes LH, Phillips C, Motta A, Burke B, Tyndall D. Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2009; 107:256-65.

21. Brown AA, Scarfe WC, Scheetz JP, Silveira AM, Farman AG. Linear accuracy of cone beam CT derived 3D images. Angle Orthod. 2009;79:150–7.

22. Lagravère MO, Gordon JM, Guedes IH, Flores-Mir C, Carey JP, Heo G, Major PW.Reliability of traditional cephalometric landmarks as seen in three-dimensional analysis in maxillary expansion treatments. Angle Orthod. 2009;79:1047-56.

23. Van Vlijmen OJ, Maal T, Bergé SJ, Bronkhorst EM, Katsaros C, Kuijpers-Jagtman AM. A comparison between 2D and 3D cephalometry on CBCT scans of human skulls. Int J Oral Maxillofac Surg. 2010;39:156-60.

24. Kumar V, Ludlow JB, Mol A, Cevidanes L. Comparison of conventional and cone beam CT synthesized cephalograms. Dentomaxillofac Radiol. 2007;36:263-9.