European Journal of Orthodontics 34 (2012) 276–286
doi:10.1093/ejo/cjq190
Advance Access Publication 12 may 2011

Repeatability and reproducibility of landmarks—a three-dimensional computed tomography study

Irem Titiz*, Michala Laubinger*, Thomas Keller**, Klaus Hertrich* and Ursula Hirschfelder*

*Department of Orthodontics, University Dental School of Erlangen-Nuremberg, Erlangen, Germany and **ACOMED statistik, Leipzig, Germany

Correspondence to: Prof. Dr Ursula Hirschfelder, Universität Erlangen-Nürnberg, Zahnklinik 3, Kieferorthopädie, Glückstraße 11, 91054 Erlangen, Germany. E-mail: sekretariat.prof.hirschfelder@uk-erlangen.de

SUMMARY The aim of this study was to investigate the repeatability and reproducibility of the placement of anthropological cephalometric landmarks on three-dimensional computed tomography (3D CT) cranial reconstructions derived from volume data sets. In addition, the influence of the observer's experience on the repeatability of landmark setting was also explored.

Twenty patients without any craniofacial deformity (11 females and 9 males; age range 6.1–16 years) were selected retrospectively from CT volume data sets already available from 695 patients of Dental Clinic 3, Orthodontics of Erlangen University Hospital. The CT examination was performed with the SOMATOM Sensation®64 (Siemens AG Medical Solutions, Erlangen, Germany). The program VoXim®6.1 (IVS Solutions AG, Chemnitz, Germany) was used for 3D reconstruction of the volume data sets. A total of 28 landmarks were examined in the skeleton module of the program VoXim®6.1. The randomly sorted data sets were analysed by two orthodontists and two postgraduate students. Each data set was analysed twice by each observer at an interval of 3 weeks. The analysis of variance regarding random effects was used to calculate the intraobserver and interobserver components of standard deviation (SD) of depiction of individual landmarks as measures of repeatability and reproducibility, respectively. Median intraserial SD and interserial SD of 0.46 mm (range: 0.14–2.00 mm) and 0.20 mm (range: 0.02–2.47 mm), respectively, were obtained depending on the landmark and plane. This study included systematic analysis of extreme values (outliers) in the assessment of the quality of measurements obtained.

Descriptive statistics revealed qualitative differences in the depiction of different landmarks. The landmarks nasion and infradentale revealed a minor SD in all three spatial coordinates with the smallest SD for infradentale (SD = 0.18 mm) in the transverse plane. However, no systematic trend was identified with regard to the influence of the observer's level of experience affecting the repeatability of landmark positioning. Thus, the repeatability and reproducibility of placements of landmarks with 3D CT were found to be acceptable for a majority of anatomical positions.

Introduction

Orthodontic diagnostics are based on clinical findings, with substantial support from imaging techniques [conventional two-dimensional (2D) X-ray images]. Orthodontics often deals with complex issues, such as craniofacial deformities in children and adults. As a result, modern imaging techniques, such as computed tomography (CT) and cone beam computerized technology (CBCT) with three-dimensional (3D) representation of the skull and teeth, play an increasingly important role in the diagnostic investigation of morphological problems and interdisciplinary treatment planning (Hirschfelder *et al.*, 2004; Kitai *et al.*, 2004; Katsumata *et al.*, 2005; Krarup *et al.*, 2005; Maeda *et al.*, 2006; Rouas *et al.*, 2007; Periago *et al.*, 2008; De Vos *et al.*, 2009; Swennen *et al.*, 2009). Chan *et al.* (2007) reviewed historical and contemporary literature and concluded that there are many potential uses for 3D CT craniofacial imaging in clinical practice. From a comparative study of 3D CT and conventional X-rays, the study group led by Kragskov *et al.* (1997) concluded that the indication for this diagnostic imaging technique would mainly be cases with pronounced craniofacial asymmetries, owing to the difficulty of landmark setting on 3D CT reconstructions.

The reproducibility of such landmarks with regard to the intraobserver and interobserver error for 2D lateral cephalograms was examined in detail as early as in 1971 by Baumrind and Frantz (1971). To date, only few comparable studies have analysed volume images, such as spiral CT. Cavalcanti *et al.* (2004) and his team studied the precision of anthropometric measurements on 13 cadaver heads using 3D CT. In a further examination, the accuracy and precision of skull landmarks were recorded *ex vivo* using a low-dose CT protocol by comparing them with a high-dose CT protocol as a reference (Connor *et al.*, 2007).

It has been established scientifically that the results of the analyses of cephalograms are strongly influenced by the

quality of landmark location (Richardson, 1966; Phillips *et al.*, 1984). Thus, the precision of reference point placement on the spatial skull representation may possibly have a crucial effect on the 3D distance and angle measurements (Piechot, 2005).

The aim of this study was to analyse whether the repeatability and reproducibility of the placement of anatomical reference points in the 3D representation using skull spiral CT images are appropriate in terms of clinical application. Repeatability can be defined as the closeness of agreement between the depiction of landmarks obtained under stipulated conditions, which are repeated measurements of one reader in our case. In contrast, reproducibility describes the closeness of agreement between the results of depiction of landmarks carried out under changed conditions of measurement. In our case, changed conditions refer to different readers (ISO VIM; JCGM, 2008).

From a statistical point of view, repeatability and reproducibility can be calculated via variance components as estimated by analysis of variances (ANOVA) of data, whereby replications and readers are considered as random effects. In this study, the two types of standard deviations (SDs) owing to intraobserver and interobserver errors were obtained, which characterize the quality of measurement in terms of repeatability and reproducibility, respectively. This study directed special attention to the hypothesis that interobserver SD may be larger than intraobserver SD.

Materials and methods

Study population

Volume data sets of 20 patients (11 females and 9 males), which had been created for 695 patients in 2007 and 2008, were selected retrospectively from CT archive of Dental Clinic 3, Orthodontics of Erlangen University Hospital. Patients with clinically symmetrical facial bones were included in this study. In order to determine symmetry of patients, two Orthodontists (Director and Assistant Director of the Orthodontic Department) along with two postgraduate students met in a conference to establish consensus about the patients' symmetry. Six hundred and ninety-five patients were individually screened using CT images and assessed. Prior to assessment, CT images had been created for differential diagnostic investigation of serious dentition disorders or osseous structural anomalies. During the assessment conference, all patients were viewed by the observers mentioned above. Only patients that were judged as symmetrical (with the overall facial skeletal structure as the criterion of symmetry) by all four observers were included in the study.

In this study, the criterion for inclusion was a sufficient examination volume, recorded earlier using CT. Based on an appraisal of a number of cases (n = 20), patients with clinically symmetrical facial bones were selected. The CT data volume sets of all 20 patients were consecutively extracted. The patients' age at the time of CT examination ranged from 6.1 to 16 years with a mean age of 12.13 years. However, the CT data sets of patients with cleft lip and palate or syndromes involving craniofacial deformities were excluded.

Method of investigation

CT examination was performed with the SOMATOM Sensation64® (Siemens AG Medical Solutions, Erlangen, Germany) at the Radiology Department of Erlangen University Hospital, using spiral CT in low-dose mode in accordance with standardized conditions (Table 1; Hirschfelder *et al.*, 2004). For further reduction of radiation exposure, 'CARE Dose4D' program was used standardwise. Exposure to radiation ranged between 0.9 and 1.5 mSv.

Based on the raw data from the CT images, axial sectional images were initially reconstructed at 0.6 mm intervals by choosing the centre and width (C/W: 700/4000) of the window and were stored in Dicom format. This image data material was subsequently used to generate 3D craniofacial reconstructions in the program VoXim®6.1 (IVS-Solutions AG, Chemnitz, Germany).

Upon importing the data into the VoXim®6.1 program, the default grey-scale gradation offered by the program was retained. It ranged between −1024 and 1942 HU. The grey-scale gradation for each data set analysed was optimized in the working window of the skeleton module used to ensure the best possible depiction of the required skeletal structures. This setting was kept identical for all the four observers to ensure identical analysis conditions. A total of 28 landmarks were set in the skeleton module of the VoXim®6.1 program by means of the virtual 3D representation of the skull (Table 2).

Placement of the 28 landmarks was carried out by four observers: two specialist orthodontists and two postgraduate students of orthodontics. The four observers assessed the sorted (according an internal randomization list provided by the biostatistician) and masked data sets of 20 patients, twice each, on two examination days at an interval of 3 weeks.

This procedure minimized the possible learning effects. For each patient, there were two measurement templates per

Table 1 Imaging record of a patient (CTDIvol, volume-related computer tomography dose index; DLP, dose length product; TI, time; cSL, slice thickness).

	Total DLP						
Total mAs 562	75						
07 July 2008							
Patient position	Scan	kV	mAs/ref.	CTDIvol	DLP	TI	cSL
H–SP							
Topogram	1	80				2.8	0.6
Mean total KC	2	100	47/80	4.51	75	1	0.6

Table 2 Definition of landmarks and severity of landmark setting in three directions.

			Severity of landmark setting		
	Landmark (abbreviation)	Definition	*x*	*y*	*z*
1	S (sella)	Centre of sella turcica	Difficult	Difficult	Difficult
2	N (nasion)	Most anterior point of the frontonasal suture	Difficult	Simple	Simple
3	D (dens)	Most superior point of odontoid process of the epistropheus	Simple	Simple	Simple
4	Or L (orbitale on the left side)	Midpoint of the left infraorbital margin	Difficult	Simple	Simple
5	Sp a (anterior nasal spine)	Tip of the anterior nasal spine	Simple	Simple	Simple
6	Pr (prosthion)	Crest of the alveolar ridge between the upper central incisors	Simple	Simple	Simple
7	Id (infradentale)	Crest of the alveolar ridge between the lower central incisors	Simple	Simple	Simple
8	Gn (gnathion)	Lower border of the mid-mandibular suture	Difficult	Difficult	Difficult
9	Por L (porion on the left side)	Superior surface of the left external auditory meatus on the left side	Difficult	Difficult	Simple
10	Co sup L (condylion superius left)	Most superior point of the left condyle	Difficult	Difficult	Simple
11	Pc L (processus coronoideus left)	Most superior point of the left coronoid process	Simple	Simple	Simple
12	Go L (gonion left)	Most inferior and posterior point at the left angle of the mandible	Difficult	Difficult	Difficult
13	U1M (OK-6 pulpa left)	Centre of the pulp cavity at the crown of the left upper first molar	Difficult	Difficult	Difficult
14	L1M (UK-6 pulpa left)	Centre of the pulp cavity at the crown of the left lower first molar	Difficult	Difficult	Difficult
15	Sp p (posterior nasal spine)	Tip of the posterior nasal spine	Simple	Simple	Simple
16	A (A-point)	Deepest point on the concave outline of the upper labial alveolar process extending from the anterior nasal spine to prosthion	Difficult	Simple	Difficult
17	B (B-point)	Deepest point on the osseous curvature between the crest of the alveolus and pogonion	Difficult	Simple	Difficult
18	Pog (pogonion)	Most anterior point on the mandibular symphysis	Difficult	Simple	Difficult
19	Co pos L (condylion posterior left)	Most posterior point of the left condyle	Difficult	Simple	Difficult
20	Co med L (condylion mediale left)	Most medial point of the left condyle	Simple	Difficult	Difficult
21	Co lat L (condylion laterale left)	Most lateral point of the left condyle	Simple	Difficult	Difficult
22	Ba (basion)	Lowest point on the anterior border of the foramen magnum	Difficult	Difficult	Simple
23	Cr sup (crista galli superior)	Most superior point of the crista galli	Simple	Simple	Simple
24	Cr (crista galli)	Most anterior point of the crista galli	Difficult	Simple	Difficult
25	Go ant L (gonion anterior left)	Most inferior point on the left lower border of the mandible behind the antigonial notch and in front of gonion	Difficult	Difficult	Simple
26	Go post L (Gonion posterior left)	Most prominent postero-superior point at the left angle of the mandible on the posterior ramus behind gonion	Difficult	Simple	Difficult
27	Inc pal L (incisura palatinal left)	Deepest point of the concave outline of the posterior border of the horizontal plate of the palatine bone	Difficult	Simple	Simple
28	Sem L (semilunare left)	Deepest point on the left sigmoid notch of the mandible	Simple	Difficult	Simple

observer, which were given an alphanumerical code to ensure accurate identification in the subsequent data analysis.

The landmarks were placed on the 3D reconstruction of the osseous cranium. Placement of each landmark was controlled in the axial sectional image as well as in the coronal and sagittal secondary reconstruction. The landmarks were corrected if necessary.

The image-defined system of data set coordinates of the VoXim®6.1 software is based on the voxel matrix of the CT data set (Figure 1). The *x*-axis and *y*-axis correspond to the reconstructed axial sectional images, and the *z*-axis is consistent with the direction of the accumulated reconstructed slices.

This coordinate system is dictated accordingly by the imaging technique. The origin of the system of data set coordinates is defined by the VoXim®6.1 program and is located in the top left-hand corner of the first image slice. The *x*–*y* direction corresponds to the transverse plane, and hence, the axial slicing. Furthermore, the *y*–*z* direction is directed along the sagittal plane, and the *x*–*z* direction corresponds to the coronal plane, thus running in the vertical direction (Figure 1).

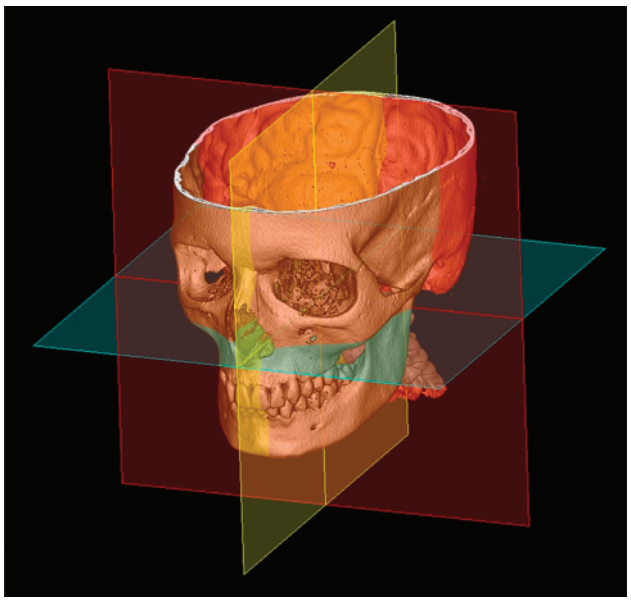


Figure 1 Diagram of the three-dimensional image-based computer tomography coordinate system in VoXim® 6.1.

Data processing and statistics

Intraobserver and interobserver variations in landmark placement were examined and plotted separately in *x*-, *y*-, and *z*-directions during the analysis. Four observers plotted 28 landmarks twice in 20 patients in three directions. This resulted in a total of 13 440 measured values. The *x*-, *y*-, and z-coordinates of the individual reference points in millimetres were first exported into the program Microsoft EXCEL® 2003 (Microsoft Deutschland GmbH, Unterschleißheim, Germany) to prepare and process the data material. The statistical analysis was performed using the SAS®9.2 program (SAS Institute Inc., Cary, North Carolina, USA).

For the purposes of descriptive statistics, individual means and SD were initially calculated for each landmark placed as a measure of the dispersion of results in the direction of *x*, *y*, and *z* axis, as well as to explore the range of values. In the second step, normalization of values was applied to counterbalance individual differences in landmark coordinates for each patient: for each subgroup defined by patient, landmark, and axis, the values were corrected by their mean, resulting in distributions with a mean of 0.

However, normal distribution of the measured results for landmark placement could not be confirmed because of a few extreme values (outliers). Hence, outliers were not included in the analysis of the intraobserver and interobserver errors (see below), if they were outside of the 4 SD range (four times SD). However, extreme values were included in the analysis by reporting the number of those events per measure and direction.

After this normalization and exclusion of outliers, dispersion of the measured values within and between the four observers was analysed through variance components as investigated by ANOVA. Thus, the intraobserver and interobserver errors were estimated in a model II ANOVA by considering repetitions and readers as random effects (random-effects model; Rasch *et al.*, 1996; Bland, 2000).

The resulting mean square errors were used to calculate the intraserial and interserial SD (SD_{intra} and SD_{inter}, respectively) as well as the total SD, $SD_{total} = (SD^2_{intra} + SD^2_{inter})^{0.5}$.

In addition, the analysis was performed separately for two subgroups of readers as defined by their experience (postgraduate students versus specialists in orthodontics). The SD_{intra} values were compared using *F*-test, whereby the hypothesis of lower SD_{intra} for more experienced readers was tested. Owing to the parallel analysis of 28 landmarks, the *P* values had to be corrected for multiple testing (Bonferroni correction: α level of 0.05/28 = 0.0018 was used in this analysis).

In addition, the influence of the severity of landmark setting was also examined. In terms of a *post hoc* analysis, the landmarks were classified into two groups based on the difficulties of setting in each direction (Table 2, simple: $N = 41$ versus difficult: $N = 43$), and Mann–Whitney *U*-test was applied to compare the medians.

Results

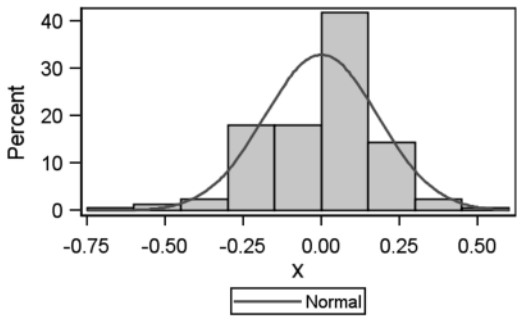
Different dispersions were found depending on the landmark and axis (*x*, *y*, and *z*). Table 3 summarizes the conventional descriptive statistics of distributions. Figure 2 shows the examples of histograms representing the distributions of measured values without and with the outliers obtained after normalization (details are described below).

The landmark- and axis-dependent results for intraserial, interserial, and total SD are presented in Table 4. As shown in Figure 3, the SD values follow skewed distributions with medians of 0.46 mm (intraserial SD), 0.20 mm (interserial SD), and 0.50 mm (total SD). About 50 per cent of the landmarks were set with a value of less than or equal to 0.50 mm of total SD, and 85 per cent of landmarks were measured with a total SD of less than or equal to 1 mm (Table 5, Figure 3). The landmarks, nasion and infradentale, displayed a small total SD in all three directions. The infradentale showed the smallest dispersion of the mean ($SD_{total} = 0.18$ mm) along the *x*-coordinate for all points (Figure 2, top). In contrast, a high SD was observed for the landmarks, gonion posterior left and incisura palatina left. The highest SD in the *z*-direction was for the landmark incisura palatina left with $SD_{total} = 3.33$ (Figure 2, bottom).

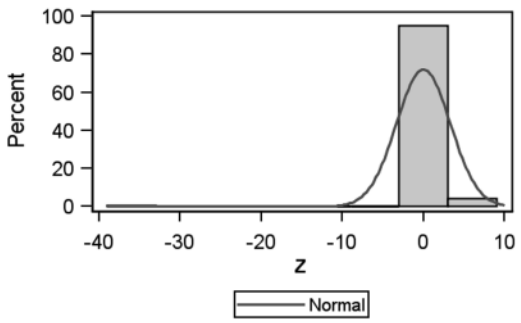
A detailed view on the repeatability and reproducibility of landmark placement in all the three spatial coordinates revealed that they are landmark dependent, and hence, dependent on the prevailing anatomical conditions, and direction dependent (*x*-, *y*-, and *z*-direction). For example, the landmark, incisura palatina left, exhibited high SD in

Table 3 Descriptive statistics of measured values used for landmark setting: corrected mean, standard deviation (SD), and range between minimum and maximum.

	x				*y*			*z*		
	Landmark	Mean (mm)	SD (mm)	Range (mm)	Mean (mm)	SD (mm)	Range (mm)	Mean (mm)	SD (mm)	Range (mm)
1	S	0.00	0.53	−1.55 to 1.72	0.00	0.44	−1.13 to 1.79	0.00	0.56	−2.26 to 1.64
2	N	0.00	0.34	−0.80 to 0.90	0.00	0.14	−0.40 to 0.52	0.00	0.49	−2.26 to 1.86
3	D	0.00	0.34	−1.45 to 0.75	0.00	0.47	−1.75 to 1.35	0.00	0.32	−0.75 to 1.91
4	Or L	0.00	3.27	−10.02 to 7.57	0.00	1.10	−2.65 to 3.95	0.00	0.47	−1.92 to 0.77
5	Sp a	0.00	0.42	−1.45 to 1.82	0.00	0.59	−1.49 to 2.67	0.00	0.33	−1.20 to 0.76
6	Pr	0.00	0.26	−0.98 to 1.39	0.00	0.65	−1.89 to 3.01	0.00	0.43	−1.95 to 1.81
7	Id	0.00	0.18	−0.69 to 0.51	0.00	0.26	−0.88 to 1.46	0.00	0.41	−1.29 to 1.16
8	Gn	0.00	0.56	−1.56 to 1.64	0.00	0.84	−3.34 to 1.66	0.00	0.26	−1.29 to 1.04
9	Por L	0.00	1.36	−3.40 to 4.16	0.00	0.76	−2.08 to 3.00	0.00	0.45	−1.55 to 1.04
10	Co sup L	0.00	1.02	−2.84 to 4.00	0.00	0.60	−1.58 to 1.98	0.00	0.36	−0.96 to 0.73
11	Pc L	0.00	0.38	−0.83 to 1.51	0.00	0.49	−1.76 to 1.49	0.00	0.30	−1.01 to 0.95
12	Go L	0.00	0.60	−1.15 to 2.34	0.00	0.95	−3.48 to 2.42	0.00	1.63	−5.50 to 3.80
13	U1M	0.00	0.48	−3.63 to 2.03	0.00	1.01	−7.13 to 6.12	0.00	1.61	−2.38 to 13.52
14	L1M	0.00	0.41	−0.77 to 3.25	0.00	0.97	−9.84 to 1.96	0.00	0.52	−2.30 to 2.80
15	Sp p	0.00	0.43	−1.86 to 1.04	0.00	0.59	−3.66 to 1.69	0.00	0.90	−8.42 to 1.88
16	A	0.00	0.70	−2.41 to 2.19	0.00	0.31	−0.88 to 1.22	0.00	0.74	−2.49 to 2.40
17	B	0.00	0.89	−3.75 to 3.79	0.00	0.51	−1.37 to 2.56	0.00	1.78	−4.70 to 4.72
18	Pog	0.00	0.51	−1.77 to 1.78	0.00	0.27	−0.73 to 1.49	0.00	0.91	−3.17 to 3.73
19	Co pos L	0.00	2.16	−8.76 to 2.64	0.00	1.05	−7.47 to 1.53	0.00	1.06	−5.88 to 2.14
20	Co med L	0.00	0.48	−4.80 to 1.00	0.00	0.69	−2.03 to 4.21	0.00	0.79	−2.69 to 1.69
21	Co lat L	0.00	0.28	−0.61 to 0.99	0.00	0.67	−1.89 to 2.90	0.00	0.94	−3.04 to 2.77
22	Ba	0.00	0.46	−1.58 to 1.56	0.00	0.40	−1.61to 0.92	0.00	0.48	−1.57 to 1.20
23	Cr sup	0.00	0.35	−2.04 to 1.86	0.00	0.97	−3.51 to 7.79	0.00	0.68	−3.45 to 6.15
24	Cr	0.00	0.64	−2.06 to 2.14	0.00	1.04	−3.74 to 5.20	0.00	1.48	−4.63 to 7.08
25	Go ant L	0.00	1.03	−2.39 to 6.36	0.00	2.08	−12.96 to 4.40	0.00	1.27	−4.61 to 4.69
26	Go post L	0.00	2.54	−4.63 to 28.98	0.00	1.16	−3.13 to 11.65	0.00	3.02	−6.34 to 8.57
27	Inc pal L	0.00	1.07	−2.54 to 4.39	0.00	0.84	−7.42 to 1.58	0.00	3.33	−38.84 to 6.06
28	Sem L	0.00	0.23	−1.59 to 0.71	0.00	0.80	−2.06 to 7.44	0.00	0.20	−0.79 to 0.66



Infradentale in x-direction,



Incisura palatina left in z-direction

Figure 2 Examples of distribution measured values: top: infradentale in *x*-direction and bottom: incisura palatina left in *z*-direction.

the *z*-direction owing to morphological criteria, whereas other points, such as orbitale left, showed a high dispersion in the *x*-direction (Table 4).

Repeatability: intraobserver landmark placement dispersion

The overall intraobserver dispersion proved to be small with its variability ranging from 0.14 to 3.32 mm (median: 0.52 mm). The landmarks, infradentale and prosthion, displayed small SD_{intra} (millimetres) for all three spatial coordinates, while gonion posterior left showed the highest variation in the *x*-direction (SD = 2.51 mm), and incisura palatina left in the *z*-direction (SD = 3.32 mm) when compared with other anatomical landmarks (Table 4).

Reproducibility: interobserver landmark placement dispersion

Comparison of the SD_{inter} [millimetre] values with those of SD_{intra} [millimetre] showed a similar small dispersion of the measured results (range: 0–2.54 mm, median = 0.22 mm). The landmarks, nasion and processus coronoideus, showed a very small SD in all the three dimensions. However, point orbitale left revealed high SD, especially in the *x*-direction (SD = 2.54 mm). All the landmarks and respective SD are presented in detail in Table 4.

Table 4 Intraobserver, interobserver, and total standard deviations (SDs) of all landmarks in the three spatial directions.

	x				*y*			*z*		
	Landmark	Intra SD (mm)	Inter SD (mm)	Total SD (mm)	Intra SD (mm)	Inter SD (mm)	Total SD (mm)	Intra SD (mm)	Inter SD (mm)	Total SD (mm)
1	S	0.47	0.22	0.53	0.42	0.14	0.44	0.53	0.19	0.56
2	N	0.33	0.04	0.34	0.14	0.02	0.14	0.46	0.17	0.49
3	D	0.31	0.13	0.34	0.45	0.15	0.47	0.32	0.03	0.32
4	Or L	2.05	2.54	3.27	0.64	0.89	1.10	0.28	0.37	0.47
5	Sp a	0.40	0.12	0.42	0.51	0.30	0.59	0.29	0.15	0.33
6	Pr	0.26	0.03	0.26	0.61	0.23	0.65	0.38	0.21	0.43
7	Id	0.16	0.08	0.18	0.24	0.90	0.26	0.30	0.27	0.41
8	Gn	0.51	0.24	0.56	0.66	0.52	0.84	0.24	0.11	0.26
9	Por L	1.07	0.84	1.36	0.64	0.42	0.76	0.33	0.30	0.44
10	Co sup L	0.80	0.64	1.02	0.50	0.32	0.60	0.31	0.18	0.36
11	Pc L	0.33	0.19	0.38	0.46	0.17	0.49	0.30	0.00	0.30
12	Go L	0.52	0.30	0.60	0.76	0.56	0.95	1.30	0.99	1.63
13	U1M	0.49	0.00	0.49	1.01	0.14	1.01	1.55	0.43	1.61
14	L1M	0.39	0.13	0.41	0.88	0.42	0.97	0.48	0.18	0.52
15	Sp p	0.42	0.10	0.43	0.59	0.08	0.59	0.89	0.13	0.90
16	A	0.64	0.31	0.70	0.29	0.09	0.31	0.70	0.24	0.47
17	B	0.76	0.47	0.89	0.43	0.28	0.51	1.10	1.40	1.78
18	Pog	0.50	0.10	0.51	0.27	0.02	0.27	0.90	0.14	0.91
19	Co pos L	1.37	1.67	2.15	0.85	0.62	1.05	0.97	0.44	1.06
20	Co med L	0.48	0.05	0.48	0.60	0.36	0.69	0.74	0.28	0.79
21	Co lat L	0.26	0.11	0.28	0.52	0.41	0.67	0.84	0.41	0.94
22	Ba	0.45	0.10	0.46	0.33	0.22	0.40	0.35	0.32	0.48
23	Cr sup	0.34	0.07	0.35	0.96	0.10	0.97	0.67	0.11	0.68
24	Cr	0.60	0.22	0.64	0.96	0.38	1.04	1.32	0.67	1.48
25	Go ant L	0.70	0.75	1.02	1.75	1.12	2.08	0.99	0.79	1.27
26	Go post L	2.51	0.41	2.54	1.34	0.24	1.16	1.66	2.52	3.02
27	Inc pal L	0.90	0.60	1.07	0.84	0.02	0.84	3.32	0.16	3.33
28	Sem L	0.20	0.10	0.23	0.78	0.17	0.80	0.18	0.08	0.20

Influence of the observer's level of experience on repeatability of landmark placement

Significant differences between the groups of specialists and postgraduate students were found for individual anatomical landmarks. For the following points in all the three spatial directions, the postgraduate students were found to place the landmarks with better repeatability than the two specialists: orbitale left (Or L) and centre of the coronal pulp of tooth 26 (U1M). In contrast, the landmarks, gnathion (Gn) and pogonion (Pog), were located with more repeatability by the specialists (Table 6). However, overall, no preferential group-specific differences in landmark placement were identified.

Influence of severity of landmark setting

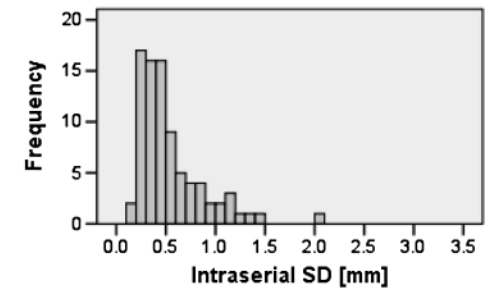
As expected, lower SD for landmarks classified as 'simple setting' were observed (intraserial SD: 0.43 versus 0.5 mm, $P = 0.08$; interserial SD: 0.19 versus 0.23 mm, $P = 0.36$; total SD: 0.46 versus 0.57 mm, $P = 0.15$), but the differences were not significant (Table 7). It should be noted that the relevance of these difference of distributions is important if their right tails are considered (maxima: intraserial SD: 1.23 versus 2.00 mm; interserial SD: 0.86 versus 2.47 mm; total SD: 1.34 versus 3.18 mm; Figure 4).

Analysis of outliers

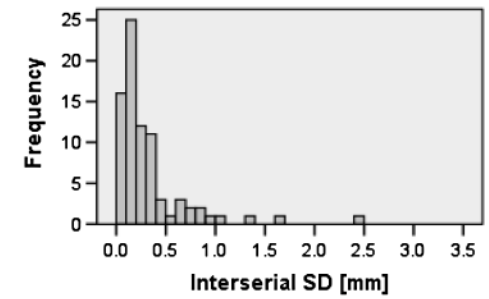
The number of outliers of each landmark is presented in Table 8. Table 9 shows that the number of outliers per rater varies. Among a total of 49 landmark settings with one or more outliers (1.1 per cent of 4480 landmark settings), 66 outliers occurred. For most of the landmarks, only few outliers were registered. However, for the landmarks, upper first molar and crista galli superior, more than five outliers were found within the 4480 measurements. In addition, some patients were more susceptible to outliers than others, as shown in Figure 5. A total of 224 settings per patient were performed, and fractions of landmark settings per patient distorted by outliers were found to range from 0 to 2.2 per cent (Figure 5). Furthermore, it was observed that the difficulty of setting does not influence the frequency of outliers (results not shown).

Discussion

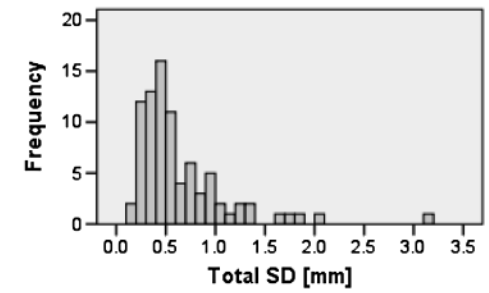
We investigated repeatability and reproducibility of the placement of anthropological cephalometric landmarks on 3D CT cranial reconstructions based on the observer's experience and severity of the landmark setting. In addition, we also carried out the analysis of extreme values.



Intraserial SD for 28 landmarks



interserial SD for 28 landmarks



total SD for 28 landmarks

Figure 3 Intraserial, interserial, and total standard deviation (SD) for 28 landmarks.

Table 5 Descriptive statistics for intraserial, interserial, and total standard deviation (SD) as found for 28 landmarks and three directions.

		SD_intra	SD_inter	SD_total
N				
	Valid	84	80	84
	Missing	0	4	0
Median		0.46	0.20	0.50
Minimum		0.14	0.02	0.14
Maximum		2.00	2.47	3.18
Percentiles				
	5	0.22	0.05	0.23
	10	0.25	0.08	0.27
	25	0.31	0.12	0.35
	50	0.46	0.20	0.50
	75	0.64	0.35	0.80
	90	1.08	0.77	1.22
	95	1.22	1.08	1.73

In our study, the SD values were used as measures of repeatability and reproducibility, whereas in other studies, the intrarater and interrater agreement was determined by using the intraclass correlation coefficient (ICC; Sayinsu *et al.*, 2007; Chien *et al.*, 2009). It should be noted that there is a direct relationship between these measurements because the ICC could directly be derived by $(SD_{inter}/SD_{total})^2$. However, the SD values expressed in millimetres describe the dispersion of the landmark setting in a direct way understandable by orthodontists. Using the properties of the underlying normal distribution of the measured values, about two-third of the values were found to be within the ±SD values, and approximately 95 per cent were observed to be within the range described by ±2 times SD.

In this study, the analysis of intraobserver and interobserver dispersion of the measured results verified the landmark- and axis-dependent variations of SD. We observed intraobserver and interobserver errors in a range of 0.14–2.00 mm and 0.02–2.47 mm (Table 5), respectively. From a clinical point of view, one can expect a method to measure the landmarks within a range of ±1 mm. Thus, by using the property of 2D range to contain approximately 95 per cent of all the measured values, a measurement with 1 mm repeatability and reproducibility, respectively, can be considered to be appropriate. The majority of the landmarks used in this study fulfil this criterion. Furthermore, the factors that might influence the repeatability and reproducibility are discussed below.

As the intraobserver and interobserver SD were of the same size, the 'inter SD more than intra SD' hypothesis could not be confirmed. A possible explanation for these similarly superior results is the precise definition of the landmarks beforehand (Table 2).

To our knowledge, this is the first report that has included systematic analysis of outliers in the assessment of the quality of measurements obtained. One has to note that the outliers occur to a relevant extent (1.1 per cent). It can be observed that the fraction of outliers depends on the raters, patients (or radiographs of the patients), as well as landmarks. As noticed earlier, some of the landmarks are prone to outliers. In particular, for the upper first molar and crista galli superior, more than 5 per 160 measurements per landmark were identified as outliers. We have referred to these outliers as errors in appropriately determining the landmarks if they have a high morphological variability. As the fractions of the outliers are not randomly distributed, it can be concluded that there is a considerable potential for lowering the extent of the outliers by quality control in terms of improved conditions with large fractions of outliers.

The results of this study are based on the analysis of consistent CT volume data sets. The high spatial resolution of less than 400 µm (Boldt *et al.*, 2009), as indicated by Siemens AG (Medical Solutions, Erlangen, Germany) for CT volume images [multispiral computer tomography (MSCT)], is likely to be an important factor in reliable

Table 6 Comparison of intraserial standard deviation (SD) of all landmarks in the three spatial directions in terms of the two observer groups of orthodontic specialists (OS) and postgraduate students (PG).

	x		*y*		*z*	
	Landmark	SD OS (mm)	SD PG (mm)	SD OS (mm)	SD PG (mm)	SD OS (mm)	SD PG (mm)
1	S	0.54	0.40	0.37	0.47	0.50	0.55
2	N	0.34	0.33	0.15	0.13	0.47	0.45
3	D	0.31	0.32	0.46	0.43	0.38	0.25
4	Or L	2.14	1.97	0.73	0.54	0.33	0.23
5	Sp a	0.36	0.43	0.55	0.47	0.30	0.29
6	Pr	0.27	0.25	0.71	0.50	0.38	0.38
7	Id	0.16	0.16	0.21	0.27	0.30	0.30
8	Gn	0.51	0.50	0.43	0.83	0.16	0.29
9	Por L	1.03	1.10	0.50	0.75	0.30	0.36
10	Co sup L	0.84	0.76	0.55	0.45	0.34	0.27
11	Pc L	0.35	0.31	0.42	0.50	0.32	0.28
12	Go L	0.41	0.61	0.70	0.82	1.19	1.40
13	U1M	0.60	0.33	1.07	0.94	1.98	0.92
14	L1M	0.30	0.45	0.38	1.18	0.35	0.59
15	Sp p	0.42	0.42	0.72	0.41	1.13	0.57
16	A	0.60	0.66	0.33	0.25	0.82	0.56
17	B	0.71	0.81	0.46	0.40	1.20	0.99
18	Pog	0.44	0.55	0.22	0.31	0.82	0.98
19	Co pos L	1.37	1.36	0.98	0.70	0.82	1.09
20	Co med L	0.62	0.29	0.63	0.55	0.64	0.82
21	Co lat L	0.25	0.27	0.44	0.60	0.71	0.95
22	Ba	0.42	0.48	0.31	0.34	0.39	0.32
23	Cr sup	0.44	0.20	1.26	0.52	0.93	0.21
24	Cr	0.62	0.58	0.98	0.95	1.26	1.38
25	Go ant L	0.83	0.53	1.98	1.48	1.13	0.82
26	Go post L	3.42	0.96	1.51	0.54	1.78	1.52
27	Inc pal L	0.73	1.02	0.95	0.70	4.52	1.28
28	Sem L	0.20	0.11	1.02	0.44	0.19	0.17

Table 7 Descriptive statistics of distribution of intraserial, interserial, and total standard deviation (SD) based on the severity of setting the landmark.

		SD_intra		SD_inter		SD_total	
	Classification	Simple	difficult	Simple	Difficult	Simple	Difficult
N							
	Valid	41	43	39	41	41	
	Missing	0	0	2	2	0	
Median		0.43	0.50	0.19	0.23	0.46	0.57
Minimum		0.14	0.22	0.02	0.03	0.14	0.23
Maximum		1.23	2.00	0.86	2.47	1.34	3.18
Percentiles							
	5	0.17	0.23	0.05	0.04	0.19	0.24
	10	0.24	0.25	0.08	0.07	0.24	0.27
	25	0.29	0.34	0.10	0.12	0.35	0.35
	50	0.43	0.50	0.19	0.23	0.46	0.57
	75	0.54	0.76	0.34	0.46	0.71	0.87
	90	0.90	1.12	0.62	1.07	1.04	1.71
	95	1.17	1.42	0.78	1.65	1.23	1.98

landmark placement. For CBCT, spatial resolution has been observed to lie in the submillimetre range (Brown *et al.*, 2009), implying that in principle, precise depiction of the osseous surface of craniofacial structures is provided for both the techniques. Furthermore, they allow for a realistic possibility of precise landmark placement, thus rendering a reliable 3D morphometric assessment of complex deformities in the maxillofacial area (Cavalcanti *et al.*, 2004; Katsumata *et al.*, 2005; Maeda *et al.*, 2006; Chan *et al.*, 2007; De Vos *et al.*, 2009

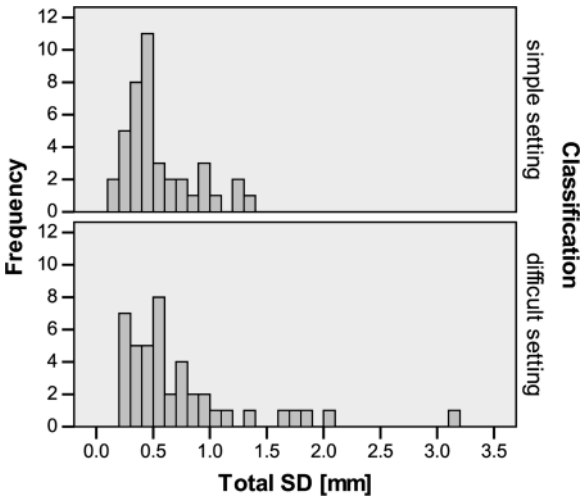


Figure 4 Histogram of total standard deviation (SD) in dependence on severity of landmark setting.

Table 8 Landmarks and number of outliers.

Outliers	Landmarks
0–2	S, N, D, Or L, Id, Gn, A, Pog, Co med L, Co lat L, Ba, Cr, Go ant L, Go post L, and Sem L
3–5	Sp a, Pr, L1M, Sp p, B, Co post L, and Inc pal L
>5	U1M and Cr sup

Table 9 Number of outliers per rater.

Rater	Number of settings with outliers	Percentage of 1120 measurements (per rater; %)
1	23	2.05
2	6	0.53
3	10	0.89
4	10	0.89
Total	49	1.09

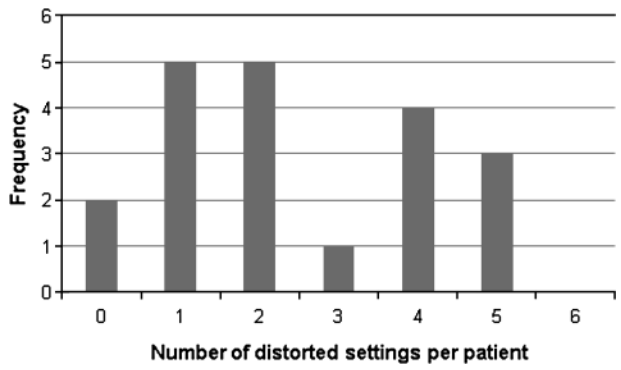


Figure 5 Number of settings distorted by outliers per patient.

Similar to the conventional 2D cephalometry based on volume data sets (Baumrind and Frantz, 1971; Jonas, 1976; Houston *et al.*, 1986; Greiner *et al.*, 2007; Moshiri *et al.*, 2007; Dvortsin *et al.*, 2008) and in view of the development of cephalometric 3D analyses (Park *et al.*, 2006), it is of utmost importance to determine the anthropometric repeatability of landmark placement in all the spatial planes (*x*-, *y*-, and *z*-direction). This was the central concern of the present study. Although numerous scientific studies regarding the accuracy of 3D measurements can be found in international publications (Kragskov *et al.*, 1997; Lagravere *et al.*, 2006; Connor *et al.*, 2007; Lou *et al.*, 2007; Brown *et al.*, 2009; Chien *et al.*, 2009; de Oliveira *et al.*, 2009; Hassan *et al.*, 2009), there are no study protocols directly comparable with our study design focusing on the repeatability and reproducibility of 3D anatomical reference point identification and placement. Chien *et al.* (2009) discovered an overall improvement in the intraobserver and interobserver reliability (which is similar to reproducibility in our study) for certain landmarks for CBCT, when compared with the 2D images. With regard to the computed tomographic quality of representation, Connor *et al.* (2007) discussed the importance of CT imaging protocol for precise landmark placement. They postulated that a higher millampere-second product could improve the precision of landmark identification and is therefore advantageous for clinical application. This aspect is certainly of high diagnostic relevance.

The dependency of repeatability and reproducibility on both the landmark and the direction has also been reported in other investigations. De Oliveira *et al.* (2009) reported that some landmarks are well identifiable in one or two planes but pose difficulties in the third plane. In addition, choosing the suitable plane for landmark placement requires time, experience, and a careful assessment of localization of the landmark. In a study to assess the lateral cephalograms reconstructed from computed tomograms, Kusnoto *et al.* (1999) found the smallest measuring error with linear measurements related to the sagittal direction (*y*-axis). Measurements in the transversal (*x*-axis) and vertical (*y*-axis) directions reportedly showed larger measuring errors. In a study by Toma *et al.* (2009), glabella and nasion showed poor reproducibility in both intra- and inter-examiner reproducibility assessment in the *y*-axis. The present study also confirmed that the reproducibility of some measuring points differed in relation to the spatial planes (sagittal, coronal, and transversal).

The imprecision of 3D measurements is found to be associated with the complexity of the localization and placement of anatomical landmarks. Similarly, in a study on landmark setting on plaster casts by means of 3D digitalization methods, Boldt *et al.* (2009) found that the repeatability and reproducibility of the measured points also depend on the complexity of the reference model surface.

Furthermore, Baumrind and Frantz (1971) described that the precision of landmark placement strongly depends on

the individual craniofacial morphology, in other words on the underlying bone structure. The value of SD was found to be significantly higher if the reference point was located in a curvature or at an anatomical prominence (such as gonion or orbitale), when compared with the points defined by the boundary structure, such as sutura frontalis for nasion or the clear transition from bone to tooth in infradentale (Baumrind and Frantz, 1971). The study group led by Legrell *et al.* (2000) also found high SD value for the landmark gonion in cephalograms.

The present study did not reveal an overall systematic trend in terms of the influence of the observer's level of experience on the repeatability of landmark placement: all the four observers had similar performance in terms of placing the landmarks in the program VoXim®6.1. The study group of de Oliveira *et al.* (2009) also found that the observer's experience only had a minimal effect regarding the error in landmark placement. A possible explanation for this minimal effect is that the landmarks were precisely defined in the run-up of the study. In contrast to these results, Jonas (1976) found a significant relation between the observer's level of experience and the degree of error in landmark setting in cephalograms with a higher degree of error in those observers with less experience.

Taken together, the insights gained about landmark repeatability from the present results coincide with those of Boldt *et al.* (2009), Olszewski *et al.* (2008), Papadopoulos *et al.* (2005), Periago *et al.* (2008), and Plooij *et al.* (2009) despite the fact that the clinical question, study design, and analysis methods in these studies differed from our study.

In this paper, the repeatability and reproducibility of specific skeletal landmarks based on 3D CT reconstructions of the osseous surface of facial parts of the skull were determined after positioning on axial primary sectional images, and 2D reconstructions were assessed. A study of the current literature revealed that the examination of the repeatability of landmark setting, using different imaging techniques, is currently of considerable clinical and scientific relevance. Surface images reconstructed from digital volume tomography can also be measured precisely in the three spatial planes (Hassan *et al.*, 2009).

For example, there are study results available on landmark identification for 3D analysis of soft tissue craniofacial surfaces performed with photogrammetric measuring techniques (Stauber *et al.*, 2008; Sawyer *et al.*, 2009). However, only accessible surfaces could be recorded by stereophotogrammetry, and hence, no skeletal structures could be studied. The study group of Toma *et al.* (2009) investigated the reproducibility of soft tissue landmarks using 3D laser scanned facial images.

In a comparative dosimetric study between dental CBCT and 64-slice CT, the 3D imaging techniques operating with ionizing radiation exposure, such as MSCT and CBCT, were cautiously assessed by taking tissue-weighting factors into consideration. It was concluded that CBCT is recommended as a dose-sparing method for general oral and maxillofacial problems (Ludlow and Ivanovic, 2008). However, with regard to the prospective development of 3D cephalometry for investigating complex and treatment-relevant 3D deformities, no conclusive recommendation could be made because sufficient volume for analysis is required in these cases and not merely the visualization of a small section.

With regard to the radiation exposure of patients, in cases where an ionizing imaging technique is used for 3D representation of osseous structures in the maxillofacial area and adequate volume is recorded, the available data sets can naturally be used to generate panoramic images of the dentition and lateral cephalograms. Thus, volume data sets can be used for 2D and 3D analysis, provided the landmark repeatability is ascertained. In the present study, we were able to confirm this for CT data sets.

Conclusions

The results of the repeatability and reproducibility of CT-based landmark setting presented in this study strongly support the idea of treatment-relevant three-dimensionally oriented cephalometric diagnostics. Therefore, it can be assumed that the repeatability of landmark setting among suitably trained observers is adequate to obtain valid test results from 3D measurements. Thus, the results achieved in this study with regard to the reproducibility of 3D landmark placement can be classified as highly relevant in a clinical context. The conceived study serves as the basis for the development and assessment of the reproducibility of 3D cephalometric analysis methods.

Funding

Luise Prell Foundation.

Acknowledgement

The authors thank Dr Alexandra Ioana Holst for her kind support in the examination of the CT data.

References

Baumrind S, Frantz R C 1971 The reliability of head film measurements. 1. Landmark identification. American Journal of Orthodontics 60: 111–127

Bland M 2000 An introduction to medical statistics, 3rd edn. Oxford University Press, Oxford, pp. 177–179

Boldt F, Weinzierl C, Hertrich K, Hirschfelder U 2009 Comparison of the spatial landmark scatter of various 3D digitalization methods. Journal of Orofacial Orthopedics 70: 247–263

Brown A A, Scarfe W C, Scheetz J P, Silveira A M, Farman A G 2009 Linear accuracy of cone beam CT derived 3D images. Angle Orthodontist 79: 150–157

Cavalcanti M G, Rocha S S, Vannier M W 2004 Craniofacial measurements based on 3D-CT volume rendering: implications for clinical applications. DentoMaxilloFacial Radiology 33: 170–176

Chan H J, Woods M, Stella D 2007 Three-dimensional computed craniofacial tomography (3D-CT): potential uses and limitations. Australian Orthodontic Journal 23: 55–64

Chien P, Parks E, Eraso F, Hartsfield J, Roberts W, Ofner S 2009 Comparison of reliability in anatomical landmark identification using two-dimensional digital cephalometrics and three-dimensional cone beam computed tomography in vivo. DentoMaxilloFacial Radiology 38: 262–273

Connor S E, Arscott T, Berry J, Greene L, O'Gorman R 2007 Precision and accuracy of low-dose CT protocols in the evaluation of skull landmarks. DentoMaxilloFacial Radiology 36: 270–276

De Oliveira A E, Cevidanes L H, Phillips C, Motta A, Burke B, Tyndall D 2009 Observer reliability of three-dimensional cephalometric landmark identification on cone-beam computerized tomography. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontics 107: 256–265

De Vos W, Casselman J, Swennen G R 2009 Cone-beam computerized tomography (CBCT) imaging of the oral and maxillofacial region: a systematic review of the literature. International Journal of Oral and Maxillofacial Surgery 38: 609–625

Dvortsin D P, Sandham A, Pruim G J, Dijkstra P U 2008 A comparison of the reproducibility of manual tracing and on-screen digitization for cephalometric profile variables. European Journal of Orthodontics 30: 586–591

Greiner M, Greiner A, Hirschfelder U 2007 Variance of landmarks in digital evaluations: comparison between CT-based and conventional digital lateral cephalometric radiographs. Journal of Orofacial Orthopedics 68: 290–298

Hassan B, van der Stelt P, Sanderink G 2009 Accuracy of three-dimensional measurements obtained from cone beam computed tomography surface-rendered images for cephalometric analysis: influence of patient scanning position. European Journal of Orthodontics 31: 129–134

Hirschfelder U, Piechot E, Schulte M, Leher A 2004 Abnormalities of the TMJ and the musculature in the oculo-auriculo-vertebral spectrum (OAV). A CT study. Journal of Orofacial Orthopedics 65: 204–216

Houston W J, Maher R E, McElroy D, Sherriff M 1986 Sources of error in measurements from cephalometric radiographs. European Journal of Orthodontics 8: 149–151

JCGM. 2008 "International vocabulary of metrology—basic and general concepts and associated terms, and associated terms, VIM", 3rd edn. (http://www.bipm.org/en/publications/guides/vim.html (4 April 2011, date last accessed))

Jonas I 1976 The influence of training on the accuracy of roentgenographic cephalometric tracings (author's transl). Radiologe 16: 427–431

Katsumata A, Fujishita M, Maeda M, Ariji Y, Ariji E, Langlais R P 2005 3D-CT evaluation of facial asymmetry. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontics 99: 212–220

Kitai N *et al.* 2004 Craniofacial morphology in an unusual case with nasal aplasia studied by roentgencephalometry and three-dimensional CT scanning. Cleft Palate-Craniofacial Journal 41: 208–213

Kragskov J, Bosch C, Gyldensted C, Sindet-Pedersen S 1997 Comparison of the reliability of craniofacial anatomic landmarks based on cephalometric radiographs and three-dimensional CT scans. Cleft Palate-Craniofacial Journal 34: 111–116

Krarup S, Darvann T A, Larsen P, Marsh J L, Kreiborg S 2005 Three-dimensional analysis of mandibular growth and tooth eruption. Journal of Anatomy 207: 669–682

Kusnoto B, Evans C A, BeGole E A, de Rijk W 1999 Assessment of 3-dimensional computer-generated cephalometric measurements. American Journal of Orthodontics and Dentofacial Orthopedics 116: 390–399

Lagravere M O, Hansen L, Harzer W, Major P W 2006 Plane orientation for standardization in 3-dimensional cephalometric analysis with computerized tomography imaging. American Journal of Orthodontics and Dentofacial Orthopedics 129: 601–604

Legrell P E, Nyquist H, Isberg A 2000 Validity of identification of gonion and antegonion in frontal cephalograms. Angle Orthodontist 70: 157–164

Lou L, Lagravere M O, Compton S, Major P W, Flores-Mir C 2007 Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: a systematic review. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontics 104: 402–411

Ludlow J B, Ivanovic M 2008 Comparative dosimetry of dental CBCT devices and 64-slice CT for oral and maxillofacial radiology. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontics 106: 106–114

Maeda M *et al.* 2006 3D-CT evaluation of facial asymmetry in patients with maxillofacial deformities. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontics 102: 382–390

Moshiri M, Scarfe W C, Hilgers M L, Scheetz J P, Silveira A M, Farman A G 2007 Accuracy of linear measurements from imaging plate and lateral cephalometric images derived from cone-beam computed tomography. American Journal of Orthodontics and Dentofacial Orthopedics 132: 550–560

Olszewski R, Reychler H, Cosnard G, Denis J M, Vynckier S, Zech F 2008 Accuracy of three-dimensional (3D) craniofacial cephalometric landmarks on a low-dose 3D computed tomograph. DentoMaxilloFacial Radiology 37: 261–267

Papadopoulos M A *et al.* 2005 Three-dimensional fetal cephalometry: an evaluation of the reliability of cephalometric measurements based on three-dimensional CT reconstructions and on dry skulls of sheep fetuses. Journal of Cranio-Maxillo-Facial Surgery 33: 229–237

Park S H, Yu H S, Kim K D, Lee K J, Baik H S 2006 A proposal for a new analysis of craniofacial morphology by 3-dimensional computed tomography. American Journal of Orthodontics and Dentofacial Orthopedics 129: 623–634

Periago D R, Scarfe W C, Moshiri M, Scheetz J P, Silveira A M, Farman A G 2008 Linear accuracy and reliability of cone beam CT derived 3-dimensional images constructed using an orthodontic volumetric rendering program. Angle Orthodontist 78: 387–395

Phillips C, Greer J, Vig P, Matteson S 1984 Photocephalometry: errors of projection and landmark location. American Journal of Orthodontics 86: 233–243

Piechot E 2005 Kondyläre und faziale Fehlentwicklungen bei Anomalien des oculo-auriculo-vertebralen Spektrums. University of Erlangen-Nürnberg: Med. Dissertation

Plooij J M *et al.* 2009 Evaluation of reproducibility and reliability of 3D soft tissue analysis using 3D stereophotogrammetry. International Journal of Oral and Maxillofacial Surgery 38: 267–273

Rasch D, Herrendörfer G, Bock J, Victor N, Guiard V 1996 Verfahrensbibliothek Versuchsplanung und—auswertung, 2nd edn. Oldenbourg, München. pp. 486

Richardson A 1966 An investigation into the reproducibility of some points, planes, and lines used in cephalometric analysis. American Journal of Orthodontics 52: 637–651

Rouas P, Nancy J, Bar D 2007 Identification of double mandibular canals: literature review and three case reports with CT scans and cone beam CT. DentoMaxilloFacial Radiology 36: 34–38

Sawyer A R, See M, Nduka C 2009 Assessment of the reproducibility of facial expressions with 3-D stereophotogrammetry. Otolaryngology—Head and Neck Surgery 140: 76–81

Sayinsu K, Isik F, Trakyali G, Arun T 2007 An evaluation of the errors in cephalometric measurements on scanned cephalometric images and conventional tracings. European Journal of Orthodontics 29: 105–108

Stauber I *et al.* 2008 Three-dimensional analysis of facial symmetry in cleft lip and palate patients using optical surface data. Journal of Orofacial Orthopedics 69: 268–282

Swennen G R *et al.* 2009 A cone-beam computed tomography triple scan procedure to obtain a three-dimensional augmented virtual skull model appropriate for orthognathic surgery planning. Journal of Craniofacial Surgery 20: 297–307

Toma A M, Zhurov A, Playle R, Ong E, Richmond S 2009 Reproducibility of facial soft tissue landmarks on 3D laser-scanned facial images. Orthodontics and Craniofacial Research 12: 33–42