Journal of Cranio-Maxillo-Facial Surgery 42 (2014) 1795–1801

Contents lists available at ScienceDirect

Journal of Cranio-Maxillo-Facial Surgery

journal homepage: www.jcmfs.com

Experimental and clinical assessment of three-dimensional cephalometry: A systematic review☆

T. Smektała [a, b, *], M. Jędrzejewski [a, b], J. Szyndel [a], K. Sporniak-Tutak [a], R. Olszewski [b]

[a] Department of Dental Surgery (Head: Trybek G, DMD), Pomeranian Medical University, Ul. Powstańców Wlkp. 72, 70-111 Szczecin, Poland
[b] Oral and Maxillofacial Surgery Research Lab (OMFS/CHEX/IREC/SSS/UCL, Head: Pr Olszewski R, PhD), Departement of Oral and Maxillofacial Surgery, Cliniques Universitaires Saint Luc, Université Catholique de Louvain, 10 Av. Hippocrate, 1200 Brussels, Belgium

A R T I C L E I N F O

Article history:
Paper received 1 March 2014
Accepted 6 June 2014
Available online 15 June 2014

Keywords:
Cephalometry
Imaging
Three-dimensional
Cone-beam computed tomography
Reproducibility of results
Dimensional measurement accuracy

A B S T R A C T

Objectives: This study provides a systematic review of the current scientific literature on three-dimensional (3D) cephalometry. The null hypothesis was that 3D cephalometry is an accurate and reproducible diagnostic technique. To examine this hypothesis, the following three research questions were proposed: 1) What is the accuracy of 3D cephalometric measurements compared to *in vitro* measurements? 2) What is the intra- and inter-observer reliability of the selection of 3D cephalometric landmarks? 3) What is the reproducibility of the linear and angular measurements?
Methods: A comprehensive database search was performed, using Medline, the Cochrane Central Register of Controlled Trials, Web of Science and Google Scholar. The titles and abstracts obtained from the search were screened and evaluated by two observers according to the inclusion and exclusion criteria.
Results: The evaluation process yielded 21 articles. A high level of agreement (<1 mm) between the *in vitro* measurements and those obtained from 3D cephalometry was observed and some landmarks provided highly reproducible results. However, the linear (0.04–7.49 mm) and angular (0.99–9.30°) measurements differed greatly.
Conclusions: The null hypothesis was rejected. This study indicates critical points regarding 3D cephalometry and provides guidance for future research in this field.

1. Introduction

Cephalography was introduced in 1931 (Broadbent, 1931). Based on this technology, a number of cephalometric analyses have been developed, which have been helpful in orthodontic diagnosis and the planning of orthognathic surgery (Chaconas and Fragiskos, 1991; Delaire et al., 1981; Steiner, 1953). However, because cephalograms are two-dimensional projections of three-dimensional (3D) structures, they have several disadvantages (Olszewski and Reychler, 2004), including a lack of perspective (Major et al., 1994), errors in projection (Ahlqvist et al., 1988), geometric distortions (Baumrind and Frantz, 1971a), imaging artefacts (Sutcher and Laskin, 1971), variations in magnification (Major et al., 1994) and head positioning errors (Baumrind and Frantz, 1971b). Introduction of computed tomography (CT) or cone-beam CT (CBCT)-based cephalometric analyses could offer solutions to these problems (Olszewski et al., 2007; Swennen et al., 2006). To date, several clinical studies of 3D cephalometry have been published (Hassan et al., 2009; Olszewski et al., 2008), including one systematic review by Rossini et al. (2011). However, Rossini et al. did not evaluate the risk of bias and did not provide any information regarding the levels of evidence of the studies included in their review. Additionally, the review presented the data in a manner that did not allow a direct comparison of the results, and it also excluded landmarks inside the skull and temporomandibular joint area, which are not typically the focus of orthodontics. While the statistical measures used to analyse inter- and intra-observer reliability across studies might not allow for comparisons, the accuracy of the angular and linear measurements should be compared and presented in a metrical manner. Additionally, because cephalometry serves both orthodontists and surgeons, all of the landmarks from the maxillofacial area should be examined. Therefore, the aim of this study was to perform a systematic review of the current

☆ This study was founded by grant from Fondation saint Luc, Fond Hervé Reychler 2012–2013 awarded to Pr R Olszewski for development of international research on computer-assisted orthognathic surgery.
* Corresponding author. Pomeranian Medical University, Departement of Dental Surgery, Ul. Powstańców Wlkp. 72, 70-11 Szczecin, Poland. Tel.: +48 663262610.
E-mail address: smektala.t@gmail.com (T. Smektała).

http://dx.doi.org/10.1016/j.jcms.2014.06.017
1010-5182/

literature on 3D cephalometry. The null hypothesis was that 3D cephalometry is an accurate and reproducible diagnostic technique. To examine this hypothesis, the following three research questions were asked:

1. What is the accuracy of 3D cephalometric measurements compared to *in vitro* measurements?
2. What is the intra- and inter-observer reliability of the selection of 3D cephalometric landmarks?
3. What is the reproducibility of the linear and angular measurements?

2. Material and methods

On March 15, 2013, a comprehensive database search was performed by a single observer using Medline (Ovid), Medline In-Process & Other Non-Indexed Citations (Ovid), PubMed, Cochrane, Web of Science and Google Scholar. No time frame was applied. Only English-language publications that were published in peer-reviewed journals were considered. A single search equation was used (Table 1). To translate the equation correctly for other databases, the help of a senior librarian was sought. The titles and abstracts obtained from this electronic search were then screened and evaluated by two observers according to the inclusion and exclusion criteria (Table 2). Studies that did not meet the inclusion criteria were excluded from further evaluation. Any discrepancies in the selection were settled through discussion. A PRISMA flow diagram illustrates the search and evaluation process (Fig. 1). All of the articles were independently screened by one reviewer to extract the necessary data. After the extraction process, all of the data were cross-checked by the second reviewer. To generate reliable evidence, a critical appraisal was conducted of all of the included articles. Specified criteria were adopted to assess the internal and external validity, reliability and objectivity of the studies (Table 3). When one or more of the key domains were determined to be high risk or unclear, the overall judgement was that the study had a high or unclear risk of bias. Conversely, when all of the key domains were low risk, the study was judged to be low risk. This review was registered in the PROSPERO International Prospective Register of Systematic Reviews under number CRD42013004996.

3. Results

Of the 579 articles initially found, following the automatic rejection of duplicates in EndNote X5 (Thomson Reuters, Philadelphia, USA) reference manager software, the search yielded a total of 412 articles. After verification of the inclusion/exclusion criteria, 389 articles were excluded. An additional search with the same search equation, performed using Google Scholar, found 6 new articles. The final number of articles was 21 (Berco et al., 2009; Bholsithi et al., 2009; Fuyamada et al., 2011; Gribel et al., 2011; Hassan et al., 2009; Lagravère et al., 2009, 2010; Ludlow et al., 2009; Medelnik et al., 2011; de Oliveira et al., 2009; Olszewski et al., 2007, 2008, 2010, 2013; Oz et al., 2011; Periago et al., 2008; Schlicher et al., 2012; Titiz et al., 2012; Varghese et al., 2010; Zamora et al., 2012, 2011). An assessment of the risk of bias revealed that none of the articles met the adopted criteria. None of the articles, except for one by Hassan et al. (2009), reported the threshold range for 3D reconstructions or the brightness and contrast values for planar selection methods. Only 10 articles of the 21 provided information about approval from local ethics committees (Fuyamada et al., 2011; Gribel et al., 2011; Lagravère et al., 2010; Ludlow et al., 2009; Medelnik et al., 2011; Olszewski et al., 2013; Oz et al., 2011; Periago et al., 2008; Schlicher et al., 2012; Zamora et al., 2012), and only these articles were assessed as low risk in the reliability domain. All of the articles met the criteria for internal validity and objectivity (Table 4). Application of the inclusion criteria for the first question resulted in 5 articles (Berco et al., 2009; Gribel et al., 2011; Olszewski et al., 2007; Periago et al., 2008; Varghese et al., 2010). Three of these articles were based on CBCT scans (Berco et al., 2009; Gribel et al., 2011; Periago et al., 2008), one on spiral CT (Varghese et al., 2010) and one on conventional CT scans (Olszewski et al., 2007). All of the authors used different acquisition protocols. Selection of cephalometric landmarks was predominantly performed on 3D reconstructions (Berco et al., 2009; Olszewski et al., 2007; Periago et al., 2008; Varghese et al., 2010), except for Gribel et al. (2011), who selected landmarks on 2D slices. The following four software programs were used in the studies: SimPlant Ortho, version 2.0 (Gribel et al., 2011); Mimics, version 11.02 (Varghese et al., 2010); ACRO 3D (Olszewski et al., 2007); and Dolphin 3D (Berco et al., 2009; Periago et al., 2008). In four of these studies (Berco et al., 2009; Gribel et al., 2011; Olszewski et al., 2007; Periago et al., 2008), comparisons of

Table 1
PubMed search equation used in the search process.

PubMed search equation
(3d[TW] OR three-dimensional[TW] OR "three dimensional"[TW]) AND (cephalometry[MeSH Terms] OR Cephalometry[TW] OR cephalometric[TW] OR cephalometr*) AND (validation[TW] OR reproducibility[TW] OR evaluation[TW] OR reliability[TW] OR repeatability[TW] OR accuracy[TW])

Table 2
Inclusion and exclusion criteria used in the study.

Inclusion criteria	Exclusion criteria
Question 1: "What is the accuracy of the measurements obtained in 3D cephalometry compared to *in vitro* measurements?"	
- Only human cadaver or skull studies - Result shown in metrical or percentage manner	- Comparative measurements (pre/post-op) which were not in the main scope of the article
Question 2: "What is the landmarks identification inter- and intra-observer repeatability?"	
- Primary publications in which inter- and intra-observer repeatability measurements are in the main scope - Detailed definitions of landmarks - Results provided for all 3 axes or in box-type approximations	- Results provided for less than 3 axes or not in box-type approximations
Question 3: "What is the linear and angular measurements repeatability in 3D cephalometric analysis?"	
- Primary publications in which linear and angular measurements repeatability are in the main scope - Detailed definitions of landmarks or/and linear/angular measurements - Results provided in standard deviation manner	
Criteria common for all questions	
- Articles focused on landmarks accuracy measurements in the area of interest to clinical orthodontists - 3D cephalometry measurements based on CT scan slices or 3D reconstructions - Description of the statistical analysis method - Detailed description of acquisition protocol	- Animal studies - Articles type other than: clinical trial, comparative study, controlled clinical trial, the *in vitro* study, randomized controlled trial, validation studies or evaluation studies - Measurements outside the maxillofacial area - Measurements associated with dental implants - Measurements associated with temporomandibular joint - Measurements after facial trauma or tumour cases in maxillofacial area