“The influence of using 2D cephalometry on orthodontic treatment outcome”

Thesis submitted in partial fulfilment of the requirements for the degree of “DOCTOR IN MEDICAL SCIENCES” presented to the Faculty of Dental Medicine of the University of Porto.

Ana Paula Oliveira dos Reis Durão

Porto, 2014

“The influence of using 2D cephalometry on orthodontic treatment outcome.”

Ana Paula Oliveira dos Reis Durão

Promotor: Prof. Dr. Reinhilde Jacobs

Full Professor, Head of the Oral Imaging Centre, Oral Imaging Center, OMFS-IMPATH research group, Department Imaging & Pathology, Faculty of Medicine, University of Leuven, Belgium.

Co-promotor: Prof. Dr. Afonso Pinhão Ferreira

Full Professor, Orthodontics Department, Director of the Faculty of Dental Medicine of the University of Porto, Portugal.

Faculty of Dental Medicine of the University of Porto

Members of the Scientific Committee

Prof. Doutor Afonso Manuel Pinhão Ferreira (Full Professor) – President

Prof. Doutor Américo dos Santos Afonso (Associate Professor with aggregation)

Prof. Doutor António Cabral Campos Felino (Full Professor) – Vice-President

Prof. Doutor César Fernando Coelho Leal Silva (Associate Professor with aggregation)

Prof. Doutor Germano Neves Pinto Rocha (Associate Professor)

Prof. Doutora Irene Graça Azevedo Pina Vaz (Associate Professor)

Prof. Doutora Inês Alexandra Costa Morais Caldas (Assistant Professor)

Prof. Doutor João Carlos Antunes Sampaio Fernandes (Full Professor)

Prof. Doutor João Carlos Gonçalves Ferreira de Pinho (Associate Professor with aggregation)

Prof. Doutor João Fernando Costa Carvalho (Full Professor)

Prof. Doutor Jorge Manuel Carvalho Dias Lopes (Full Professor)

Prof. Doutor José António Macedo Carvalho Capelas (Associate Professor with aggregation)

Prof. Doutor José Carlos Reis Campos (Assistant Professor with aggregation)

Prof. Doutor José Mário Castro Rocha (Assistant Professor)

Prof. Douto Manuel José Fontes de Carvalho (Associate Professor)

Prof. Doutora Maria Cristina P. C. M. Figueiredo Pollmann (Associate Professor)

Prof. Doutora Maria Helena Guimarães Figueiral da Silva (Full Professor)

Prof. Doutora Maria Helena Raposo Fernandes (Full Professor)

Prof. Doutora Maria Lurdes Ferreira Lobo Pereira (Assistant Professor)

Prof. Doutor Mário Augusto Pires Vaz (Associate Professor - invited)

Prof. Doutor Mário Jorge Rebolho Fernandes Silva (Full Professor)

Prof. Doutor Mário Ramalho Vasconcelos (Associate Professor with aggregation)

Prof. Doutor Miguel Fernando Silva Gonçalves Pinto (Full Professor)

Prof. Doutor Paulo Rui Galrão Ribeiro Melo (Associate Professor with aggregation)

Prof. Doutor Ricardo Manuel Lobo Faria Almeida (Associate Professor with aggregation)

Emeritus Professors

Prof. Doutor Adão Fernando Pereira (Full Professor)

Prof. Doutor Amílcar Almeida Oliveira (Associate Professor)

Prof. Doutor António Manuel Machado Capelas (Associate Professor †)

Dr. António Ulisses Matos dos Santos (Invited Assistant)

Prof. Doutor Durval Manuel Belo Moreira (Associate Professor with aggregation)

Prof. Doutor Francisco António Rebelo Morais Caldas (Full Professor)

Dr. José Maria Vaz Osório (Invited Assistant)

Prof. Doutor José Serra Silva Campos Neves (Full Professor)

Prof. Doutor Manuel Desport Marques (Invited Associate Professor †)

Prof. Doutor Manuel Guedes de Figueiredo (Associate Professor)

Retired Professors

Prof. Doutor António Manuel Guerra Capelas (Assistant Professor)

Prof. Dr. Artur Manuel Osório de Araújo (Invited Associate Professor)

Prof. Doutor Fernando Jorge Morais Branco (Full Professor)

Prof. Doutor Fernando José Brandão Martins Peres (Full Professor †)

Prof. Doutor José Albertino Cruz Lordelo (Associate Professor with aggregation)

Prof. Doutor José Carlos Pina Almeida Rebelo (Full Professor)

Prof. Doutor Manuel Pedro da Fonseca Paulo (Full Professor)

Prof. Doutora Maria Adelaide Macedo Carvalho Capelas (Associate Professor †)

Prof. Doutora Maria Purificação Valenzuela Sampaio Tavares (Full Professor)

Prof. Doutor Rogério Serapião Martins Aguiar Branco (Full Professor)

Dedications:

To Hugo, Pedro and Miguel.

To my mother and father.

To the memory of my grandparents.

ACKNOWLEDGEMENTS

My sincere gratitude is owed to my promoter, Prof. Dr. Reinhilde Jacobs, Head of the Oral Imaging Centre and coordinator of the OMFS-IMPATH research group of the Faculty of Medicine of the University of Leuven in Belgium for her scientific and friendly support. Thank you for always believing in this project and for your invaluable help.

My appreciation must be extended to Prof. Dr. Afonso Pinhão Ferreira, my co-promoter and Director of the Faculty of Dental Medicine, of the University of Porto. It has been an honor and pleasure to have your irreplaceable help.

I express my thanks to Prof. Dr. João Carlos Antunes Sampaio Fernandes, as the director of the PhD Doctoral Program.

I must thank Prof. Dr. Jorge Dias Lopes and Prof. Dr. Maria Cristina Figueiredo Pollman, for allowing me to use the patient files available at the Orthodontic Department.

I am also in indebted with Prof. Dr. José Albino Teixeira Koch and to Prof. Dr. António Marcelo Miranda for their comprehension and friendly support while I was in Leuven and during these three years.

I should express my appreciation to Prof. Dr. António Felino for the encouragement for enrolling in the PhD program.

I would like to recognize the contribution of Dr. Ana Gião, Dr. Eugénio Martins, Dr. Mónica Pinho, Dr. Nídia Rocha, Dr. Nelso Reis, Dr. Saul Castro, Dr. Susana Silva, Dr. Telmo Moreira, and Dr. Vanda Urzal for the patient files evaluation. The supply of the patient files by the post-graduate students Ana

Margarida Nunes, Andreia Fontes, Maria Inês Correia, Maria Passos and Rita Carvalho. I must recognize the great effort in gathering all the files by Maria Passos. And also Radio Memory, for providing the software for cephalometric analysis.

Likewise need to express my thanks Prof. Dr. Cristina Santos and Dr. Cláudia Camila Dias for statistical support.

I must thank Dr. Carla Pinto and Dr. Margarida Lessa for their friendship and important words at the right moment. And similarly to Delfina Alves for her librarian help.

This research was carried out at the Faculty of Medicine of the University of Leuven in Belgium and Faculty of Dental Medicine, of the University of Porto. My sincere appreciation to the team of Prof. Dr. Reinhilde Jacobs at the University of KU Leuven, for their hospitality during my stay there. A especial thanks to Pisha Pittayapat, Ali Alqerban, Napat Bolstad and Maria Ivete Bolzan Rockenbach.

To Márcia Provenzano, a great friend for all the teachings and her friendship. Also for her precious orientations that contributed to my professional growth.

To all my friends, especially to Joana Alves.

My special thanks to my mother and father for all the values transmitted, unending love and for always encouraging me during this time.

Also, to my sister, and best friend, for her friendly support and protection. To my lovely Teresinha and Manuel. To my grandparents, even they are not present, for all that they represent.

I want to thank my in-laws, especially my mother in law for her unequal dedication to Pedro and Miguel.

My honest gratefulness is owed to my husband, Hugo, and children, Pedro and Miguel. For their love, support and motivation given during these years. For all the precious moments that I have missed, and for your understanding.

TABLE OF CONTENTS

LIST OF ABREVIATIONS

2D	Two-dimensional
3D	Three-dimensional
LCR	Lateral cephalometric radiograph
DMFR	Dentomaxillofacial radiologists
mAs	Product of tube current (mA) and exposure time (s)
mGy	milligray
kV	kilovoltage
m	meters
mm	millimetres
PA	postero-anterior
CCD	Charged Couple Device
ALARA	As low as reasonably achievable
N	Nasion
Or	Orbital
S	Sella
Po	Porion
Co	Condylion
Go	Gonion
Me	Menton
Pog	Pogonion
Gn	Gnathion
B	B point

A	A point
ANS	Anterior Nasal Spine
PNS	Posterior Nasal Spine
LIA	Lower incisor apex
LIB	Lower incisor border
UIB	Upper incisor border
UIA	Upper incisor apex

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

INTRODUCTION

Etymology

“Cephalometric” is originated from two Greek words: *kephalo,* meaning “head” and *metron,* meaning “measurement”. Cephalometrics or cephalometry is concerned with measuring the dimensions of the head (hard and soft tissues) (Finlay, 1980). “Orthodontics” is originated from two Greek words: *orthos,* meaning “straight, proper or perfect” and *odous* meaning “tooth”. Orthodontics is the specialty of dentistry that is concerned with the study of growth of the craniofacial complex, development of occlusion, and treatment of dentofacial abnormalities (AlBarakati. *et al*., 2012; Moyers, 1988).

The origin of cephalometry

Cephalometrics did not begin with orthodontics, but it was initiated with the study of human growth and development of craniofacial anatomy (Wahl, 2006). The art of measuring skulls of animals became known as craniometrics. This method has been studied in the area of physical anthropology before the discovery of X-ray in 1895 by Wilhelm Conrad Röntgen. Hippocrates, a pioneer in physical anthropology (460-375 BC), left numerous descriptions on the existent variations in the skulls. Leonardo da Vinci (1452-1519) and Albrecht Dürer (1471-1528) demonstrated the first metrical studies of the head. They established proportions between lines and segments and explained why the proclined facial contour differed from the retroclined configuration by changing the angle between vertical and horizontal axes (Figure 1 and 2).

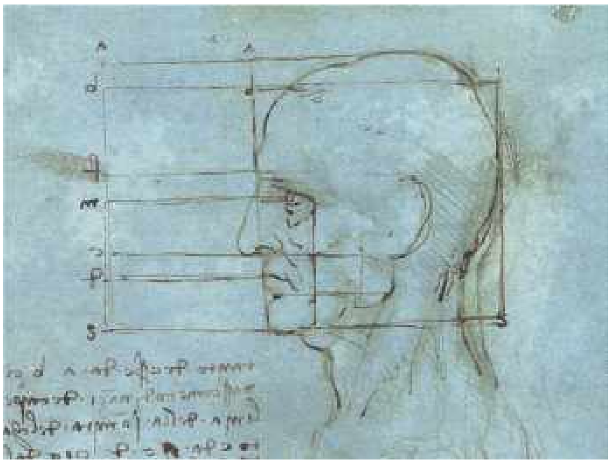

Figure 1. Drawing of a human's head with measurements, by Leonardo da Vinci (1488-9).

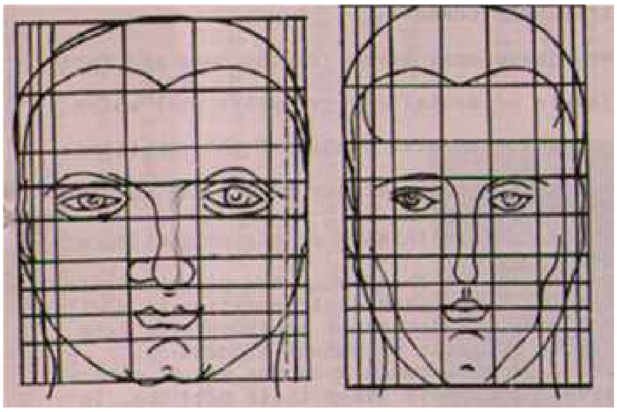

Figure 2. Representation of measurements of two human heads, by Albrecht Dürer.

Anders Retzius (1796–1860) first used the *cephalic index* in physical anthropology to classify ancient human remains found in Europe. The *cephalic index* is a rating scale used to calculate the size of the head, expressing the ratio of the maximum breadth of a skull to its maximum antero-posterior length. It is calculated by multiplying the maximum width of the head by 100 and dividing that number by the maximum length of the head. He classified skulls in three main categories: "dolichocephalic" (from the Ancient Greek *kephalê*, head, and *dolikhos*, long and thin), "brachycephalic" (short and broad) and "mesocephalic" (intermediate length and width) (Figure 3). Nowadays, it is used to classify individual head appearance.

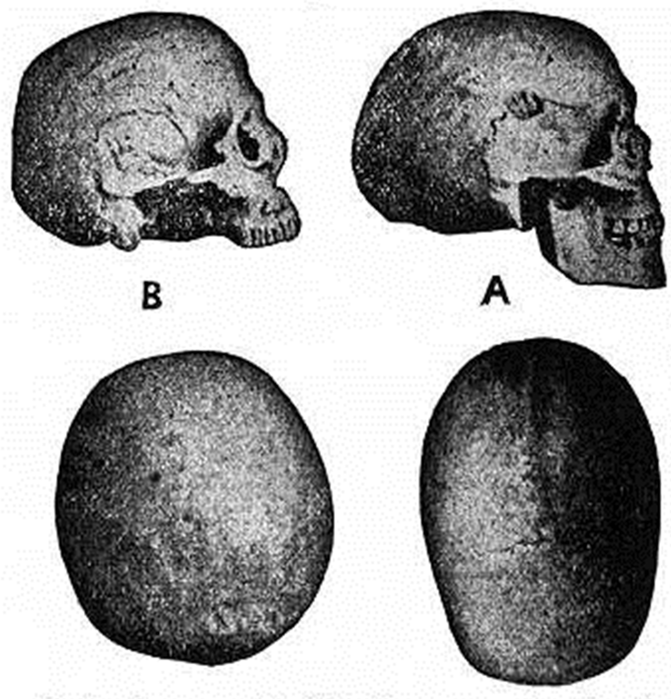

Figure 3. The pictures shows a superior and a lateral view of two skulls, one is brachycephalic (B) and the other is dolicocephalic (A).

Later, Petrus Camper (1722-1789), physician, anatomist, and painter, was possibly the first to employ angles in measuring faces (Wahl, 2006; Finlay, 1980). He defined the “facial line” (*linea facialis).* It became the universal measurement for the study of the human face. In 1780, he did measurements on human skulls and primates, describing the Camper’s facial angle, which was formed by the intersection of a facial line and a horizontal plane. The facial line was tangential to the most prominent part of the frontal bone and the convexity of the upper teeth. The horizontal plane passed through the lower part of the nasal aperture, backwards along the line of the zygomatic arch, and through the center of the external auditory meatus. Two years after the death of Petrus Camper, his well-known work on natural variants of the face was published (Finlay, 1980; Wahl, 2006). The facial angle, according to Camper, was of 80 degrees for European, 70 degrees for African, 58 degrees for orangutan and 42 degrees for monkeys (Figure 4) (Finlay, 1980).

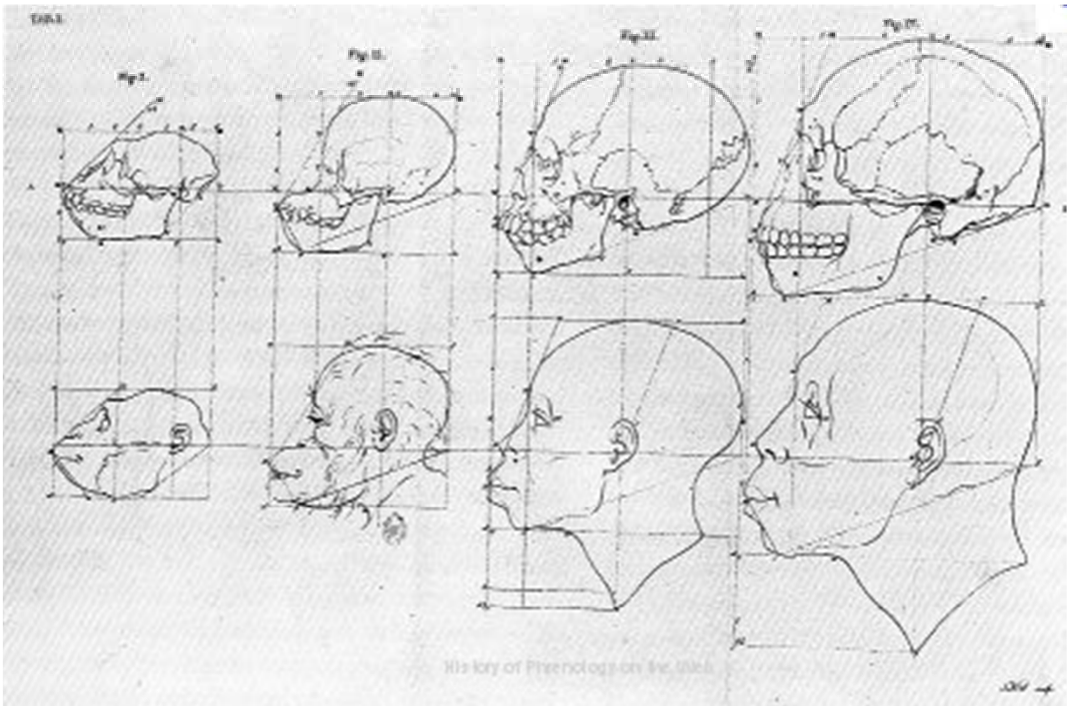

Figure 4 – Method of determining the facial angle by Petrus Camper.

The Frankfort Plane

In the XIII General Congress of the Society of German Anthropology (performed in Frankfurt-am-Main, 1884) the plane of Von Iheming was approved, which now serves as a universal method of cranium orientation. Observation of the cranium should be performed with the skull in a standard orientation, whereby the

Frankfort plane is horizontal, i.e. parallel to the floor. The Frankfort plane can be determined on the dry skull, on patients and on radiographs. The Frankfort plane (Figure 5) is a transverse plane through the skull and it is perpendicular to the mid-sagittal plane. The plane runs through a line joining the uppermost point of the bony left external auditory meatus (anatomic Po) and the lowest point on the left infraorbital margin (Or) (Whaites, 2007; Finlay, 1980).

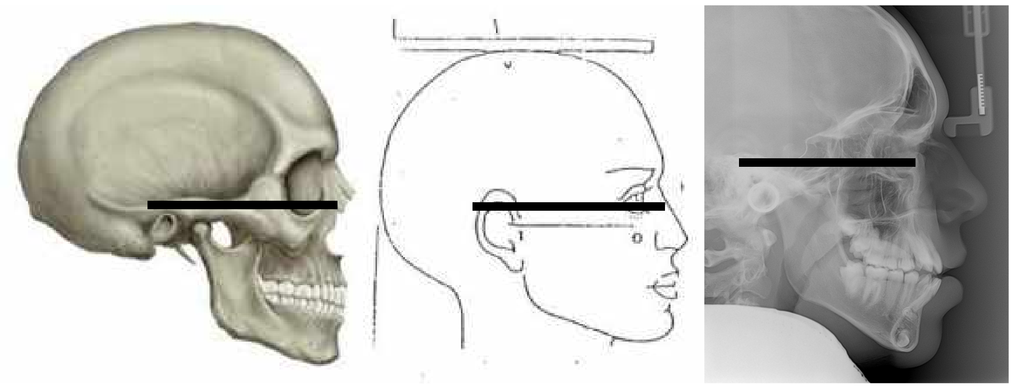

Figure 5. Illustration of the Frankfort Plane on a skull, a patient and on a LCR.

However, the orientation of the Frankfort plane may be difficult due to the identification of the Po and Or on radiographs. An alternative reference line, which is easier to identify is the Sella-Nasion (SN) plane (Figure 6). It runs from the landmarks S to N. On average this plane is orientated 6 to 7 degrees to the Frankfort Plane. The SN plane was defined in 1920 by Broadbent. This reference plane would become more used after its inclusion in Steiner's cephalometric analysis (Steiner, 1953). The definition of either the Frankfort or SN plane, presents some problems (Houston, 1991). Regarding the SN plane, the point S can vary both antero-posteriorly and vertically. These intracranial reference planes can also diverge, in the same patient within time. Natural head position (NHP) can also be used as a reference. It provides an extracranial reference line, defined as a physiologic position and it is relatively constant over time. The concept of NHP was introduced in orthodontics in the 1950s by Downs (1956), Bjerin (1957), and Moorrees and Kean (1958). NHP has been found to be highly reproducible in adults and children, males and females, Caucasians and non-Caucasians, with a variance of only about 4°. Some authors believe that the analysis based on NHP should have a greater clinical application than traditional methods in describing morphology (Bansal *et al*., 2012).

Figure 6. Illustration of the SN Plane on a lateral cephalometric radiography.

Cephalometric Radiography

Pacini immobilized the patient's head with bandages or gauze, taking radiographs with the sagittal plane parallel to the radiographic film. The equipment had an arm with a distance of two meters between the X-ray source and the film (Wahl, 2006; Athanasios and Athanasiou, 1995; Moyers, 1988). He used craniometric points available for anthropology studies and evaluated the development and deviations of the normality in structures of the skull. In 1922, he was the first to use skull radiographs for craniometrical measurements, and demonstrated that cranioskeletal measurements could be made from skull radiographs more easily than from the skull itself.

Hofrath (Figure 7) in Germany used a cephalostat of Korkhaus. He described in detail its radiographic technique and cephalometric analysis which was published in Germany in 1931.

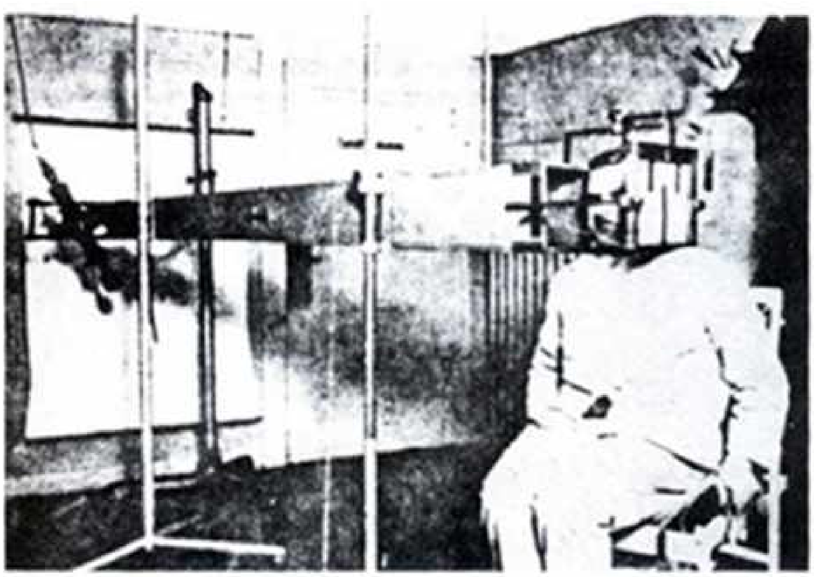

Figure 7. Cephalostat used by Hofrath in Germany.

Broadbent-Bolton cephalostat

Broadbent designed a head-holder, of excellent accuracy. The basic principles of this cephalostat are still in use today (Athanasios and Athanasiou, 1995). The cephalostat design was named Broadbent-Bolton (Figure 8), due to the financial support that he was given by the Bolton foundation. It was first used in children. This cephalostat used two X-ray sources separately and two film receptors, to take one posterior-anterior (PA) radiograph and one lateral radiograph. By using two X-ray sources in different locations, the patient head position did not have to be moved or changed between the two exposures (Athanasios and Athanasiou, 1995; Moyers, 1988) (Figure 9). From this moment, the method of performing measurements from radiographs of the skull, as a scientific assessment for orthodontic problems, has become possible. The serial x-rays, which previously were taken with imprecise cephalostat and therefore of questionable value, were modified after the Broadbent invention. These radiographs are now routinely used in the observation of skull growth and in the evaluation of orthodontic treatment.

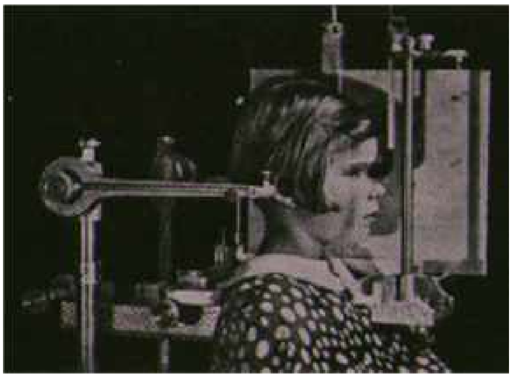

Figure 8. The Broadbent-Bolton cephalostat (Athanasios and Athanasiou, 1995).

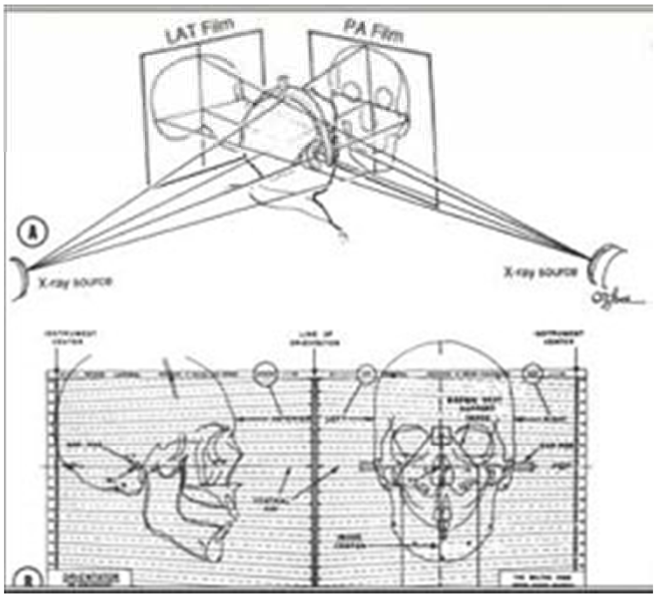

Figure 9. Two X–ray sources were positioned at mutually perpendicular locations, the patient head position did not have to be moved or changed to take a lateral view or a posterior-anterior view (Raju et al., 2010).

Teleradiography

In 1940, Higley presented a cephalostat with only one x-ray source and a movable head fixer. In the same year, Margolis developed a cephalostat with the same principles but with less distortion, which solved some problems of the existent technique. At the First Congress of Cephalometric Radiography in 1957, the teleradiography technique was standardized and a distance of 1.524 meters from the focal spot to the plane of the image receptor was determined to be the standard, as well as the positioning of the left (as opposed to the right) side of the patient’s head near to the image receptor. The distance of the head to the image receptor is standardized being of 20 cm from the sagittal plane of the patient to the image receptor.

By having a relatively large distance between the X-ray source and the head, it helps to minimize magnification errors (Sánchez and Filho, 2009).

The introduction of the head positioning device and the technique of radiographic cephalometry were pioneered by Broadbent in the United States and by Hofrath in Germany in 1931, simultaneously but independently (AlBarakati. *et al*., 2012; Devereux. *et al*., 2011; Nijkamp. *et al*., 2008). Until 1931, diagnosis was performed with clinical examination. After 1931, possibilities emerged for orthodontists, with LCR providing invaluable help in treatment planning, analysis of growth, mid-treatment monitoring and prediction of possible treatment outcomes. The main clinical indications of this radiographic technique can be considered in two major areas: orthodontics and orthognatic surgery (Whaites, 2007).

Lateral cephalometry and orthodontics

Since the introduction of lateral cephalometric radiograph (LCR) (also denoted as "lateral cephalogram", "lateral cephalometry" or "lateral teleradiograph") in 1931, this radiograph and its related analysis has become a standard tool in orthodontic assessment and treatment planning (AlBarakati *et al*., 2012; Devereux *et al*., 2011; Nijkamp *et al*., 2008). Lateral cephalogram is different from a lateral skull view by the standardized projection geometry using a cephalostat, to enable standardized measurements of jaw bones, teeth and skeletal relationships. Apart from lateral cephalometry, posterior-anterior (PA) projections can also be carried out using standardized projection geometry, particularly when skull asymmetry does apply. However, the indication for these PA cephalograms is far below that of the lateral ones, and so not very much used in Orthodontics. The present review, will therefore only focus on lateral cephalograms. Indeed, nowadays, orthodontic treatment is performed in many children in Europe, with many of them receiving a lateral cephalogram during the initial diagnostic phase and many also later on, at the end of the treatment period.

Notwithstanding the fact that it is widely used, the real value of lateral cephalometry for the diagnosis and planning of the orthodontic treatment remains uncertain (Bourriau *et al*., 2012; Devereux *et al*., 2011; Nijkamp *et al*., 2008, Pae *et al*., 2001; Bruks *et al*., 1999; Atchison *et al*., 1991). Some authors stated that in many instances an adequate orthodontic diagnosis and treatment plan cannot be done without comparing cephalograms before and after orthodontic treatment. For that reason a lateral cephalogram is needed. They reinforced by stating that to treat skeletal malocclusions without a cephalometric radiograph is a serious error (Graber and Vanarsdall, 1994). However, Atchison *et al.* in 1991, reported that many radiographic techniques used in orthodontics are often not useful or are ineffective. According to Atchison *et al*., approximately three quarters of the radiographs exposed for orthodontic treatment purposes did not provide unexpected information which might lead to a change in the orthodontic diagnosis or treatment planning. In 1992, the same authors stated that the decision of taking a cephalogram prior to orthodontic treatment may be influenced by several factors, such as the suspicion by the clinician of underlying disease or medico legal reasons. According to the European Commission guidelines on radiation protection in dental radiology in 2004, only a small percentage of diagnosis and treatment plan changed after evaluating radiographs, alternating from 16% to 37% and 4% to 20% respectively (European Commission, 2004).

Dose reduction in lateral cephalometric radiography

The International Commission on Radiological Protection (ICRP) recommends that any practice involving ionizing radiation, or irradiation of patients with ionizing radiation, should be justified in relation to other diagnostic methods and produces a positive benefit to the patient (ICRP, 2007). The benefit should overcome any possible risk of damage that may occur associated with the use of ionizing radiation, taking into account social and economic factors, among others. The appropriate justification and imaging technique selection is also crucial in orthodontics. That is due to the fact that the patients are usually children and because the treatment period is usually 18 months or more. Radiographs are often taken at different time intervals during treatment, and young children are more vulnerable to radiation exposure (Tsuji *et al.*, 2006). Therefore, it is a basic premise of radiological practice that patient exposure should be kept "As Low As Reasonably Achievable" (ALARA principle), while at the same time producing images of sufficient diagnostic quality.

Dose reduction in lateral cephalometric radiography may be achieved by several means, which include:

- reduction of the field by collimating the beam to shield the thyroid gland and/or the brain tissue;
- use of collar shielding for the thyroid gland;
- using a more sensitive detector than conventional film, such as a photo-stimulable phosphor plate or a direct-digital scanning system;
- remove the anti-scatter grid;
- introduction of the air-gap technique;
- lowering the mAs yields the lowest effective dose and is therefore preferred;

- use of distance: For a point source of radiation, the dose rate falls off as the inverse of the square of the distance from the source. A true teleradiographic cephalostat would introduce a gap of 4 meters between the head of the patient and the x-ray source, because the radiation dose is reduced exponentially with increased distance. An equivalent dose of 1 intra-oral radiograph can be reached at 4 meters (ICRP, 2007; Tsiji *et al*., 2006; European Commission, 2004; Gijbels *et al*.,2003). Kaeppler *et al.,* 2007, referred that the most frequently used kilovoltage is of 70kV. The use of a digital imaging receptor (phosphor-stimulated computed plates) can also substantially reduce radiation exposure, when compared to conventional film radiography (Chen *et al*., 2004; Lim and Foong, 1997; Seki and Okano, 1993). At some important organs of the head and neck region, the absorbed dose from conventional radiography was approximately 2-fold higher than for the digital radiography. On the side of the head closer to the tube, Visser *et al.* in 2001, measured 81 mGy *versus* 34 mGy at the level of the lens of the eye, 103 mGy *versus* 45 mGy at the parotid gland, 53 mGy *versus* 34 mGy at the level of the submandibular gland, and 3 mGy *versus* 2 mGy at the level of the thyroid gland. The absorbed dose was about 9 times less on the side of the head nearer to the film than on the other side.

Digital image receptors can be classed as indirect (using phosphor-stimulated computed plates) and direct (using charged couple device-CCD), according to whether the receptor is physically linked to the computer and is capable of converting the ionising radiation into electrical signals directly.

CCD sensors are relatively small. Large ones, as large as a patient's head would be very expensive and difficult to make. Therefore, in the direct digital (CCD) imaging method the head is "scanned" rather than imaged using a one-shot approach (Gijbels *et al*., 2001). By contrast, the one shot approach can be applied with the indirect digital (phosphor plate) technique, giving it the advantages of reducing exposure time and therefore also minimising movement artefact (Chen *et al*., 2004).

Beam Collimation

Beam collimation is recommended by the European guidelines on radiation protection in dental radiology in order to restrict the irradiated field to the minimum area required for diagnosis. In the past, this was performed with true teleradiographic machines, but cephalometric arms in modern multimodal units often have little collimation potential. Wedge collimation is possible but not available on any kind of digital cephalometric equipments. Gijbels *et al*., 2003, suggested that the use of a wedge-shaped collimator mounted on the X-ray tube could reduce the dose to more than 40%. Tsuji *et al*. in 2006, suggested a triangle-shaped collimation to reduce the effective dose to the thyroid gland and also avoid scatter radiation (Figure 10).

Later, in 2012, Lee *et al*, advocated a dose reduction of approximately 60%.

Radiation protection is especially important for children. Some authors, state that since the brain and thyroid receive high radiation doses, wedge-shaped collimation should be considered (Gijbels *et al*., 2001).

Radiation hazards from cephalometry examinations have been reported since the fifties (Tyndall *et al*., 1988). In orthodontics the area of interest is the facial skeleton, which is situated below the level of the base of the skull

(European Commission, 2004). Imaging structures superior to the superior orbital rim, posterior to the occipital condyles, and inferior to the hyoid bone are clinically unnecessary (Mupparapu, 2005). However, some authors believe that beam collimators do not ensure complete protection and also involve a major change with high costs in cephalometric equipment (Sansare *et al.*, 2011). Moreover in some machines this modification is not possible. Besides the known advantages of using beam collimation, its use in orthodontics is not a current practice. Hoogeveen *et al.*, in 2014, suggested two reasons for that, one is due to anatomical variability of the area below the mandible and the fact that the use of wedge collimation covers the cervical vertebrae, disabling the determination of bone maturation. Another reason is because these collimators were not designed for today's combination panoramic–cephalometric imaging systems.

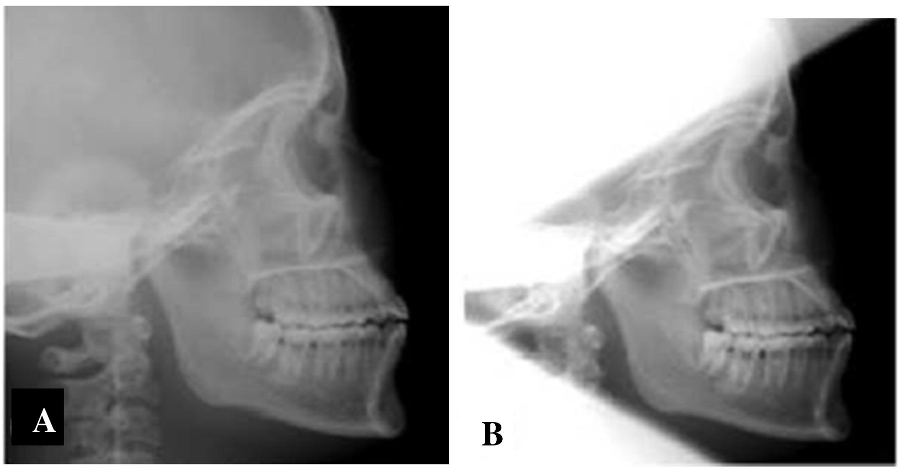

Figure 10. Images performed without (A) and with (B) a triangular shaped-collimation (Tsuji et al., 2006).

Therefore, Hoogeveen *et al.*, suggest the use of an “anatomically shaped cranial collimator” (ACC) (Figure 11). It should be attached to the cephalostat and shield the cranial area of the skull. This ACC produced a smaller dose reduction than previously reported for wedge-shaped collimators, with a reduction of 27–35%. This collimator does not protect the thyroid gland and thyroid shielding is recommended.

Figure 11. Patient in cephalostat and radiography with anatomically shaped cranial collimator (ACC) attached, proposed by Hoogeveen et al., 2014.

Thyroid shielding

The thyroid gland is one of the most radiosensitive organs in the head and neck region. It is often exposed in cephalometric radiography, if the beam is not collimated (European Commission, 2004). Despite the fact that the amount of radiation needed to cause thyroid cancer is big, it is advisable to reduce radiation exposure, especially in children (Sinnot *et al*., 2010). Lead collar thyroid shielding is currently the most efficient way to reduce radiation to the thyroid gland. Although, using a lead collar for orthodontic/orthognathic radiographs can partly or fully cover the soft tissues of the lower chin contour. The chin and soft tissue profile are needed for evaluation in the radiograph and must not be obscured by a thyroid shield (Sansare *et al*., 2011). Taking into account, as previously stated, that this radiograph is often taken in children or young adults, who have greater risk of radiation induced thyroid cancer than older individuals, the use of thyroid collar is strongly encouraged. However, in young patients it can be difficult to use the collar since it may obscure the soft tissue contour of the mandible leading to repeat radiographs or retakes. Collimating the beam does not completely protect the thyroid gland due to rays which are backscattered, due to secondary radiation, and due to unfocused primary rays (Sansare *et al*., 2011). Nevertheless, it is still safer trying to avoid thyroid exposure in the first place by using appropriate beam collimation (Sansare *et al*., 2011; European Commission, 2004).

Orthodontic diagnostic guidelines

The radiographs commonly used for an initial assessment, during and after orthodontic treatment include a panoramic radiograph and a lateral cephalogram. The selection of the adequate radiographic technique should be based on clinical common sense, taking into account patient's age and stage of treatment. Although in evidence based dentistry, guidelines other than clinical common sense should exist. It is crucial that the radiographs contribute to add relevant information that could not be obtained by other diagnostic procedures such as medical and dental history, dental casts or photographs (European Commission, 2004; Bruks *et al.*, 1999). In 1999 Bruks *et al.*, reported that only a small percentage of the provisional orthodontic treatment plan made without cephalometry was changed after the clinicians evaluated the cephalometric radiographs. Back in 1992, Atchison *et al.*, concluded that the orthodontists involved in their study ordered radiographs in most of the cases for medico-legal proposes, so they proposed an algorithm in an attempt to suggest what unnecessary radiographs were (Atchinson *et al.*, 1992). In many European countries, prior to starting orthodontic treatment, records of the patient such as dental cast, extra and intra-oral photographs, panoramic and cephalometric radiographs are collected (Atchison *et al.*, 1991; Nijkamp *et al.*, 2008). Some authors inferred that clinical examinations and dental casts should be adequate in 55% of the cases to plan orthodontic treatment, demonstrating that panoramic and cephalometric radiography were unnecessary for making the treatment plan (Bruks *et al.*, 1999).

This controversy is also present in orthodontic textbooks where selection criteria and guidelines for orthodontic radiographs are not referred to, while the available techniques are described, leaving open the interpretation of when to use radiography, or even advocating it for general use (Atchison *et al.*, 1991).

Therefore, guidelines for orthodontic radiographs should be created, to identify the optimal clinical circumstances for ordering radiographs necessary for the diagnosis and treatment planning (Issacson and Thom, 2001; Atchison *et al.*, 1992). Although, the creation of orthodontic radiographic selection criteria are difficult due to the heterogeneity of patients treated (Nijkamp *et al.*, 2008).

In 2004, the European guidelines on radiation protection in dental radiology recommended the use of cephalometry in specific situations (Figure 12):

- At the end of functional appliance treatment to see the position to which the lower anterior teeth have been proclined.
- At the end of presurgical treatment for orthognathic cases.
- Just prior to the end of active fixed appliance treatment to assess the position of lower incisors.

When assessing lower incisors position, lateral cephalogram is endorsed if the information is believed to change the orthodontist's decision on their finishing or retention mechanics. In some occasions, lateral cephalogram can change some of aspects of the treatment plan, such as teeth extraction and anchorage features. Although, after evaluating orthodontic radiographs, the diagnosis and treatment plan may not be changed (European Commission, 2004).

Restricted guidelines for orthodontic radiographs of the British Orthodontic Society have been created stating as selection criteria for cephalometry:

- patients with skeletal discrepancy when functional appliances or fixed appliances will be used for labio-lingual movement of the incisors;
- patients with a moderate skeletal discrepancy treated with fixed appliances who are being followed at a teaching environment;
- assessment of unerupted, malformed or misplaced teeth.

Other clinical indications for cephalometry that are not listed above should have a clear justification (Issacson and Thom, 2001).

The European Society of Lingual Orthodontics, in 2010, in line with the School of Orthodontist of the Portuguese Dental Association states that cephalometry should be performed before and after treatment. There is still lack of scientific evidence about the validity and reliability of cephalometric imaging for orthodontic treatment planning. Till present neither cost-benefit analysis, nor evidence about the benefit in relation of treatment time reduction, quality performance or prediction of results have been demonstrated (Nijkamp *et al.*, 2008).

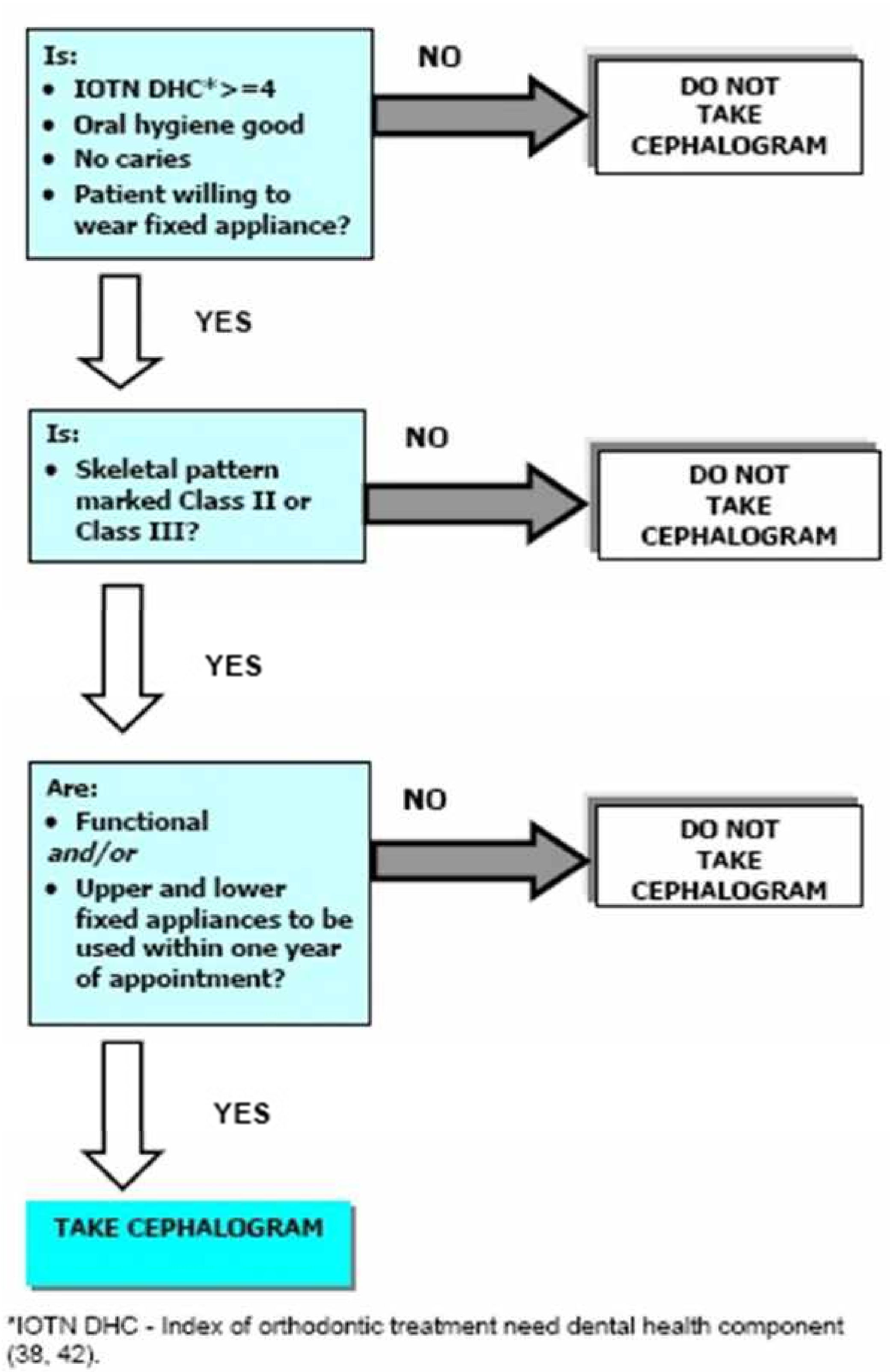

Figure 12. Flow chart showing clinical decision making, regarding lateral cephalograms (European Commission, 2004).

Technique and equipment

Cephalometry produces standardized images of the entire head and a portion of the cervical spine. It is used to identify skeletal and dental landmarks for orthodontic and craniofacial analysis (Chien *et al.*, 2009). It is a standardized and reproducible lateral skull radiograph used to assess the relationship of teeth to the jaws and the jaws to the facial skeleton (Whaites, 2007). The fact that this is a standardized technique is of extreme importance. It is sometimes necessary to perform these radiographs at different periods of time during the orthodontic treatment. A comparison is possible by superimposing the cephalometry tracings.

This technique requires three components: 1) a fixed X-ray point source, 2) a cephalostat where the patient's head is fixed at three points (external auditory meatus bilaterally and bridge of the nose), and 3) an image receptor (Athanasios and Athanasiou, 1995; Graber and Vanarsdall, 1994). The sagittal plane of the patient should be perpendicular to the central ray of the beam and parallel to the plane of the image receptor. The Frankfort plane should be horizontal. The patient is positioned with one side toward the image receptor, conventionally it is the left side which should be nearest to the image receptor. Patient should bite in centric occlusion position and the lips should be relaxed (Albarakati *et al.*, 2012; Athanasios and Athanasiou, 1995; Moyers, 1988).

Exact superimposition of the right and left sides is impossible due to magnification of the structures further away from image receptor and the slightly lesser magnification of the structures nearer to the image receptor. Structures close to the midsagittal plane should be nearly exactly superimposed. Bilateral structures near to the midsagittal plane show less discrepancy in size compared with bilateral structures further away from the midsagittal plane (Bourriau *et al.*, 2012; Duarte *et al*., 2009; White and Paroah, 2009; Whaites, 2007; European Commission, 2004; National Radiological Protection Board, 2001; Ahlqvist *et al*., 1986). When lateral cephalometry was used for the very first time, the distance from the x-ray source to the film was much greater, and conversely the magnification was smaller being of 3% at a distance of 5 meters, 3.5% at a distance of 4 meters and 11.5% at a distance of 1.5 meters (Bourriau *et al*., 2012; Ahlqvist *et al*., 1986). Nowadays, only the equipments with a focus-to-film distance of 1.5 to 1.8 meters are in use. Although to minimise the magnification effects, the focus-to-film distance should be greater than 1 meter and ideally within the range 1.5 to 1.8 meters. There is always a minimal enlargement that still creates discrepancies between left/middle/right sides of the skull. The equipment should provide a perfect alignment between patient, X-ray source and image receptor, to reduce errors on the radiography. A light beam diaphragm, or other suitable means, should be used to help collimate the x-ray beam to include only the area that would be used for orthodontic proposes (National Radiological Protection Board, 2001).

Visualisation of the soft tissue profile is necessary, therefore, an aluminium wedge filter should be provided at the anterior part of the x-ray tube head between patient and the X-ray tube, to absorb some radiation (White and Paroah, 2009; Whaites, 2007). The aluminium wedge filter attenuates the X-ray beam in the region of the facial soft tissues (Whaites, 2007). In the beginning of cephalometry, two images were taken at different kilovoltages, first one to visualize soft tissue the second one to visualize hard tissues. Nowadays a soft tissue filter is used to overcome this double irradiation to the patient.

Analysis of the cephalograms

After obtaining a good quality lateral cephalogram, it is possible to perform a cephalometric analysis, which allows angle and linear measurements to be made, including:

- the outline and inclination of the anterior teeth;
- the positional relationship of the mandibular and maxillary dental bases to the cranial base;
- the positional relationship between maxillary and mandibular dental bases;
- the relationship between the bones of the skull and the soft tissue profile of the face (Bourriau *et al.*, 2012; Deveraux *et al.*, 2011; Sánchez and Filho, 2009; Arpoen *et al.*, 2008; Whaites, 2007; McIntyre and Mossey, 2003).

In 1951, Downs published the first article on cephalometric analysis. Until recently, cephalometric analysis could only be done manually and laboriously. A sheet of tracing paper or transparent acetate was placed directly over the radiograph on top of a lightbox, and the anatomical landmarks are identified using pencil or pen onto the paper or acetate, producing the “orthodontic tracing”. After this step, the various angles and all the measurements and other calculations are performed manually from the tracing. Nowadays, there are numerous computer software programmes available that allow a faster identification of the anatomical landmarks, calculating the data and indicating the most suitable treatment plan. The software requires a digital image, which may be digitally acquired radiographic image or obtained after digitizing a conventional film radiograph on an optical scanner (Lim and Foong, 1997). There are many analyses available and the choice may be based on clinician's preference or patients’ conditions. Some authors compared the accuracy of digital cephalometric measurements with the hand-tracing method (Bruntz *et al*., 2006; Santoro *et al*.,2006; Chen *et al*., 2004). Computerized cephalometric measurement using direct digital imaging is better than digitized conventional radiographs. However the principle of the digital cephalometric analysis is the same. The observer needs to identify each landmark. All the values are then compared with reference values. In 1982, De Abreu found a lack of agreement in the four cephalometric analyses he studied. Despite his observation, few authors have afterwards investigated the importance and usefulness of the different existing landmarks (Chen *et al*., 2004).

Definitions of anatomical landmarks used in 2D lateral cephalometry

Anatomical points or landmarks identified on lateral cephalometric radiographs to allow precise linear and angular measurements. The points are recorder either on an overlying sheet of paper or acetate or digitally. The definition of the main cephalometric landmarks is listed below (Figure 13):

- Porion (Po): Most superior point of left external auditory meatus.
- Sella (S): Geometric centre of the *sella turcica.*
- Orbitale (Or): Most inferior point of the infraorbital margin.
- Nasion (N): Most anterior point on frontonasal suture.
- Basion (B): Lowest point on anterior rim of foramen magnum.
- Pogonion (Pog): Most anterior midpoint of the bony chin.
- Gnathion (Gn): Most anterior and inferior point on the bony outline of the chin, situated equidistant from pogonion and menton.
- Menton (Me): Lowest point on the bony outline of the mandibular symphysis.

- Gonion (Go): Point on curvature of the angle of the mandible located by bisecting the angle formed by lines tangent to the posterior ramus and the inferior border of the mandible.
- Anterior Nasal Spine (ANS): The tip of the anterior nasal spine.
- Posterior Nasal Spine (PNS): The tip of the posterior spine of the palatine bone in the hard palate.
- Point A (A): Deepest midline point between the anterior nasal spine and prosthion.
- Prostion (Pr): Most anterior point of the alveolar crest in the premaxilla, usually between the upper central incisors.
- Point B (B): Deepest point in the bony outline between the infradental and the Pogonion.
- Infradental (Id): Most anterior point of the alveolar crest, situated below the lower central incisors.

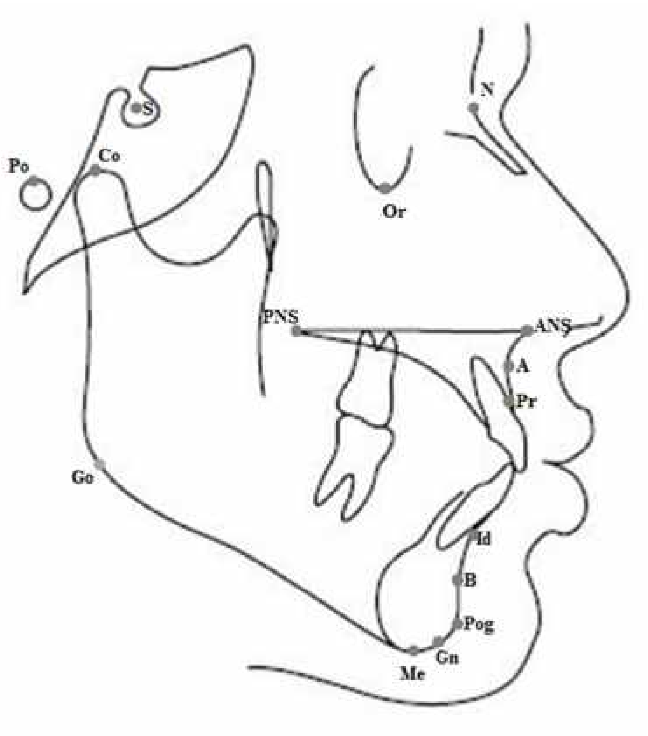

Figure 13. Cephalometric tracing of a lateral cephalometric radiography showing the main cephalometric landmarks.

Accuracy of cephalometric measurements

The accuracy of cephalometric measurements is of great interest. Many studies have been published on the errors associated with landmark identification, errors arising from the registration of landmarks, and errors due to measurement procedures (Chen *et al*., 2004). Errors due to the projection of a three-dimensional object on a two-dimensional film have been studied less extensively (Albarakati *et al*., 2012; Bruks *et al*., 1999; Ahlqvist *et al*., 1986). Few studies, however, have attempted to assess the accuracy of cephalometric measurements as applied three-dimensionally (3D) because of known intrinsic limitations of these images, such as distortion and magnification. Lateral cephalograms have intrinsic limitations that result in distorted images, enlarged in some areas and reduced in others.

When doing the tracing, precise landmark identification is important for the diagnosis and treatment plan (Sánchez and Filho, 2009). A trained person should do the tracing, it can be done by orthodontists or dentomaxillofacial radiologists.

Measurements based on cephalometry may involve errors, which are classified by Baumrind and Frantz as “errors of projection” and “errors of identification” (Baumrind and Frantz, 1971).

- **Projection errors**

Projection errors result from imaging 3D structures in a two dimensional (2D) radiographic image. Projection magnification of objects is the result of varying the distance between individual structures and the film or imaging receptor, resulting in variable enlargement of some structures depending on proximity to the image receptor. The positioning of the patient’s head is also of extreme importance, since a slight rotation of the head may lead to distortion and errors in linear and angulation measurements. Ahlqvist *et al.* (1986), reported that a +/- 5º of head rotation from the ideal position resulted in an insignificant error, however if the head rotation increased the probability of an error occurring was greater and may become significant even at rotations of a few degrees more than +/-5º.

- **Identification errors**

Errors of identification are those that can occur in the landmark identification process, such as the porion, condylion, orbitale, basion, gonion, anterior and posterior nasal spine, and lower incisor apex. Adenwalla *et al.* in 1988, studied the reliability of the Po and Co identification on lateral cephalogram, and concluded that these two anatomical landmarks could not be accurately located on lateral cephalograms taken with the patient in the mouth closed position. Therefore, they suggested an open-mouth cephalogram should be taken and superimposed on the respective cephalogram in the centric occlusion position to obtain the most accurate and reliable measurements. The main problem with these two landmarks is that the ear rods are superimposed on the patient skull region of interest. These errors are due to overlapping structures that are superimposed on landmarks of interest, as well as the resolution and quality of the acquired images. Inherent cephalometric errors can lead to variations in orthodontic and surgical treatment planning (Chien *et al*., 2009). The errors in cephalometric analysis are composed of systematic errors and random errors. The latter involves tracing, landmark identification, and measurements errors (Chen *et al*., 2004).

Previously, landmark identification and measurements were done by tracing outlines on the radiograph and measuring by hand. Nowadays, many cephalometric analysis software programmes are available and only landmark identification has to be done by hand whilst the analysis is done automatically. This means that identification errors may still occur. Computer-aided cephalometric analysis can totally eliminate the mechanical errors in drawing lines between landmarks and in measurements with a protractor, although it does not introduce more measurement errors than hand tracing, as long as the landmarks are identified manually (Chen *et al*., 2004). Digitally acquired cephalometric imaging presents numerous advantages, as the possibility of enhancement imaging techniques that allow improved landmark identification, faster cephalometric data acquisition and analysis, more efficient storage and archiving and easier transfer of the image to distant sites.

Recently, automatic cephalometric landmark identification is possible using cephalometric software can be used directly on a digitally acquired image or after digitizing a conventional film with a scanner or a digital camera (AlBarakati *et al*., 2012). For this modality the mean success rate for identifying landmark positions was 88% with a range of 77% to 100% (Tanikawa *et al*., 2009).

OBJECTIVES AND HYPOTHESIS

The overall aim of this thesis was to validate the accuracy and reliability of 2D cephalometric radiograph in orthodontic diagnosis and treatment planning. The outcome of this study is mandatory to further judge any potential and additional role of 3D cephalometric analysis.

The various chapters and topics address the following hypotheses:

1. 2D cephalometrics suffers a poor accuracy when compared to real skull analysis (*Chapter 2*).
2. 2D cephalometrics has a poor intra- and inter-observer variability, thus influencing planning and treatment decisions (*Chapter 3*).
3. Landmark identification on the point Sella as a reduced variability, and does interfere with the angles SNA and SNB (*Chapter 4*).
4. The availability of a 2D lateral cephalometric radiograph influences the orthodontic treatment plan and decision in some but not all cases. (*Chapter 5*).

CHAPTER 1. Systematic Review

Ana R Durão, Pisha Pittayapat, Ivete B Rockenbach, Raphael Olszewski, Suk Ng, Afonso P Ferreira, Reinhilde Jacobs. **Validity of 2D lateral cephalometry in orthodontics: a systematic review.** Progress in Orthodontics 2013, 14:31 (20 September 2013) DOI: 10.1186/2196-1042-14-31.

SYSTEMATIC REVIEW

1.1 Introduction

The aim of this systematic review was to evaluate the available scientific literature and existing evidence about the validation of lateral cephalometric radiograph in orthodontics. This review also studied the accuracy and reliability of lateral cephalograms and its cephalometric analysis.

We did not attempt to evaluate the value of this radiographic technique for other purposes.

1.2 Information sources

A comprehensive electronic database search to identify relevant publications was conducted, and the reference lists in relevant articles were searched manually for additional literature. We set no language limitations, although we did not attempt to explore the informally published literature: conference proceedings and abstracts of research presented at conferences and dissertations. The following databases were searched: Ovid Medline (1946 to 11 January 2012), Scopus (to 11 January 2012) and Web of Science (1899 to 11 January 2012).

1.3 Observers

Two trained observers, participated in this study, the author and one other observer. Both are experienced dentomaxillofacial radiologists with an active academic research function.

1.4 Search strategy

We developed the search strategy with the help of an information specialist. The searches did not have a date limit and were not restricted to particular types of study design. The search strategy focused on the following terms:

Cephalometr* AND (orthodontic* OR "orthodontic treatment planning") AND (“efficacy” OR “reproducibility” OR “repeatability” OR “reliability” OR “accuracy” OR “validity” OR “validation” OR “precision” OR “variability” OR “efficiency” OR “comparison”) NOT ("Cone-Beam Computed Tomography" OR "Three-Dimensional imaging" OR "Cone Beam Computed Tomography" OR "Cone Beam CT" OR "Volumetric Computed Tomography" OR "Volume Computed Tomography" OR "Volume CT" OR "Volumetric CT" OR "Cone beam CT" OR "CBCT" OR "digital volume tomography" OR "DVT" OR "Spiral Computed Tomography" OR "Spiral Computer-Assisted Tomography" OR "Spiral Computerized Tomography" OR "spiral CT Scan" OR "spiral CT Scans" OR "Helical CT" OR "Helical CTS" OR "Helical Computed Tomography" OR "Spiral CAT Scan" OR "Spiral CAT Scans" OR “3D” OR “3-D” OR "three dimension*)

1.5 Study selection

At the first stage, the two reviewers independently screened the titles of the retrieved records, and only the titles related to 2D cephalometry, radiographs for orthodontic treatment and tracings were included. Next, the abstracts of the retrieved publications were read by the two observers and categorised according to the study topic. An article had only to be justified by one observer to be included for the second selection phase. Two articles of interest in languages other than English were included. Of these included, one article was written in Portuguese and another in French. Eligibility of potential articles was determined by applying the following inclusion criteria to the article abstracts: (1) technical efficacy, (2) diagnostic accuracy efficacy, (3) diagnostic thinking efficacy, (4) therapeutic efficacy, (5) patient outcome efficacy or any combination of the previous items as published by Fryback and Thornbury in 1991. The other inclusion criteria were (1) accuracy, (2) reliability, (3) validity of lateral cephalometric radiograph, (4) landmark identification on tracings (intra- and inter-observer errors) and (5) the effect of using 2D cephalometry on the orthodontic treatment plan.

Diagnostic accuracy efficacy was defined as follows:

1. Observer performance expressed as overall agreement, kappa index or correlation coefficients
2. Diagnostic accuracy as percentage of correct landmark identification and further tracing analysis, validity and effectiveness of cephalometry in orthodontic treatment planning
3. Sensitivity, specificity or predictive values of landmark identification

Diagnostic thinking efficacy was defined as follows:

1. Percentage of cases in a series in which images were judged ‘helpful’ for the diagnosis
2. Difference in clinicians' subjective estimated diagnosis probabilities before and after evaluation of the cephalogram

Therapeutic efficacy was defined as follows:

1. Percentage of times the image was judged helpful in planning management of the patients in a case series
2. Percentage of times therapy-planned pre-visualization of a lateral cephalogram needed to be changed after the image information was obtained
3. Percentage of times clinicians prospectively stated therapeutic choices needed to be changed after evaluating a cephalogram
4. Whether different analyses lead to different decisions on treatment planning
5. Intra- and inter-observer identification errors
6. Reliability of landmark identification

The analysis had to be based on primary materials or comprise a review on efficacy. When an abstract was considered by at least one author to be relevant, it was read in full text. At the second stage, the full texts were retrieved and critically examined. Reference lists of publications that had been found to be relevant in the first stage were hand-searched, and articles containing the words 'cephalometry', 'lateral cephalometric radiography', together with 'treatment planning', 'orthodontic radiographs', 'landmark identification' and 'error' were selected. Book chapters and reviews were excluded since the aim of this systematic review was to evaluate primary studies.

1.6 Data extraction

Data was extracted with the aid of protocol 1 (Table 1.1). It was established by reading the relevant literature on how to critically evaluate studies about diagnostic methods. To minimise bias, two observers independently evaluated the quality and validity of original studies according to the quality assessment of diagnostic accuracy studies tool using protocol 2 (quality assessment of studies of diagnostic accuracy included in systematic reviews - QUADAS) (Table 1.2) (Whiting *et al,* 2003). When there was any disagreement concerning the relevance of an article, it was resolved by a discussion between the two reviewers. Each observer presented their arguments, and further discussion was held until a consensus was reached. Before the assessment, the protocols were tested for ten publications. A further five publications were read to calibrate the two reviewers regarding the criteria in protocol 2. Only publications that were found to be relevant to the reviewer in both protocols 1 (diagnostic efficacy) and 2 (level of evidence) were ultimately included. The quality and internal validity (level of evidence) of each publication was judged to be high, moderate or low according to the criteria in the following subsection.

Table 1.1. Protocol 1, Selection for inclusion of publications.

First author: **Title:** **Journal; Year; Volume; Pages:**			
		Yes	**No**
1.	Is there a well-defined hypothesis?		
2.	Are the accuracy, reliability, validity of cephalometry studied?		
3.	Is the contribution of cephalometry in determining the treatment plan evaluated?		
4.	Reliability of landmark identification in cephalometry?		
5.	Errors that occur in cephalometry?		
6.	What is the level according to Fryback and Thornbury?		
7.	Is the publication relevant for the review?		

Table 1.2. Protocol 2, based on the QUADAS-2 tool for evaluation of methodology of included studies.

Observer initials _______________ Date ________

Paper nº / ______
First author; Title; Journal; Year; Volume; Pages
____

1. Are the results of the study valid?

Yes No Unclear
2. Was the spectrum of patient's representative of the patients who perform orthodontic treatment?
Yes No Unclear
3. Were selection criteria clearly described?
Yes No Unclear
4. Is the reference standard likely to correctly classify the target condition?
Yes No Unclear
5. Were the methods for performing the radiographic examination described in sufficient detail to permit replication?
Yes No Unclear
6. Was the execution of the reference standard described in sufficient detail to permit its replication?
Yes No Unclear
7. Were the index test results interpreted without knowledge of the results of the reference standard?
Yes No Unclear
8. Were the reference standard results interpreted without knowledge of the results of the index test?
Yes No Unclear
9. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
Yes No Unclear
10. Were uninterpreTable / intermediate test results reported?
Yes No Unclear
11. Were withdrawals from the study explained?
Yes No Unclear
12. Was the number of observers sufficient to evaluate the influence of observer reproducibility and diagnostic efficacy?
Yes No Unclear
13. Was observer reproducibility described?
Yes No Unclear
14. Were appropriate results presented (percentage of correct diagnosis, sensitivity, specificity, predictive values, measurements of ROC, likelihood ratios, or other relevant measurements) and were these calculated appropriately?
Yes No Unclear
Comments

Levels of evidence and criteria for evidence synthesis:

- ***High level of evidence***

A study was classified with high level of evidence if it fulfilled all of the following criteria:

- There was an independent blind comparison between test and reference methods.
- The population was described so that the status, prevalence and severity of the condition were clear. The spectrum of patients was similar to the spectrum of patients on whom the test method will be applied in clinical practice.
- The results of the test method being evaluated did not influence the decision to perform the reference method(s).
- Test and reference methods were well described concerning technique and implementation.
- The judgments (observations and measurements) were well described considering diagnostic criteria applied and information and instructions to the observers.
- The reproducibility of the test method was described for one observer (intra-observer performance) as well as for several (minimum 3) observers (inter-observer performance).
- The results were presented in terms of relevant data needed for necessary calculations.

- ***Moderate level of evidence***

A study was assessed to have a moderate level of evidence if any of the above criteria were not met. On the other hand, the study was assessed not to have deficits that are described below for studies with a low level of evidence.

- ***Low level of evidence***

A study was assessed to have a low level of evidence if it met any of the following criteria:

- The evaluation of the test and reference methods was non-independent.
- The population was not clearly described, and the spectrum of patients was distorted.
- The results of the test method influenced the decision to perform the reference method.
- The test or the reference method or both were not satisfactorily described.
- The judgments were not well described.
- The reproducibility of the test method was not described or was described for only one observer.
- The results could have a systematic bias.
- The results were not presented in a way that allowed efficacy calculations to be made.

Rating conclusions according to evidence grade

The scientific evidence of a conclusion on diagnostic efficacy was judged to be strong, moderately strong, limited or insufficient depending on the quality and internal validity (level of evidence) of the publications assessed (CBEM, Jaeschke *et al.,* 1994).

- Strong research-based evidence: at least two of the publications or a systematic review must have a high-level of evidence.
- Moderately strong research-based evidence: one of the publications must have a high level of evidence and two more of the publications must have a moderate level of evidence.
- Limited research-based evidence: at least two of the publications must have a moderate level of evidence.
- Insufficient research-based evidence: scientific evidence is insufficient or lacking according to the criteria defined in the present study.

1.7 Synthesis of evidence

The results of this review were described narratively. No meta-analyses were attempted because of lack of original studies.

1.8 Results

The number of articles reviewed in each phase to perform this systematic review is presented in the PRISMA flow diagram (Figure 1.1) (Moher *et al*., 2009). The initial search revealed 784 articles listed in Medline (Ovid), 1,034 in Scopus and

264 articles in the Web of Science. The second stage of the search protocol was to retrieve the reference lists of the selected articles, which yielded 14 additional articles of interest. After excluding 1,128 duplicates, 968 articles remained for review. In the first phase selection, the observers screened the articles by reading titles and abstracts. Articles that were not eligible because of irrelevant aims and were not directly related to this systematic review were excluded, thus 203 articles remained for further reading. Thirty-five articles were assessed for eligibility.

After screening all the articles using protocols 1 and 2, 17 articles met the inclusion criteria and were selected for qualitative synthesis and appraised to present some level of evidence. All articles that remained after screening passed the qualitative synthesis.

These 17 articles were categorised by topics as follows: 7 studies on the role of cephalometry on the orthodontic treatment planning, 8 studies on cephalometric measurements and landmark identification and 2 studies on cephalometric analysis.

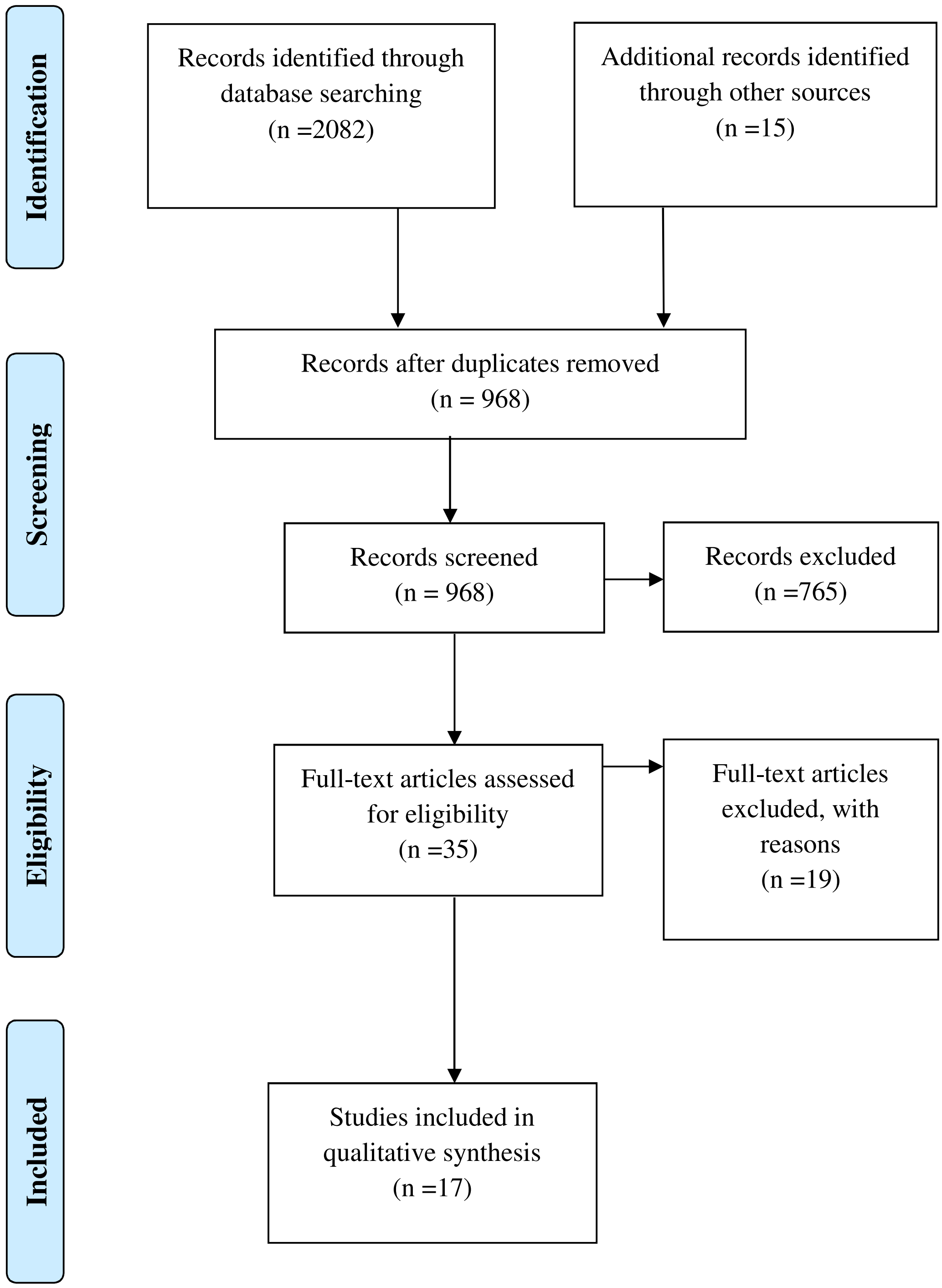

Figure 1.1. Methodology followed in the article selection process (adapted from: Moher et al., 2009).

1.8.1 Role of cephalometry on the orthodontic treatment planning

Seven articles related to the importance and contribution of cephalometry to orthodontic treatment planning was found (Table 1.3). Six of the publications were found to have low levels of evidence (Deveraux *et al.,* 2011; Nijkamp *et al.,* 2008; Bruks *et al.,* 1999; Atchinson *et al.,* 1992; Atchinson *et al.,* 1991; Silling *et al.,* 1979) and one classified as moderate level of evidence (Pae *et al.,* 2001).

Table 1.3. Publications related to the importance and contribution of cephalometry on the orthodontic treatment planning.

Authors (year)	Aim of the study	Observers	Subjects	Design of the study	Statistical method	Results according to authors	Level of evidence
Silling *et al.*, 1979	Assess usefulness of cephalometric analysis	24 orthodontists	6 patients	Stratified random design: 12 orthodontists analysed 6 patients with cephalograms and 12 orthodontists studied 6 patients without cephalogram	Not referred	Class I patient: disagreement on extractions, anchorage and growth potential decisions No need for lateral cephalometry, except for atypical class II division 1 patients, by 4 orthodontists Anchorage problems S between patients with and without lateral cephalogram	Low

Bruks *et al*., 1999	Evaluation of lateral cephalometric and panoramic radiography	4 dentists and senior orthodontist	70 patients	Clinical evaluations and treatment plan by 4 dentists:	Descriptive statistics and statistical analyses with computer software. Kruskal-Wallis test to evaluate differences between groups	Impact on diagnosis relating to the ordering sequence of cephalogram: first choice, 68%; second choice, 73%; third choice, 80%	Low
				1. Study casts + photographs 2. Adding radiographs		93% of cases: same treatment plan before and after radiographic analysis	
Pae *et al*., 2001	Examine the link between lateral cephalograms and occlusal trays	16 orthodontists	80 patients	T1: casts evaluated; T2 (1 week later): casts + lateral cephalograms	Rash model, regression plots, two-way ANOVA, *post hoc* multiple comparison Bonferroni and paired *t* test	Class II division 2 patients: 126 extractions planned at T1; 80 at T2 A lateral cephalogram influenced degree of severity, but not the difficulty of treatment	Moderate

Nijkamp *et al*., 2008	Influence of lateral cephalometry on treatment plan	10 post-graduate trainees and 4 orthodontists	48 patients	Randomised crossover design - T1: casts, T2 (1 month after): with lateral cephalometry and tracing, and T3 and T4 (repeated after 1 and 2 months)	Overall proportion of agreement	Consistency of treatment plan was NS between the use only of dental casts or with additional cephalometry Influence of cephalometrics on orthodontic treatment planning: NS	Low
Devereux *et al*., 2011	Influence of lateral cephalometry on treatment plan	114 orthodontists	6 patients	3 groups: (a) no lateral cephalogram and tracings, (b) some with lateral cephalogram and tracings and (c) all with lateral cephalogram and tracings	Chi-square and binary logistic regression	Treatment plan changed for extraction pattern (42.9%), anchorage reinforcement (24%) and decision to extract (19.7%) Class I patient: lateral cephalogram less times ordered. Only patients where treatment plan changed after its analysis	Low

						NS impact of cephalometrics on treatment plan	
Atchison *et al.*, 1991	Determine quantitatively the diagnosis and treatment plan information after radiograph evaluation	39 orthodontists	6 patients	A 2-h interview for diagnosis and treatment planning of 6 cases. Study cast, intra- and extra-oral photographs, tracing and clinical findings available. A radiograph only if judged helpful	Analysis of variance with repeated measurements and covariance, homogeneity value and descriptive statistics	98% of cases: at least one of the radiographs unproductive 3/4 of radiographs did not provide information to change diagnosis and treatment plan	Low
Atchison *et al.*, 1992	Identify selection criteria for ordering orthodontic radiographs	39 orthodontists	6 patients	A 2-h interview for diagnosis and treatment planning of 6 cases. Study cast, intra- and extra-oral photographs, tracing and clinical findings available	Not referred	14.4% of radiographs ordered for skeletal relationship of the jaws Lateral cephalograms accounted for 34% of required information	Low

	26% of all ordered radiographs produced modifications on diagnosis or treatment plan Pretreatment lateral cephalogram required in all patients needing orthodontic treatment

1.8.2 Cephalometric measurements and landmark identification

Only eight articles were selected as eligible in this category (Table 1.4). Five publications presented a moderate level of evidence (Kamoen *et al.,* 2001; Tng *et al.,* 1994; Haynes and Chau, 1993; Houston *et al.,* 1986; Baumrind and Frantz, 1971), while the other three were identified as having a low level of evidence (Bourriau *et al.,* 2012; Ahlqvist *et al.,* 1986; Kvam and Krogstad, 1969).

Table 1.4. Publications concerning landmark identification.

Authors (year)	Aim of the study	Observers	Subjects	Design of the study	Statistical method	Results according to authors	Level of evidence
Baumrind and Frantz, 1971	Quantification of errors in landmark identification Effects of errors on angular and linear measurements	5 observers	20 lateral skull radiographs	Observer identified 16 cephalometric landmarks on a transparent plastic template	Mean, standard deviation and standard errors	Least reliable landmarks: Gonion and lower incisor apex	Moderate
Kvam and Krogstad, 1969	Evaluation of measurements in lateral cephalograms. Assess influence of knowledge and impact of angular errors	18 observers	3 lateral skull radiographs	Hand cephalometric analysis made by each participant, 8 angles measured	Mean and standard deviation	16 out of 24 angular measurements: less variability in post-graduates than students In 7 measurements, no difference was observed Post-graduates' tracings used for diagnostic purposes Standard deviation of students greater than post-graduates	Low
Haynes and Chau, 1993	Evaluation of landmark identification on Delaire analysis	2 observers	28 lateral skull radiographs	Establish a co-ordinate system for measurement on tracings	Mean deviation	Intra-observer: NS differences between values of T1 and T2 tracings	Moderate

	Comparison with data of conventional cephalometry			Radiographs were traced twice by each observer (3 to 4 weeks)		Inter-observer: differences between the averaged mean values on tracings were NS for either *x* or *y* co-ordinates	
Ahlqvist *et al*., 1986	Study the magnitude of projection errors on measurements in cephalometry Study the effects of incorrect patient position on linear measurements	1 observer	A patient was modelled	Computer software designed to allow movement of model on the 3 axes. The magnitude of errors was studied by a diagram	Measurement errors studied by a diagram with the relative length of distances between modelled landmarks	Less than 1% error on length measurements if head is rotated up to 5° Head rotated more than 5° the error is increased	Low
Houston *et al*., 1986	Evaluate errors at various stages of measurements in cephalometric radiograph	4 observers	24 lateral cephalograms	2 radiographs of the same patient Radiographs traced on acetate sheet by each observer at T1/T2 (1-week interval)	Analysis of variance	Error variance is small (radiograph and tracing) when compared with the variance among groups SNA has a higher tracing variance than SNB due to the difficulty to identify point A	Moderate

Kamoen *et al*., 2001	Determine errors involved in landmark identification and its consequence to treatment results	4 observers	50 lateral cephalograms	Items studied: (1) accuracy of digitiser, (2) intra- and inter-observer digitising errors and (3) intra- and inter-observer tracing errors	(1) Levene's test for homogeneity of variances, (2) one-way ANOVA and (3) Levene's test for homogeneity	(1) NS variances of co-ordinates for landmark at different positions on the digitiser. (2) NS intra- and inter-observer differences in digitisation. (3) S differences in landmarks and in the same landmark on different cephalograms and between observers	Moderate
Tng *et al*., 1994	Evaluate the validity of dental and skeletal landmarks. Effect on angles and distances.	1 observer	2 lateral cephalograms of 30 dry skulls	Steel balls placed in 15 dental and skeletal landmarks Two radiographs taken with and without the markers and digitised. Measurements compared	Mean and standard deviation	7 out of 10 skeletal and 5 dental landmarks were NS ($p < 0.05$) 4 angles (SNA-SN/MnP, MxP/MnP and LI/MnP) and 3 distances (N-Me, MxP-Me and Lie to APg) were invalid ($p < 0.05$) Major errors in angles with dental landmarks	Moderate
Bourriau *et al*., 2012	Analyse the influence of film-object distance and type of receptor on landmark identification	53 orthodontists	4 lateral cephalograms of the same patient	19 cephalometric landmarks on each film	Mean	NS difference between 2 imaging receptors neither between 2 cephalograms achieved by 2 equipments ($p > 0.99$)	Low

		2 radiographs performed at an equipment with a 4-m arm and 2 in a 1.50-m arm equipment with 2 different imaging receptors (digital and indirect digital)	Results obtained by cephalometric analysis was judged: ‘very important’ for 20.5%, ‘important’ for 70%, ‘less important’ for 8% and ‘accessory’ for 1 participant

NS, non-significant; S, significant.

1.8.3 Cephalometric analysis

Two publications with low-level evidence were found (Abdullah *et al.,* 2006; De Abreu, 1982). The studies did not use any reference standards, and the number of observers was not stated. The study designs were also not clearly explained (Table 1.5).

Table 1.5. Publications on cephalometric analysis.

Authors (year)	**Aim of the study**	**Observers**	**Subjects**	**Design of the study**	**Statistical method**	**Results according to authors**	**Level of evidence**
De Abreu, 1982	Assessment criteria of unanimity for different cephalometric analyses	Not referred	129 patients	Diagnosis performed based on Ricketts, Steiner, Cervera and Coutand cephalometric analyses	Not referred	3 out of 61 cases with similar diagnosis. In 23 cases, 4 analyses achieved similar diagnosis. In 13 cases, 3 different diagnoses were obtained. In 8 cases, the diagnosis was different for class II and class III	Low
Abdullah *et al*., 2006	Examine accuracy and precision of Steiner analysis for changes on ANB angle, the Pg-NB distance and upper and lower incisor positions	Different orthodontists (not reference to the number)	275 patients	Radiographs traced and analysed by orthodontists according to the Steiner analysis Radiographs at the end of treatment (T2) were traced by one observer	Paired *t* test, mean and standard deviation	The predicted change in L1 (lower incisor) to NB was underestimated by 0.8 mm. Only the prediction for Pogonion and NB showed improvement of the precision (30%)	Low

1.9 Discussion

The validity, efficacy and contribution of cephalometry in orthodontic treatment planning remain questionable (Deveraux *et al.,* 2011). In 2002, 90% of orthodontists in the USA routinely performed cephalometric radiographs (Nijkamp *et al.,* 2008). This systematic review was performed to assess the validity and reliability of 2D lateral cephalometry used for orthodontic treatment planning as well as the errors that can occur on 2D tracing. Despite the abundant amount of articles found on lateral cephalometry (*n* = 968), it is surprising that the present systematic review could only identify very few studies (*n* = 17, 1.6%) on its validity and reliability. This finding underlines the need for the present study and is an important cross point, considering the fact that we are flooding into 3D cephalometric studies nowadays. Apart from our findings, 2D cephalometry has other specific limitations, such as orthognatic surgery, airway and growth assessment and skeletal maturation. In order to be included in this systematic review, publications had to satisfy pre-defined methodological criteria. Two protocols were used regarding the search strategy, one based on diagnostic methods and the second based on the QUADAS tool (Whiting *et al.,* 2003). The 'levels of evidence' for assessing the quality and internal quality of each publication included in this review - how well the study was designed, how reliable its results appeared to be and the extent to which it addressed the questions posed - were modified according to the Oxford Centre for Evidence-Based Medicine levels of evidence for diagnostic methods (CBEM, 2012). Only publications assessed to present a high or moderate level of evidence can form the basis for any scientific conclusions. Ten articles were identified as low level of evidence, five had moderate level and only one showed high level of evidence.

All retrieved articles, assessing the importance and contribution of lateral cephalometric radiograph in orthodontic treatment, concluded that there is no significant difference on treatment planning decision with or without the evaluation of the lateral cephalogram. However, it should be considered that the suitable studies in this review were based on small samples rather than large cohorts representing the entire population. In one study, the sample used was restricted (six patients) (Deveraux *et al.,* 2011). Furthermore, the short time lapse between observations in some studies did not allow a full washout effect, which could lead to the repetition of the results (Pae *et al.,* 2001; Atchison *et al.,* 1992; Atchison *et al.,* 1991). The latter bias is further strengthened by the fact that recognition factors were often included, e.g. the possibility of identifying patient by photographic visualisation as part of the examination. On the other hand, in one paper, only dental casts were presented to the observers, which might also lead to error since it does not mimic the clinical situation. Sample bias is also suspected based on the fact that selection of subjects is often poorly described or unclear (Deveraux *et al.,* 2011; Bruks *et al.,* 1999; Silling *et al.,* 1979), like the questions made to the observers that were not stated by any questionnaire (Bruks *et al.,* 1999), and in one article, observers were forced to choose yes/no answers, which again do not perfectly simulate the reality (Nijkamp *et al.,* 2008).

In the two articles by Atchison *et al*., there was the possibility to identify patients as well as sample size was very restricted (six patients). There was no repetition of the questionnaire to test the variability between answers (Atchison *et* *al.,* 1992; Atchison *et al.* 1991). When it comes to the validity and reliability of cephalometric analysis, several errors should be considered: landmark identification, tracing and measuring, and magnification of certain anatomical structures.

Landmarks placed in anatomically formed edges are easier to identify, while some landmarks placed on curves are more prone to error. The gonion and lower incisor apex are the least consistent landmarks (Baumrind and Frantz, 1971). Furthermore, landmarks such as point A have a higher variance than others like point B because of wider variation and anatomical localisation of point A (Houston *et al.*, 1986). Dental landmarks tend to have poorer validity than skeletal landmarks. Also, when landmarks are located on a curve like point A, point B or Pogonion, the error is larger (Tng *et al.,* 1994). The evidence shows that landmark identification is a great source of error in 2D lateral cephalometry (Kamoen *et al.*, 2001). Major errors in angles with dental landmarks may occur (Tng *et al.,* 1994). In addition, different levels of knowledge and experiences between the observers also lead to varying results on landmark identification. In a study using 18 observers, in which 13 were dental students and 5 were post-graduate's, the lasts revealed lower intra-observer tracing variance than dental students (Kvam and Krogstad, 1969). Patient positioning during the procedure is also very important to avoid errors on measurements and landmark identification (Houston *et al.*, 1986; Ahlqvist *et al.*, 1986). The publication of Ahlqvist *et al.*, 1986 was assessed with a low level of evidence because there was only one observer. A similar classification occurred for Bourriau *et al.*, 2012, intra-observer agreement could not be evaluated and the number of radiographs (n = 4) used was very low. Kvam and Krogstad's (1969), publication also used a limited number of subjects (*n* = 3). The choice of the observers also plays an important role on the results. Eighteen observers, in which 13 were dental and 5 were post-graduate students, participated in their study (Kvam and Krogstad, 1969). The latter can also bias results because of the distinct level of education and expertise due to the lack of experience of the observers.

Regarding the influence of magnification, Bourriau *et al*., 2012 could not identify significant differences between equipment with a 4-m distant cephalometric machine and a 1.5-m distant cephalometric arm. Despite that, it should be considered that distance varying between the X-ray source and the image receptor will always cause a degree of magnification, the larger the distance, the lower the magnification. A focus object distance of 4 m in 2D cephalometric equipment is usually favoured for the reduced radiation burden and lack of enlargement, while equipment with 1.5-m arm has a direct advantage of being compact and integrated in a multimodal system as well as having an increased resolution. On the other hand, panoramic equipment with a cephalometric arm at a 1.5-m distance may present shortcomings in enlargement factors and superimposition of the bilateral structures more distant from the midsagittal plane, considering the less magnified structures on the side nearby the image receptor (White and Paroah, 2009). We were not able to identify studies correlating landmark identification errors in lateral cephalograms and their influence on the outcome of patient treatment.

Finally, in 1982, De Abreu showed that different 2D cephalometric analysis may lead to different diagnosis of the same patient, varying the diagnosis between class II and class III in 8 out of 129 cases. Also, Abdullah *et al.*, 2006 found that Steiner's cephalometric analysis is not accurate enough to plan orthodontic treatment. Both publications were assessed with low levels of evidence. In both publications, the number of observers was not referred. Furthermore, the statistical method used was not mentioned in (De Abreu, 1982).

The accuracy in the evaluation of the results, as well as producing changes in the treatment compared with clinical evaluation, seems to be one of the major benefits of 2D cephalometry. Risk-benefit analysis should be carefully evaluated.

1.10 Conclusions

The existing literature suggested that lateral cephalometric radiographs have been used without adequate scientific evidence of its usefulness and are often used prior to treatment. There is a need for diagnostic accuracy studies on 2D lateral cephalometric radiograph where standardised methodological criteria for diagnostic thinking efficacy and therapeutic efficacy are incorporated. This systematic review has shown that the evidence to agree or disagree on the usefulness of this radiographic technique in orthodontics today is limited. Lateral cephalograms are used in many occasions for reasons other than clinical diagnosis or treatment, such as medico-legal reasons in a teaching environment or due to a lack of experience in the field. These conclusions are rather worrying. The use of radiation in children should be even better justified, and scientific evidence of that justification seems lacking. At present, there is a need for further studies on larger patient populations, focusing on the therapeutic efficacy of lateral cephalograms.

CHAPTER 2. Accuracy and reliability of 2D cephalometric analysis in orthodontics as compared to the gold standard measurement on skull

ACCURACY AND RELIABILITY OF 2D CEPHALOMETRIC ANALYSIS IN ORTHODONTICS AS COMPARED TO THE GOLD STANDARD MEASUREMENT ON SKULL

2.1 Introduction

Human form measurements have been based on self-portrait, sculpture or drawing throughout the history. Likewise, craniofacial measurements have been intensely investigated by anthropologists, especially the proportions and relationships between anatomical craniofacial structures. By means of craniometrics, direct measurement on dry skulls was used extensively to determine their characteristic relationship to sex, body type, or genetic population, until the discovery of x-rays and the introduction of cephalometry. Lateral cephalometry radiography (LCR) was introduced simultaneously by a German dentist, Hofrath, and an American dentist, Broadbent, in 1931 (Wahl, 2006). It has been tremendously used in craniofacial analysis, and as a standard tool in orthodontics (Broadbent, 1931). It is used to define the morphology and predict the facial skeleton's growth, treatment planning and evaluation of treatment outcome (Baumrind and Frantz, 1971). Moreover, specific identification of anatomical landmarks can be performed on cephalometric radiographs. It allows measurements of various angular and linear variables. Nevertheless, the scientific value of cephalometric analyses is still questioned due to its lack of validity and reliability as a diagnostic tool. Several errors in landmark identification, linear and angle measurements and magnification of certain anatomical structures should be considered (Chen *et al.*, 2004). In addition, magnification of the radiograph, patient positioning or occasional different levels of knowledge and experience between observers may also lead to different results and interfere with the reliability of measurements (Kamoen *et al*., 2001; Tng *et al*., 1994; Ahlqvist *et al*., 1986; Houston *et al*., 1986; Kvam and Krogstad, 1969).

Previous studies have indicated that one of the major errors in cephalometric studies is caused due to inconsistency in landmark identification. Each landmark exhibits a characteristic pattern of error which contributes to measurement inaccuracy (Haynes and Chau, 1993). Only two studies reported the validity of skeletal landmarks: one performed by Mattila and Haataja in 1968, and the other by Tng *et al.* in 1994. Mattila and Haataja studied the validity of eight skeletal landmarks in the cranium and maxilla, but no statistical test was used to evaluate their results. Tng *et al.* investigated true anatomical landmarks in comparison with landmarks identified on cephalograms, and found that there is a trend for a minor degree of error for cephalometric angles and distances involving only skeletal landmarks compared to those involving skeletal and dental landmarks. They stressed that landmarks identified on cephalograms differed from true anatomical landmarks (Tng *et al*., 1994). Even though the validity of landmarks has been examined, the former studies did not cover its effect on linear measurement between anatomical landmarks. Therefore, the present aim o was to evaluate the reliability of some linear measurements commonly used in 2D lateral cephalometric analysis and its accuracy when compared to the gold standard measurements performed on skulls.

2.2 Materials and Methods

Twenty dry mixed dentate human skulls from the Anatomy Department of the University of Hasselt were used. These were selected according to the following inclusion criteria: reproducible occlusion, presence of permanent upper and lower incisors, and presence of at least one molar on either side to maintain the vertical dimension. The mandibles were stably connected to the maxillae through occlusal interdigitation at the maximum occlusion, with the condyles located in the glenoid fossa. The mandibles were attached to the skulls with broad tape attached from the temporal bone of one side, crossing the inferior border of the mandible, to the temporal bone of the opposite side.

Radiographs

Lateral cephalograms were acquired by positioning the skulls in a standard panoramic-cephalometric device (Veraviewepocs 2D®, J. Morita, Kyoto, Japan). The magnification ratio of the lateral cephalometry was 1.1. The skulls were stabilized in the cephalostat on an aluminum filter box (thickness of 400 aluminum foils sheets). It had 18.5 cm of diameter and 2.5 cm of thickness. The purpose of using the aluminum filter was to simulate a real situation, mimicking soft tissue attenuation, and not facilitate the identification of bony landmarks on radiographs.

The radiographic settings we used were 77 kV, 7.2 mA and 3.2 s. All the images were then exported in TIFF format, and imported to Adobe Photoshop® CS3 software (Adobe Systems Incorporated, California, USA). Before the radiographic evaluation, the skull position was adjusted to allow the Frankfort horizontal plane to be parallel with the horizontal plane for further measurements.

Analysis

Two experienced observers (dentomaxillofacial radiologists) performed this study with a session of calibration prior to the analysis. Ten commonly used skeletal landmarks were identified on twenty skulls and radiographs according to figure 2.1 (Proffit *et al.*, 2006). Both observers had been informed about all the anatomical landmarks, identification methods used on radiographs, and also craniometric measurement of the skulls. Five skulls and its radiographs were used for calibration. At the end of the calibration, both observers were in agreement and any remaining doubt was clarified. In case of any uncertainty between the two observers, an additional advice from a third observer was essential to reach agreement.

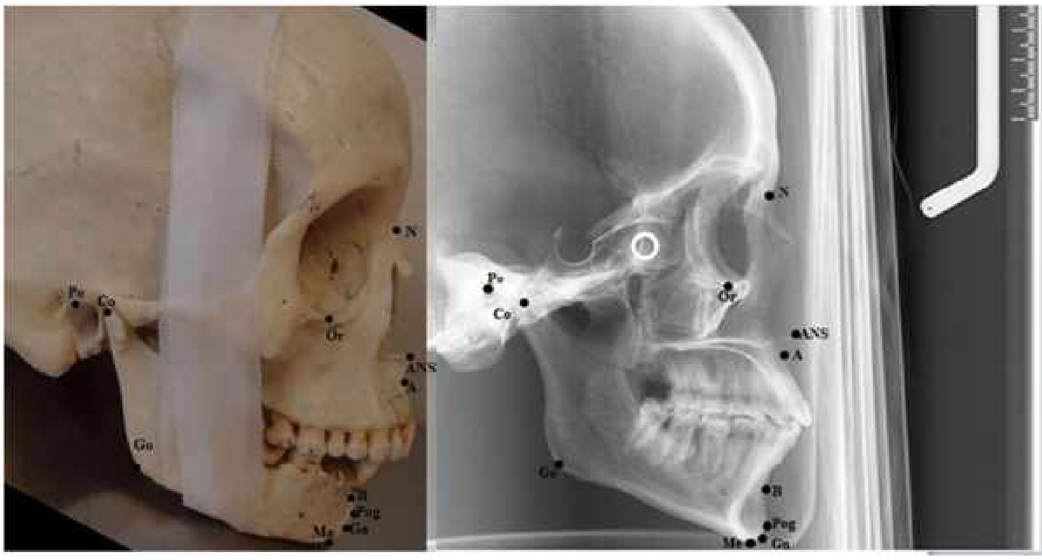

Figure 2.1. Cephalometric landmarks used in the study. N – Nasion; Me – Menton; ANS – Anterior Nasal Spine; Co – Condylion; Gn – Gnathion; A – Point A; B - Point B; Pog – Pogonion; Po – Porion; Or – Orbitale; Go – Gonion.

Craniometric measurements considered to be the gold standard were done on 20 dry dentate skulls by using a digital caliper (Absolute Digimatic Caliper No. 500-161U; Mitutoyo America Corp., Aurora, IL). The same measurements were performed by digital determining the landmarks on the viewing monitor in a dim-lighted room without any interruption. All measurements were repeated one month later, both on skulls and radiographs. The results of the intra- and inter-observer reliability were analysed. The linear measurements were chosen according to the vertical and anteroposterior dimensions of the craniofacial form (Table 2.1). The landmarks on which these measurements were based represented both midsagittal and bilateral anatomical structures.

Table 2.1. Linear measurements evaluated on human skulls and lateral cephalometric radiographs in this study (mm).

Linear Measurements
Total anterior face height: N-Me
Upper face height: ANS-N
Lower face height: ANS-Me
Mandibular unit length: Co-Gn
Maxillary unit length: Co-ANS
AN: A to N with respect to true vertical
BN: B to N with respect to true vertical
PogN: Pog to N with respect to true vertical
Po-Or (Frankfort plane)
Go-Me (mandibular plane)

Statistical analysis

Variables were described through its mean, standard deviation and measurements of dispersion. Intra- and inter-observer variation was studied using the intraclass correlation coefficient (ICC) with a confidence interval of 95%. General guidelines for this measure rate an ICC > 0.90 as excellent, an ICC of 0.75–0.90 as good, and an ICC < 0.75 as representing poor to moderate reliability (Shrout and Fleiss, 1979). Differences between the measurements performed on skulls and on radiographs were evaluated by the Bland-Altman limits of agreement (Bland and Altman, 1986). One sample *t*-test was used to evaluate if the mean of the differences between the two measurements was different from 0 (Moore and McCabe, 2006).

The Statistical Package for Social Sciences 20.0 for Windows (SPSS Inc., Chicago, Illinois, USA) was used for statistical analysis. The level of statistical significance for all tests was set at $\alpha = 0.05$.

2.3 Results

Intra-observer consistency is shown on Table 2.2. On Table 2.3 the inter-observer reliability is presented.

Craniometric measurement revealed ICC values in general, above 0.90, for the intra-observer reliability, with exception of the A-N measurement for observer 2, which showed an ICC of 0.76 (Table 2.2).

For the inter-observer reliability seen in craniometric measurement, the ICC was also, in general, above 0.90, with exception of ANS-N for the second observation; A-N and Po-Or for both observations (Table 2.3).

Intra-observer reliability for the linear measurement on radiographs revealed ICC values above 0.90, except for ANS-N and Co-ANS for the second observer, and A-N for both observers (Table 2.2).

There was an overall good agreement regarding inter-observer reliability for linear measurement performed on radiographs, when comparing between linear measurements, with the exception of ANS-N, Co-ANS, A-N and Po-Or for both observations (Table 2.3).

With regards to accuracy of 2D cephalometric radiographs, the mean differences between linear measurements (mm) when performed by both observers on skulls and radiographs were investigated and the results are shown in Table 2.4.

Radiograph and craniometric measurements presented statistically significant differences between them, with $p < 0.05$, implying that there was a difference in landmark identification between these two modalities.

We found that seven of the ten linear measurements on radiographs were on average significantly higher. Only three of the linear measurements were on average significantly higher when performed directly on the skulls (Co-Gn, Co-ANS, and Go-Me). It was seen that these three measurements had at least one bilateral landmark. The widest deviation between the two methods was seen on the measurement N-Me, with a difference of 0.96 mm. The lowest value was detected on the measurements between Co-Gn (0.14) and Po-Or (0.14). Bland-Altman limits of agreement showed the 95% differences between measurements performed on the skulls and on radiographs. All the differences found between the two methods were inferior to two units of measurement (mm), which is, generally, within one standard deviation of the norm values in cephalometric analysis (Chen *et al*., 2004).

Table 2.2. Mean differences between the first and second observations with regards to intra-observer agreement (mm).

	Observation 1				Observation 2			
	Mean (SD)	ICC	CI 95%	LA	Mean (SD)	ICC	CI 95%	LA
N-Me								
Skull	10.08 (0.96)	0.999	0.997-0.999	-0.10;0.09	10.08 (0.96)	0.998	0.995-0.999	-0.11;0.12
Radiograph	11.02 (1.01)	0.978	0.948-0.991	-0.47;0.36	11.03 (1.02)	0.999	0.998-1.000	-0.06;0.09
ANS-N								
Skull	4.41 (0.32)	0.949	0.810-0.978	-0.19;0.21	4.43 (0.34)	0.926	0.832-0.969	-0.26;0.26
Radiograph	4.79 (0.35)	0.905	0.786-0.960	-0.36;0.25	4.82 (0.32)	0.831	0.636-0.926	-0.49;0.39
ANS-Me								
Skull	5.87 (0.72)	0.997	0.94-0.999	-0.14;0.06	5.84 (0.76)	0.980	0.952-0.991	-0.34;0.26
Radiograph	6.38 (0.82)	0.984	0.961-0.993	-0.34;0.24	6.43 (0.83)	0.973	0.937-0.989	-0.49;0.26
Co-Gn								
Skull	10.87 (0.89)	0.989	0.974-0.996	-0.31;0.20	10.85 (0.87)	0.994	0.985-0.997	-0.25;0.13
Radiograph	10.72 (0.93)	0.989	0.973-0.995	-0.28;0.27	10.71 (0.90)	0.982	0.957-0.992	-0.36;0.32
Co-ANS								
Skull	9.19 (0.60)	0.981	0.954-0.992	-0.24;0.22	9.22 (0.60)	0.972	0.934-0.988	-0.34;0.22
Radiograph	8.54 (0.57)	0.935	0.851-0.973	-0.40;0.42	8.61 (0.50)	0.845	0.663-0.933	-0.72;0.43
A-N								
Skull	4.97 (0.35)	0.911	0.798-0.962	-0.27;0.32	4.90 (0.35)	0.763	0.512-0.895	-0.43;0.59
Radiograph	5.31 (0.36)	0.797	0.573-0.911	-0.51;0.45	5.39 (0.35)	0.619	0.276-0.822	-0.58;0.76
B-N								
Skull	8.49 (0.74)	0.982	0.957-0.993	-0.26;0.29	8.57 (0.75)	0.959	0.905-0.983	-0.53;0.32
Radiograph	9.25 (0.76)	0.991	0.979-0.996	-0.20;0.20	9.39 (0.82)	0.984	0.962-0.993	-0.27;0.31
Pog-N								
Skull	9.39 (0.88)	0.991	0.979-0.996	-0.24;0.22	9.45 (0.87)	0.982	0.958-0.993	0.24;0.41
Radiograph	10.29 (0.95)	0.982	0.956-0.992	-0.34;0.38	10.29 (0.97)	0.991	0.978-0.996	-0.26;0.25
Po-Or								
Skull	7.24 (0.38)	0.957	0.901-0.982	-0.13;0.32	7.40 (0.41)	0.910	0.082-0.745	-0.78;1.14
Radiograph	7.42 (0.40)	0.957	0.900-0.982	-0.28;0.18	7.50 (0.38)	0.906	0.789-0.960	-0.36;0.30
Go-Me								
Skull	7.43 (0.57)	0.955	0.895-0.981	-0.30;0.37	7.55 (0.65)	0.931	0.841-0.971	-0.44;0.53
Radiograph	7.05 (0.55)	0.936	0.853-0.973	-0.42;0.36	7.03 (0.54)	0.952	0.889-0.980	-0.23;0.44

SD – standard deviation; ICC- Intraclass correlation; CI (5% - 95%) confidence interval; LA- Limits of agreement

Table 2.3. Inter-observer agreement (mm).

	Observer 1				Observer 2			
	Mean (SD)	**ICC**	**CI 95%**	**LA**	**Mean (SD)**	**ICC**	**CI 95%**	**LA**
N-Me								
Skull	10.08 (0.96)	0.997	0.993-0.999	-0.14;0.14	10.07 (0.95)	0.999	0.998-1.000	-0.07;0.08
Radiograph	11.02 (1.00)	0.972	0.934-0.988	-0.52;0.43	11.04 (1.01)	0.996	0.900-0.998	-0.16;0.20
ANS-N								
Skull	4.42 (0.32)	0.954	0.893-0.981	-0.20;0.19	4.42 (0.32)	0.852	0.677-0.936	-0.38;0.34
Radiograph	4.78 (0.32)	0.855	0.684-0.937	-0.40;0.32	4.83 (0.39)	0.861	0.694-0.940	-0.45;0.38
ANS-Me								
Skull	5.83 (0.74)	0.992	0.982-0.997	-0.15;0.21	5.88 (0.74)	0.980	0.951-0.991	-0.27;0.32
Radiograph	6.36 (0.82)	0.953	0.890-0.980	-0.51;0.49	6.44 (0.82)	0.985	0.965-0.994	-0.36;0.20
Co-Gn								
Skull	10.83 (0.87)	0.982	0.957-0.992	-0.27;0.29	10.89 (0.89)	0.994	0.986-0.998	-0.18;0.20
Radiograph	10.71 (0.90)	0.978	0.947-0.991	-0.36;0.39	10.72 (0.92)	0.990	0.977-0.996	-0.25;0.26
Co-ANS								
Skull	9.18 (0.61)	0.982	0.957-0.992	-0.23;0.23	9.22 (0.60)	0.990	0.976-0.996	-0.22;0.12
Radiograph	8.55 (0.55)	0.857	0.688-0.938	-0.59;0.60	8.61 (0.52)	0.866	0.706-0.942	-0.72;0.43
A-N								
Skull	4.96 (0.34)	0.857	0.687-0.938	-0.33;0.41	4.91 (0.37)	0.867	0.707-0.943	-0.29;0.48
Radiograph	5.36 (0.36)	0.673	0.361-0.850	-0.77;0.49	5.33 (0.35)	0.740	0.470-0.883	-0.55;0.51
B-N								
Skull	8.51 (0.74)	0.954	0.892-0.980	-0.47;0.42	8.55 (0.73)	0.947	0.877-0.978	-0.62;0.33
Radiograph	9.33 (0.79)	0.977	0.945-0.990	-0.49;0.19	9.32 (0.79)	0.984	0.962-0.993	-0.41;0.14
Pog-N								
Skull	9.44 (0.88)	0.991	0.980-0.996	-0.34;0.11	9.40 (0.87)	0.980	0.952-0.992	-0.36;0.33
Radiograph	10.29 (0.96)	0.972	0.933-0.988	-0.44;0.46	10.29 (0.95)	0.989	0.973-0.995	-0.26;0.26
Po-Or								
Skull	7.35 (0.38)	0.805	0.116-0.706	-1.07;0.66	7.25 (0.41)	0.804	0.586-0.914	-0.64;0.41
Radiograph	7.45 (0.39)	0.944	0.871-0.976	-0.35;0.16	7.48 (0.39)	0.873	0.720-0.945	-0.47;0.32
Go-Me								
Skull	7.51 (0.61)	0.919	0.816-0.966	-0.62;0.36	7.47 (0.63)	0.925	0.829-0.968	-0.50;0.42
Radiograph	7.06 (0.51)	0.901	0.778-0.958	-0.50;0.42	7.02 (0.57)	0.950	0.883-0.79	-0.27;0.45

SD – standard deviation; ICC- Intraclass correlation; CI (5% - 95%) confidence interval; LA- Limits of agreement

Table 2.4. Mean of differences and level of agreement between measurements performed on the skull and radiography.

	Mean of differences (mm)	***p***	**LA**
N-Me	-0.96	<0.001	-1.710;-0.742
ANS-N	-0.39	<0.001	-0.712;-0.067
ANS-Me	-0.58	<0.001	-0.869;-0.294
Co-Gn	0.14	<0.001	-0.191;0.477
Co-ANS	0.62	<0.001	0.252;-0.986
A-N	-0.41	<0.001	-0.753;-0.074
B-N	-0.79	<0.001	-1.179;-0.409
Pog-N	-0.87	<0.001	-1.148;-0.602
Po-Or	-0.15	0.001	-0.860;0.566
Go-Me	0.45	<0.001	0.038;0.859

p - *One-sample t- test*; LA- Limits of agreement

2.4 Discussion

Evidence shows that landmark identification is a great source of error in 2D cephalometric analysis because of the uncertainty in recognizing accurately where the landmark is located. Some landmarks also show a wider variation in localization than others (Tng *et al.*, 1994; Baumrind and Frantz, 1971). Superimposition between bilateral anatomical structures and anatomical localization may hinder its identification, as for example of landmarks Co, Go, Po, Or, and lower incisor apex (Tng *et al.*, 1994; Baumrind and Frantz, 1971). Therefore, it is essential to accurately determine anatomical landmarks in order to reduce linear measurement error in cephalometric analysis. Moreover, it is important to assess the quantitative differences between craniometric measurement and the corresponding radiographic measurements.

The observers' agreement is another factor that influences the measurement error. Chen *et al*. (2004) found that in general the inter-observer error presents greater values than the intra-observer error. We confirmed that, on average, there was a higher rate of inter-observer error. We found that intra- and inter-observer reliability of linear measurements performed on skulls were on average significantly lower than on radiographs (Table 2.2 and 2.3).

Table 2 shows that intra-observer reliability for skull linear measurement A-N was the least consistent for observer 2, with an ICC of 0.76. When comparing intra-observer reliability on radiographs, the lowest agreement was seen in A-N, Co-ANS and ANS-N, respectively, for both observers. Linear measurement A-N showed a lower agreement between observers both on skulls and on radiographs. This might be due to the localization of point A, Co and ANS (Tng *et al.*, 1994; Baumrind and Frantz, 1971). The evidence shows that bilateral anatomical landmark identification, such as Co, is a great source of error in 2D lateral cephalometry (Tng *et al.*, 1994). Relating to points A and ANS, they might appear more radiolucent on radiograph, which may lead to uncertain position of these landmarks.

Intra- and inter-observer SD for the skulls and radiographs were lower (value inferior to 0.5) for linear measurements ANS-N, A-N and Po-Or on observations 1 and 2.

On average, in a 12-years old male, the Harvold linear measurement ANS-Me presents a SD of approximately 3.7 mm (Proffit *et al.,* 2006), which is a value higher than the ones we found (maximum 0.83).

The results revealed that, in general, craniometric measurements tended to be shorter than linear measurement on radiographs, except for Co-Gn (mandibular unit), Co-ANS (maxillary unit), and Go-Me (mandibular plane) (Table 2.4). This may be related with the fact that on these linear measurements, at least one of the landmarks is placed on bilateral structures (Co and Go), which may have increased this variability. Also, it is more difficult to establish a middle point directly on the skull than on the radiograph. Validity of cephalometric distances depended on the validity of individual landmarks involved.

In the case of a linear measurement, it is known that the shorter the line segment measured, the greater the percentage of error produced by a given measurement error (Chen *et al.*, 2004).

Our results contrast with the study from Farkas *et al.* (2002), where they found that singular and paired cephalometric distances were significantly shorter than the craniometric distances on postero-anterior cephalometric radiographs. Our ten measurements were statistically significant ($p<0.05$), even though intervals oh the limits of agreement were on average low (see Table 3.4).

The mean difference was significant and presented the highest variance for total anterior face height linear measurement (on average, N-Me at 0.956 mm). This means that there is a 95% chance that the value varies from -1.71 to -0.74, which is within the clinically acceptable limits, since it is inferior to 1 mm (Table 2.4).

The McNamara cephalometric analysis, published in 1983, estimated an error of +/- 2 mm for linear measurement A-N (Proffit *et al.,* 2006), while in the present study we found a confidence interval of -0.753 to -0.074, which shows that the confidence interval presents values much lower than 2 mm.

The shortest mean differences were observed in Co-Gn (0.14 mm) and Po-Or (-0.15 mm), which showed an extremely low value. Considering Po-Or, even though the mean difference was low, there was no significant difference between the two measurement methods. This could be explained by measurement errors from equipment, observers, or both. Therefore, these results should be investigated and taken into consideration. However, this might also have happened because of being easier to identify the Co and Gn on radiographs than on skulls.

Regarding radiographs, when landmarks were located at superimposed structures or placed on curves, they tended to have poorer validity, for example for linear measurements that contained A-point, Co, Gn and Po. Superimposition of adjacent structures complicates the identification of certain landmarks, such as Co, Or and Po, on radiographs.

There is always a degree of magnification on radiographs, caused by the variable distance between the X-ray source and the image receptor. Thus, exact superimposition of right and left sides is impossible due to magnification of structures further from the image receptor and the slightly lesser magnification of structures nearer to the image receptor (Duarte *et al.*, 2009; European Commission, 2004). Panoramic equipment with a cephalometric arm at a 1.5 m distance, as used in this study, may present shortcomings in enlargement factors and in superimposition of bilateral structures more distant from the midsagittal plane. In former studies where equipment with a 4 m arm was used, the long distance allowed radiation at a much lower dose and a parallel bundling of the x-rays, as to guarantee a magnification of x1, eliminating any left/right magnification differences (Bourriau *et al.,* 2012). Nevertheless, the logistic requirements for such a cephalometric machine with a 4 m distance separation, made companies and dentists favor cephalometric arms (1.5 m focus-object distance) integral to panoramic equipment. The latter are more compact, but may present a differential enlargement between the left and right sides and contrasts within the midline enlargement.

It has been suggested that observed differences should represent at least twice the standard deviation of the estimating error in order to be significant (Baumrind and Frantz, 1971). The current differences are usually shorter than +/-1 mm, which is less than the estimated standard deviation for each linear measurement. Besides, cephalometric analysis finally reports on relative relations. The presently found significant deviations may thus have rather limited interference on the orthodontic diagnosis and treatment planning. However, one should perform a thorough cephalometric analysis on a large sample, with subsequent treatment planning, in order to exclude any occurrence of a significant clinical effect.

2.5 Conclusions

In the present study, linear radiographic measurements systematically and significantly overestimated the gold standard measurements of the skulls, while intra- and inter-observer reliability was also significant. Nevertheless, the differences found were most often inferior to 1 mm, which is generally within the accepted standard deviation. Some linear measurements were more reliable than others. Further studies focusing on the impact of deviating cephalometric analysis on a larger sample may be required to determine its clinical impact. Considering the use of relative rather than absolute data analysis, the impact of this discrepancy on clinical analysis may be expected to be low.

CHAPTER 3. Reproducibility of 2D cephalometric landmark identification by orthodontists and dentomaxillofacial radiologists

REPRODUCIBILITY OF 2D CEPHALOMETRIC LANDMARK IDENTIFICATION BY ORTHODONTISTS AND DENTOMAXILLOFACIAL RADIOLOGISTS

3.1 Introduction

Since its introduction by Broadbent in 1931, lateral cephalometric radiography has been widely used in orthodontics (Broadbent, 1931). It is used to define morphology and predict growth of the facial skeleton, treatment planning as well as evaluation of treatment outcome (Baumrind and Frantz, 1971). Cephalometric analyses provide angular and linear measurements useful to perform a diagnosis and a treatment plan in orthodontics. Errors in cephalometric analyses may occur by numerous reasons. One of the most important errors happens due to inconsistent and imprecise landmark identification. Inaccurate landmark identification may lead to erroneous diagnoses and treatment plans for orthodontic cases (Chen *et al.*, 2004; Tng *et al*, 1994). The identification of certain anatomical landmarks such as Porion (Po), Condylion (Co), Orbitale (Or), Basion (Ba), Gonion (Go), Anterior Nasal Spine (ANS), Posterior Nasal Spine (PNS) and Lower inferior apex (LIA) might be more prone to error due to overlapping structures superimposed on the landmark or its location (Baumrind and Frantz, 1971). Likewise, quality of the radiographic image interferes with correct identification of some landmarks, such as Po, Co, Or, ANS, point B, Pogonion (Pog), Go and glabella (Miloro *et al.*,2013; Kamoen *et al.*, 2001). Moreover, some authors believe that different levels of knowledge and observers background play an important role on landmark identification (Miloro *et al.*, 2013; Kamoen *et al.*, 2001; Gravely and Benzies, 1974; Kvam and Krogstad, 1969). Other authors believe that errors can be caused by different individual conceptions of landmark definition and its perception, rather than education and training (Kamoen *et al.*, 2001; Lau *et al.*, 1997). Inconsistency of landmark identification can increase the degree of error (Silveira and Silveira, 2006; Chen *et al.*, 2004; Chen *et al.*, 2000). Inter-observer reproducibility of landmark identifications was found to be very low among dentomaxillofacial radiologists (DMFR) (Silveira and Silveira, 2006).

Some dentomaxillofacial radiologists, as well as orthodontists, are trained to perform 2D cephalometric analyses. There are no previous reports on the reliability of landmark identification that compare orthodontists and dentomaxillofacial radiologists. Therefore, the aims of the present study were to evaluate the reproducibility of 17 commonly used cephalometric landmarks by orthodontists and dentomaxillofacial radiologists, and to assess the impact of different landmark identifications on patient diagnosis.

3.2 Materials and Methods

Twenty digital lateral cephalometric radiographs were selected from the database at the Oral Imaging Center, University of Leuven. Lateral cephalograms were acquired by positioning the patients in a standard digital cephalometric device and using a charged couple device sensor (Veraviewepocs 2D®, J. Morita, Kyoto, Japan). Exposure values were set at 77 kV and 7.2 mA, with an exposure time of approximately 1.6 s, according to each patient. Inclusion criteria were:

- No evidence of current orthodontic treatment.
- Digital cephalometric image were of good quality to allow the identification of landmarks, and the ruler on the radiograph was clearly visible, allowing calibration of the images in the cephalometric analysis software program.
- There were no unerupted or partially erupted incisors that could have compromised landmarks identification.
- No gross skeletal asymmetry.

All selected images were then exported in TIFF format, and subsequently imported to the computerized program for cephalometric analysis (*Radiocef Studio 2*, Radio Memory Ltd., Belo Horizonte, Brazil).

Analysis

Seventeen commonly used cephalometric landmarks were included in this analysis; these are shown in Figure 3.1 (Proffit *et al*., 2006). Landmark identification was carried out on the digital image, using a mouse-driven cursor in a predetermined sequence.

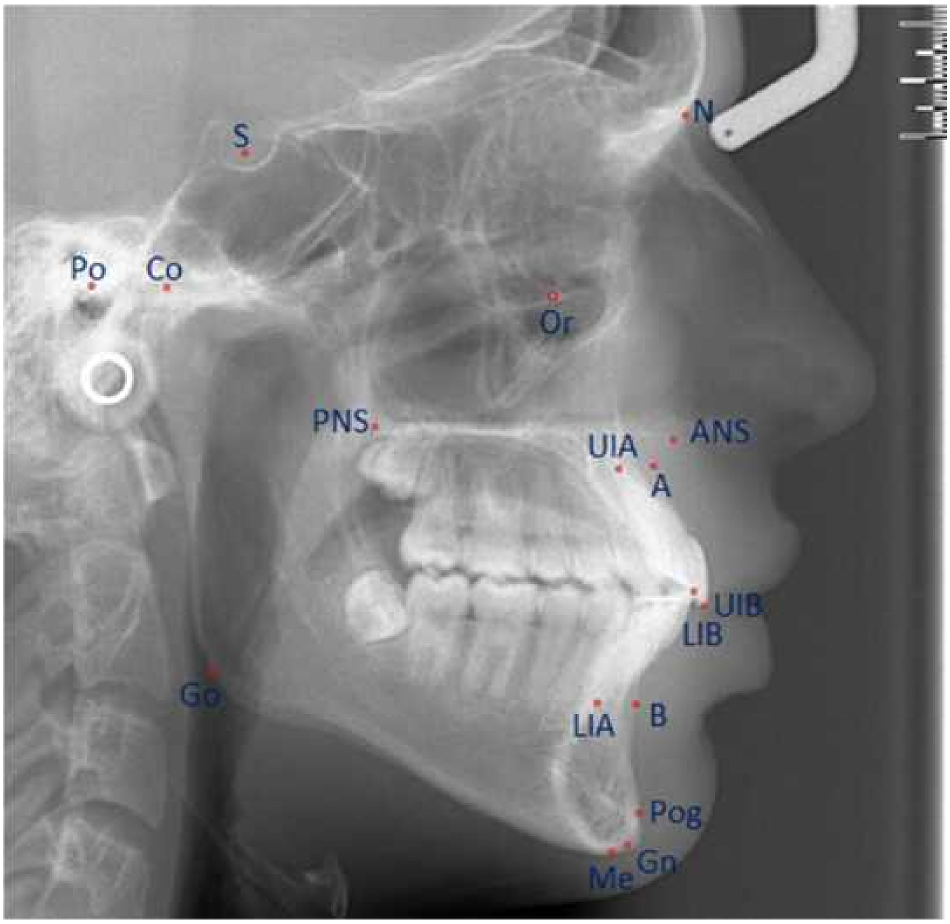

Figure 3.1 - Cephalometric landmarks used in the study. N – Nasion; Or – Orbitale; S – Sella; Co – Condylion; Po – Porion; PNS- Posterior Nasal Spine; ANS – Anterior Nasal Spine; A – Point A; UIA – Upper incisor apex, UIB – Upper incisor border; LIB – lower incisor border; LIA – Lower incisor apex; B - Point B; Pog – Pogonion; Gn –Gnathion; Me – Menton; Go – Gonion.

Eight experienced observers (four orthodontists and four dentomaxillofacial radiologists) performed this study. Experience of the observers ranged from eight to 15 years. An initial training and calibration session was attended by all observers, including an explanation of the anatomical structures and required landmark identification. At the end of the session, the main author clarified any remaining doubt. Thus, all observers followed the same definitions of landmarks in the identification process. For optimal visualization, landmark identification was performed in a dim-lighted room without any interruption. Intra-observer reliability was assessed by one dentomaxillofacial radiologist, repeating the same procedure 3 months after.

After selecting a landmark with the mouse cursor, a dot on the monitor-displayed image indicated its position. Landmark position could be corrected until the operator was satisfied. Vertical and horizontal positions of each landmark were recorded in the format of *x* and *y* coordinates. Landmarks' digitized coordinates were then imported into the Excel software (Version 2003; Microsoft, Redmond, Washington, USA). Statistical analysis was performed using Statistical Package for Social Sciences, version 20.0 for Windows (SPSS Inc., Chicago, Illinois, USA). The level of statistical significance for all tests was set at $\alpha = 0.05$. We also compared which group was closer to the gold standard measurements. Some linear and angular measurements used in Ricketts and McNamara's cephalometric analysis were performed by all observers on 20 cephalometric radiographs. Of these, three radiographs were classified as borderline cases, in between orthognathic surgery and orthodontics. An example, showing the differences on landmark identification by two observers is seen in Figure 3.2. In general, the differences ranged between 1 and 2 mm.

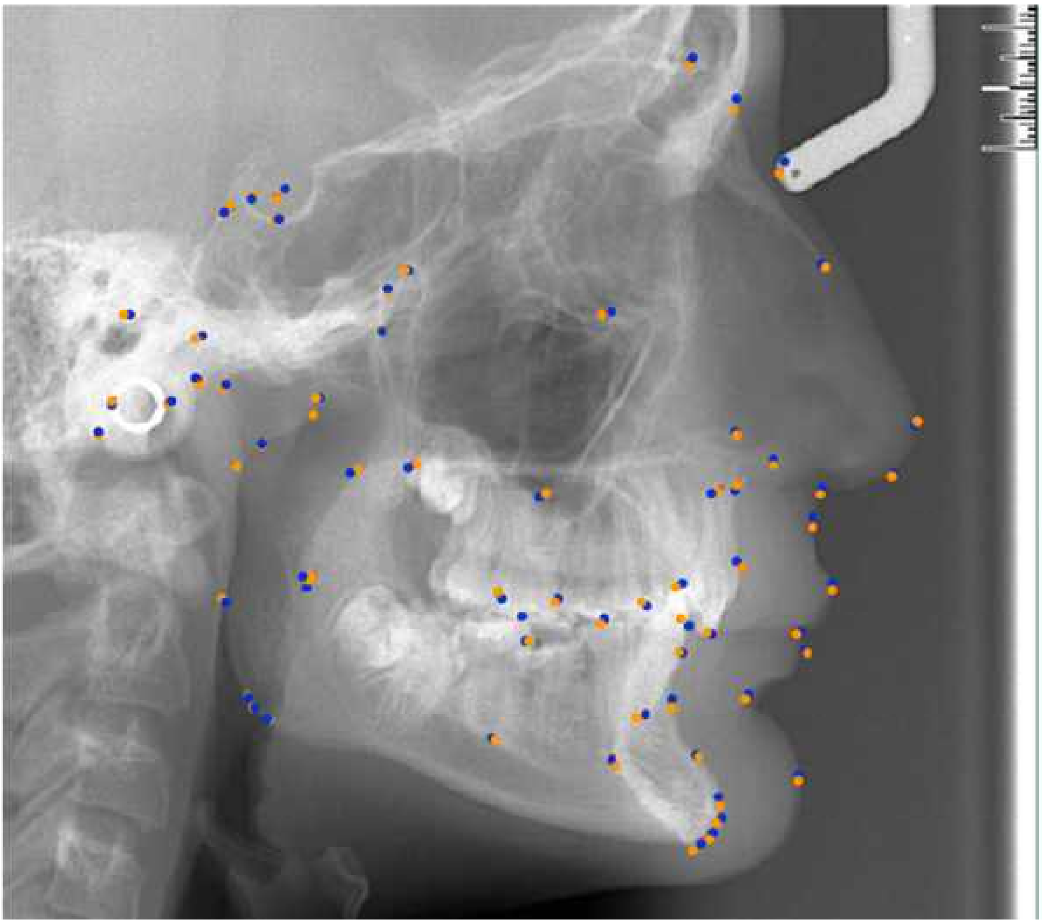

Figure 3.2 – Example of a lateral cephalometric radiography with identification of landmarks by two observers.

The same computer software was used to access differences on angular and linear measurements. The angular and linear measurements used were the following: A-N; Co-Gn (Mandibular unit length); Co-A; (Po-Or).(Go-Me); Pog-N (Facial plane); LIB (A-Pog); Convexity of Point A; Go-Me (Mandibular Plane) and S-Go.

Statistical analysis

To analyse the precision in landmark identification, each landmark's mean, standard deviation and measurements of dispersion were calculated. Intra- and inter-observer reliability for each landmark in the x and y directions were studied, using intraclass correlation coefficients (ICC) with a confidence interval of 95%. General guidelines for this measure rate an ICC > 0.90 as excellent, an ICC of 0.75–0.90 as good, and an ICC < 0.75 as representing poor to moderate reliability (Shrout and Fleiss, 1979).

Therefore, the "best estimate" of landmarks identification was obtained from the mean value of each landmark identified by the observers, and defined as the gold standard. Inter-observer reliability was assessed using the Euclidean distances. The average distance between the mean positions pointed by an observer was calculated and presented as the "intra-observer error", which was an indicator of reliability. Differences in landmark location were analysed by Student's t-test with the significance level of $p < 0.05$.

3.3 Results

The ICC was calculated for the intra- and also for inter-observer repeatability in the two groups (Table 3.1). In general, the ICC from the intra-observer ranged above 0.90, which implied an excellent agreement, with exception of the x direction of points Po, Me and point B. y component of landmarks N, Or and S, was considered good (ICC between 0.75 and 0.90). Furthermore, the vertical components of landmarks Go and point B revealed a poor or moderate agreement (ICC <0.75) seen for the intra-observer reproducibility.

The ICC for the inter-observer reliability was in general above 0.90 for all observers, with exception of the *x* component of points N and Or, which presented a good agreement (Table 3.1). Overall, in both groups, the highest variation found was associated with the vertical component of point Go (1.73 mm) and the lowest was seen in the vertical component of point Po (0.04 mm).

For the orthodontists, the ICC showed an overall lower value when compared to DMFRs. The ICC varied from 0.75 to 0.90 regarding landmarks Or, Po, Gn, point B and UIA. Likewise, the *x* coordinates of landmarks N, Me, Pog, point A, PNS, LIA and LIB also showed a good agreement. Only the *x* coordinate of ANS was classified as having poor or moderate agreement.

For dentomaxillofacial radiologists, the overall ICC was higher than 0.90. The exceptions were the *x* coordinates of landmarks Or and Po and the *y* coordinates of Go and point B, for which the agreement was good. Poor or moderate agreement was observed in the *y* component of Or. Overall in both groups, there was a high variation related to point Co, in the *x* direction. Between two DMFRs observers there was a difference of 5.05 mm; and between orthodontists, a difference of 3.56 mm was found. The horizontal component of point Or was less reproducible for DMFRs. Point Go in the *x* and *y* directions, points Me and PNS in the *x* direction and point B in the *y* direction were less precise among orthodontists.

Table 3.1. ICC for inter- and intra-observer evaluation.

	Inter-observer*		Intra-observer**	
	ICC	95% CI	ICC	95% CI
N – Nasion				
x	0.78	0.310-0.848	0.99	0.985-0.998
y	0.99	0.985-0.998	0.85	0.301-0.845
Or – Orbitale				
x	0.88	0.726-0.951	0.91	0.797-0.965
y	0.96	0.923-0.987	0.78	0.202-0.812
S – Sella				
x	0.99	0.984-0.997	0.99	0.976-0.996
y	0.98	0.962-0.994	0.87	0.701-0.946
Po- Porion				
x	0.97	0.916-0.986	0.80	0.569-0.917
y	0.95	0.889-0.982	0.94	0.845-0.974
Co – Condylion				
x	0.91	0.797-0.965	0.94	0.856-0.976
y	0.97	0.917-0.986	0.96	0.899-0.983
Go – Gonion				
x	0.98	0.956-0.993	0.98	0.964-0.994
y	0.93	0.830-0.971	0.71	0.394-0.872
Me – Menton				
x	0.97	0.918-0.987	0.84	0.631-0.932
y	0.99	0.970-0.995	0.97	0.932-0.989
Pog – Pogonion				
x	0.98	0.941-0.990	0.98	0.949-0.992
y	0.98	0.951-0.992	0.98	0.951-0.992
Gn – Gnathion				
x	0.97	0.924-0.988	0.97	0.936-0.990
y	0.99	0.963-0.994	0.99	0.954-0.979
B point				
x	0.97	0.930-0.989	0.89	0.754-0.957
y	0.96	0.894-0.983	0.72	0.418-0.879
A point				
x	0.97	0.916-0.986	0.92	0.804-0.966
y	0.98	0.960-0.994	0.95	0.890-0.982
ENA				
x	0.93	0.823-0.970	0.95	0.868-0.978
y	0.99	0.968-0.995	0.97	0.939-0.990
ENP				
x	0.98	0.952-0.992	0.94	0.852-0.975
y	0.99	0.972-0.996	0.99	0.969-0.995
LIA				
x	0.97	0.932-0.989	0.97	0.917-0.987
y	0.97	0.937-0.990	0.99	0.969-0.995
LIB				
x	0.98	0.961-0.994	0.98	0.960-0.994
y	0.99	0.968-0.995	0.99	0.981-0.997
UIB				
x	0.98	0.951-0.992	0.98	0.960-0.994
y	0.98	0.947-0.991	0.98	0.963-0.994
UIA				
x	0.97	0.925-0.988	0.98	0.956-0.993
y	0.97	0.924-0.988	0.98	0.944-0.991

* Between the 8 observers; ** a dentomaxillofacial radiologist

The Euclidean distance was used to test differences in landmark identification among observers and regarding the gold standard. The mean location differences of all landmarks for the orthodontists ranged from 5.92 mm to 0.99 mm. Generally, the landmark with least location differences was LIB (0.99 mm; SD 0.65 mm) and the one with most differences was point Gn (5.92 mm; SD 4.59 mm). The minimal and maximal variations on the horizontal component were associated with point ANS (0.53 mm; SD 3.74 mm and 2.97 mm; SD 2.02 mm). Regarding the reproducibility of the vertical component, point A presented the minimum variation (0.10 mm; SD 1.86 mm), while Gn was the most variable (4.60 mm; SD 3.67 mm). Table 3.2 shows the Euclidean distances between the "best estimate" of each landmark and orthodontist observers, defined as the inter-observer error of landmark identification. In general, orthodontists revealed errors inferior to 1 mm in points S, Pog, LIB, and UIB in both horizontal and vertical directions.

Table 3.2. Minimum and maximum euclidean distances (in mm) for orthodontists, defined as absolute differences in millimetres between the mean values and standard deviations of each landmark and the averaged for all observers.

Landmark	Horizontal component (x)			Vertical component (y)		
	Mean	**SD**	***p***	**Mean**	**SD**	***p***
N – Nasion						
Minimum	0.19	3.30	0.804	-0.34	0.74	0.054
Maximum	1.43	2.74	0.031	0.21	1.89	0.631
Or – Orbitale						
Minimum	0.24	1.46	0.650	0.20	1.74	0.902
Maximum	2.04	1.85	<0.001	-1.29	2.13	0.014
S – Sella						
Minimum	0.24	0.92	0.263	-0.32	0.67	0.048
Maximum	-0.42	0.80	0.029	-0.76	1.09	0.006
Po- Porion						
Minimum	0.53	1.59	0.155	0.44	1.47	0.194
Maximum	-2.60	2.08	<0.001	-1.42	2.28	0.012
Co – Condylion						
Minimum	0.23	1.69	0.549	-2.61	2.11	<0.001
Maximum	-2.66	1.13	<0.001	1.15	2.68	0.069
Go - Gonion						
Minimum	0.50	1.34	0.110	-1.09	2.47	0.063
Maximum	2.20	1.83	<0.001	2.14	2.36	0.001
Me - Menton						
Minimum	0.73	3.77	0.395	-0.42	1.84	0.323
Maximum	-1.57	2.38	0.008	1.44	2.64	0.025
Pog - Pogonion						
Minimum	0.27	2.30	0.430	-0.16	1.99	0.984
Maximum	0.42	3.31	0.576	0.61	2.23	0.238
Gn - Gnathion						
Minimum	-0.31	2.22	0.025	0.21	2.00	0.847
Maximum	2.27	4.11	0.023	4.60	3.67	0.000
B point						
Minimum	0.20	1.84	0.725	0.21	2.87	0.759
Maximum	0.87	3.79	0.319	3.38	2.73	<0.001
A point						
Minimum	0.17	2.38	0.754	0.10	1.32	<0.001
Maximum	1.27	2.40	0.029	2.08	2.25	0.001
ANS						
Minimum	0.16	3.06	0.897	0.22	1.73	0.599
Maximum	2.97	3.74	0.002	-0.81	1.10	0.004
PNS						
Minimum	0.54	1.44	0.113	0.34	1.26	0.240
Maximum	-2.66	1.43	<0.001	-0.80	1.23	0.009
LIA						
Minimum	-0.38	2.38	0.495	-0.37	2.51	0.516
Maximum	-1.44	2.01	0.005	2.61	2.53	0.010
LIB						
Minimum	-0.10	2.01	0.781	0.24	2.19	0.631
Maximum	0.40	2.39	0.461	1.00	2.59	0.048
UIB						
Minimum	-0.23	1.71	0.556	-0.21	2.13	0.659
Maximum	-0.68	1.59	0.069	1.39	2.16	0.010
UIA						
Minimum	0.33	2.50	0.566	0.21	2.41	0.896
Maximum	-1.20	1.87	0.010	-2.41	1.61	0.000

SD – standard deviation; ICC- Intraclass correlation; CI (5% - 95%) confidence interval

Overall, dentomaxillofacial radiologists revealed errors inferior to 1 mm in both horizontal and vertical directions related to points N, S, A and LIA.

Smaller variations were seen between the "best estimate" of each landmark and points identified by dentomaxillofacial radiologists (Table 3.3). The average greatest Euclidean distance was observed in point Or (4.41 mm; SD 2.04 mm) and the lowest in point LIB (0.84 mm; SD 0.46 mm). Furthermore, the landmarks presenting the minimal and the maximal horizontal component variations were point LIB (0.08 mm; SD 0.81) and point Or (3.94 mm; SD 2.51 mm), respectively. The landmark with least vertical component location differences was point LIB (0.10 mm; SD 1.08 mm) and the one with most differences was point Gn (2.28 mm; SD 1.65 mm). Only the errors of points S, LIB and A point were, overall, inferior to 1 mm in both directions. The "best estimate" for each landmark was defined as the mean position identified by 8 observers.

Despite an overall value inferior to a variation of 2 mm (Kamoen et al., 2001; Lau et al, 1997), some landmarks presented higher deviations. Some DMFRs observers presented variations higher that 2 mm on the horizontal component of Or, Po, ANS, as well as for the Go and Gn. Orthodontists showed differences higher than 2 mm for the *x* coordinates of Or, Po, Co, Go, Gn, ANS, PNS and UIA.

Concerning reliability of 2D cephalometric landmark identification between observers, expressed as "inter-observer error", this presented statistically significant differences for some landmarks ($p < 0.05$).

Table 3.3. Minimum and maximum Euclidean distances for dentomaxillofacial radiologists, defined as absolute differences in millimetres between the mean values and standard deviations of each landmark and the averaged for all observers (mm).

Landmark	**Horizontal component (*x*)**			**Vertical component (*y*)**		
	Mean	**SD**	***p***	**Mean**	**SD**	***p***
N – Nasion						
Minimum	-1.94	2.20	0.001	0.19	0.50	0.101
Maximum	0.15	2.30	0.952	0.33	0.97	0.148
Or – Orbitale						
Minimum	-0.47	1.78	0.248	0.20	0.71	0.737
Maximum	3.94	2.51	<0.001	-2.20	6.73	0.160
S – Sella						
Minimum	-0.20	0.63	0.022	-0.66	1.02	0.009
Maximum	0.45	0.29	<0.001	1.90	3.24	0.017
Po- Porion						
Minimum	0.26	1.07	0.284	0.21	0.89	0.453
Maximum	2.08	2.69	0.003	1.65	1.13	<0.001
Co – Condylion						
Minimum	0.48	1.69	0.216	-0.18	1.34	0.545
Maximum	1.66	1.30	<0.001	0.80	1.39	0.018
Go - Gonion						
Minimum	-1.19	0.71	<0.001	-2.14	1.95	0.000
Maximum	0.25	0.97	0.751	0.33	1.97	0.965
Me - Menton						
Minimum	-0.37	1.45	0.268	-1.55	1.67	0.001
Maximum	1.51	1.21	<0.001	0.49	1.67	0.205
Pog - Pogonion						
Minimum	0.60	0.73	0.961	0.19	1.26	0.608
Maximum	-1.10	1.40	0.002	-1.65	1.57	<0.001
Gn - Gnathion						
Minimum	-0.20	0.94	0.353	0.17	1.59	0.880
Maximum	-1.36	1.57	0.001	-2.28	1.65	<0.001
B point						
Minimum	0.39	1.97	0.086	0.99	2.21	0.060
Maximum	-1.33	1.43	0.001	-1.50	2.15	0.005
A point						
Minimum	-0.32	0.72	0.706	0.21	1.28	0.898
Maximum	-1.00	1.39	0.001	-1.00	1.03	<0.001
ANS						
Minimum	0.25	1.37	0.421	0.57	0.91	0.012
Maximum	-2.62	1.14	<0.001	-0.63	1.71	0.115
PNS						
Minimum	0.16	1.31	0.818	0.14	0.83	0.448
Maximum	-0.17	1.51	0.616	-0.81	1.05	0.003
LIA						
Minimum	0.93	1.42	0.008	-0.39	1.29	0.189
Maximum	-1.66	1.39	<0.001	-2.09	1.37	<0.001
LIB						
Minimum	0.08	0.81	0.664	0.10	1.08	0.691
Maximum	-1.03	1.51	0.002	-1.02	1.30	0.002
UIB						
Minimum	0.30	1.06	0.215	-0.19	1.08	0.431
Maximum	-1.21	1.41	0.001	-1.25	1.47	0.001
UIA						
Minimum	0.67	1.06	0.011	0.48	0.88	0.025
Maximum	-0.98	1.29	0.003	0.76	1.06	0.005

SD – standard deviation; ICC- Intraclass correlation; CI (5% - 95%) confidence interval

For both groups, reproducibility of points N, Or, Me, ANS and UIA was better in the vertical direction. On the other hand, the consistency of point Go was greater on the horizontal direction. Also, orthodontists revealed less variance on the horizontal component of point Gn and point B, while points S and Pog were more reproducible. Regarding the impact on linear and angular measurements, we found that, in general, the SD was relatively small and did not exceed the SD proposed by the analysis in each measurement (Table 3.4). We saw the largest SD in the linear measurement Co-Gn (mandibular unit length) (4.43 mm) and the lowest range of variation in the A-Pog (0.10 mm). In fact, Co and Gn were the least reliable landmarks. We also found changes in the SNA angle for three patients. Two patients' diagnoses were modified from maxillary retrusion and protrusion to normal, and other was altered from normal position of the maxilla to protruded maxilla. Moreover, patients' skeletal classification did not seem to change.

Table 3.4. Standard deviation for each linear and angular measurement, identified by two observers on 20 radiographs.

Patient	**A-N (±2.70mm)**	**Co-Gn**	**Co-A (±6.00mm)**	**(Po-Or).(Go-Me) (±3.90º)**	**Pog-N (±3.80mm)**	**A-Pog (±2.40mm)**	**Convexity of Point A**	**Go-Me (±5.00mm)**	**S-Go (±6.00mm)**
P1	0.05	0.47	1.19	1.86	1.03	0.93	0.04	3.61	2.52
P2	0.78	1.90	0.58	0.18	0.71	0.31	0.41	3.26	1.63
P3	0.31	0.97	1.75	0.91	0.29	0.38	0.07	1.03	0.40
P4	0.03	0.71	1.84	0.87	0.02	0.89	0.14	2.07	0.17
P5	1.20	0.25	0.45	1.84	1.68	0.15	0.23	1.12	0.08
P6	1.23	0.34	1.42	0.29	0.13	1.28	1.09	0.54	1.22
P7	3.13	0.68	2.49	0.60	2.11	0.25	0.42	0.54	0.76
P8	0.49	1.33	0.08	1.66	2.06	0.96	0.61	0.78	0.01
P9	0.93	1.10	0.08	1.82	0.26	0.09	0.95	0.69	1.73
P10	0.46	0.23	0.78	1.04	1.40	0.43	0.27	1.81	1.36
P11	0.01	0.35	1.34	0.99	0.52	0.18	0.29	0.08	0.79
P12	0.41	3.73	1.14	0.42	0.23	0.31	0.30	4.18	2.60
P13	0.54	4.43	3.16	0.94	1.44	0.09	0.36	1.85	2.90
P14	0.51	2.60	0.52	0.64	1.03	0.18	0.01	2.15	0.88
P15	1.82	3.45	0.59	1.96	0.01	0.58	0.57	0.20	1.34
P16	2.65	0.30	0.85	2.36	0.27	0.37	0.72	3.74	2.76
P17	1.50	3.45	1.72	1.15	1.53	0.01	1.15	0.45	2.14
P18	1.07	1.15	1.35	1.15	1.43	0.35	0.37	2.93	1.67
P19	0.63	3.15	1.76	1.55	2.20	0.21	0.77	0.78	3.78
P20	0.16	1.47	2.34	0.26	0.64	0.10	0.49	1.46	1.26

3.4 Discussion

The main errors occurring in 2D cephalometric analysis include projection and tracing errors. The most important source of tracing errors occurs in landmark identification. It is known that intra-observer error is generally less frequent than the inter-observer one (Chen *et al.,* 2004; Chen *et al.,* 2000). We have found low variability and high agreement in landmark identification for the intra-observer evaluation (ICC > 0.90). Results have shown a higher rate of inter-observer errors in the identification of landmarks. Silveira and Silveira, 2006, revealed a very low reproducibility among dentomaxillofacial radiologists in the identification of landmarks (Silveira and Silveira*,* 2006). Similarly, we found that inter-observer variation may influence the reliability of landmark identification. Some authors believe that an individual perception of the landmark definition can lead to variations on angular and linear measurements (Lau *et al*., 1997; Miloro *et al.*, 2013). Other authors' state that, even in severe cases, accuracy of cephalometric analysis is not affected (Wah, 1995). The average value of measurements performed by all observers was used as the gold standard for a specific landmark to quantify the degree of error. Inter-observer error was used as a variable when determining reliability, i.e., the dispersion of error around the "best estimate" for each landmark. Differences in landmark identification were seen. Nevertheless, this may be considered to have a low clinical impact. In general, we found significant statistical differences for some landmarks horizontal component in both groups. DMFRs revealed, overall, considerable variations when identifying the horizontal components of landmarks Or, Me, ANS and LIA. Orthodontists showed significant differences in points ANS, Or, Po, Co and Me. Previous studies report identification errors greater than 1 mm for the following landmarks: point A, ANS, Ba, Co, Or and Po (Wah, 1995, McClure *et al.* 2005; Chien *et al*, 2009). Miethke, in 1989, stressed that the most reproducible landmarks were LIB and UIB, and that the majority of the landmarks revealed a SD of 2.0 mm. We found that the majority of the studied landmarks varied more than 1 mm. Accordingly, McClure *et al*., in 2005, observed a reduced reliability in the horizontal direction of ANS, PNS and Me; the vertical component of point Pog; and both components of points Ba, Go, Co, Or and Po.

The line formed by Po and Or (Frankfort horizontal) is important, since it establishes the horizontal standard reference plane. Some linear and angular measurements are evaluated according to this plane, such as the results given by Ricketts' cephalometric analysis. We assume that projection and tracing errors might be the reason for these results. Many factors can interfere with the reliability of cephalometric landmark identification, including the nature of cephalometric landmarks, resolution and quality of digital images, and also training level or experience of observers (Chen *et al*., 2004; Houston *et al.,* 1986). All observers had significant experience in cephalometric analysis. Previous studies have shown that observer's experience can lead to a wider variation, however the degree of error is similar among observers with the same training background (Gravely and Benzies 1974; McClure *et al,* 2005). Other major cause of error regarding reliability of cephalometric landmark identification is the specific nature of a landmark. Superimposition of adjacent structures on the radiograph may complicate the identification of certain landmarks, such as Co and Po. Chien *et al.,* in 2009, found a high variation for the vertical component of point Go. Comparing to our findings, the variation of this landmark was smaller than 1 mm for one orthodontist and one DMFR. This could have happened due to difficulty in establishing the landmark along broadly curved structures, such as the mandible.

Some cephalometric landmarks are more reliable in either the horizontal or vertical plane, making the distribution of errors follow a pattern of a non-circular envelope (Baumrind and Frantz, 1971). Differences in identification of landmarks were found in both groups along both axes. Although greater differences, were seen on the horizontal axis. Orthodontists and DMFRs revealed more variations when identifying the horizontal component of some landmarks.

Despite that the majority of landmarks revealed low reproducibility, the *x* and *y* components of points Or, Go, Gn and LIA, the *x* coordinates of points Po, ANS, Co, PNS and the *y* component of point B showed a mean value, for at least one observer, higher than 2 mm. Landmark identification errors inferior to 1 mm are considered accurate (Richardson, 1981; Chen *et al*, 2000). Other authors showed that a difference of 2 mm is considered acceptable and does not have any influence on orthodontic diagnosis and treatment plan (Kamoen *et al*., 2001; Lau *et al,* 1997). We found differences in the SNA angle on three patients. Apart from that, we did not see differences between diagnoses associated with a variation of 1 to 2 mm in landmark identification. A greater difference regarding all landmarks would probably have an impact on diagnosis and possibly, in treatment planning. The latter is important, since variations that might exceed the SD proposed for a predetermined linear or angular measurement performed by one observer could demonstrate lack of knowledge and/or experience. This is particularly important as we considered that variations higher than 2 mm can relate to lack of knowledge and/or experience of the observers. Although inter-observer variations occurred heavily on landmark identifications, they may not have an impact in patient diagnosis. This is one of the reasons why landmark identification reproducibility is quite low. Thus, reliability of cephalometric analysis should be questioned. Depending on the observer and on the type error, different results with no impact on diagnosis and treatment planning may appear. The existing literature suggests that lateral cephalometric radiographs have been used without adequate scientific evidence of its utility, and that it is often used prior to treatment. The evidence to agree or disagree with the usefulness of this radiographic technique in orthodontics is limited (Durão *et al*., 2013).

Many variables contribute to the final diagnosis and treatment plan in orthodontics, such as face-bow recording, clinical examination, intra- and extra-oral photographs. Therefore, it is difficult to predict if a single error on landmark identification will have an impact on clinical practice. A combined error on dental casts and cephalometric analyses may lead to erroneous decisions about teeth extraction (Silveira and Silveira, 2006). The patient should be treated with maximum accuracy in every steps of diagnosis and treatment.

3.5 Conclusions

We verified that some landmarks were not as reproducible as others, either on the horizontal or vertical component. The most consistent landmark identified in both groups was the LIB, while the least reliable points were Co, Gn, Or and ANS. Furthermore, the greatest variation was found in Co-Gn plane. Our results suggest a low reliability in the identification of cephalometric landmarks and lower agreement between orthodontists. In the presence of a range of variation from 1 to 2 mm on landmark identification, the patient's diagnosis was altered. Moreover, we found changes in the SNA angle. Further studies focusing on the impact of deviating cephalometric analysis on a larger sample and in borderline cases may be needed to determine the real clinical impact.

CHAPTER 4. Variations in Sella landmark identification and its effect in angles SNA and SNB in lateral cephalometric radiographs

VARIATIONS IN SELLA LANDMARK IDENTIFICATION AND ITS EFFECT IN ANGLES SNA AND SNB IN LATERAL CEPHALOMETRIC RADIOGRAPHS

4.1 Introduction

The *Sella turcica* is routinely traced for cephalometric analysis. It is defined as a depression in the skull base, where the pituitary gland is situated. Since the introduction of lateral cephalometric radiography by Broadbent in 1931, this radiographic technique has been widely used in orthodontics to evaluate cranial and dentofacial growth (Broadbent, 1931). Cephalometric analyses are based on angular and linear measurements, which might present some errors (Hussels *et al.*, 1984). Nevertheless, it is widely used in orthodontics. One of the major causes of error in cephalometric analyses occurs in the identification of landmarks, moreover certain cephalometric points are more difficult to identify. The Sella (S) point, which is located at the midpoint cavity of the *sella turcica*, is an example (Proffit *et al.,* 2006). This point is considered to be a floating landmark because it is identified by visual criteria and is not situated on a specific structure.

In 1953, Steiner developed a cephalometric analysis, known nowadays as the first of the modern cephalometric analyses. He indicated some craniofacial norms "which expressed the concept of a normal average American child of average age" (Steiner, 1953). An analysis based on dentoalveolar compensatory mechanism was proposed, in order to determine the nature of malocclusion. Steiner created his cephalometric analysis based on the analyses of Downs, Margolis, Riedel, Thompson and Wylie, combined with some of his own cephalometric values. The major influence to his work was from Riedel. He studied several patients' cephalograms for the relationship of the maxilla to the cranium and mandible. In 1952, Riedel defined the ANB angle, which is based on A and B points - the deepest bony outline points of the maxilla and mandible, respectively. He established his analysis considering the SN line, which refers to the anterior cranium base, and used as reference, angles SNA and SNB to provide information on the upper and lower facial prognatism. He indicated that the arithmetic difference between SNA and SNB would result in the ANB angle. It indicates the magnitude of skeletal-jaw discrepancies, and was the major reference for Steiner, since it is an expression of the dental apical base relationships. ANB angle may vary according to the vertical distance between landmarks N and points A and B (vertical height of the face). If this distance increases, the ANB angle decreases. Incorrect identification or growth may create different positions of point N, and will also affect angle ANB (Proffit *et al*., 2006). Anteroposterior jaw relation can be determined either by angle ANB or Wits appraisal. Studies have shown that these two methods present some limitations. Angle ANB can vary according to cranial base length and/or jaws rotation, and Wits appraisal can change with the occlusal plane. Therefore, some authors suggested that both methods should be used (Ishikawa *et al*., 1998). Other factors can affect angle ANB, including: patient's age (ANB decreases with age), position of point N, SN plane rotation, occlusal plane, and maxillary or facial prognatism (Oktay, 1991).

Steiner indicated that, besides knowing where the discrepancy was, the most important factor was to know its magnitude (Proffit *et al*., 2006). The normal value for the angle ANB in a Caucasian should be of 2º (Proffit *et al*., 2006).

These angles evaluation plays an important role in the diagnosis and treatment of malocclusions. Anteroposterior position of the maxilla in relation to the anterior cranial base is determined by angle SNA. Its standard value is 82° ± 3°. The SNB defines the anteroposterior position of the mandible, for which the standard value is 79° ± 3°. If the SNA or the SNB is greater or lower than the standard value, this indicates that the mandible or maxilla is either positioned anteriorly or posteriorly to the cranial base. If SNA is greater than 85°, it indicates a maxillary protrusion, and if it is lower than 79°, it reveals a maxillary retrusion. Likewise, if SNB is lower than 76°, it suggests a mandibular retrusion; and if it is greater than 82°, it indicates a mandibular protrusion. This interpretation is only valid if the SN plane is normally inclined to the true horizontal (Po-Or) and the N point position is normal (Proffit *et al.*, 2006). SN plane represents the anterior cranial base. Variability in S landmark identification may modify angles SNA and SNB. On our previous study, we found that landmark S had low intra- and inter-observer variability, which was consistent with other studies (Miloro *et al.,* 2013, Oz *et al.,* 2011; Chen *et al.*, 2004). Errors in cephalometric analyses may occur by numerous reasons. One of the most important errors happens due to inconsistent and imprecise landmark identification. Inaccurate landmark identification may lead to erroneous diagnoses and treatment plans for orthodontic cases (Chen *et al.*, 2004; Tng *et al.*, 1994). Moreover, some authors stated that different levels of knowledge and observers background play an important role in landmark identification (Miloro *et al.*, 2013; Kamoen *et al.*, 2001; Gravely and Benzies, 1974; Kvam and Krogstad, 1969). Other authors believe that errors can be caused by different individual conceptions of landmarks' definitions and its perception, rather than education and training (Chen *et al*., 2004; Kamoen *et al*., 2001; Lau *et al*., 1997). Skeletal landmarks, like points A, B and N, play an important role in patient's skeletal diagnosis. The aims of this study were to determine intra- and inter-observer precision in identification of the landmarks Sella (S), Nasion (N), point A and B, as well as to determine how it can interfere with angular measurements of SNA and SNB by orthodontists and dentomaxillofacial radiologists.

4.2 Materials and Methods

Twenty digital lateral cephalometric radiographs were selected from the database at the Oral Imaging Center, University of Leuven. Lateral cephalograms were acquired by positioning the patients in a standard cephalometric device (Veraviewepocs 2D®, J. Morita, Kyoto, Japan). The exposure values were set at 77 kV and 7.2 mA, with an exposure time of approximately 1.6 s, according to each patient. The radiographs were considered to have good quality. Inclusion criteria were:

- No evidence of current orthodontic treatment.
- Digital cephalometric image were of good quality to allow the identification of landmarks, and the ruler on the radiograph was clearly visible, allowing calibration of the images in the cephalometric analysis software program.
- There were no unerupted or partially erupted incisors that could have compromised landmarks identification.
- No gross skeletal asymmetry.

All of images selected were then exported in TIFF format and introduced in the PowerPoint software (Version 2010; Microsoft, Redmond, Washington, USA). The Sella-Nasion horizontal plane was used as a reference and all images were orientated accordingly.

Analysis

A PowerPoint file with 20 lateral cephalometric radiographs was sent by e-mail to ten experienced observers (five orthodontists and five dentomaxillofacial radiologists [DMFR]). Each observer identified the following landmarks on each radiograph: Sella (S), Point A (A), Point B (B) and Nasion (N), by placing a predefined red dot (Figure 4.1). A detailed explanation of the procedure and definitions of the 4 landmarks were given (orally and on paper) to all observers. Thus, observers followed the same landmarks definitions during identification process. The same procedure was repeated 8 weeks after to test the intra-observer variance.

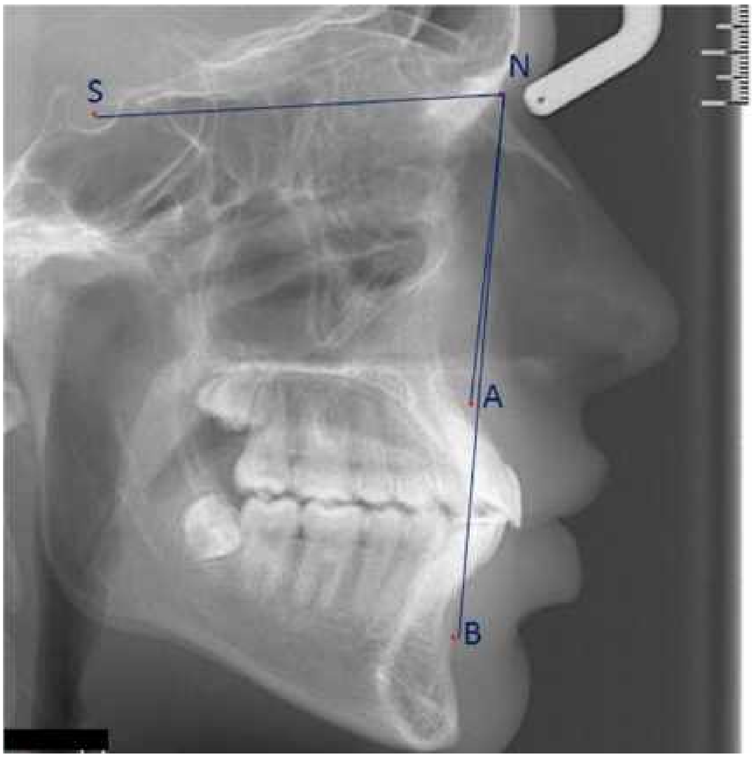

Figure 4.1. Lateral cephalometric radiograph showing the identified landmarks and the measured angles.

After receiving all files, the main observer exported the images in TIFF format, and subsequently imported them to the computerized program for cephalometric analysis (*Radiocef Studio 2*; Radio Memory Ltd., Belo Horizonte, Brazil). Calibration of the actual size of each image, in millimetres, was based on the measurement of the known distance (10 mm) between the two fixed points of the LCR. The vertical and horizontal positions of each landmark were recorded in the format of *x* and *y* coordinates. The angles SNB (indicates whether the mandible is normal, prognathic, or retrognathic) and SNA (indicates whether the maxilla is normal, prognathic, or retrognathic) were measured.

The digitized landmarks' coordinates and measured angles were then copied into the Excel software (Version 2003; Microsoft, Redmond, Washington, USA).

Statistical analysis

Variation of landmark identification and angle measurements differences, mean, standard deviation and intraclass correlation coefficients (ICC) was analysed. Intra- and inter-observer reliability for each landmark in the *x* and *y* directions were calculated using ICC with a confidence interval of 95%. General guidelines for ICC rate as excellent an ICC of >0.90, an ICC of 0.75–0.90 as good, and an ICC of <0.75 as poor to moderate reliability (Shrout and Fleiss, 1979). Angles SNA and SNB were categorized according to standard values, defined by Steiner's cephalometric analysis (Proffit *et al.,* 2006). Weighted kappa was calculated as well as the percentage of agreement (agreement measurements/total measurements). SNA and SNB values were observed and patient diagnosis was performed for each observation. The Statistical Package for Social Sciences 20.0 for Windows (SPSS Inc., Chicago, Illinois, USA) was used for statistical analysis with statistical significance for all tests at $\alpha = 0.05$.

4.3 Results

For the intra- and inter-observer reproducibility regarding identification of the horizontal and vertical components of landmark S, the ICC ranged between 0.75 and 0.90, implying that there was an intra-observer good agreement, as well as between dentomaxillofacial radiologists and orthodontists (Table 4.1). Two observers obtained an ICC on the horizontal component as poor to moderate (<0.75). One observer also achieved a poor-to-moderate agreement on the vertical component of the S landmark. In general, orthodontists tended to identify the S point more to the right (*x* direction -0.22 mm) and lower (*y* direction -028 mm) than dentomaxillofacial radiologists.

Table 4.1. Intra- and inter-observer differences in S landmark identification (mm).

	Horizontal component (*x*)				**Vertical component (*y*)**			
	Mean	**SD**	**ICC**	**[CI 95%]**	**Mean**	**SD**	**ICC**	**[CI 95%]**
DMFR	0.08	0.40	0.984	0.961-0.994	-0.17	0.52	0.995	0.986-0.998
Orthodontists	0.25	0.54	0.943	0.861-0.977	-1.09	1.26	0.961	0.905-0.984
DMFR-Ortho	-0.22	0.95	0.443	0.863-0.977	-0.28	1.47	0.944	0.865-0.978

SD – standard deviation; ICC- Intraclass correlation; CI (5% - 95%) confidence interval

Intra- and inter-observer reliability for the SNA and SNB angles is shown on Table 4.2. We found, in general, ICC values superior to 0.90, which shows an excellent agreement intra- and inter-observer for the two angles. An ICC between 0.75 and 0.90 was identified by three observers regarding angle SNA, and for one observer regarding SNB. Orthodontists tended to produce larger SNA (-0.18º) angles than dentomaxillofacial radiologists. In contrast, angle SNB tended to be lower when measured by orthodontists (0.55º) (Table 4.2).

Table 4.2. Intra and inter-observer differences in SNA and SNB angles (in °).

	SNA				SNB			
	Mean	**SD**	**ICC**	**[CI 95%]**	**Mean**	**SD**	**ICC**	**[CI 95%]**
DMFRs	-0.10	0.60	0.905	0.770-0.961	-0.01	0.52	0.976	0.948-0.997
Orthodontists	0.31	0.70	0.927	0.824-0.970	0.23	0.67	0.967	0.920-0.994
DMFR-Ortho	-0.18	0.65	0.977	0.942-0.991	0.55	1.26	0.937	0.845-0.994

SD – standard deviation; ICC- Intraclass correlation; CI (5% - 95%) confidence interval

To analyse the differences in landmark identification and its effect on angles SNA and SNB, a correlation method was used (Table 4.3).

Table 4.3. Correlation between point S identification of and its effect on SNA and SNB angles.

	SNB		SNA	
	r	***p***	**R**	***p***
DMFRs	-0.308	0.186	0.078	0.743
Orthodontists	-0.092	0.699	-0.074	0.075

r- Pearson correlation; *p* – 0.05

On the one hand, we found that if there were higher differences in landmark identification, the SNA angle was smaller (DMFRs: -0.308º and orthodontists: -0.092º). On the other hand, for the SNB, we saw that if the distance between landmarks was higher, the angles tended to be higher (DMFR: 0.078º and orthodontists: -0.074º). Nevertheless, these results were not statistically significant.

We found that, in a number of cases, maxilla and mandibular positions diagnosis was changed due to differences in SNA and SNB angles (Table 4.4). SNA angle revealed changes between 0.50º and 7.92º, while the minimum difference on SNB was of 0.31º and the major was of 8.42º.

Table 4.4. Number of cases (n) in which diagnosis was changed, regarding the SNA and SNB standard values, according to each observer. Minimum and maximum degree variations are indicated.

<table>
<tr><th></th><th>SNA (n)</th><th>Minimum (º)</th><th>Maximum (º)</th><th>SNB (n)</th><th>Minimum (º)</th><th>Maximum (º)</th></tr>
<tr><td>Observer 1</td><td>4</td><td>1.33</td><td>6.83</td><td>1</td><td colspan="2">1.96</td></tr>
<tr><td>Observer 2</td><td>1</td><td colspan="2">1.62</td><td>3</td><td>0.31</td><td>5.21</td></tr>
<tr><td>Observer 3</td><td>6</td><td>0.50</td><td>5.99</td><td>1</td><td colspan="2">0.90</td></tr>
<tr><td>Observer 4</td><td>6</td><td>1.07</td><td>6.07</td><td>1</td><td colspan="2">6.00</td></tr>
<tr><td>Observer 5</td><td>0</td><td colspan="2">-</td><td>0</td><td colspan="2">-</td></tr>
<tr><td>Observer 6</td><td>1</td><td>1.31</td><td>7.24</td><td>7</td><td>2.45</td><td>8.42</td></tr>
<tr><td>Observer 7</td><td>5</td><td>0.64</td><td>7.92</td><td>4</td><td>0.40</td><td>8.71</td></tr>
<tr><td>Observer 8</td><td>6</td><td>1.01</td><td>4.07</td><td>2</td><td>1.01</td><td>2.73</td></tr>
<tr><td>Observer 9</td><td>8</td><td>1.58</td><td>6.00</td><td>3</td><td>1.21</td><td>2.39</td></tr>
<tr><td>Observer 10</td><td>5</td><td>1.81</td><td>4.95</td><td>7</td><td>1.07</td><td>5.55</td></tr>
</table>

Amongst orthodontists, the degree of these angles was different changing patients' diagnosis 25 times, regarding maxillary position (SNA angle). Most of the times, position of the maxilla diagnosis was changed from normal to retruded. Only in one case diagnosis was modified from protruded to retruded maxilla.

A slightly higher percentage of agreement was evident for orthodontists (70%) in relation to dentomaxillofacial radiologists (55%).

When measured by dentomaxillofacial radiologists, angle SNB, was changed in six patients. The percentage of agreement for the SNB was equal in both groups (80%). The value in which the SNB is considered normal was changed by orthodontists 23 times. Major differences were noted between the diagnoses of normal to retruded mandible. The diagnosis was modified from retruded to protruded mandible only in one case.

4.4 Discussion

Inaccurate landmark identification is one the most frequent source of error that occurs in cephalometric analysis (Chen *et al.*, 2004; Tng *et al.*, 1994). Also, observers training experience and background may lead to errors (Miloro *et al.*, 2013; Kamoen *et al.*, 2001; Gravely and Benzies, 1974; Kvam and Krogstad, 1969). All orthodontists involved in this study had the same background experience, since they were trained at the same institution. Some differences in background education existed between dentomaxillofacial radiologists. To eliminate possible errors that could arise from this, a detailed explanation of landmark's definition was given by the main author. Furthermore, a low variability in the identification of the point S was found.

Depending on the magnitude of the error landmarks identification, patients diagnosis can change. We studied how an imprecise identification of point S could lead to different SNA and SNB angles. According to some authors, landmark identification errors inferior to 1 mm are considered accurate (Chen *et al.*, 2000; Richardson A., 1981). Other authors believe that a difference of 2 mm is considered acceptable and does not have any influence in orthodontic diagnosis and treatment plan (Kamoen *et al.*, 2001; Lau *et al.*, 1997). The variation in identification of the S landmark was relatively low, presenting a deviation of -0.22 mm in the x direction and of 0.28 mm in the y direction. Some cephalometric landmarks are more reliable in either the horizontal or vertical plane (Baumrind and Frantz, 1971). We had previously revealed a low variability for the S landmark. In that previous study, we had suggested that, with a small range of variation (1 to 2 mm) in landmark identification, patient diagnosis could change. Nevertheless, in general, in the present study, intra- and inter-observer agreement for the SNA and SNB angles was good (ICC 0.75-0.90).

Steiner used angles SNA and SNB for patients' diagnoses. We revealed that minor changes in S landmark identification could change both SNA and SNB classifications. Dentomaxillofacial radiologists, showed differences in patients diagnosis in 17 cases out of 100 observations. Regarding SNB, mandible position diagnosis was changed in 5 cases. Overall, a higher variability was found amongst orthodontists. Between orthodontists, maxillary position diagnosis was changed in 25 cases, while the mandible's position diagnosis was changed in 23 cases. Larger variations were found in both groups on the SNA angle. This could happen due to point A identification, which is more difficult to identify than point AB. We should remember that "standard" values for these two angles were defined by a small number of individuals that were supposed to be representative of a population. Due to the inflexible interval given for these angles, patients who deviate just slightly from the standard value may have an erroneous diagnosis. We suggested that changes of 0.50º in SNA and of 0.31º in SNB could alter patients diagnosis concerning mandible and the maxilla positions. An incorrect diagnosis may lead to erroneous orthodontic treatment. To perform a diagnosis and treatment plan in orthodontics, many variables are taken into account; therefore, each step of the process should be performed with maximum accuracy. The results of this study question the validity of cephalometric analysis in orthodontics, since a small variation on these landmarks' identification can lead to different diagnosis and, thus, different treatment plans.

4.5 Conclusions

In conclusion, identification of the Sella landmark revealed a better agreement amongst dentomaxillofacial radiologists. Orthodontists, however, showed a larger variability in S identification and, consequently, the SNA and SNB angles drifted significantly. Small modifications in identifications of the S, N, A and B points may lead to differences in angles SNA and SNB. Therefore, patient diagnosis and treatment can vary. More differences existed regarding SNA than SNB. Further studies on a larger patient sample with inclusion of more borderline cases may be needed to determine the real clinical impact on treatment planning.

CHAPTER 5. Influence of using lateral cephalometric radiography in orthodontic diagnosis and treatment planning

Ana R Durão, Ali Alqerban, Afonso P Ferreira, Reinhilde Jacobs. **Influence of Lateral cephalometric radiography in orthodontic diagnosis and treatment planning**. Accepted for publication in “Angle Orthodontist” on the 5th of April 2014. Angle Orthodontist, Vol 86, No 4, 2014. DOI.10.2319/011214-41.1.

INFLUENCE OF USING LATERAL CEPHALOMETRIC RADIOGRAPHY IN ORTHODONTIC DIAGNOSIS AND TREATMENT PLANNING

5.1 Introduction

Lateral cephalometric radiography (LCR) is widely used in orthodontic assessment and treatment planning. Despite that, its usefulness in orthodontics remains questionable. Silling *et al*., in 1979, stressed that lateral cephalometric radiography was only needed for Class II division 1 patients. Later, in 1991, Han *et al*. stated that patient's examination together with dental casts provided sufficient information to perform a diagnosis. According to them, only 55% of treatment plans were changed after the LCR evaluation. In the same vein, Bruks *et al*., in 1999, suggested that in 93% of the cases, treatment plans remained unachanged after the LCR evaluation. They evaluated the patient, dental casts, and extraoral photographs. In contrast Pae *et al*., in 2001, revealed that in patients with Class II division 2 malocclusion and bimaxillary protrusion, this radiography could change the decision with regard to teeth extraction. In 2008, Nijkamp *et al.* reinforced that LCR does not seem to have any impact on orthodontic treatment plan for Class II division 1 patients. Recently, in 2011, Deveraux *et al.* concluded that only in one out of six patients' orthodontists decide to change their treatment decisions regarding with regard to tooth extraction. In contrast with the previous study, they suggested that LCR may be justified for orthodontic treatment. Considering the controversy in the literature, the present aim was to further explore the impact of additional LCR in orthodontic diagnosis and the treatment planning.

5.2 Materials and Methods

Forty-three patients with pretreatment diagnostic records were randomly selected. All patients were seeking orthodontic treatment at the Faculty of Dental Medicine of the University of Porto. The study was approved by the Ethics Committee of the Faculty of Dental Medicine of the University of Porto (900079). The patients' ages ranged from ten to 42 years-old (24 female and 19 male). Orthodontic diagnostic records included: three photographs of the angle of trimmed dental casts, digital lateral cephalometric and panoramic radiographs, as well as standard clinical photographs comprising seven intra- and four extraoral pictures (Figure 5.1). The patients' identification was blurred to avoid recognition. All the blinded information was saved in a PDF file and recorded in a compact disk and given to each observer. Ten qualified orthodontists were involved in this study, with experiences ranging from five to 24 years. Patients' records were evaluated during two sessions. The time interval between observations was at least eight weeks. At the first session orthodontists evaluated records without LCR. In the second session the same information was presented, but this time the LCR was added. Between the two sessions the order in which the cases were presented was altered to avoid bias.

The evaluation process for the two sessions involved the use of a questionnaire concerning the diagnosis and treatment plan; the questionnaire contained the following questions:

1. Skeletal relationship: neutro-, disto-, or mesio-relation?

2. Angle classification of the occlusion based on molar relationships: Class I, Class II or Class III

3. Detection of any abnormality?

4. The treatment plan will be: orthopedic growth modification; orthognathic surgery; dentoalveolar compensation?

5. Is there enough space for all teeth to erupt?

6. Would you extract any teeth in this patient? If yes, which one?

7. Would you expand the upper arch?

8. Would you use anchorage in the maxilla, mandible, or both?

9. Do you expect any complications during the treatment?

10. How long do you expect the treatment to last?

11. Would you need any additional information to make a decision? Which information?

12. How long has it been since you were qualified as an orthodontist?

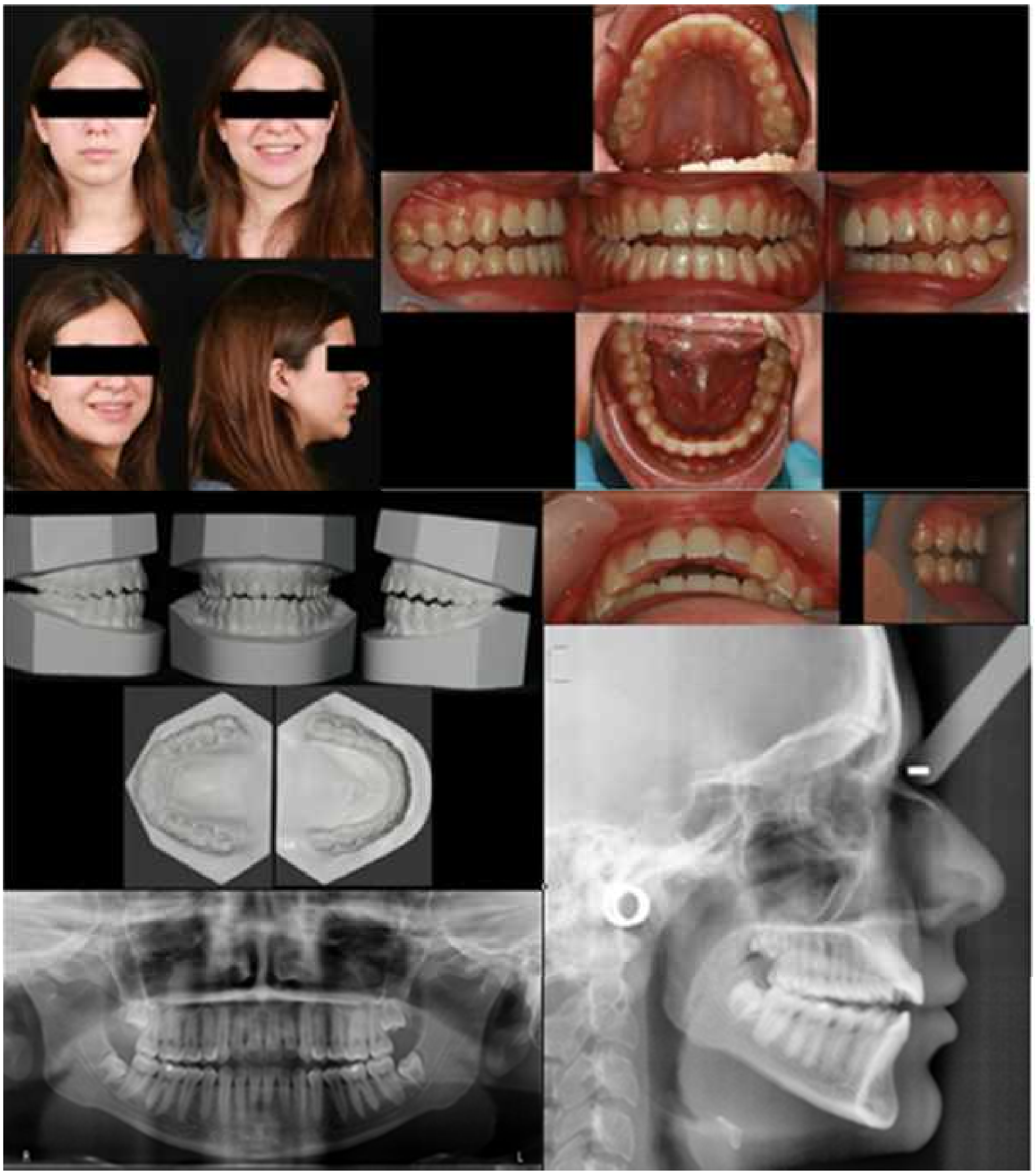

Figure 5.1. Example of the information given to orthodontists.

<u>Statistical analysis</u>

The percentage of agreement of the answers between the two sessions (ratio of agreement between cases and total cases used) was evaluated. This was carried out for each patient to test for differences in the percentages of changed decisions regarding diagnosis and treatment planning.

5.3 Results

The percentage of agreement between sessions was lower with regard to diagnosis than it was with regard to treatment planning (Table 5.1). Treatment planning seemed to be changed, on average, in 36% of the cases by adding a LCR. In addition, the skeletal classification diagnosis was changed, on average, in 56% of the cases, and, in general, in 52% of the cases the malocclusion classification seemed to be altered. The most frequent changes appeared in Class II malocclusion patients.

With regard to skeletal classification, the least experienced observer was the least consistent (28%), while the more experienced observer was the more reliable (67%). On average, ten cases were classified in the first session as Class II, and after evaluating the LCR the diagnosis of the skeletal classification changed to Class I. In nine cases skeletal classification was altered from Class I to Class II. Overall only in a single case did the orthodontists change from Class III to Class I. The presence of abnormality revealed a very good agreement between the two sessions (overall 87%). With regard to treatment modalities, in general there was an agreement of 64%. The most experienced observer revealed 80% of agreement between sessions, changing the treatment plan in only 8 cases, while the lower percentage was of 37%, seen in an observer with ten years of practice. In 26 cases the treatment modality was changed in the majority of cases, being altered from dentoalveolar compensation to surgery. The most frequent modifications in treatment modalities were seen in Class II patients. One observer changed the decision to extract in 19% of the cases after evaluating the LCR. Table 5.2 demonstrates the comparisons with regard to treatment duration, in months, between the first and second sessions. Only two observers revealed statistically significant differences. After viewing the LCR, one observer suggested that the treatment should be longer. On the second occasion another observer, proposed shorter treatment duration. Two orthodontists stated that LCR was needed for a correct evaluation of all cases. At the second observation, one still needed the LCR analysis (in 27 out of 43 cases) and the other was satisfied. One revealed that to perform a precise diagnosis, dental casts together with LCR was necessary for all cases. The others judged the LCR helpful only for some cases, varying between Class I and Class II (Table 5.3). Consensus was achieved related to clinical examination. In general, the orthodontists stressed the need to examine the patients personally.

Table 5.1. Mean percentage of agreement between the first and second sessions for all observers.

Questions	**Percentage of agreement**
Q1	43%
Q2 right	47%
Q2 left	50%
Q3	87%
Q4	64%
Q5	58%
Q6	56%
Q7	58%
Q8 maxilla	58%
Q8 mandible	67%
Q9	65%
Q11	63%

Table 5.2. The mean differences in proposed treatment plan duration (months) between the two sessions.

	Mean (months)	**SD**	***p****
Observer 1			0.297
1st session	28.23	15.06	
2nd session	30.39	13.65	
Observer 2			0.077
1st session	24.42	3.79	
2nd session	25.57	3.26	
Observer 3			0.366
1st session	22.59	2.97	
2nd session	23.30	1.95	
Observer 4			0.142
1st session	27.07	6.72	
2nd session	25.26	5.16	
Observer 5			0.328
1st session	26.38	5.32	
2nd session on	25.58	6.10	
Observer 6			0.979
1st session	20.93	4.61	
2nd session n	21.00	4.82	
Observer 7			0.234
1st session	30.28	9.88	
2nd session	31.26	4.64	
Observer 8			**0.033***
1st session	25.26	3.60	
2nd session	32.09	15.87	
Observer 9			0.726
1st session	28.47	5.93	
2nd session	28.09	7.09	
Observer 10			**0.044***
1st session	28.50	7.29	
2nd session	28.09	7.09	

SD- Standard Deviation; *$p < 0.05$

Table 5.3. Number of additional information required for each observer in the 1st and 2nd observations.

Observer	Additional information required							
	1st occasion				2nd occasion			
	LCR	DC	LCR + DC	LCR + DC + CBCT	LCR analysis	DC	LCR analysis + DC	Facial and LCR analysis
Observer 1	2	18	16	2		28	10	
Observer 2	43				27			
Observer 3	21				2			
Observer 4	16		11		10	10	8	
Observer 5	29				7	9		
Observer 6	27				6			
Observer 7	28				25		5	7
Observer 8	43							
Observer 9		43*				43*		
Observer 10	29				29			

DC –Dental casts; CBCB – cone-beam computed tomography. * plus intra-oral x-rays, natural head position.

5.4 Discussion

We performed this study, to highlight the usefulness of two-dimensional (2D) cephalometric imaging for orthodontic treatment planning. LCR has been routinely used since its discovery, although, major concerns arise when patients are exposed to radiation when not clearly justified. According to the ALARA principle, there is a need to reduce radiation exposure and eliminate unnecessary radiographies. We selected the patients at random to allow our sample to be representative of a population, rather that choosing any particular malocclusion or specific age. Forty-three patient files were selected. At first an experimental observational setup was performed with three orthodontists who evaluated five patient files and validated the questionnaire. After that the study proceeded. Patient records were reordered at the second observation so that orthodontists could not recognize the sequence. We performed two observation sessions, with a minimum of 8 weeks between sessions. Observers had some differences in terms of their background experience. The most experienced observer had completed 24 years of practice while the least had only five years of experience. The observer's background plays an important role, with regards to the necessity of having additional diagnostic tools to perform a diagnosis. It was suggested that the need for LCR or its analysis was more dependent on background rather than on years of experience. For example, observer 8, who was the most experienced thought that LCR would be helpful for all cases and observer 5, with only six years of experience, only judged it necessary to use LCR in 27 of the cases. However, after viewing the LCR, observer 8 ascertained that cephalometric analysis was not necessary. In contrast, observer 6 judged that the cephalometric analyses would be helpful. In general, the biggest complaint from orthodontics was the absence of (1) clinical examination and (2) the reason why the patient sought orthodontic treatment. Today digital records are accepted for diagnosis and treatment planning for professional examinations. Two orthodontists revealed that in order to perform a correct diagnosis and treatment planning they needed LCR for all cases. Another orthodontist ascertained that for all cases the natural head position, dental casts in centric relation, and LCR together with clinical examination of the patient would be important to render a diagnosis and develop a treatment plan. The need of cephalometric analysis was also asserted by some orthodontists. Two orthodontists revealed that they did not need a cephalometric analysis, while the radiographic examination was useful. One orthodontist required a Cone Beam Computed Tomography (CBCT) for two cases; in these the patients had impacted canines.

The questionnaire involved 12 questions; the first three questions concerned diagnosis. Questions number 4 and 6-10 related to treatment planning. In general, the percentage of agreement was higher regarding treatment planning. Some authors have ascertained that experienced orthodontists can achieve a correct diagnosis and treatment plan without viewing LCR (Atchinson et al., 1991; Silling et al., 1979). Other authors believe that diagnosis based on clinical examinations together with photographs and dental casts can provide sufficient information to develop a treatment plan. In this study, we found a moderately high percentage of agreement for treatment planning between the two sessions. This could suggest that LCR may not have influence on orthodontic treatment planning. With regard to skeletal pattern classification, our sample contained 19 patients with Class I occlusion; 19 patients with Class II occlusion and five patients with Class III occlusion. For that reason, it is impossible to ascertain that LCR is not needed for all patients since there is a great variation in malocclusions. To define strict selection criteria to perform a LCR is difficult. Even text books do not express this issue very clearly. The indication for LCR must be must be constructed on an individual basis rather than based general conditions (Nijkamp et al., 2008; Bruks et al., 1999; Atchison et al., 1991). Regarding treatment duration between the two sessions, the only statistically significant was found for two observers. Further studies focusing on this subject are encouraged.

5.5 Conclusions

The results of our study suggest that the majority of Portuguese orthodontists judge that LCR is important to producing a treatment plan. Despite that, it does not seem to have an influence on orthodontic treatment planning.

GENERAL DISCUSSION AND CONCLUSIONS

The present thesis addresses the usefulness of 2D cephalometry in orthodontic treatment. In fact, the use of an imaging modality without being clearly justified is worrying. Apart from that, once justified a basic principle in radiation protection is ALARA, stating that the radiation exposure should be “As Low As Reasonably Achievable”, balancing the benefit of the exposure with its detrimental effects. In radiography, this benefit is reflected as an improvement in diagnosis and/or treatment planning of the patient, owing to the additional information obtained from the radiological examination. Once clearly justified, a correct interpretation and analysis of the radiograph is mandatory. One of the major errors in cephalometric analysis was seen in the identification of landmarks. We also saw that observers’ background plays an important role, not only in cephalometric analysis, but also in the requirement of this radiography as a diagnostic tool to perform a diagnosis and/or a treatment plan. Certain landmarks are more reproducible than others, either on the horizontal or vertical component. The most consistent landmark identified amongst orthodontist and dentomaxillofacial radiologist was the LIB, while the least reliable points were Co, Gn, Or and ANS. Furthermore, the greatest variation was found in Co-Gn plane. Our results suggest a low reproducibility in the identification of cephalometric landmarks. In general, orthodontists revealed a lower agreement on the identification of landmarks than dentomaxillofacial radiologists. We saw that for a range of variation between 1 to 2 mm on the identification of landmarks, the diagnosis of the patient was altered. Nevertheless, small modifications in S, N, A and B points localization can lead to differences in the angles SNA and SNB. Thus, differences on patient diagnosis and treatment planning may happen. How can we rely on a diagnostic tool that has been proven to have so many factors that can interfere with its analysis? How can we justify its use too all patients undergoing orthodontic treatment? Minor changes in cephalometric landmarks identification can produce differences on angular as well as on linear measurements. Apart from that, Portuguese orthodontists indicated that cephalometric radiography was helpful for the majority of cases. Nevertheless, this may be rather related to the educational background, as an additional cephalometric radiography changed the treatment plan in 36% of the cases.

From the present thesis it can be concluded that the routine use of LCR for orthodontic treatment should be questioned, considering that a low percentage of LCR showed an impact in orthodontic treatment plan. And considering that, small variations (1 to 2 mm) in the identification of certain landmarks can lead to different angular measurements. Regarding diagnosis, in 56% of the cases skeletal classification was changed and in 52% of the cases mallococclusion classification was altered, after evaluating the LCR. We found that orthodontic diagnosis and treatment plan based on dental casts, intra- and extra-oral photographs and panoramic radiograph provided in the majority of cases, sufficient information to perform a treatment plan in orthodontics. Although establishing strict guidelines is impossible due to the variety of malocclusions, the indication to perform a LCR should be based on individual criteria.

Hypotheses

A series of hypotheses were defined in the Introduction & Hypotheses section of this thesis. Our study allowed to confirm or refute them, or to formulate interesting tracks for future research.

The various chapters and topics address the following hypotheses:

2D cephalometrics suffers a poor accuracy when compared to real skull analysis.

This topic was covered by *Chapters 2*. Although, radiographic measurements systematically overestimated the gold standard measurements on skulls, the differences found were most often less than 1 mm, which is generally within the accepted standard deviation. Yet, one should realize it can still create bigger deviations in angular measures.

2D cephalometrics has poor intra- and inter-observer variability, thus influencing planning and treatment decisions.

This hypothesis was addressed in *Chapter* 3, and was confirmed. Small variations in landmark identifications can change the SNA angle which can lead to a change in treatment plan. Some landmarks were more reproducible than others, either on the vertical or horizontal components.

Landmark identification on the point Sella as a reduced variability, and does not interfere with the angles SNA and SNB.

This hypothesis was refuted in *Chapter 4.* Besides the low variability Sella landmark identification we suggested that it can interfere with the SNA and SNB. After the evaluation of 20 LCR by ten observers, according to the mandibular position we saw that 42 patients had different proposed diagnosis by different observers. Mandibular position diagnosis was changed in 29 patients in relation to the SNB angle.

The availability of the 2D lateral cephalometric radiograph influences the orthodontic treatment plan and decision in some but not all cases.

Chapter 5 addressed this hypothesis. The majority of Portuguese orthodontists suggest that LCR is needed to produce a treatment plan. Yet, in 36% of the cases, it was changed by adding a lateral cephalometric radiograph. Regarding diagnosis, in 56% of the cases, the skeletal classification was modified after viewing the lateral cephalometric image. In the same way, 52% patients' occlusal classification was modified.

Conclusions and future prospects

In this thesis, a variety of topics regarding the usefulness of 2D lateral cephalometric radiography (LRC) were addressed. In our systematic review, we saw that many factors can contribute to the lack of scientific value for cephalometric analysis. It was showed that the use of LCR could be justified under specific clinical conditions. Despite that, we found low variability in orthodontic diagnosis and treatment plan performed with or without LCR. The present study was initiated by the fact that three-dimensional (3D) cephalometric analysis is emerging, while there is still lack of scientific evidence on the validity and reliability of two-dimensional (2D) cephalometric imaging for orthodontic treatment planning. We therefore recommend that LCR radiography should be justified on an individual basis.

SUMMARY

Since the introduction of lateral cephalometric radiography in 1931 by Broadbent in the United States and by Hofrath in Germany, this radiograph and its analysis have become standard tools in the orthodontic assessment and treatment planning. Notwithstanding the fact that it is widely used, the real value of lateral cephalometric radiography for diagnosis and planning of orthodontic treatment remains uncertain.

The various chapters of this thesis cover various aspects on the validity and usefulness of lateral cephalometric radiographs. The aim was to validate its usefulness and accuracy. An extensive systematic review was performed to evaluate the existent literature regarding the real contribution of this radiographic technique to the diagnosis and treatment plan in orthodontics (*Chapter 1*). The reliability of some linear measurements commonly used in 2D lateral cephalometric analysis and its accuracy in comparison with the gold standard measurements on skulls were appraised (*Chapter 2*). Furthermore, the reproducibility of commonly used cephalometric landmarks identified by orthodontists and dentomaxillofacial radiologists and the impact of different landmark identifications in patient diagnosis were assessed (*Chapter 3*). We also evaluated the impact of S landmark identification of different localization effects on the angles SNA and SNB (*Chapter 4*). Finally, we evaluated the impact that lateral cephalometric radiography has in the orthodontic diagnosis and treatment plan (*Chapter 5*). The results from Chapter 1 revealed a very low number (n=17) of manuscripts related to this matter. When comparing measurements performed

on radiographs and skulls, we saw that measurements performed on radiographs systematically overestimated the ones executed on skulls (*Chapter 2*). We found low reproducibility on landmark identification and concluded that minor changes on landmark identification (1 to 2 mm) can in fact change the diagnosis of a patient (*Chapters 3 and 4*). Portuguese orthodontists seem to indicate that lateral cephalometric radiography is an important tool for the diagnosis and treatment plan in orthodontics. However, LCR seems to have a limited influence on treatment plan decisions (*Chapter 5*).

The various chapters in this thesis seem to contribute to the validation of LRC, and report on its accuracy, at the same time pointing towards its shortcomings when used in orthodontics.

RESUMO

Desde a introdução da telerradiografia de perfil em 1931 por Broadbent nos Estados Unidos e por Hofrath na Alemanha, esta radiografia, assim como a sua análise, tornou-se numa ferramenta muito utilizada no diagnóstico e planeamento em Ortodontia. Não obstante o facto de ser amplamente utilizada, o real valor da telerradiografia de perfil para o diagnóstico e planeamento do tratamento ortodôntico permanece desconhecido.

Os vários capítulos desta tese abrangem vários aspetos sobre a validade e a utilidade da radiografia cefalométrica da face em incidência lateral. O objetivo foi validar a sua precisão e utilidade na Ortodontia. Uma extensa revisão sistemática foi realizada para avaliar a literatura existente sobre a real contribuição desta técnica radiográfica para o diagnóstico e plano de tratamento ortodônticos (*Capítulo 1*). A confiabilidade de algumas medidas lineares comummente usadas em análise cefalométrica e a sua precisão em comparação com as medidas realizadas em crânios foi avaliada (*Capítulo 2*). Além disso, foram avaliadas a reprodutibilidade na identificação de pontos cefalométricos por ortodontistas e radiologistas dentomaxilofaciais e o consequente impacto de diferenças nas medidas lineares no diagnóstico do paciente (*Capítulo 3*). Avaliou-se ainda o impacto da identificação do ponto cefalométrico S e o seu efeito sobre os ângulos SNA e SNB (*Capítulo 4*). Finalmente, investigou-se o impacto que a radiografia cefalométrica lateral tem no diagnóstico e no planamento do tratamento em Ortodontia (*Capítulo 5*). Os resultados revelaram um número muito baixo (n = 17) de artigos relacionados com este tema. Quando comparadas determinadas medidas lineares realizadas em radiografias e crânios, verificou-se que as medidas feitas

nas radiografias foram geralmente maiores do que as efetuadas nos crânios. Contudo, não foi encontrada nenhuma diferença estatisticamente significativa (Capítulo 2). Detetou-se uma baixa reprodutibilidade na identificação de vários pontos cefalométricos e concluiu-se que pequenas diferenças na identificação de pontos cefalométricos (1 a 2 mm) podem alterar o diagnóstico do paciente (Capítulo 3 e 4). Os ortodontistas portugueses parecem achar que telerradiografia da face em incidência de perfil é uma ferramenta essencial para o diagnóstico e plano de tratamento em Ortodontia. No entanto, verificou-se uma baixa variabilidade no plano de tratamento após a avaliação desta radiografia. O plano de tratamento foi alterado em apenas 36% dos casos (Capítulo 5).

Os vários capítulos desta tese parecem contribuir para a validação da telerradiografia da face em incidência de perfil, mostrando também as suas limitações quando utilizada em Ortodontia.

REFERENCES

Abdullah RTH, Kuijpers MAR, Bergé SJ, Katsaros C. Steiner cephalometric analysis: predicted and actual treatment outcome compared. Orthod Craniofac Res. 2006;9:77-83.

Abu-Tayyem HM, Alshamsi AH, Hafez S, Eldin EM. Cephalometric norms for a sample of Emirates adults. Open Stomatol. 2011;1:75-83.

Adenwalla S, Kronman J, Attarzadeh F. Porion and condyle as cephalometric Landmarks - An error study. Am J Orthod Dentofac Orthop. 1988;94:411-415.

AlBarakati SF, Kula KS, Ghoneima AA. The reliability and reproducibility of cephalometric measurements: a comparison of conventional and digital methods. Dentomaxillofacial Radiol. 2012;41:11-17.

Ahlqvist J, Eliasson S, Welander U. The effect of projection errors on cephalometric length measurements. Eur J Ortho. 1986;8:141-148.

Arponen H, Elf H, Evälahti M, Waltimo-Sirén J. Reliability of cranial base measurements on lateral skull radiographs. Orthod Craniofac Res. 2008;11:201-210.

Atchison K, Luke L, White SC. Contribution of pretreatment radiographs to orthodontists' decision making. Oral Surg; Oral Med; Oral Pathol. 1991;71:238-245.

Atchison K, Luke L, White SC. An algorithm for ordering pretreatment orthodontic radiographs. Am J Orthod Dentofacial Orthop. 1992:102;29-44.

Athanasios, Athanasiou. Orthodontic Cephalometry. 1995:9-19, Mosby-Wolfe.

Bansal N, Singla J, Gera G, Gupta M, Kaur G. Reliability of natural head position in orthodontic diagnosis: A cephalometric study. Contemp Clin Dent. 2012;3(2):180-183.

Baumrind S, Frantz RC. The reliability of head film measurements. 1. Landmark identification. Am J Orthod. 1971;60:111-127.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986;327:307−310.

Broadbent BH. A new x-ray technique and its application to orthodontia. Angle Orthod. 1931;1:45-66.

Bourriau J, Bidange G, Foucart JM. Les erreurs de mesure en céphalométrie 2D. Orthod Fr. 2012;83:23-36.

Bruks A, Enberg K, Nordqvist I, Hansson AS, Jansson L, Svenson B. Radiographic examinations as an aid to orthodontic diagnosis and treatment planning. Swed Dent J. 1999;23:77-85.

Bruntz LQ, Palomo JM, Baden S, Hans MG. A comparison of scanned lateral cephalograms with corresponding original radiographs. Am J Orthod Dentofacial Orthop. 2006;130:340-348.

Celik E, Polat-Ozsoy O, Memikoglu TUT. Comparison of cephalometric measurements with digital versus conventional cephalometric analysis. European Journal of Orthodontics. 2009;31:241-246.

Chien PC, Parks ET, Eraso F, Hartsfield JK, Roberts WE, Ofner S. Comparison of reliability in anatomical landmark identification using two-dimensional digital cephalometrics and three-dimensional cone beam computed tomography in vivo. Dentomaxillofac Radiol. 2009;38(5):262-273.

Chen YJ, Chen SK, Huang HW, Yao CC, Chang HF. Reliability of landmark identification in cephalometric radiography acquired by a storage phosphor imaging system. Dentomaxillofacial Radiol. 2004;33:301-306.

Chen YJ, Chen SK, Chang HF, Chen KC. Comparison of Landmark Identification in Traditional Versus Computer-Aided Digital Cephalometry. The Angle Orthodontist. 2000;70(5):387-392.

Devereux L, Moles D, Cunningham SJ, McKnight M. How important are lateral cephalometric radiographs in orthodontic treatment planning? Am J Orthod Dentofacial Orthop. 2011;139:175-181.

De Abreu JL. O diagnóstico da classe esquelética. Comparação dos Resultados obtidos pelos métodos de: Steiner, Ricketts, Cervera e Coutand. Rev Port Estomatol Cir Maxilofac. 1982;23:89-101.

Diário da República, 1.ª série – Nº. 94 – 16 of May of 2005.

Downs WB. The role of cephalometrics in Orthodontic case analysis and diagnosis. Presented at the meeting of the American Association of Orthodontics, Louisville, April, 1951.

Duarte HEM, Vieck R, Siqueira DF, Angelieri F, Bommarito S, Dalben G, Sannorniya EK. Effect of image compression of digital lateral cephalograms on the reproducibility of cephalometric points. Dentomaxillofacial Radiol. 2009;38:393-400.

Durão AR, Pittayapat P, Rockenbach MI, Olszewski R, Ng S, Ferreira AP, Jacobs R. Validity of 2D lateral cephalometry in orthodontics: a systematic review. Prog Orthod. 2013;14(1):31.

European guidelines on radiation protection in dental radiology. Issue Nº 136.

European Society of Lingual Orthodontics Guidelines. 2010.

http://ec.europa.eu/energy/nuclear/radioprotection/publication/doc/136_en.pdf, accessed 18 March 2012. Pp 24, 25, 26, 45, 46, 49, 62 e 68.

Farkas LG, Tompson BD, Katic MJ, Forrest CR. Differences between direct (anthropometric) and indirect (cephalometric) measurements of the skull. Journal of Craniofacial Surgery. 2002;13(1):105-8.

Finlay L. Craniometry and cephalometry: A history prior to the advent of radiography. Angle orthod. 1980;50:312-321.

Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making. 1991;11:88-94.

Gijbels F, Sanderink G, Wyatt J, Van Dam J, Nowak B, Jacobs T. Radiation doses of collimated vs non-collimated cephalometric exposures. Dentomaxillofacial Radiol. 2003;32:128-133.

Gijbels F, Serhal CB, Willems G, Bosmans H, Sanderink G, Persoons M, Jacobs R. Diagnostic yield of conventional and digital cephalometric images: a human cadaver study. Dentomaxillofacial Radiol. 2001;30:101-105.

Graber TM, Vanarsdall RL. Orthodontics- current principles and techniques. 2nd edition. 1994:48-52. The Mosby Company.

Gravely JF, Benzies PM. The clinical significance of tracing error in cephalometry. Br J Orthod. 1974;1:95-101.

Gul E, Fida M. A Comparison of Cephalometric Analyses for Assessing Sagittal Jaw Relationship. J Coll Phys Surg P. 2008;18:679-683.

Han UK, Vig KW, Weintraub JA, Vig PS, Kowalski CJ. Consistency of orthodontic treatment decisions relative to diagnostic records. Am J Orthod Dentofacial Orthop. 1991;100:212-219.

Haynes S, Chau MN. Inter- and intra-observer identification of landmarks used in the Delaire analysis. Eur J Orthod. 1993;15:79-84.

Hoogeveen RC, Van der Stelt PF, Berkhout WER. Anatomically shaped cranial collimation (ACC) for lateral cephalometric radiography: a technical report. Dentomaxillofac Radiol. 2014;43(1). doi: 10.1259/dmfr.20130203.

Houston WJB, Maher RE, McElroy D, Sherriff M. Sources of error in measurements from cephalometric radiographs. Eur J Orthod. 1986;8:149-151.

Hussels W, Nanda RS. Analysis of factors affecting angle ANB. Am J Orthod. 1984;May;85(5):411-423.

ICRP, 2007. 2007 Recommendations of the International Commission on Radiological Protection (Users Edition). ICRP Publication 103 (Users Edition). Ann. ICRP 37 (2-4).

Isaacson KG, Thom AR. Guidelines for the use of radiographs in clinical orthodontics. British Orthodontic Society, London. 2001.

Ishikawa H, Nakamura S, Kim C, Iwasaki H, Satoh Y, Yoshida S. Individual growth in Class III malocclusions and its relationship to the chin cap effects. Am J Orthod Dentofacial Orthop. 1998:114(3):337-346.

Kaeppler G, Dietz K, Reinert S. Possibilities of dose reduction in lateral cephalometric radiographs and its effects on clinical diagnostics. Dentomaxillofacial Radiol. 2007;36:39-44.

Kamoen A, Dermaut L, Verbeeck R. The clinical significance of error measurement in the interpretation of treatment results. Eur J Orthod. 2001;23:569-578.

Kvam E, Krogstad O. Variability in tracings of lateral head plates for diagnostic orthodontic purposes. A methodology study. Acta Odontol Scand. 1969;27:359-369.

Lau PY, Cooke MS, Hagg U. Effect of training and experience on cephalometric measurement errors on surgical patients. Int J Adult Orthod Orthognath Surg. 1997;12:204-213.

Lee B, Shin G, Kang S, Shin B, Back I, Park H. Dose evaluation of selective collimation effect in cephalography by measurement and Monte Carlo simulation. Radiat Prot Dosimetry. 2012;148:58-64.

Lim KF, Foong KW. Phosphor-stimulated computed cephalometry: reliability of landmark identification. Br J Orth. 1997;24:301-308.

Mattila K, Haataja J. On the accuracy of determining certain reference points in cephalometric radiography. Odontologisk Tidskrift. 1968;76:249-259.

McClure SR, Sadowsky PL, Ferreira A, Jacobson A. Reliability of digital versus conventional cephalometric radiology: a comparative evaluation of landmark identification error. Seminar in Orthodontics. 2005;11:98-110.

McIntyre GT, Mossey PA. Size and shape measurement in contemporary cephalometrics. European Journal Orthodothontics. 2003;25(3):231-242.

Miethke RR. Zur Lokalisationsgenauigkeit kephalometrischer Referenzpunkte. Prakt Kieferorthop. 1989;3:107-122.

Miloro M, Borba AM, Ribeiro-Junior O, Naclério-Homem MG, Jungner M. Is there consistency in cephalometric landmark identification amongst oral and maxillofacial surgeons? Int J Oral Maxillofac Surg. 2013. http://dx.doi.org/10.1016/j.ijom.2013.08.007.

Moore D, McCabe G. Introduction to the practice of statistics. 4th edition. New York: Freeman. 2006.

Moyers RE. Handbook of Orthodontics. 4th edition. 1988;3:249-255.

Mupparapu M. Radiation protection guidelines for the practicing orthodontist. Am J Orthod Dentofacial Orthop. 2005;128:168-72.

National Radiological Protection Board. 2001: 23.

Nijkamp P, Habets L, Aartman I, Zentner A. The influence of cephalometrics on orthodontic treatment planning. Eur J Orthod. 2008;30:630-635.

Oktay H. A comparison of ANB, WITS, AF-BF, and APDI measurements. Am J Orthod Dentofacial Orthop. 1991;99:122-128.

Oz U, Orhan K, Abe N. Comparison of linear and angular measurements using two dimensional conventional methods and three-dimensional cone beam CT images reconstructed from a volumetric rendering program in vivo. Dentomaxillofacial Radiol. 2011;40:492-500.

Pae EK, McKenna GA, Sheehan TJ, Garcia R, Kuhlberg A, Nanda R. Role of lateral cephalograms in assessing severity and difficulty of orthodontic cases. Am J Orthod Dentofacial Orthop. 2001;120:254-262.

Proffit WR, Fields HW, Sarver DM. Contemporary Orthodontics. 4th edition. Chapter 6. Mosby Elsevier. 2006.

Raju S, Jayadeb CV. The digital era unravels – An Insight into 3-D digital applications in Orthodontics. Orthodontic Cyber journal. 2010. (http://orthocj.com/2010/10/the-digital-era-unravels-an-insight-into-3-d-digital-applications-in-orthodontics/.)

Richardson A. A comparison of traditional and computerized methods of cephalometric analysis. Eur J Orthod. 1981;3:15-20.

Sánchez JF, Filho O. Atlas de Cefalometría y Análisis Facial. 2009:41-44,47. Ripano Editora.

Sansare KP, Khanna V, Karjodkar F. Utility of thyroid collars in cephalometric radiography. Dentomaxillofacial Radiol. 2011;40:471-475.

Santoro M, Jarjoura K, Cangialosi TJ. Accuracy of digital and analogue cephalometric measurements assessed with the sandwich technique. Am J Orthod Dentofac Orthop. 2006;129:345-351.

Seki K, Okano T. Exposure reduction in cephalography with a digital photostimulable phosphor imaging system. Dentomaxillofacial Radiol. 1993;22:127-130.

Silling G, Rauch MA, Pentel L, Garfinkel L, Halberstadt G. The significance of cephalometrics in treatment planning. Angle Orthod. 1979;49:259-262.

Silveira HL, Silveira HE. Reproducibility of cephalometric measurements made by three radiology clinics. Angle Orthodontist. 2006;76:394-399.

Sinnot B, Ron E, Schneider AB. Exposing the thyroid to radiation: a review of its current extent, risks, and implications. Endocrine Reviews. 2010;31:756-773.

Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin. 1979;86:420-428.

Steiner, CC. Cephalometrics for you and me. Am J Orthod. 1953;39(10):729-755.

http://www.cleber.com.br/historia.html. Assessed in 3 February of 2012.

Tanikawa C, Yagi M, Takada K. Automated Cephalometry: System Performance Reliability Using Landmark-Dependent Criteria. The Angle Orthodontist. 2009;79(6):1037-1046.

Tng TT, Chan T, Hägg U, Cooke M. Validity of cephalometric landmarks. An experimental study. Eur J Orthod. 1994;16:110-120.

Tsuji Y, Araki K, Endo A, Okano T. Scatter radiation in cephalometric radiography: the effects of grid and collimation. Dentomaxillofacial Radiol. 2006;35:278-282.

Tyndall DA, Matteson SR, Soltmann RE, Hamilton TL, Proffit WR. Exposure reduction in cephalometric radiology: A comprehensive approach. Am J Orthod Dentofac Orthop. 1988;93:400-412.

Visser H, Rödig T, Hermann KP. Dose Reduction by Direct-Digital Cephalometric Radiography. Angle Orthod. 2001;71:159-163.

Wah PL, Cooke MS, Hägg U. Comparative cephalometric errors for orthodontic and surgical patients. Int J Adult Orthodon Orthognath Surg. 1995;10(2):119-126.

Wahl N. Orthodontics in 3 millennia. Chapter 7: Facial analysis before the advent of the cephalometer. Am J Orthod Dentofacial Orthop. 2006;129:293-298.

Walker G, Kowalski C. The Distribution of the ANB Angle in “Normal” Individuals. The Angle Orthodontist. 1971;41(4):332-335.

Whaites E. Essentials of Dental Radiography and Radiology. 4th edition. Chapter 15. Churchill Livingstone Elsevier. 2007.

White S, Pharoah M. Oral Radiology – Principles and Interpretation. Mosby Elsevier. 2009;6:191-195.

Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3:25-37.

APPENDICES

APPENDIX 1

Porto, 21st of June 2011

President of the Ethics Committee

of the Faculty of Dental Medicine, University of Porto

Prof. Doutor Fernando Morais Branco

Subject: Opinion on the research related to the PhD thesis "**The Influence of using 2D Cephalometry on orthodontic treatment outcome**" of the student Ana Paula Oliveira dos Reis Durão.

Within the realization of the project of Doctorate Program in Dental Medicine of the Dental Medicine Faculty of the University of Porto, we will prepare a research paper entitled "**The Influence of using 2D Cephalometry on orthodontic treatment outcome**".

Accordingly, we will develop appropriate systematic review of the literature, to define research methodologies and the treatment of data. I write this letter to the Council of the Ethics opinion on the protocol established for carrying out this research work, which is attached to this letter. We request the Ethics Board of the Faculty of Dental Medicine of the University of Porto, to authorize the protocol of this work.

Kind regards,

____

Ana Paula Reis Durão

APPENDIX 2

U.PORTO
FACULDADE DE MEDICINA DENTÁRIA
UNIVERSIDADE DO PORTO

Exma. Senhora
Mestre Ana Paula Oliveira dos Reis Durão
Estudante do Curso de Doutoramento em Medicina Dentária da
Faculdade de Medicina Dentária da
Universidade do Porto

000079

Assunto: Avaliação pela Comissão de Ética do projecto de investigação subordinado ao tema: *"The influence of using 2D cephalometry on orthodontic treatment outcome"*.

Informo V. Exa. que o projeto supra citado foi:

- **Aprovado** na reunião da Comissão de Ética do dia 25 de Janeiro de 2012.

Com os melhores cumprimentos,

O Presidente da Comissão de Ética

António Felino
(Professor Catedrático)

APPENDIX 3

Questionnaire

Name:______________________ Date:________________Case nº___________

1. How would you classify the skeletal problem of this patient?

Class I ☐

Class II ☐

Class III ☐

2. Classification of the occlusion (molar relationship):

Class I
Class II
Class III

R	L

3. Detection of abnormality

Yes ☐

No ☐

4. The treatment planning will be:

Orthopedic growth modification ☐

Orthognathic Surgery ☐

Dentoalveolar compensation ☐

5. Is there enough space for all teeth to erupt?

Yes ☐

No ☐

6. Would you extract teeth in this patient?

Yes ☐

No ☐

Extraction of the 2 premolars _________
Extraction of the 4 premolars _____________
Extraction of 1 premolar______
Others: ☐ which teeth_________________

7. Would you expand the upper arch?

Yes ☐

No ☐

8. Would you use anchorage?

In the maxilla:

Yes ☐

No ☐

In the mandible:

Yes ☐

No ☐

9. Do you expect any complications during the treatment process?

Yes ☐

No ☐

10. How long you expect the treatment will take?

…………month……….year

11. How long has it been since you qualified as an orthodontist?

---

12. Would you need additional information to make a decision?

Yes ☐

No ☐

If yes, which?

____