Original article

Reproducibility and speed of landmarking process in cephalometric analysis using two input devices: mouse-driven cursor *versus* pen

Alice Cutrera, DDS[1]
Ersilia Barbato, DMD, MSD[2]
Francesco Maiorana, PhD[3]
Daniela Giordano, MS, PhD[3]
Rosalia Leonardi, DMD, MSD[4]

[1] Department of Orthodontics, University of Catania, Italy
[2] Department of Oral and Maxillofacial Sciences, School of Dentistry, "Sapienza" University of Rome, Italy
[3] Department of Electrical, Electronic and Computer Engineering, University of Catania, Italy
[4] Department of Medical and Surgical Science, Section of Dentistry, University of Catania, Italy

Corresponding author:
Alice Cutrera
Department of Orthodontics, University of Catania
Via S. Sofia, 78
97019 Vittoria (Catania), Italy
E-mail: alice.cutrera@gmail.com

Summary

Aims. **To define if the new portable appliances, like smartphone, iPad, small laptop and tablet can be used in cephalometric tracing without dropping out the validity of any measurement.**
Methods. **We investigated and compared the reproducibility and the speed of landmarks identification process on lateral X-rays in two input devices: a mouse-driven cursor and a pen used as input means in mobile devices. One expert located 22 landmarks on 15 lateral X-rays in a repeated measure design two times, at time T1 and T2, after at least one month. The Intraclass Correlation coefficient was used to evaluate the reproducibility for each landmark tracing and the agreement between the value derived from both input devices. Also, the mean errors in measurements, the standard deviation and the Friedman Test significans (P < 0.05) between both input were statistically evaluated.**
Results. **All landmarks had a high agreement and the Friedman Test indicated statistically significant differences (P<0.05) for the identification of Na, Po, Pt, PNS, Ba, Pg, Gn, UIE, UIA, APOcc and PPOcc landmarks.**
Conclusions. **Even if the mouse input give higher agreement for landmark tracing the differences are really minimal and they can be ignored in private practice. We suggest the adequacy of pen input in clinical setting.**

Key words: computed aided cephalometric landmark tracing, landmark identification errors, lateral cephalometry, intraexaminer reliability.

Introduction

Cephalometric analysis is one of the major diagnostic tool in orthodontic, since 1931 when Broadbent introduced a standard diagnostic method to analyze the lateral head film (1). The constant progress in technologies has allowed the orthodontists to perform the cephalometric analysis on digital head film by means of computer software simplifying the landmark identification process, because of the image enhancement (2-5), and eliminating the majority of the errors that usually occurs during the hand tracing (drawing lines and use rulers and protractors for measuring distances and angles) (6, 7). Additionally, the cephalometric analysis performed on digital images lets the decreasing of time spent during the private practice and lets to achieve many different analysis at one time (8).
However, landmark identification on lateral X-ray still is the main source of error in cephalometric analysis regardless of the kind of input devices used. Abelson compared the reliability of cephalometric landmarking using both mouse and pen input focusing on micro parameters differences (9, 10-18). The Author claimed that the landmarking process and cephalometric tracing soft tissue and hard tissue outlines is faster and more accurate using a pen tablet than a mouse. Vogel (19) evaluated the hand occlusion with tablet sized direct pen input. But, only few Authors studied how new technological devices with their small monitors size (iPad, smartphone, tablets) could affect the efficiency of the accuracy of cephalometric tracing process in orthodontics. Goracci and Ferrari (20) analyzed the reproducibility of cephalometric measurements performed with software for tablet (where the clinician had to identify the landmarks with a stylus pen on the tablet screen) and for personal computers. Those Authors concluded that Pc-aided and tablet-assisted cephalometric tracing had good agreement, but they did not evaluate if the landmarks are more affected to errors during their detection in both devices. However, in that study, the speed efficiency for both digital appliances was not analyzed.

The aim of this study, therefore, is to compare two classical input devices, mouse and pen, by considering macro results and performance on a real clinical task such as a complete cephalometric landmarking analysis requiring great precision in pointing. Hence, we wanted to understand if the two different input devices affect:

1. accuracy of landmarks detection in lateral cephalometric X-ray;
2. to quantify time spent to detect the landmarks.

Materials and methods

All the cephalometric landmarks in both devices were recordered by the same investigator, just graduated from the Department of Orthodontic at the University of Catania, performed complete cephalometric analysis on 15 X-rays, randomly collected from the archives of the Orthodontics Department from 2010 till 2014. To simulate common clinical practice, we didn't deliberately focus on gender, type of occlusion and skeletal patterns. The subjects (5 males and 10 females) were aged between 10 and 15 (mean age 12.9 ± 1.7). Exclusion criteria were: (1) unerupted or missing incisors or (2) unerupted teeth overlying the incisor apices; (3) obvious malpositioning of the head in the cephalostat; (4) no unerupted or partially erupted teeth that would have hindered landmark identification; (5) patients with severe cranio-facial deformities; (6) posterior teeth not in maximum intercuspation and (7) facial asymmetries. Image quality was not used as an exclusion criteria since the selection was made from a pool of images representative of the quality levels found in daily clinical practice. A sample collection was approved by the Local Research Ethics Committee and informed consent was obtained from each patient's parents before the study. The cephalometric radiographs were scanned (Epson Expression 1680 Twain 2.10 Pro, Epson Company) at a resolution of 300 dpi with 256 gray levels to transform the analogue image into tiff digital format and stored blinded in a Personal Computer (Intel Pentium IV, 3.2 GH with 2 GB RAM, 300 GB Hard Disk, ASUSTeK Computer Incorporated). An extension of a software developed in other analysis (14) was used to record landmark coordinates and their Euclidean distances in millimeters. Twenty-two commonly used cephalometric landmarks were included in this study (10) (Tab. 1). The sequence of landmarks was enforced by the software interface. The software tool also logged the time elapsing between any two consecutive landmarks, the total time spent recording the landmarks and the whole set up of mouse movements.

The observer recorded the 22 landmarks on the images displayed on a Toshiba Portege M205 with a 1200 TFT polysilicon display with a native 1400 X 1050 resolution (Toshiba America Inc., New York, USA). The equipment was used in tablet mode with the pen parallel to the desk and in laptop mode when the mouse was used as landmarking device. The scanned images maximum size was 2700 X 2500 pixels and they were resized keeping the original proportions by a resampling procedure available in Adobe Photoshop CS4 software (Adobe, Inc. San Jose, California, USA) in order to obtain images with a vertical side of 900 pixels maximum in order to display all the

Table 1. The twenty-two commonly used cephalometric landmarks included in this study.

	Landmarks	
Name	Abbreviation	Definition
Nasion	Na	Junction of the frontal nasal bones at the naso-frontal suture
Sella	S	The midpoint od Sella Turcica
Orbitale	Or	The most inferior point of the infraorbital margin
Porion	Po	The most superior point of infraorbital margin
Basion	Ba	The lowest point on the anterior margin of the foramen magnum in the midline
Pterigo- Maxillary Fessure	Pt	The intersection of the posterior border of the foramen rotundum with the posterior wall of the pterigomaxillary fessure
Anterior Nasal Spine	ANS	Tip of the Anterior Nasal Spine
Point A	A	The deepest point in the concavity of the anterior maxilla between the anterior nasal spine and the alveolar crest
Posterior Nasal Spine	PNS	Tip of the Posterior Nasal Spine
Point B	B	The deepest in the concavity of the anterior mandible between the alveolar crest and Pogonion
Menton	Me	The most inferior point of the Chin
Gonion	Go	The most outward point on the angle of the mandible formed by the conjunction of the rams and the body of the mandible
Condilion	Co	The most upper-posterior point of the condile
Upper Incisal Edge	UIE	Tip of the crown of the upper central incisor
Lower Incisal Edge	LIE	Tip of the lower central incisor
Upper Incisal Apex	UIA	Tip if the apex of the upper central incisor
Lower Incisal Apex	LIA	Tip of the apex of the lower central incisor
Anterior Occlusal Plane	APOcc	Contact point between the first upper and lower premolars
Posterior Occlusal Plane	PPOcc	Contact point between the first upper and lower molars

X-ray without scrolling bars. The participant didn't have any previous experience with software tool and he was briefed about the purpose of the experiment. The landmarking process was performed directly on the monitor displayed X-ray with a mouse-controlled cursor. This cursor consisted of an empty arrow. A red dot on the selected pixel was used to signal the selected landmark on the screen. The landmark position could be corrected until the operator was satisfied.
A repeated measure design was chosen. The first time (T1), the observer performed the identification of 22 landmarks on all the 15 X-rays, without any time limit, on each of the two input devices. A rotation scheme was used to ensure that any X-ray head film was displayed to the examiner at the same frequency on each one of the two input devices, in order to avoid any learning effect. No more than 5 radiographs per day were analyzed to avoid the examiner fatigue (21). The landmarking process was repeated a second time (T2) for all X-rays and in both advices at least one month after the first session. Overall in the experiment 3,960 data points from the investigator, two input devices, 15 radiographs, 22 landmarks and 2 repeated measures were collected. The Euclidean distance between each cephalometric point located at T1 and T2 for both devices was used as the gold standard measurement to assess repeatability.
The Intraclass Correlation Coefficient was calculated for each Euclidean distance landmark detection to define the reliability of the point identification for each used input device. The ICC was computed using the more restrictive index of absolute agreement on the basis of the vector distance of the landmark position from the origin of the coordinates system (10). Overall the rate of agreement was considered low for an ICC $\leqq$ 0.80 and an ICC > 0.80 was indicative for good agreement.
After outliers removal the mean and standard deviations of the distances between the landmark coordinates at T1 and T2 for each landmark, and the time spent to locate each landmark were computed for each of the two experimental conditions.
For each input devices were calculated the mean errors and the standard deviations (in millimeters) of the Euclidean distance in landmark pointing between T1 and T2. It was also detected the Friedman's test significance between T1 and T2 for mouse cursor and pen input device where the significance level was 5% (P<0.05). Non-parametric Friedman's test was used to analyze the variance in a repeated measures design because the data sets had a non-normal distribution with non homogeneous variance. All statistical analysis were done with the software MATLAB version 7.10.0.499 (R2010a) (The Mathworks, Natik, MA, USA) and its Statistics toolbox.

Results

Table 2 shows the ICC values of landmarking process with the mouse-cursor devices for both times (T1 and T2) and Table 3 shows the ICC of landmarking process between T1 and T2 using the pen device. All the pointing processed had ICC $\geqq$ 0.95: this result was indicative of a very high agreement among the landmarking process for both devices. The mouse-cursor device had the highest values of correlation, ICC > 0.996, for Menton (X-axis), Orbitale (X-axis) and Basion (X-axis). The pen device had the highest values of correlation, ICC > 0.996 for Sella (X-axis), Porion (X-axis) and Pterigo-maxillary fissure (X-axis).
After outliers removal the mean and standard deviations of the distances between the landmark coordi-

Table 2. The Intra-Class Correlation for the mouse cursor devise evaluates the agreement among the landmarking process in T1 and T2. The agreement is high when ICC>0.80.

Mouse-Cursor Device	X-axis	Y-axis
Nasion	0,997535404229079	0,99752295715014
Sella	0,999401854224947	0,987099486029854
Orbitale	0,999666897017603	0,99005469272275
Porion	0,999569537007793	0,980427303078359
Basion	0,999156109755965	0,996381499987584
Pterigoid maxillary fessure	0,999697903832193	0,999454418826513
Anterior Nasal Spine	0,998665572003988	0,992379285485212
A Point	0,9971841105861	0,98115227811275
Posterion Nasal Spine	0,989655916459857	0,989261997063066
B Point	0,999502641163362	0,998076974380408
Pogonion	0,994347093257402	0,983123759705164
Gnation	0,987763576302059	0,988314973962179
Menton	0,998168211697634	0,99575280126277
Gonion	0,99611348877175	0,969128526136086
Condylion	0,992494024483712	0,995058661875066
Occlusal Point Upper Incisor	0,988628251924127	0,998653318730979
Occlusal Point Lower Incisor	0,952687899194511	0,986275922263936
Apex Upper Incisor	0,953401022916771	0,991830699536617
Apex Lower Incisor	0,99625081352735	0,983498206000574
Anterior Occlusal Point	0,993671053793834	0,955790067372581
Posterior Occusal Point	0,997407175137818	0,988157332078623

Table 3. The Intra-Class Correlation for the pen devise evaluates the agreement among the landmarking process in T1 and T2. The agreement is high when ICC > 0.80.

Pen device	X-axis	Y-axis
Nasion	0,997535404229079	0,99752295715014
Sella	0,999401854224947	0,987099486029854
Orbitale	0,999666897017603	0,99005469272275
Porion	0,999569537007793	0,980427303078359
Basion	0,999156109755965	0,996381499987584
Pterigoid maxillary fessure	0,999697903832193	0,999454418826513
Anterior Nasal Spine	0,998665572003988	0,992379285485212
A Point	0,9971841105861	0,98115227811275
Posterion Nasal Spine	0,989655916459857	0,989261997063066
B Point	0,999502641163362	0,998076974380408
Pogonion	0,994347093257402	0,983123759705164
Gnation	0,987763576302059	0,988314973962179
Menton	0,998168211697634	0,99575280126277
Gonion	0,99611348877175	0,969128526136086
Condylion	0,992494024483712	0,995058661875066
Occlusal Point Upper Incisor	0,988628251924127	0,998653318730979
Occlusal Point Lower Incisor	0,952687899194511	0,986275922263936
Apex Upper Incisor	0,953401022916771	0,991830699536617
Apex Lower Incisor	0,99625081352735	0,983498206000574
Anterior Occlusal Point	0,993671053793834	0,955790067372581
Posterior Occusal Point	0,997407175137818	0,988157332078623

Table 4.The mean error, the standard deviation and the Friedman's test significance between T1 and T2 in millimeters for the cephalometric landmarks detected with the two input devices: mouse and pen. NS = not significant. P < 0.05.

Landmark	Mouse	Pen	Significance
Na	0,37±0,21	0,57±0,31	*
S	0,30±0,10	0,38±0,19	NS
Or	0,91±0,58	0,91±0,48	NS
Po	0,68±0,21	0,72±0,35	NS
Ba	0,59±0,28	1,34±0,90	*
Pt	0,92±0,54	1,63±1,00	*
ANS	0,76±0,42	0,76±0,42	NS
A	0,44±0,18	0,65±0,26	*
PNS	0,79±0,42	0,99±0,57	NS
Ba	0,81±0,40	0,73±0,40	NS
PM	0,57±0,25	0,66±0,40	NS
Pg	0,39±0,21	0,54±0,31	*
Gn	0,54±0,24	0,80±0,42	NS
Me	0,37±0,21	0,72±0,26	*
Go	1,10±0,65	1,48±0,43	NS
Co	1,22±0,63	1,56±0,91	NS
UIE	0,27±0,12	0,29±0,18	NS
LIE	0,32±0,17	0,37±0,16	NS
UIA	0,98±0,45	0,95±0,34	NS
LIA	0,97±0,41	1,24±0,52	NS
APOcc	0,84±0,24	1,63±1,00	NS
PPOcc	0,97±0,37	1,58±0,63	*

nates at T1 and T2 for each landmark, and the time spent to locate each landmark, were computed for the two experimental conditions.

The mean errors and the standard deviations (in millimeters) of the Euclidean distance in landmark detection between T1 and T2 and the significance of the Friedman's test are reported in Table 4 for each input device. The Friedman test indicates that, except for Na, Ba, Pt, A, Pg, Me and PPocc there was no statistically significant difference, with P < 0.05, in landmark detection repeatability under the two experimental conditions. In all the cases with statistically significant difference the lowest error was obtained with the mouse as input device.

Table 5 shows the mean values, standard deviations and the significance of the Friedman's test of the time employed by the observer to locate each cephalometric landmark (averaged over all users and images

Table 5. The mean, standard deviation and Friedman's test significance for the time, in hundreds of seconds, required to digitize each landmark. Data have been averaged over all users, all X-rays at T1 and T2 for each of the two input devices: mouse and pen. NS = not significant. p < 0.05.

Landmark	Mouse	Pen	Significance
Na	439,29±253,51	718,94±652,18	*
S	252,25±92,88	298,74±91,28	NS
Or	164,72±48,81	222,61±26,37	NS
Po	166,94±73,64	214,42±65,45	*
Ba	257,50±97,64	286,44±95,92	NS
Pt	148,04±67,66	202,52±71,82	*
ANS	256,03±84,73	300,62±68,93	NS
A	265,88±124,29	225,54±139,94	NS
PNS	248,25±121,93	312,24±132,17	*
Ba	158,39±64,60	278,90±52,46	*
PM	338,63±115,25	359,37±209,16	NS
Pg	216,21±109,94	398,53±152,87	*
Gn	192,96±86,22	288,58±135,58	*
Me	204,41±107,19	223,60±117,58	NS
Go	346,25±106,06	465,50±220,38	NS
Co	203,58±79,22	264,84±34,85	NS
UIE	98,73±71,75	170,77±90,20	*
LIE	247,46±77,80	228,44±94,90	NS
UIA	106,58±48,84	227,19±108,98	*
LIA	271,24±104,59	277,52±98,89	NS
APOcc	146,45±63,95	250,65±33,13	*
PPOcc	368,68±130,36	489,57±156,23	*

during both experiments i.e. T1 and T2), for the two input devices. Statistically significant differences (P<0.05) were found for landmarks Na, Po, Pt, PNS, Ba, Pg, Gn, UIE, UIA, APOcc and PPOcc.

Discussion

The development of technology introduced the orthodontists to a new approach for diagnosis by the use of digital systems for tracing and analyzing cephalometric head films. Digital radiology has many advantages: reduced radiation dose, easier information access and image manipulation (22, 23), however the reproducibility of digital cephalometric analysis has to be high and has to reduce to the minimum the errors.
To avoid the inter examiner errors and to standardize the protocol for a comparative study, all the tracings were done by the same investigator.
Based on the Friedman test outcomes, there is no statistically significant difference on the repeatability of landmark detection for the two input device, except for Nasion, Basion, Pterigo-Maxillary Fessure, A Point, Pogonion, Menton and Posterio Occlusal Plane. For all those landmarks with statistically difference, the lowest error was obtained using the mouse. This can be due to the parallax error using the pen. In some cases locating some landmarks, such as Me, requires to move the hand down to avoid hand occlusions of the area where the landmark lies.
Our findings are not in line with Chen, Polat-Ozsoy and Celik's results (22-24), who analyzed the reproducibility of the landmarking process during conventional hand-tracing. These Authors claimed that Gonion, Porion, Orbitale, and Lower Incisor Apex are unreliable because of the superimposition of many anatomic structures that causes great variation. This difference could be because the digital images can be manipulated to improve the quality of the lateral head films so to reduce errors due to the superimposition of other anatomical structures or to the low contrast of the X-ray (4, 5, 7).
About the time spent to point the landmarks on lateral X-rays, even if a statistically significant difference exist for half of the landmarks, the difference of the sum of average time spent is 51 seconds when the mouse is used as input device and 67 seconds when the pen is used as input device. This difference is reduced if we avoid considering the first landmark that usually requires more time to be located due to the habit of the expert to obtain an overall impression of the X-ray. In this case the time required using the mouse is 47 seconds and the time required using the pen is 60 seconds. In any case a difference between 16 and 13 seconds for an entire analysis is not so relevant in clinical practice.
The difference in landmarking using the pen can be due to the parallax error experienced and reported by the expert in a post interview and in a questionnaire that requires an inspection to precisely locate the landmarks. However, the characteristic method of landmark detection by touching the screen needed to be tested in an independent study.
In conclusion this work is a quantitative study to evaluate the differences between two different input devices, mouse and pen, in accuracy and speed landmarking process. Hence we claim:

- Landmarks, in most of the cases, were pointed with more accuracy by mouse device respect the pen;

- Less time was needed to identify the landmarks with the mouse device.

The difference in accuracy with the advent of more sophisticated devices will probably decrease and it could be interesting to extend the analysis to new pen devices as well as to evaluate the performance when touch input is used.

References

1. Liu JK, Cheng KS. Accuracy of computerized automatic identification of cephalometric landmarks. Am J Orthod Dentofacial Orthop. 2000;118:535-540.
2. Major PW, Johnson DE, Hesse KL, Glover KE. Landmark identification error in posterior anterior cephalometrics. Angle Orthod. 1994;64:447-54.
3. Sicurezza E, Greco M, Giordano D, Maiorana F, Leonardi R. Accuracy of landmark identification on postero-anterior cephalograms. Prog Orthod. 2012;13;132-140.
4. Leonardi RM, Giordano D, Maiorana F, Greco M. Accuracy of cephalometric landmarks on monitor-displayed radiographs with and without image emboss enhancement. Eur J Orthod. 2010;32;242-247.
5. Giordano D, Maiorana F, Leonardi R. Effects of monitor size on accuracy and time needed to detect cephalometric radiographs landmarks. Displays. 2012;33;206-213.
6. Tsorovas G, Karsten AL. A comparison of hand-tracing and cephalometric analysis computer programs with and without advanced features-accuracy and time demands. Eur J Orthod. 2010;32:721-728.
7. Leonardi R, Giordano D, Maiorana F. An evaluation of cellular neural networks for the automatic identification of cephalometric landmarks on digital images. J. Biomed. Biotechnol. 2009;1-12. Article ID 717102.
8. AlBarakati SF, Kula KS, Ghoneima AA. The reliabiity and the reproducibility of cephalometric measurements: a comparison of conventional and digital methods. Dentomaxillofac Radiol. 2012;41:11-17.
9. Trpkova B, Major P, Prasad N, Nebbe B. Cephalometric landmarks identification and reproducibility: a meta analysis. Am J Orthod Dentofacial Orthop. 1997;112:165-70.
10. McClure SR, Sadowsky PL, Ferreira A, Jacobson A. Reliability of digital versus conventional cephalometric radiology: a comparative evaluation of landmark identification error. In Seminars in Orthodontics. 2005;11:98-110. WB Saunders.
11. Giordano D, Leonardi R. Web-trace and the learning of visual discrimination skills. 1st International Workshop on Pen-Based Learning Technologies. PLT. 2007.
12. Leonardi R, Giordano D, Caltabiano M. Interactive online program to improve cephalometric tracing skills. Am J Orthod Dentofacial Orthop. 2004;126;256-258.
13. Abelson M. Potpourri II. Digital Imaging Update. Am J Orthod Dentofacial Orthop. 1999;115;599-601.
14. ISO, Ergonomic requirements for office work with visual display terminals (VDTs), Part 9: Requirements for non-keyboard input devices. International Organization for Standardisation. 2000.
15. Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. J Exp Psychology. 1954;47;381-391.
16. MacKenzie IS, Sellen A, Buxton W. A comparison of input devices in elemental pointing and dragging tasks. Proc CHI, New York: ACM.1991;161-166.
17. Accot J, Zhai S. Beyond Fitts' law: models for trajectory-based HCI tasks. In Proceedings of the ACM SIGCHI Conference on Human factors in computing systems. ACM. 1997;295-302.
19. MacKenzie IS, Kauppinen T, Silfverberg M. Accuracy measures for evaluating computer pointing devices. In Proceedings of the SIGCHI conference on Human factors in computing. New York, ACM. Systems. 200;9-16.
20. Vogel D, Cudmore M, Casiez G, Balakrishnan R, Keliher L. Hand occlusion with tablet-sized direct pen input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM. 2009;557-566.
21. Goracci C, Ferrari M. Reproducibility of measurements in tablet-assisted, PC-aided, and manual cephalometric analysis. Angle Orthod. 2014;84;437-442.
21. Erkan M, Gurel HG, Nur M, Demirel B. Reliability of four different computerized cephalometric analysis programs. Eur J Orthod. 2012;34:318-321.
22. Chen YJ, Chen SK, Chan HF, Chen KC. Comparison of landmark identification in traditional versus computer-aided digital cephalometry. Angle Orthod. 2000;70:387-392.
23. Celia E, Polat-Ozsoy O, Ufuk Toygar Memikoglu T. Comparison of cephalometric measurements with digital versus conventional cephalometric analysis. Eur J Orthod. 2009;31:241-246.
24. Polat-Ozsoy O, Celik E, Ufuk Toygar Memikoglu T. Differences in cephalometric measurements: a comparison of digital versus hand-tracing methods. Eur J Orthod. 2009;31:254-259.