Assessment of the Reliability of Automatic Cephalometric Analysis Software

Niwat Anuwongnukroh[1*], Surachai Dechkunakorn[1], Suchaya Damrongsri[2], Chayawat Nilwarat[3], Natthasit Pudpong[3], Watcharapon Radomsutthisarn[3], Silinda Kangern[3]
[1] Department of Orthodontics, Faculty of Dentistry, Mahidol University, Bangkok Thailand
[2] Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Mahidol University, Bangkok Thailand
[3] Undergraduate Students, Faculty of Dentistry, Mahidol University, 6 Yothee Road, Rajthevee, Bangkok, Thailand
Email: *niwat.anu@mahidol.ac.th, surachai.dec@mahidol.ac.th, Suchaya.drs@mahidol.ac.th
Chaiyawat12@hotmail.com, Monsterfarm200x@hotmail.com, Kyo_inw@hotmail.com,
nn.numnim@hotmail.com

***Abstract*—The aim of this study was to evaluate and compare the reliability of a fully automatic cephalometric analysis software with manual cephalometric tracing. The lateral cephalograms of 108 orthodontic patients were selected. Eight linear (Pg-NB, Co-A, Co-Gn, U1-NA, L1-NB, Lower lip to E-plane, S-Go, and N-Me) and 9 angular (NS-Ba, SNA, SNB, NS-MP, FH-FO, U1-NA, L1-NB, L1-MP, and U1-L1) measurements were used in this study. The cephalometric analyses were performed by both manual method and automatic software. The differences between two methods were compared by paired t-test with $p<0.05$. Analysis of interexaminer calibration of the measurement revealed a high reliability. The result showed that there were statistically significant differences ($p<0.05$) in 13/17 parameters between the two methods, which consisted of 6 linear parameters (Pg-NB, Co-A, Co-Gn, U1-NA, L1-NB, and Lower lip to E-plane) and 7 angular parameters (SNA, NS-MP, FH-FO, U1-NA, L1-NB, L1-MP, and U1-L1). Only 4 parameters (NS-Ba, SNB, S-Go, and N-Me) did not show any significant differences. It is summarized that 76.47% (13/17 parameters) of cephalometric measurements performed automatically by the dental imaging software showed statistically significant differences when compared with the manual method. The automatic software could not reliably locate all cephalometric landmarks. Hence, the clinicians should not rely on the fully automatic analysis mode since the algorithm of the software still needs improvement for the higher accuracy in locating the cephalometric landmarks. Thus, to obtain accurate results, manual adjustments to the automatically located cephalometric landmarks are recommended.**

***Index Terms*— Automatic cephalometric analysis, Lateral Cephalogram, Orthodontics**

I. Introduction

Cephalometric analysis is a tool for orthodontic diagnosis and treatment planning. Manual tracing is time consuming method for measuring linear and angular parameters of cephalograms. However, this method is still considered as a gold standard in cephalometric analysis [1, 2]. Due to the rapid progress in science and technology, the field of dentistry is constantly evolving. One of the applications of digital technology in orthodontics is the use computer programs for analyzing lateral cephalograms (computer-aided cephalometric analysis), which is aimed as a time saving alternative to manual tracing.

This approach uses manual identification of landmarks, based either on an overlaid tracing of a radiograph followed by the transfer of the tracing to a digitizer linked to a computer, or a direct digitization of the lateral skull radiograph using a direct digitizer linked to a computer, and then locating landmarks on the monitor [3-6]. For an automatic cephalometric analysis, a scanned or digital cephalometric image is stored in the computer and loaded by a software. The software then automatically locates the landmarks and performs the measurements for cephalometric analysis. However, there are also be errors in the software algorithms leading to faulty identification of cephalometric landmarks. Therefore, the purpose of this study was to evaluate the reliability of the cephalometric analysis using the dental imaging software (Carestream Dental, Version 6.14) which is a fully automatic cephalometric analysis program.

II. Materials and Methods

The ethics approval for this study was obtained from the Faculty of Dentistry/Faculty of Pharmacy, Mahidol University, Institutional Review Board. One hundred and eight lateral cephalograms of patients undergoing orthodontic treatment were randomly selected from database of the Oral and Maxillofacial Radiology Clinic, Faculty of Dentistry, Mahidol University.

The inclusion criteria were: (1) the radiographs taken from the same x-ray unit (CS 9000C) with magnification ration of 1:1, (2) the radiograph size 30x30 cm, (3) high quality radiographs without any artifacts that could interfere with locating anatomical points, (4) lateral cephalogram with fully intact permanent central incisors and first permanent molars and no craniofacial deformities, such as cleft lip and cleft palate, etc.

For the manual cephalometric method, acetate papers were overlaid on lateral cephalograms and the outline of skull and facial structure were traced by one examiner.

Manuscript received July 27, 2017; revised December 25, 2017.

doi: 10.18178/ijmerr.7.1.61-65

The anatomical landmarks were used in this study :

Sella(S):	the center of the fossa of the sphenoid bone as seen in the lateral cephalometric radiograph,
Nasion (N):	the most anterior point of the frontonasal suture in the midsagittal plane,
Subspinale (A-point):	the deepest midline point in the curved bony outline from the base to the alvelolar process of the maxilla (the deepest point between the anterior nasal spine and prosthion),
Supramentale (B-point):	the most posterior point in the outer contour of the mandibular alveolar process in the median plane (the deepest point between the infradentale and Pogonion),
Basion (Ba):	the lowest point on the anterior margin of the foramen magnum in the median plane,
Condylion (Co):	the most posterosuperior point on the condyle of condyle,
Pogonion (Pg):	the most anterior point of the bony chin in the median plane,
Gnathion (Gn):	the intersection of the facial and the mandibular planes (the most downward and forward point on the profile curvature of the symphysis of the mandible),
Gonion (Go):	the intersection of the lines tangent to the posterior margin of the ascending ramus and the mandibular plane (the most posterior and inferior point on the angle of the mandible that is formed by the junction of the ramus and the body of the mandible),
Menton (Me):	the lowest point of the symphysis of the mandible in the midsagittal plane,
Porion (Po):	the most superiorly point of the external auditory meatus,
Orbitale (Or):	the lowest point on the inferior rim of the orbit,
Upper incisor (U1):	the long axis of the upper incisor,
Lower incisor (L1):	the long axis of the lower incisor.

The reference planes were used in this study :

Rickett's E-line (E-line):	The line joins soft tissue chin and the tip of the nose,
Mandibular plane (MP):	The plane joins Gonion (Gn) and Menton (Me),
Functional occlusal plane (FO):	A plane drawn between the cusp tips of the permanent molars and premolars,
Frankfort Horizontal plane (FH):	This is the plane joining Porion (Po) and Orbitale (Or).

The definitions of each variables of in this study were :

Pg-NB(mm):	the distance from Pogonion to N-B plane,
Co-A(mm):	the distance from Condylion to A-point,
Co-Gn(mm):	the distance from Condylion to Gnathion,
U1-NA(mm):	the distance from incisor edge of the upper incisor to N-A plane,
L1-NB(mm):	the distance from incisor edge of the lower incisor to N-B plane,
Lower lip to E-line(mm):	the distance from lower lip to E-line,
S-Go(mm):	the distance from Sella to Gonion,
N-Me(mm):	the distance from Nasion to Menton,
NS-Ba(dg):	the angle between NS plane and NBa plane,
SNA(dg):	the angle between NS plane and NA plane,
SNB(dg):	the angle between NS plane and NB plane,
NS-MP(dg):	the angle between NS plane and MP plane,
FH-FO(dg):	the angle between FH plane and FO plane,
U1-NA(dg):	the angle between the axis of upper incisor and NA plane,
L1-NB(dg):	the angle between the axis of lower incisor and NB plane,
L1-MP(dg):	the angle between the axis of lower incisor and MP plane,
U1-L1(dg):	the angle between the axis of upper incisor and the axis of lower incisor (Interincisal angle).

Then, the anatomical landmarks were defined and consensus-approval was given by 2 experienced orthodontists a,b in order to avoid professional bias. Reference lines were drawn for measuring all the linear and angular parameters. Eight linear parameters (Pg-NB, Co-A, Co-Gn, U1-NA, L1-NB, Lower lip to E-plane, S-Go, and N-Me [Fig. 1]) and 9 angular parameters (NS-Ba, SNA, SNB, NS-MP, FH-FO, U1-NA, L1-NB, L1-MP, and U1-L1(interincisal angle) [Fig. 2]) were used in this study. All parameters were manually measured with the same cephalometric protractor by 2 examiners e,f. Each measurement of each parameter from the 2 examiners were calculated for mean and recorded as manual measurement.

For the automatic imaging software, all lateral cephalograms were automatically analyzed by the software. Then, the analyzed values were printed out. Dependent paired t-test was used to compare the differences of the linear and angular parameters between the two methods.

To assess the reliability of the manual measurement, 10 randomly selected lateral cephalograms were repeatedly measured by the same examiner 2 weeks following the first measurements. Interclass and intraclass correlation coefficients were used for reliability analysis.

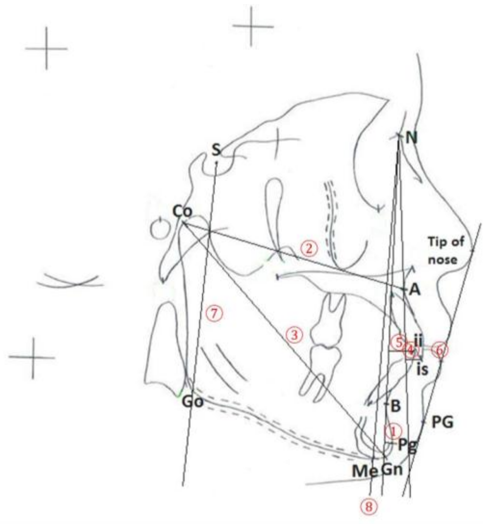

Figure 1. Eight linear parameters used in this study.(1) Pg-NB, (2) Co-A, (3) Co-Gn, (4) U1-NA, (5) L1-NB,(6) Lower lip to E-plane, (7) S-Go (8) N-Me

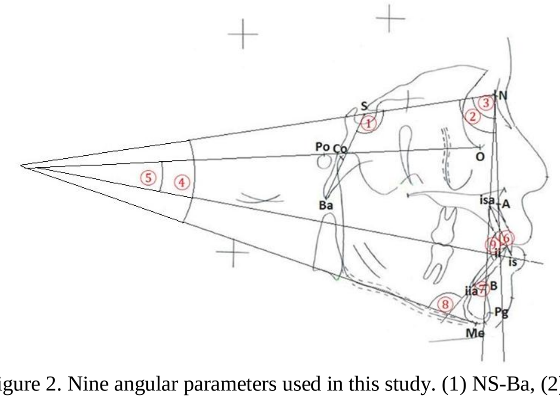

Figure 2. Nine angular parameters used in this study. (1) NS-Ba, (2) SNA, (3) SNB, (4) NS-MP, (5) FH-FO, (6) U1-NA, (7) L1-NB, (8) L1-MP, (9) U1-L1 (interincisal angle)

III. Results

For the interexaminer reliability, interclass correlation coefficient ranged from 0.98 to 1.00, and for the intraexaminer reliability, correlation coefficient ranged from 0.98 to 1.00. These results showed a high reliability of measurement in both and between the examiners.

The comparison of measurements between the manual and automatic tracings revealed that 13/17 parameters (76.47%), which included 6 linear parameters (Pg-NB, Co-A, Co-Gn, U1-NA, L1-NB, and Lower lip to E-plane) and 7 angular parameters (SNA, NS-MP, FH-FO, U1-NA, L1-NB, L1-MP, and U1-L1) showed statistically significant ($p<0.05$) differences. Only 4/17 parameters (23.53%) which included 2 linear parameters (S-Go and N-Me) and 2 angular parameters (NS-Ba, SNB) did not show any statistically significant differences (Table I).

TABLE I. Mean Differences of Each Parameter between the Two Methods

Parameter	Manual			Automatic Software			Mean difference (SD)	P value
	n	Mean	SD	n	Mean	SD		
Pg-NB (mm)	108	1.49	1.054	67	1.10	0.75	0.38955 (0.12028)	0.002 *
Co-A (mm)	108	81.76	5.57	108	85.33	5.50	-3.56713 (4.33714)	0.000 *
Co-Gn (mm)	108	109.63	7.84	108	108.38	8.27	1.25648 (3.99990)	0.001 *
U1-NA (mm)	108	6.61	2.59	106	2.70	1.72	3.91226 (2.72494)	0.000 *
L1-NB (mm)	108	7.44	2.48	106	6.35	2.94	1.08349 (2.00367)	0.000 *
L lip to E-plane(mm)	108	2.73	2.12	92	3.25	2.31	-0.52283 (1.55912)	0.002 *
NS-Ba (degree)	108	129.93	5.85	108	130.75	4.92	-0.81806 (5.33720)	0.114
SNA (degree)	108	83.67	4.14	108	86.71	3.65	-3.03704 (4.02464)	0.000 *
SNB (degree)	108	81.72	4.96	108	82.27	5.23	-0.54907 (3.52659)	0.109
NS-MP (degree)	108	30.40	6.39	108	32.90	6.53	-2.49769 (6.65439)	0.000 *
FH-FO (degree)	108	8.60	4.52	108	13.67	5.19	-5.06713 (6.74447)	0.000 *
U1-NA (degree)	108	28.69	7.93	102	19.77	5.98	8.91422 (7.52492)	0.000 *
L1-NB (degree)	108	33.62	6.39	61	29.74	9.85	3.88525 (7.24462)	0.000 *
L1-MP (degree)	108	99.56	10.03	108	92.40	10.92	7.17361 (7.65986)	0.000 *
U1-L1 angle(degree)	108	117.52	11.08	108	125.82	14.10	-8.29630 (10.84422)	0.000 *
S-Go (mm)	108	75.00	6.62	108	74.13	6.30	0.86296 (5.34576)	0.096
N-Me (mm)	108	111.54	7.43	108	112.39	8.54	-0.84815 (5.61488)	0.119

*Significant at $p<0.05$

IV. Discussion

The challenging problem in an automated cephalometric analysis is landmark detection, given that the calculations have already been automated with success. The first attempt at automated landmarking of cephalograms was made by Cohen in 1984[7], followed by more studies on this topic. Automatic identification of landmarks has been undertaken in different ways that involve computer vision and artificial intelligence techniques. All in all, these approaches can be classified into four broad categories, based on the techniques, or combination of techniques that have been employed. These categories are: (1) image filtering plus knowledge-based landmark search [8-10]; (2) model-based approaches [10-12]; (3) soft-computing approaches [13-15]and (4) hybrid approaches [16-18]. Advances and affordability in digital radiographic imaging have recently increased the demand for the medical profession to automate analysis and diagnostic tasks that were once performed manually. In this respect, several attempts to automate cephalometric analysis have been carried out.

For this study was used an automated cephalometric analysis software which is the application in the X-ray equipment for testing the efficiency of software.

The cephalometric radiographs used were randomly selected and the variables used in this study were commonly used for orthodontic diagnosis and treatment planning. In Table 1, it is noted that the sample sizes of some parameters in the automatic method were less than 108. The reason for the inequality of sample size arose from the inability of the software to interpret the negative value in the 4 parameters [Pg-NB, U1-NA (mm), L1-NB (mm), Lower lip to E-plane (mm)] and the misinterpretation of the linear pair of angles in 2 parameters (U1-NA and L1-NB). Therefore, those values that did not correspond to the manual measurement were excluded.

Errors in cephalometric analysis may come from systematic and random errors [19]. Systematic errors can occur when obtaining cephalograms if the geometry of the system varies and no compensation is made [20]. Random errors involve tracing, landmark identification, and measurement errors [19]. The greatest cause of random errors is difficulty in identifying landmarks. Influential factors for the identification of landmarks include image thickness and resolution, anatomical complexity, superimposition of the structures, the experience of the observers when locating a landmark, and the manual measurement errors [21]. To reduce random errors in this study, landmark identification in manual tracing were identified by the consensus of two experienced orthodontists, and the accuracy of measurement was tested by inter- and intra-examiner reliability.

In comparison, there were multiple differences in the values obtained from the automatic mode, when compared to those of manual analysis. Differences were observed in 76.47% of measurements. These differences were most likely a result of inaccurate landmark identification by the software. The inaccurately identified landmarks by the fully automatic software were the followings:

Bilateral structures: As a rule, when there is an overlapping of the right and left anatomical structure such as inferior border of mandible, condyle, porion, orbitale, and teeth, the observer should trace the average part of bilateral structures before locating the landmark on the tracing line. However, the automatic cephalometric software could not accurately trace the average part of the structure which, in turn, led to misidentification of the landmark. This error resulted in the statistically significant differences between the two methods in those measured parameters that related to bilateral structures landmarks which were Co-A, Co-Gn, U1-NA (mm), L1-NB (mm), FH-FO, U1-NA (degree), L1-NB (degree), L1-MP and interincisal angle.

Porion (Po) point: The automatic cephalometric software identified porion at the upper point of the ear rod (mechanical porion) which did not incongruent with the anatomic porion (the superior point of the external auditory meatus) from manual tracing. By using the different landmark, it, then, caused a statistically significant difference in the FH-FO angle between the two methods. It was also reported that porion and orbitale are commonly misidentified by automatic software [22].
Central incisors: There were significant differences between the two methods in U1-NA (degree), L1-NB (degree), L1-MP, and interincisal angle. The difference may be due to difficulties in identifying landmarks, as a result of the superimposition of incisal edges and root tips. Previous studies [3,22,23] have also reported that incisor apices had low reliability because they were superimposed by the surrounding structures leading to blurred image and tracing difficulties.

Molar teeth: The automatic software had low efficiency in locating the molar teeth. In most samples, the software traced the second molar instead of the first molar. This in turn led to unreliable functional occlusal plane. As a consequence, the FH-FO parameter showed a significant difference between the two methods.

Soft tissue profile: Although the automatic cephalometric software enabled to trace the outline of the soft tissue profile outline correctly, it had an error when locating the tip of nose, lower lip point, and soft tissue pogonion. These three points related to lower lip to E-plane parameter that showed statistically significant ($p<0.05$) differences when compared with the manual method.

During the evaluation of 72 cephalographs, Sommer et al [24] found the mean absolute angle differences between the hand-based and fully automatic methods exceeded 2 degrees (allowed tolerance limit of 2 degrees). In this study, significant mean differences of the angular parameters ranged from -2.49 (NS-MP) to 8.91 (U1-NA). The result obtained in this study is consistent with that of Sommer and coworkers.

Overall, the results in this study showed that the fully automatic software used in this study had difficulty identifying the landmarks that were involved in calculating the measurements. However, it should be noted that the manufacturer (Carestream Dental, USA) recommends checking and correcting all automatically located landmarks prior to completing the analysis.

V. Conclusion

The fully automatic mode of cephalometric analysis software is not as reliable as manual analysis. It should only be used to support a diagnosis and not as a diagnostic tool. The operator must check, review, and change all landmarks that are inaccurately identified by the software before completion of cephalometric analysis.

References

[1] Y. J. Chen, S. K. Chen, H. G. Chang, K. C. Chen, "Comparison of landmark identification in traditional versus computer-aided digital cephalometry," *Angle Orthod*, vol. 70, no. 387-392, 2000.

[2] S. Baumrind, D. M. Miller, "Computer-aided head film analysis: The University of California San Francisco method," *Am J Orthod*, vol. 78, pp. 41-65, 1980.

[3] O. Watcharinporn, S. Dechkunakorn, N. Anuwongnukroh, C. Sinthanayothin, P. Santiwong, W. Bholsithi, et al,. "Assessment of the accuracy and reliability of cephsmile v.2 in cephalograms and model measurements," *LNEE*. vol. 143, pp. 413-422, 2012.

[4] S. Baumrind, R. C. Frantz, "The reliability of head film measurements," *1. Landmark Identification. Am J of Orthod*, vol. 60, pp. 111-127, 1971.

[5] A. Richardson," A comparison of traditional and computerized methods of cephalometrics analysis," *Eur J Orthod*., vol. 3, pp. 15-20, 1981.

[6] P. J. Turner, S. Weerakone, "An evaluation of the reproducibility of landmark identification using scanned cephalometric images," *J Orthod*. vol. 28, pp. 221-229, 2001.

[7] A. M. Cohen, Ip HH, A. D. Linney, "A preliminary study of computer recognition and identification of skeletal landmarks as a new method of cephalometric analysis," *Br J Orthod*. vol. 11, pp. 143-154, 1984.

[8] S. Parthasarathy, S. T. Nugent, P. G. Gregson, D. F. Fay, “Automatic landmarking of cephalograms,” *Comput Biomed Res*., vol. 22, pp. 248-269, 1989.
[9] D. B. Forsyth, D. N. Davis, “Assessment of an automated cephalometric analysis system,” *Eur J Orthod*. 996; no.1 8, pp. 471-478.
[10] J. Ren, D. Lid, D. Feng, J. Shao, “A knowledge-base automatic cephalometric analysis method,” in *Proc. Conference Proceedings: 20th International Conference of the IEEE Engineering in Medicine and Biology Society*. *IEEE*;, vol. 20, no. 2, pp. 723-727, 1998.
[11] T. J. Hutton, S. Cunningham, and P. Hammond, “An evaluation of active shape models for the automatic identification of cephalometric landmarks,” *Eur J Orthod*. vol. 22, pp. 499-508, 2000.
[12] D. J. Rudolph, P. M. Sinclair, and J. M. Coggins, “Automatic computerized radiographic identification of cephalometric landmarkings,” *Am J. Orthod Dentofacial Orthop*. vol. 113, pp. 173-179, 1998.
[13] V. Ciesielski, A. Innes, J. Sabu, and J. Mamutil, “Genetic programming for landmark detection in cephalometric radiology images,” *Int J Knowl Based Intell Eng Syst*. vol. 7, pp. 164-171, 2003.
[14] S. Sanei, P. Sanei, M. Zahabsaniesi, “Cephalograms analysis applying template matching and fuzzy logic,” *Image Vision Comp*. vol. 18, pp. 39-48, 1999.
[15] I. El-Feghi, M. A. Sid-Ahmed, M. Ahmedi, “Automatic localization of craniofacial landmarks for assisted cephlometry,” *Pattern Recognit*. vol. 37, pp. 609-621. [33,34,37], 2004.
[16] V. Grau, M. Alcaniz, M. C. Juan, C. Monserrat, C. Knoil, “Automatic localization of cephalometric landmarks,” *J Biomed Inform*. vol. 34, pp. 146-156, 2001.
[17] J. Yang, X. Ling, Y. Lu, M. Wei, G. Ding, “Cephalometric image analysis and measurement for orthognathic surgery,” *Med Biol Eng Comput.*, vol. 39, pp. 279-284, 2001.
[18] D. Giordano, R. Leonard, F. Maiorana, G. Cristaldi, M. Distefano, “Automatic landmarking of cephalograms by CNNS,” *Lect Notes Artif Int.*, no. 3581, pp. 342-352, 2005.
[19] R. Leonardi, D. Giordano, F. Maiorana, C. Apampinato, “Automatic cephalometric analysis: A systematic review,” *Angle Orthod*. vol. 78. pp. 145-151, 2008.
[20] W. J. B. Houston, “The analysis of errors in orthodontic measurements,” *Am J Orthod*. vol. 83, pp. 382-390, 1983.
[21] B. Trpkova, P. Major, N. Prasad, B. Nebbe, “Cephalometric landmarks identification and reproducibility: A meta-analysis,” *Am J Ortho Dentofacial Orthop*. vol. 112, pp. 165-70, 1997.
[22] P. C. Chien, E. T. Parks, F. Eraso, J. K. Hartsfield, W. E. Roberts, and S. Ofner, “Comparison of reliability in anatomical landmark identification using two-dimensional digital cephalometrics and three-dimensional cone beam computed tomography in vivo,” *Dentomaxillofacial Radiology*, vol. 38, pp. 262-273, 2009.
[23] H. L. da Silveira, H. E. Silveira, “Reproducibility of cephalometric measurements made by three radiology clinics,” *Angle Orthod*. vol. 76. pp. 394-399, 2006.
[24] T. Sommer, R. Ciesielski, J. Erbersdobler, W. Orthuber, and H. Fischer-Brandies, “Precision of cephalometric via fully and semiautomatic evaluation of digital lateral cephalographs,” *Dentomaxillofacial Radiology*, vol. 38, pp. 401-406, 2009.

Niwat Anuwongnukroh, born on December 09, 1961 is an Associate Professor in Department of Orthodontics, Faculty of Dentistry, Mahidol University, Thailand. He received his M.S. degree from Saint Louis University, USA and American Board of Orthodontics & Thai Board of Orthodontics. He used to be the Head of Orthodontic Department, Assistant Dean for Research and Technology, Director of Residency Program in Orthodontics. His major research interest is Clinical Research and Orthodontic Materials.