

두부방사선사진의 계측오류에 관한 연구

나 광 천¹⁾ · 윤 영 주²⁾ · 김 광 원³⁾

본 연구는 교정 치료를 위한 진단과 치료 계획의 수립에 널리 이용되는 진단 수단의 하나인 두부 방사선 사진을 계측하는 연구에서 볼 수 있는 오류의 중요성을 인식하고, 반복 계측의 방법으로 체계적 오차(systematic error)를 찾아내고 임의적오차(random error)의 범위를 산출하는 과정을 통해 기준점의 식별에 보다 주의가 기울여져야 할 해부학적 기준점들을 찾아보고자 시행되었는데, 통법에 의해 촬영된 100장의 측모 두부 방사선 사진을 연구대상으로 하여 61개의 기준점과 130개의 계측항목을 선정하였고 반복계측을 위해 25개의 표본을 무작위로 선정, 계측오차를 분석하여 다음과 같은 결론을 얻었다.

1. 반복계측에 따른 계측치의 비교에서 130개 전체 계측항목의 18.4%인 24개 항목에서 유의성 있는 차이를 보였다.
2. 반복계측에서 기준점 식별에 차이를 보이는 빈도는 상악 전치의 치근단점(as), 상악골의 최후방벽(tu), 연조직 nasion과 전두정(ft) 그리고 기도의 ad3점에서 매우 높았다.
3. 5% 이하의 유의도를 보이는 수준까지 기준점 마킹을 수정한 후, 임의적오차의 범위는 0.67-1.71mm 또는 도를 보였다.
4. 가장 임의적오차를 보인 계측항목은 상, 하악 전치가 이루는 각(interincisal angle ; ILs-ILi)이었다.
5. 계측 오차의 문제는 주로 해부학적 정의에 대한 정확성의 부족과 불분명한 방사선사진상에 기인되었다.

이상의 연구 결과로 두부 방사선 사진의 계측에는 많은 오류의 가능성이 있음을 확인할 수 있었고 이에 따라 두부 방사선 계측에서 뿐 아니라 계측이 이루어지는 모든 연구에서는 항상 계측오류에 대한 분석이 함께 이루어져야 할 것이며, 조사자로서 뿐만 아니라 다른 연구자의 연구결과를 올바르게 인식하기 위한 목적으로도 오류에 대한 인식은 필수적인 것으로 생각된다.

(주요단어 : 두부방사선 사진, 계측오류)

I. 서 론

교정 치료를 위한 진단과 치료 계획의 수립에 널리 이용되는 진단적 수단의 하나로 측모 두부 방사선 사진을 들 수 있는데, 기준점을 설정하고 이를 계측, 분석하는 과정에서 많은 오류가 발생할 가능성이 있다. 이러한 오류에 관련된 많은 선학들의 지적과 보고들

을 볼 수 있는데^{8,9,13,20}, 매우 신중하게 여기에 관한 분석이 포함되어 있는 연구 보고를 볼 수 있는 한편, 오류 분석의 중요성에 대한 인식이 부족하거나, 오류의 분석이 이루어졌다 하더라도 결과의 해석에 그 오류가 어떻게 영향을 미치는가에 대한 인식이 부족한 것들을 볼 수 있다. 오류에 대한 평가나 인식이 명확히 이루어지지 않았다면, 어떤 연구에 의해 얻어진 결과는 매우 제한적인 가치를 지닐 수 밖에 없는데, 그 이유는 얻어진 분석치가 편향적인 계측에 따른 결과일지 모른다는 것, 그리고 어떤 분석치의 실체가 임의적 오차(random error)에 의해 불분명해질 수 있다는 것

1) 조선대학교 치과대학 교정학 교실, 치의학석사

2) 조선대학교 치과대학 교정학 교실, 조교수

3) 조선대학교 치과대학 교정학 교실, 부교수

에 대한 구분을 명확하게 해줄 수 없기 때문이다¹⁶⁾.

두부 방사선사진에서 있을 수 있는 오류에 대한 인식은 계측을 시행하는 연구자 뿐만 아니라 계측과 관련된 연구 결과를 파악하는 독자에게도 매우 중요한 문제가 되며, 이는 단지 두부방사선 사진의 계측에서 뿐만 아니라 계측이 이루어지는 다른 모든 연구에도 관련이 있다.

Houston¹⁶⁾는 교정학적 계측치의 오류 분석에 관한 연구에서 오류의 분석에는 계측으로부터 얻어진 값이 관심을 가지고 보고자 하는 실제의 값과 일치하는 정도를 파악하는 정확성의 개념과, 같은 대상을 반복하여 측정할 때 측정된 값의 유사성을 가리키는 재현성의 개념이 필요하다고 하였다.

방사선사진은 언제나 실제의 크기에 비해 확대가 일어나며 왜곡되어 인기될 가능성이 있기 때문에 정확성의 측면에서 계측치와 실측치간의 차이에 대한 평가가 필요하며 최소한 계측치를 임상적으로 적용시킬 때, 적어도 이를 인식하고 신중하게 사용되어야 한다.

재현성의 문제에 관계되는 계측의 오류에는 크게 체계적 오차(systematic error)와 임의적 오차(random error)의 두 가지로 나누어 볼 수 있는데, 두부 방사선 사진에서는 서로 다른 장비를 사용하여 촬영된 일련의 사진을 확대율의 조정 없이 비교할 때 발생할 수 있는 오류, 기준점의 정의나 개념을 달리하는 서로 다른 계측자에 의해 얻어진 계측치를 비교하는 경우에 볼 수 있는 오류, 동일한 계측자라 할지라도 시차를 두고 측정된 계측치간의 오류 그리고 연구 결과를 염두에 두고 측정된 계측치에서도 오류의 가능성이 있는데 이러한 오류는 체계적 오차(systemic error)의 범주에 넣을 수 있는 한편, 세팔로스테트(cephalostat)내에 두부를 고정시키는 과정에서의 문제를 비롯하여 얻어진 방사선 사진의 농도 및 선명도에 기인되는 문제, 정확성이 결여된 해부학적 기준점의 정의 때문에 이들을 정확히 인식해 내기 힘든 문제로 구분해 볼 수 있는 임의적 오차(random error) 등 대단히 많은 오류의 가능성을 내포하고 있다.

본 연구는 두부 방사선 계측에서 실측치와 방사선 사진의 계측치간의 차이를 다루는 정확성(validity)에 초점을 맞추기보다는 동일 표본에 대한 동일 계측자의 반복 계측에서 계측치의 재현성에 중점을 두고 이러한 과정에서 나타날 수 있는 오류의 정도를 파악하여 교정학적 계측에서 있을 수 있는 오류의 중요성을 인식하고 기준점 식별에 보다 주의가 기울여져야 할

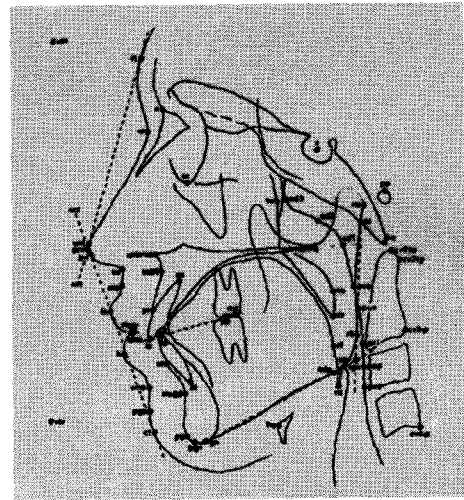


Fig. 1. Anatomical landmarks used in this study.

해부학적 구조를 찾고자 하는 목적으로 이루어졌으며 그 결과 다소의 지견을 얻었기에 보고하는 바이다.

II. 연구자료 및 연구방법

가. 연구자료

조선대학교 치과대학 학생(남자 51명, 여자 49명)을 연구 대상으로 하였으며, 남녀의 구별이나 두개 안면 구조의 형태는 고려되지 않았다.

나. 연구방법

(1) 두부 방사선 계측 사진의 촬영

거울상을 이용한 자연 상태의 두부 위치에서 통법에 의한 측모 두부 방사선사진을 Finland의 Planmeca사 PM2002 CC PROLINE cephalostat 촬영기를 이용하여 촬영하였는데 Focus-Median평면, Film-Median평면이 각각 153.5cm, 13.5cm를 지남으로써 8.8%의 확대율을 보였다.

(2) 해부학적 기준점 및 계측항목의 설정

두개, 안면, 연, 경조직에 대하여 61개의 기준점을 설정하였으며(Fig 1), 1) 골격 및 치아, 치조 관련 항목 2) 두부자세 관련 항목 3) 설골 관련 항목 4) 기도 관련 항목 5) 연조직 관련 항목으로 나누어 130개의 계측항목을 설정하였다(Table 1).

Tabel 1. List of variables used in this study.

A. SKELETAL & DENTOALVEOLAR MORPHOLOGY					
1. n-s	2. s-ba	3. n-ba	4. n-ar	5. ss-pm	6. sp-pm
7. pg-tgo	8. n-sp	9. n-ss	10. sp-gn	11. n-gn	12. s-pm
13. s-ar	14. ar-tgo	15. s-tgo	16. n-gm/s-tgo	17. pg-tgo/n-s	18. n-tgo
19. s-pgn	20. pgn-ar	21. pgn-rls	22. sp-is	23. pr-pr'	24. is-FH
25. is-NSL	26. is to n-p	27. ii to n-pg	28. pr-is	29. id-ii	30. id-id'
31. gn-ii	32. is-as/n-ss	33. is to n-ss	34. ii-ai/n-sm	35. ii to n-sm	36. pg to n-sm
37. ii to ss-pg	38. oj	39. ob	40. n-s-ba	41. n-s-ar	42. n-s-pgn
43. s-ar-tgo	44. s-nss	45. s-n-sm	46. ss-n-sm	47. po-n	48. n to FH
49. ss to FH ver through po		50. sm to FH ver through po		51. s-n-pg	52. n-pg/FH
53. n-ss/pg-ss	54. NL/F	55. ss-sm/ML	56. ss-sm/n-pg	57. ss-n-pg	58. NSL/NL
59. NSL/ML	60. NSL/RL	61. FH/ML	62. NL/ML	63. NL/RL	64. ML/RL
65. n-tgo/RL	66. n-tgo/ML	67. n-s-pm	68. pm-s-ba	69. OLs/NL	70. OLi/ML
71. OLx/ML	72. ILs/NL	73. ILs/ML	74. ILi/FH	75. ILs/ILi	76. pr-n-ss
77. CL/ML	78. cv2ap-cv4ip				
B. POSTURE					
79. NSL/VER	80. NL/VER	81. ML/VER	82. RL/VER	83. OPT/HOR	84. CVT/HOR
85. NSL/OPT	86. NL/OPT	87. ML/OPT	88. RL/OPT	89. NSL/CVT	90. NL/CVT
91. ML/CVT	92. RL/CVT	93. OPT/CVT	94. NSL/HOR	95. FH/HOR	96. FH/VER
97. HF/OPT	98. HF/CVT				
C. HYOID POSITION					
99. hy to NSL	100. hy to NL	101. hy to ML	102. hy to CVT	103. hy to pm	
D. AIRWAY DIMENSION					
104. tu-ad3	105. pm-ad2	106. m-ad1	107. ve-pve	108. uv-puv	109. rl-prl
110. va-pva	111. ep-pep	112. pm-uv			
E. SOFT TISSUE MORPHOLOGY					
113. n _s -sn(mm)		114. n _s -lnt(mm)		115. lnt to n-ss(mm)	
116. pm to n-ss(mm)		117. n-n _s -unt(°)		118. n-n _s -pm(°)	
119. sto to NL(mm)		120. n-n _s -ss _s		121. sn to lnt-ls(mm)	
122. ls to NCL(mm)		123. sto to ML(mm)		124. n-n _s -sm _s (°)	
125. sm _s to li-pg _s (mm)		126. li to NCL(mm)		127. ss _s -n _s -n _s (°)	
128. sto to OL _s (mm)		129. s-n _s -pg _s (°)		130. NFL/NCL(°)	

(3) 자료의 입력

100개의 표본에 대하여 선정된 61개의 기준점을 필름 상에 연심연필로 직접 마킹한 다음, 디지털타이저를 이용, 정해진 순서에 따라 입력하였으며(Fig. 2), 입력이 완료된 후, 플로터에 의한 출력을 통해 정확성을 확인하는 절차를 거쳤다.

(4) 계측오차의 분석을 위한 표본의 추출 및 자료의 재입력

100개의 전체 표본으로부터 무작위로 25개의 표본

을 추출하였으며 이를 계측 오차의 분석을 위한 표본으로 이용하였다. 1차적으로 마킹된 기준점을 모두 지운상태에서 2차적 마킹을 한 후 자료입력의 같은 과정을 반복하였다.

(5) 계측치의 산물

130개의 계측항목에 대해 1,2차 마킹에 의한 계측치를 각각 산출하였으며 SAS 통계프로그램을 이용, t-test를 시행하여 이들간의 차이에 대한 유의성을 검증하였다.

Table 2. Variables shown significant difference between 1st and 2nd marking

Variables	N	1st marking		2nd marking		P-value
		Mean	S.D.	Mean	S.D.	
ss-pm	25	51.52	2.52	51.92	2.48	0.0317
n-ss	25	68.43	3.75	67.55	3.85	0.024
s-pm	25	53.76	3.11	53.43	2.98	0.0424
pgn-ar	25	129.90	5.08	121.45	4.87	0.0154
ii to n-pg	25	5.91	2.51	6.10	2.62	0.0236
lls/n-ss	25	24.31	6.09	21.87	5.46	0.0001
ii to ss-pg	25	3.90	1.62	4.22	1.96	0.0240
n-s-ar	25	126.24	3.79	125.64	3.85	0.0197
s-ar-tgo	25	147.89	5.17	148.75	5.06	0.0060
n-s-pm	25	75.18	2.67	75.89	3.00	0.0063
pm-s-ba	25	55.88	3.50	55.22	3.41	0.0158
lls/NL	25	116.45	5.41	113.87	4.69	0.0001
lls-/ML	25	96.31	5.20	95.82	5.01	0.0486
lli/FH	25	59.93	5.85	60.58	5.34	0.0435
lls/ILi	25	125.76	5.74	128.94	3.24	0.0001
ML/CVT	25	74.99	6.45	75.27	6.22	0.0350
tu-ad3	25	8.89	2.08	7.57	1.69	0.0122
pm-ad1	25	28.39	3.66	27.44	3.55	0.0004
pm to n-ss	25	31.01	2.23	31.44	2.07	0.0081
s-ns-unt	25	119.49	3.86	118.50	3.48	0.0366
s-ns-sss	25	98.01	2.42	97.15	2.27	0.0032
s-ns-sms	25	90.48	2.63	89.83	2.64	0.0177
sss-ns-sms	25	7.54	1.58	7.33	1.61	0.0265
s-ns-pgs	25	90.35	2.88	89.83	2.98	0.0255

Table 3. Variables shown significant difference after correction of marking

Variables	N	1st marking		2nd marking		P-value	S
		Mean	S.D.	Mean	S.D.		
s-ar-tgo	24	147.93	5.2	148.75	5.06	0.0042	0.71
n-s-pm	24	75.56	2.76	76.04	3.15	0.0231	0.67
lls/ML	24	114.65	4.68	113.87	4.69	0.0390	1.68
lls/MLi	25	127.5	5.63	128.94	5.24	0.0088	1.71
tu-ad3	24	8.49	1.69	7.57	1.69	0.0091	1.20
s-ns-sss	25	97.70	2.57	97.15	2.27	0.0336	0.86

(6) 기준점 마킹의 수정

1,2차 마킹에 의한 계측치의 유의성있는 차이를 보이는 항목과 관련된 기준점을 중심으로 기준점의 마킹을 수정하였는데 마킹의 수정은 전체 표본 모두에

서 이루어졌다.

(7) 임의적 오차의 산정

25개의 무작위표본에서 전체 계측항목에 대해 유

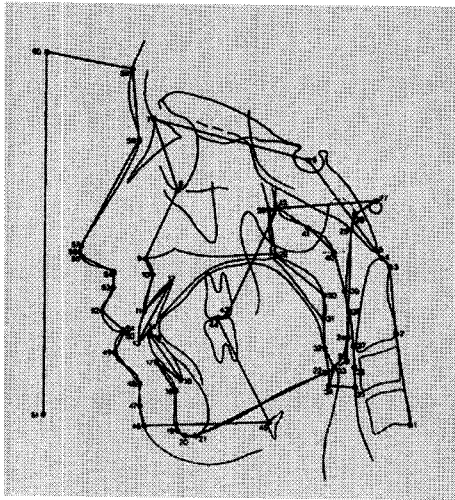


Fig. 2. Digitizing sequence.

의차를 보이는 항목의 비율이 5% 유의수준 이하로 얻어질 때까지 전체표본에 대한 수정이 반복되었으며 최종적으로

$$Se = \sqrt{\frac{Sd^2}{2}} \quad (se = \text{standard error, } sd = \text{standard deviation})$$

의 공식에 따라 유의성 있는 차이를 보이는 항목에 대한 임의적 오차의 범위를 산정하였다.

III. 연구성적

(1) 기준점 마킹의 수정전 계측치

기준점 마킹의 수정전 1,2차 반복 마킹에 의한 계측치의 비교에서 전체 130개 계측항목의 18.4%의 24항목에서 통계적으로 유의성 있는 차이를 보였으며 이들 계측 항목과 관련된 기준점의 마킹을 방사선사진을 중첩시켜 비교한 결과, 차이를 보였던 빈도는 상악 전치의 치근단점(as), 상악골의 최후방벽(tu), 연조직 nasion(N')와 전두정(ft) 그리고 기도의 ad3에서 매우 높았다. Table 2에 반복계측에서 유의성 있는 차이를 보였던 항목의 목록을 나타내었다.

(2) 마킹의 수정후 계측치

전체 계측항목에 대하여 5% 이내에 유의성을 보이는 수준까지 기준점 마킹의 반복한 결과, 수정후 전체 계측 항목의 4.6%인 6개 항목에서 반복계측 군의 계

측치간에 유의성 있는 차이를 보였으며 임의적 계측 오차는 0.67-0.71mm 또는 도의 범위를 보였다. 가장 큰 계측 오차를 보인 항목은 상, 하악 전치가 이루는 각(interincisal angle ; ILs-ILi)이었다. Table 3에 이들 항목에 대한 목록과 계측오차의 정도를 표시하였다.

IV. 총괄 및 고안

두부계측학은 인간의 두부를 기하학적 도식으로 요약하는 방사선적 기술이라고 할 수 있는데¹⁹⁾ 교정학 영역에서는 두개안면구조의 형태학적인 관찰과 성장이나 치료에 의한 변화를 평가할 수 있으며, 비정상적 상황을 진단함과 동시에 미래의 결과를 예측하여 치료계획을 세우거나 치료의 결과를 평가하는데 이용되고 있다. 이러한 과정은 대체로 특정 기준점을 선정하여 이들이 이루는 각과 선의 변수를 계측하게 되는데 잠재적으로 많은 오류에 의한 영향을 받고 있으며 다양한 양상으로 최종적인 결과의 평가에 영향을 미치게 된다.

그러나 이러한 오류의 원인은 쉽게 경계지를 수 없을 정도로 많은 요소가 관련되며 상호 관련되어 있다. 계측오차는 일반적으로 체계적 오차(systematic error)와 임의적 오차(random error)로 구분될 수 있는데, 본 연구는 주로 기준점 식별의 오류에 초점을 맞추어 이에 따라 나타나는 계측오차를 분석함으로써 계측에 있어서 그 중요성을 인식하고 계측치의 신뢰도를 높이기 위한 과정을 보여주는 목적으로 이루어졌다.

방사선사진은 촬영시 조사선이 투과되는 물체의 모든 부분에 평행하게 진행되지 않기 때문에 필연적으로 상의 확대가 발생되며, 또한 서로 다른 평면간의 서로 다른 비율의 확대에 기인되어 상의 왜곡이 발생된다. 두부의 계측 분석에 사용되는 대부분의 기준점들은 정중시상면에 위치하나 일부 구조물의 기준점은 정중시상면과 다른 깊이에 있기 때문에 왜곡의 현상이 발생됨에 따라 선과 각의 계측치는 다양하게 영향받을 수 있다. Mitgaard등¹⁸⁾은 25쌍의 방사선 사진을 통해 계측 거리간의 부조화를 보고하였으며, Broadbent¹⁰⁾는 정중시상면과 그것에서 떨어진 면 모두에서 기준점의 투사 왜곡을 수정하기 위해 Bolton Orientator를 사용하였고, Adams⁴⁾는 계측 기준점의 투사 왜곡의 양에 대한 표준화를 시도한 바 있다. 본 연구에서 방사선사진상의 계측거리는 실측치에 비해 8.8%의 확대율을 보였는데 이는 초점-물체간 거리와

물체-필름간 거리에 기인되는 현상으로 모든 방사선 사진상에서 공통적으로 볼 수 있는 현상이나, 이를 최소로 하기 위해서는 긴 초점-물체간 거리와 짧은 물체-필름간 거리를 지닌 기계적 배열이 제안되고 있다¹²⁾.

Houston¹⁶⁾은 방사선사진의 계측오차를 최소화하기 위해서는 방사선사진의 촬영에 관련된 요인, 기준점 식별과 관련된 요인, 측정 방법의 설계와 관련된 요인 등에 대한 고려가 되어야 한다고 하였다. 방사선사진의 촬영과 관련된 요인으로는 장비와 관련된 모든 부분이 확실히 고정되어야 한다고 하였으며, 방사선사진상에서 선과 각의 왜곡은 부적절한 환자 위치에 결정된다는 Ahlqvist^{5,6,7)}의 보고나 세팔로스테트(cephalostat)내에서 잘못된 두부의 위치는 방사선 사진상의 비대칭적 왜곡을 발생시킨다는 보고⁹⁾ 등에 비추어 볼 때 이어로드(earrod)내에 존재하는 금속링은 좌우상이 일치야 이루어져야 하며 확대율 판정을 위해 정중시상면에 눈금자를 위치시키는 것 등이 제안되고 있다. 또한 방사선량의 조절이나 필름 및 증감지의 선택과 같은 방사선사진의 질을 향상시킬 수 있는 모든 노력이 기울여져야 함을 제시하고 있다. 본 연구에서도 이러한 방사선 촬영과 관련된 모든 요인들을 고려하여 촬영이 이루어지도록 노력하였다.

기준점 식별과 관련된 요인으로는 적절한 조명하에서 판독이 이루어져야 한다는 것과 세부적인 소견이 불분명해지는 저질의 투사용지 사용과 같은 경우를 들 수 있다. 그러나 기준점 식별을 증진시키는데 무엇보다도 가장 중요한 것은 이에 관한 풍부한 경험이 필요하다는 것이며, 특별히 한사람 이상의 조사자가 관련된 경우라면 조사자간의 기준점식별에 대한 인식을 같이 하는 것이 대단히 중요하다¹⁵⁾. 본 연구에서는 동일인에 의한 반복계측이 이루어지도록 계측 오차의 분석 방법을 설계하였기에 조사자간의 기준점 식별에 대한 같은 인식의 문제는 없었으나 이는 매우 중요한 과정으로 사료된다.

측정방법의 설계와 관련된 요인으로는 우선 비정상적으로 벗어난 수치에 대한 평가가 필요하다. 일반적으로 어떤 계측치가 다른 보고 등을 통하여 알려진 표준편차의 범위로부터 3 S.D. 이상 벗어난 경우는 기준점 식별이 잘못되었을 가능성이 높다고 알려져 있다¹⁾. 따라서 계측치에 대한 일차적인 검증이 필요할 것으로 사료되며, 극단적인 계측치를 임시로 제거한 상태에서 표준편차를 산출하여 이를 기준으로 극단적인 계측치를 평가해보는 것이 보다 쉽게 잘못된

계측치를 찾아낼 수 있는 방법으로 제안되기도 하였다¹⁴⁾.

체계적 오차와 임의적 오차의 평가는 반복 측정함으로써 얻어지는데 하나의 방법으로 동일 개체로부터 반복 촬영된 방사선 사진을 이용하는 방법을 들 수 있다. 이 방법은 불필요한 방사선에의 노출과 관련된 윤리적인 문제로 일반적인 방법이라고 할 수는 없으나, Houston 등¹⁷⁾에 의해 25명의 표본을 대상으로 시행된 연구를 살펴보면 큰 오류를 보이는 경우란 주로 계측점 식별과 관련된 문제에 기인되었음을 보고하고 있다. 그들은 이 연구에서 환자를 세팔로스테트내에 위치시킬 때 충분한 주의가 기울여지지 않은 경우에 기인된 오류도 있을 수 있으나, 투사의 반복이나 계측의 반복을 통해 기준점 식별에 특별한 주의가 기울여져야 함을 강조하고 있다.

방사선 사진의 반복 촬영을 대체할 수 있는 방법으로는 한 장의 필름을 반복 계측함으로써 오류의 평가할 수 있는데, 이상적으로는 모든 표본에 대한 반복 계측이 이루어져야 하나 표본의 숫자가 큰 경우에는 전체의 표본으로부터 일부를 무작위로 추출하여 오류에 대한 분석을 할 수 있다²⁾. 특별히 질이 좋은 방사선사진을 선택하거나, 추출된 표본에 대해서만 특별한 주의를 기울여 기준점의 식별 및 마킹이 이루어져서는 안되며 본래의 연구에 행해지는 것과 같은 정도의 주의가 이루어져야 한다. 만약 어떤 계측이 장시간에 걸쳐 이루어진다면 무작위로 추출된 표본과 본래의 표본에 대한 특정을 통해 시간경과에 따른 편향적인 요소 역시 검출되어야 한다.

본 연구는 필름을 반복 마킹하는 방법을 이용하여 계측함으로써 체계적 오차를 찾아내고 임의적오차를 산출해내는 과정으로 이루어졌다.

1) 체계적 오차의 검출

체계적 오차의 검출은 반복 측정된 표본간의 차이에 대한 t-검증을 시행함으로써 이루어질 수 있다. 여기에는 다음의 몇가지 요소가 관련되는데, 우선 충분한 표본의 숫자가 필요하다. 그렇지 않다면 오직 매우 큰 체계적 오차만이 검출될 수 있기 때문이다. t-값을 산출하는 공식 ($t = \frac{\text{mean difference}}{s} \sqrt{n}$)에서 볼 수 있듯이 표본의 숫자가 너무 작아지면 주어진 평균의 차이와 표준편차에 대하여 t-값이 작아지기 때문에 매우 중요한 체계적 오차까지도 통계학적으로 유

의한 것으로 판명되지 않기 때문이다. 요구되는 표본의 숫자는 표본간 계측치의 차이에 있어서 그 표준편차의 크기에 따라 달라지나, 일반적으로 무작위로 선정된 25개 이상의 표본이 이용되어야 한다고 한다³⁾.

큰 임의적 오차에 기인된 표본간 계측치의 차이에 있어서 그 표준편차가 큰 경우에는 체계적 오차가 불분명해지는데, 이는 표준오차가 커짐에 따라 역시 t-값이 작아지기 때문이다. 유의성의 수준도 문제가 될 수 있는데 유의성의 수준을 적절히 설정함으로써 실제로 차이를 보이는 문제가 t-검증에서 기각되는 일이 없도록 해야 할 것이다. 일반적으로 10% 수준에서의 유의성을 보이는 차이는 의심스러운 것으로 평가되고 있다¹⁶⁾.

본 연구에서는 이러한 조건들을 고려하여 전체로부터 25개의 무작위표본을 추출하여 5%의 유의수준에서 t-검증을 시행한 바 전체 130개 측정항목의 18.4%에 해당하는 24개의 측정항목에서 유의성 있는 차이를 관찰할 수 있었는데, 이는 체계적 오차에 의한 부분과 임의적오차의 부분이 모두 관련되어 있는 결과라 할 수 있으며, 이와같은 반복계측을 통해 볼 때 계측치간에는 많은 차이를 보이고 있기 때문에 오류에 대한 검증이나 감소를 위한 노력없이 이를 그대로 어떤 결과치로 판단하여 이용한다는 것은 잘못된 결론에 이를 수 있을 것으로 생각된다. 차이를 보이는 계측항목의 기준점 마킹을 비교하였을 때 마킹의 실수에 의한 빗나감도 관찰되었지만, 주로 기준점의 해부학적 정의가 불명확하거나 방사선사진의 식별이 어려운 경우가 많았으며 상악전치의 치근단점(as), 상악골의 최후방벽(tu), 연조직 nasion(n'), 전두정(ft) 그리고 기도상의 기준점인 ad3와 같은 기준점에서 많은 차이를 보였다. 이러한 결과는 이와같은 기준점들에서는 보다 신중한 마킹이 이루어져야 함을 시사하는 것으로 생각되었다.

유의성있는 차이가 관찰된 계측항목에 해당하는 기준점을 중심으로 전체표본에 대한 마킹의 수정을 시행하였는데, 이는 임의적오차의 정도를 산출하기 위한 과정이라 할 수 있으며, 전체계측항목의 5%이하에서 유의성있는 차이가 관찰될 때까지 마킹의 수정을 계속하였다.

2) 임의적오차의 산출

임의적오차는 표본의 자연적인 다양성에 더하여 비교군간의 실제적 차이를 불분명하게 하는 결과를

초래하기 때문에 역시 중요한 문제로 고려하지 않을 수 없다. 또한 임의적오차는 비교군간의 상관관계에 영향을 미쳐 상관계수를 통계학적으로 산출할 때 그 상관도를 감소시키는 결과를 초래한다.

반복 계측군간의 차이에 대한 표준편차가 임의적 오차의 크기가 되는데 반복 측정차에 대한 분산의 정도는 그 표준편차의 제곱으로, 이는 단일 측정치의 2배가 되므로 이 값의 절반이 단일측정치에 대한 분산이라 할 수 있다. 따라서 두 계측군 모두의 오차가 여기에 기여를 하였기에 단일표본의 오차는 이 값의 제곱근으로 산출되며 $Se = \sqrt{\frac{Sd^2}{2}}$ 의 공식으로 표현될 수 있다¹⁶⁾. Dahlberg¹¹⁾는 오차 분산의 산출 공식으로 $S_e^2 = \sum d^2 / 2n$ 을 제안하였으며 이는 많은 교정학적 연구에 인용되어 왔다.

본 연구에서 마킹의 수정에 따라 t-검증을 시행한 결과, 전체 계측항목의 4.6%인 6개 계측항목에서 유의성있는 차이가 관찰되었고 그 범위는 0.67-1.71도 또는 mm를 보였다. 이 수치는 신뢰의 한계를 표시해주는 것으로써 증례의 68%에서는 임의적 오차의 1 S.D.이내에 존재하며 95%의 구간에서는 2 S.D.이내에 존재함을 의미하는 것이다.

본 연구에서 가장 큰 임의적오차를 보였던 상, 하악 전치가 이루는 각은 95%의 유의수준에서 볼 때, 실제의 계측치는 $127.55 \pm 2 \times 1.71(124.13-130.97)$ 도의 범위내에 있는 것으로 해석할 수 있는 것이다.

계측오차에 대한 분석은 매우 지루하고 어떤 특별한 보상이 주어지는 작업은 아니나, 계측이 이루어지는 모든 연구에서는 이러한 과정이 시행되어야 하고, 또한 보고서에 포함되어야 한다는 것에 많은 이들이 공감하고 있다¹⁶⁾. 이는 연구결과의 질적인 면을 향상시키는 것이 될 것이다. 교정학 분야에서 성장발육의 결과나 치료에 의한 어떤 변화를 방사선사진의 증첩에 의한 연구방법 등으로 평가하는데 있어서 계측오류에 관한 평가나 보고가 덧붙여지지 않는다면 그 결과는 신뢰하기 어려운 것이라 할 수 있겠다. 따라서 추후 시행되는 모든 계측과 관련된 모든 연구에는 오류의 분석이 함께 이루어져야 할 것으로 생각된다.

V. 결 론

본 연구는 교정 치료를 위한 진단과 치료 계획의 수립에 널리 이용되는 진단 수단인 두부 방사선 사진을 계측하는 연구에서 볼 수 있는 오류의 중

요성을 인식하고, 반복 계측의 방법으로 체계적 오차(systematic error)를 찾아내고 임의적오차(random error)의 범위를 산출하는 과정을 통해 기준점의 식별에 보다 주의가 기울여져야 할 해부학적 기준점들을 찾아보고자 시행되었는데, 통법에 의해 촬영된 100장의 측모 두부 방사선 사진을 연구대상으로 하여 61개의 기준점과 130개의 계측항목을 선정하였고 반복계측을 위해 25개의 표본을 무작위로 선정, 계측오차를 분석하여 다음과 같은 결론을 얻었다.

1. 반복계측에 따른 계측치의 비교에서 130개 전체 계측항목의 18.4%인 24개 항목에서 유의성 있는 차이를 보였다.
2. 반복계측에서 기준점 식별에 차이를 보이는 빈도는 상악 전치의 치근단점(as), 상악골의 최후방벽(tu), 연조직 nasion과 전두정(ft) 그리고 기도의 ad3점에서 매우 높았다.
3. 5% 이하의 유의도를 보이는 수준까지 기준점 마킹을 수정한 후, 임의적오차의 범위는 0.67-1.71mm 또는 도를 보였다.
4. 가장 임의적오차를 보인 계측항목은 상, 하악 전치가 이루는 각(interincisal angle ; ILs-ILi)이었다.
5. 계측 오차의 문제는 주로 해부학적 정의에 대한 정확성의 부족과 불분명한 방사선사진상에 기인되었다.

이상의 연구 결과로 두부 방사선 사진의 계측에는 많은 오류의 가능성이 있음을 확인할 수 있었고 이에 따라 두부 방사선 계측에서 뿐 아니라 계측이 이루어지는 모든 연구에서는 항상 계측오류에 대한 분석이 함께 이루어져야 할 것이며, 조사자로서 뿐만 아니라 다른 연구자의 연구결과를 올바르게 인식하기 위한 목적으로도 오류에 대한 인식은 필수적인 것으로 생각된다.

참고문헌

1. 김동희, 박양중, 손중권, 허명희 : 통계학, 서울 : 자유아카데미, 1994, pp.162-165.
2. 김동희, 박양중, 손중권, 허명희 : 통계학, 서울 : 자유아카데미, 1994, pp.5-12.

3. 김동희, 박양중, 손중권, 허명희 : 통계학, 서울 : 자유아카데미, 1994, pp.12-18.
4. Adams, J.W. : Correction of error in cephalometric roentgenograms. Angle Orthod. 10, 3-13, 1940.
5. Ahlqvist, J., Eliasson S., and Welander U. : The cephalographic projection Part II. Principles of image distortion in cephalography. Dentomaxillofacial Radiol. 12, 101-8, 1983.
6. Ahlqvist, J., Eliasson S., and Welander U. : The effect of projection errors on cephalometric length measurements. Eur. J. Orthod. 8, 141-148, 1986.
7. Ahlqvist, J., Eliasson S., and Welander U. : The effect of projection errors on angular measurements. in cephalometry. Eur. J. Orthod. 10, 353-361, 1988.
8. Baumrind, S. and Frantz, R. : The reliability of head film measurements. 1. Landmark identification. Am. J. Orthod. 60, 111-127, 1971.
9. Baumrind, S. and Frantz, R. : The reliability of head film measurements. 2. Conventional angular and linear measurements. Am. J. Orthod. 60, 505-517, 1966.
10. Broadbent, B.H.Sr., Broadbent, B.H.Jr. and Golden, W.H. : Bolton standards of dentofacial developmental growth, St. Louis, The C.V. Mosby company, 1975, pp.33-44.
11. Dahlberg, G. : Statistical methods for medical and biological students, New York, Interscience Publication, 1940
12. Frankln J.B. : Certain factors of aberration to be considered in clinical roentgenographic cephalometry. Am. J. Orthod. 38, 351-368, 1952.
13. Gravely, J.F., and Benzie P.M. : The clinical significance of tracing error in cephalometry. Br. J. Orthod. 1, 95-101, 1974.
14. Healy, M.J.R. : Outliers in clinical chemistry quantity control schemes. Clin. Chem. 25, 675-677, 1979.
15. Houston, W.J.B. : A comparison of the reliability of measurement of cephalometric radiographs by tracings and direct digitization. Swed. Dent. J. 15, 99-103, 1982.
16. Houston, W.J.B. : The analysis of errors in orthodontic measurement. Am. J. Orthod. 83, 382-90, 1983.
17. Houston, W.J.B. and Maher, R. : Sources of errors in cephalometric measurement. Br. J. Orthod. 1983
18. Mitgaard, J., Bjork, A., and Linder-Aronson, S. : Reproducibility of cephalometric Landmarks and errors of measurement of cephalometric cranial distances. Angle Orthod. 44, 55-61, 1974.
19. Moyer, R.E., Bookstein, F.L., and Hunter W.S. : Analysis of the craniofacial skeleton : Cephalometrics, In : Moyers RE (ed) Handbook of Orthodontics, Chicago, 1988, pp.247-309.
20. Richardson, A. : An investigation into the reproducibility of some points, planes, and lines used in cephalometric analysis. Am. J. Orthod. 52, 637-651, 1966.

- ABSTRACT -

A STUDY ON THE ERRORS IN THE CEPHALOMETRIC MEASUREMENTS

Kwang-Cheon Na, Young-Jooh Yoon, Kwang-Won Kim

Department of Orthodontics, College of Dentistry, Chosun University

This study was done to recognize the importance of errors in measurements of cephalometric radiograph and to find the anatomical structures those need special care to select as a reference points through the detection of the systematic errors and estimation of random errors.

For this purpose, 100 cephalometric radiographs were prepared by usual manner and 61 reference points, and 130 measurement variables were established.

Measurement errors were detected and estimated by the comparison of the 25 randomly-selected samples for repeated measurements with the main sample.

The following results were obtained :

1. In comparison of the repeated measurements, there were statistical significant differences in 24 variables which were 18.4% of 130 total variables.
2. The frequency of the difference in identification of the reference points between the repeated measurements was very high in the root apex of upper incisor(as), the most posterior wall of maxilla(tu), soft tissue nasion(n'), soft tissue frontal eminence(ft), and ad3 in airway.
3. After correction of reference points marking until the level of below 5% significance, the range of random errors were from 0.67 to 1.71 degree or mm.
4. The variable shown the largest random error was the interincisal angle(ILs-ILi).
5. Measurement errors were mainly caused by the lack of precision in anatomic definitions and obscure radiographic image.

From the above results, the author could find the high possibility of errors in cephalometric measurements and from this point, we should include error analysis in all the studies concerning measurements.

It is essential to have a concept of error analysis not only for the investigator but also for a reader of other articles.

KOREA. J. ORTHOD. 1998 ; 28 : 75-83

※ **Key words** : Cephalometrics, Measurement error