Observations on Inter-Rater Agreement in Assessing Fusion Activities at Elbow, Wrist and Pelvis by Conventional and Digital X-Ray Films

Aniruddha Das[1], Srimoyee Mukherjee [2], Arani Majumder[3], Debasis Sarkar[4] , Partha Pratim Mukhopadhyay[5]

[1]Assistant Professor, [2] Junior Resident, [3] Tutor, [4] Tutor, [5]Ex- Professor Department of Forensic & State Medicine, Burdwan Medical College, West Bengal

Abstract

The present study was designed to measure the inter-rater reliability between two observers in opining and grading the degree of fusion of the secondary centres of ossification using digital and conventional X ray. Attempt was made to examine whether the digital x-ray is superior to conventional X-ray as regards the interrater reliability. It was found that The Kappa score ranged from 0.560 to 0.761 for conventional x-ray films; whereas it ranged from 0.840 to 0.946 in case of digital x-rays. Interrater agreement was noted to be near perfect using digital x ray for all the centres of ossification applied in the study .The accuracy of digital radiography in age determination by observing the fusion activity at elbow, wrist and pelvis (Using Schmeling's staging) will be superior than conventional radiography

Keywords*: Inter-rater agreement, Cohens Kappa, Digital-Conventional X-ray elbow-wrist-pelvis, ossification*

Introduction

The human skeleton undergoes age related morphological and structural changes. Of these changes, the secondary centre of ossification and their fusion activity follow a definite chronological pattern. This phenomenon is utilized in estimating age from radiological examination of bones .

Radiography is the commonest modality used to determine bone age. Radiological methods of assessing fusion activity may be by conventional or digital x ray. Non-radiological techniques of visualizing bone maturity (i.e. Ultrasonography,) have also been reported, but are not as accurate and reliable as radiography. Forensic Age Estimation constitutes a field of expertise in clinical forensic medicine. It aims at defining as accurately as possible the chronological age of individuals by radiological examination for bone development. This is often applied in legal proceedings, immigration cases and competitive sports to determine age of the subject

Examining the X ray film is it conventional or digital, requires ample amount of experience and expertise. The interpretation of the degree of fusion activity is often wrought with subjectivity and can depend on several factors including the type and quality of the film. Therefore, the repeatability, standardization and inter-observer agreement are import factors to reckon with.

The concept of inter-rater reliability or agreement pervades many areas of modern research and it becomes an important facet of accuracy of observation. Inter-rater reliability can be measured by various statistical methods. Cohen's Kappa is a commonly employed method to explore the agreement between two raters.

G. Galstaun [(1)] is considered as the pioneer of

Corresponding Author :
Dr. Srimoyee Mukherjee
Junior Resident, Department of Forensic & State Medicine, Burdwan Medical College, Purba Bardhaman, West Bengal,Pin 713104
E mail id - fsmdrani1982@gmail.com,
Mobile :- 9831437298, 9748883743

Radiological age estimation of Indian subjects. He prepared a chart containing the ages of appearance and fusion of several secondary centres based on conventional radiological examination on specific regional Indian population. It is widely used as reference and accepted in judicial proceedings.

Hjren et al. (2011) [2] conducted a study to show the urgent need of better legal procedures for age determination. In an earlier research conducted on a population of 114 immigrant Moroccan males in 2003, Garamendi et al. [3], found Carpus X-ray (Greulich and Pyle method) to be the most useful method to predict chronological age of over or under 18.

In 2000, Schmeling et al [4] in their study showed the importance of developing a standardized method of estimation of age. They established that combination of physical and secondary sexual characteristics, dental X-ray, X-ray of left hand along with X-ray and/or CT Scan of clavicles would increase the diagnostic accuracy for estimation of age.

Numerous researches have established that x ray remains the modality of choice in estimating age in almost all populations [5-14]. With the advent of digital X-ray its gradual increase in clinical application was noteworthy. It was followed that digital images would far out match the conventional ones in quality and clarity. Likewise it raises several questions: What would be the rational choice? Given the cost involved, what would be the option in a setting with fewer resources? Besides, how far did the observers agree on the findings using digital images compared to conventional X-ray? These pertinent questions prompted us to examine the interrater agreement in assessing the fusion activity in several bones using conventional and digital X-ray.

The present study was designed to measure the inter-rater reliability between two observers in opining and grading the degree of fusion of the secondary centres of ossification using digital and conventional x-ray. Also attempt was made to examine whether the digital x-ray is superior to conventional X-ray as regards the interrater reliability.

Materials and Methods

The present research was conducted in Department of F.S.M. of Burdwan Medical College, a tertiary Medical Teaching Institution of West Bengal, India. Thirty sets of conventional X-rays (each set consisting of 3 radiographs one each of elbow, wrist and pelvis in AP view), and 30 similar sets of digital X-rays (each set consisting of 3 radiographs) were used in the study. All the X-ray plates from the conventional group were examined using standard view box by first two authors of this study separately after blinding. The raters gave their scoring on the degree of fusion using Schmelings score for all the centres separately. Then the scores were examined for inter-rater agreement by calculating Cohens Kappa for each centre. Similar scoring and calculations were done for the digital X-ray group on all those 8 centres. Finally the ratings were compared and designated according to the degree of agreement. Prior to starting the experiments a series of workshops on the detection of fusion activity and scoring system was held by the authors under the guidance of the senior-most author. Thereafter two were randomly chosen to conduct the final scoring. This was done to exclude the bias that might arise out of difference in experience and expertise of the observers.

Analysis and calculation for Cohen's Kappa statistics was done by SPSS software for windows version 19. P value of less than 0.05 was considered significant when applicable.

Inclusion Criteria- All 30 sets of digital X-rays of elbow joint in the departmental archive, irrespective of sex, age and side.

Exclusion Criteria- All the X-ray films which were very old, worn out and inconspicuous (with improper exposure)

The identity of the subjects was not disclosed.

Schmeling's Score [15] :-

Schmeling differentiated five stages of bony ossification:

- Stage 1: ossification centre is NOT ossified
- Stage 2: ossification centre is ossified, BUT epiphyseal plate is not ossified
- Stage 3: epiphyseal plate is partly ossified

• Stage 4: epiphyseal plate is fully ossified and epiphyseal scar is visible

• Stage 5: epiphyseal plate is fully ossified and epiphseal scar is NO longer visible'

Kappa Scoring [16,17,18] :- The Kappa values vary from 0 to 1, where.

0 = agreement equivalent to chance.

0.1 – 0.20 = slight agreement.

0.21 – 0.40 = fair agreement.

0.41 – 0.60 = moderate agreement.

0.61 – 0.80 = substantial agreement.

0.81 – 0.99 = near perfect agreement

1 = perfect agreement

X-rays were taken from the departmental archive. Each set comprised the following:--

1) AP view of the wrist joint

2) AP and Lateral views of Elbow joint (in the same plate)

3) AP view of Pelvis.

Centres of ossification used in the study

1.	Lateral epicondyle (LE)	2.	Lower end of radius (LR)
3.	Medial epicondyle (ME)	4.	Lower end of ulna (LU)
5.	Tip of the olecranon (OL)	6.	Iliac crest (IC)
7.	Head of radius (HR)	8.	Ischial tuberosity (IT)

Results

Table 1- :- Symmetric Measures of Lateral Epicondyle (Digital X-ray)

LATERAL EPICONDYLE DIGITAL		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	.847	.083	7.498	.000
N of Valid Cases		30			

Table- 2 :- Symmetric Measures of Lateral Epicondyle (Conventional X-ray)

LATERAL EPICONDYLE CONVENTIONAL		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	.560	.127	4.378	.000

Table- 3 :- Symmetric Measures of Ischial Tuberosity (Digital X-ray)

ISCHIAL TUBEROSITY DIGITAL		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	.840	.089	6.187	.000
N of Valid Cases		30			

Table- 4:- Symmetric Measures of Ischial Tuberosity (Conventional X-ray)

ISCHIAL TUBEROSITY CONVENTIONAL		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	.761	.107	4.984	.000
N of Valid Cases		30			

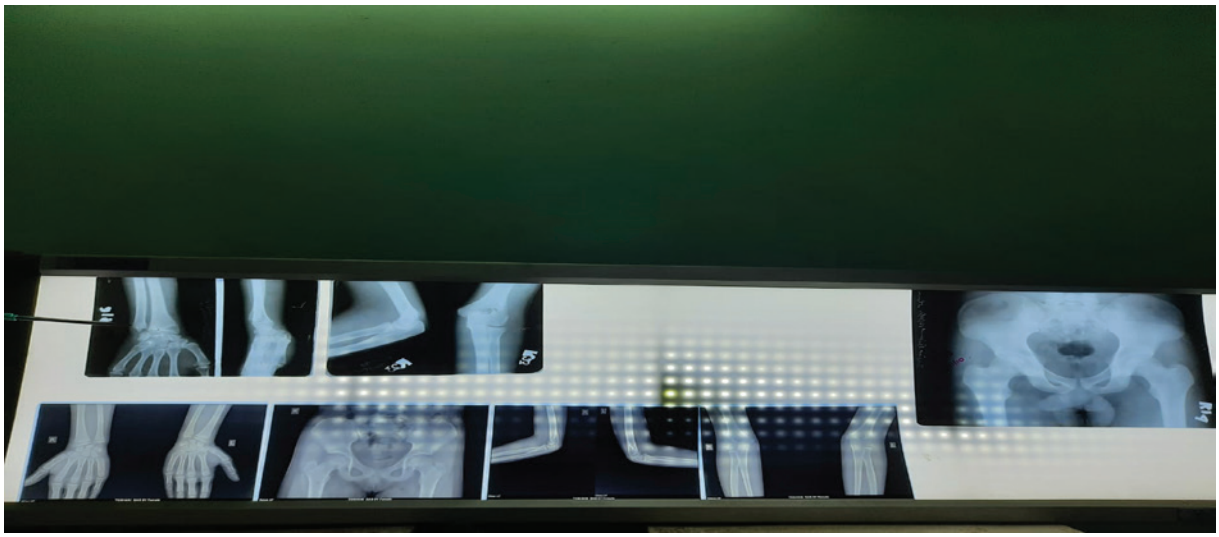
TABLE 5:- SHOWING INTER-RATER AGREEMENT IN THE FUSIONAL ACTIVITIES OF THE 8 BONY PARAMETERS WHILE STUDYING CONVENTIONAL AND DIGITAL X-RAYS

SL. NO.	BONY PARAMETERS	OBSERVED KAPPA SCORING	
		CONVENTIONAL X-RAY	DIGITAL X-RAY
1	LE	0.560	0.847
2	ME	0.670	0.946
3	OL	0.733	0.891
4	HR	0.747	0.870
5	LR	0.691	0.861
6	LU	0.670	0.857
7	IC	0.760	0.844
8	IT	0.761	0.840

1.	Lateral epicondyle=LE	2.	Lower end of radius= LR
3.	Medial epicondyle=ME	4.	Lower end of ulna=LU
5.	Tip of the olecranon=OL	6.	Iliac crest= IC
7.	Head of radius=HR	8.	Ischial tuberosity= IT

Table 6 Cross tabulation of scores for lateral epicondyle using digital x ray						
Count						
		DLE2				Total
		2.00	3.00	4.00	5.00	
DLE1	2.00	4	0	0	0	4
	3.00	0	14	3	0	17
	4.00	0	0	2	0	2
	5.00	0	0	0	7	7
Total		4	14	5	7	30

Photograph 1:- Digital & Conventional X-rays in 4 in 1 view box in Dept .of FSM, BMCH



Discussions

The present investigation, perhaps for the first time, was contemplated to examine the interrater agreement in assessing the fusion activity of centres of ossification of different bones of upper limbs and pelvis using conventional and digital X-ray.

The degrees of fusion activity of 8 ossification centres were assessed by using Schmeling's Score. The commonly assessed centres around elbow, wrist and pelvis were used in the study. The inter-rater agreement was calculated by estimating Cohen's Kappas.

The Kappa score was found to range from 0.560 to 0.761 for conventional x-ray films; whereas it ranged from 0.840 to 0.946 in case of digital x-rays.

The highest inter-observer agreement using conventional X-ray was noticed in the fusion activities of Ischial Tuberosity with value of K= 0.761. The lowest inter-rater reliability was noted in scoring the fusion activity of Lateral Epicondyle using conventional x-ray (K=0.560)

For digital X ray the highest agreement was noted in medial epicondyle (K=.946) while the lowest agreement was found in the ischial tuberosity (K=0.840)

For the lateral epicondyle the maximum inter observer agreement was noticed in the fusion activities graded as 4 and 5 as per Schmeling's stage. However, the observers agreed less when the Schmeling's stage was 3 while using digital X ray films. (as ascertained from the cross-tabulation Table 6).

The results with conventional X-ray for the same centre showed that the agreement was far less. This may be due to difference in opinion regarding presence / disappearance of the conjoint epiphyseal Cross tabulation of all other centres similarly showed more agreement at stage 3 and 4. Moreover, the epiphyseal scar also leads to difference in opinion and level of agreement.

There was difference in level of agreement on the fusion activities of the above noted 8 bony parameters between digital and conventional x-ray. The former showed near perfect agreement for all the centres while the latter showed moderate to significant agreement. This study clearly indicates that the digital x-rays produce much more precision than conventional x-rays. Earlier works have also showed that measurements using digital x ray can be successfully used for age assessment in adolescents.[17]

The secondary ossification centres play an important role in assessment of age from x ray so their activity needs to be assessed correctly. This is the basis of scientific evidence for determination of age [4, 5-7, 11]. Though several interrelated factors influence the ossification, estimates based on chronology help arrive at the age of the subject. This has been agreed upon by several researchers [9-12, 14].

Agreement between observers serves as the essential perquisite of the efficacy and utility of a method.

Assessment of fusion from films or comparing with known atlases for given population has been the age-old practice of determining age. With newer techniques, more research and suitable charts need to be evaluated for estimating age.

In the Indian context, especially for the Bengali population, the Galstaun [(1)] chart needs to be re assessed and refined using digital X- ray. More studies are required prior to formulating database for a given population. Herein lays the importance of the present investigation where agreement was noted to be near perfect using digital X-ray.

Cohen's Kappa was run to test the inter-rater agreement between two observers. Kappa statistics measures the agreement over and above that by chance. However, the correctness of the observation in this context cannot be tested. Both observers agreeing to large extent does not however mean that they were both correct in their interpretation. This is the methodological issue that needs to be considered. Later there is a scope to design the study with two or more raters who are randomly selected and applying robust technique like Fleiss18 kappa

Conclusions

Interrater agreement was noted to be near perfect using digital X-ray for all the centres of ossification applied in the study.

The accuracy of digital radiography in age determination by observing the fusion activity at elbow, wrist and pelvis (Using Schmeling's staging) will be superior to conventional radiography.

The advantages of digital x-ray are higher resolution, low exposure, data storage, processing, and retrieval and near perfect agreement between raters.

To the best of our knowledge this is perhaps the first reported study to assess and compare interrater agreement on fusion activity of ossification centres of upper extremity and pelvis using digital and conventional x ray.

Ethical Clearance:- Taken from Institutional Ethics Committee (IEC) of Burdwan Medical College, Purba Bardhaman.

Source of Funding :- Nil

Conflict of Interest :- Nil

Declaration

A shortened version of this study was presented by Second Author as oral presentation in Forensic Medicon-2020 at Hyderabad, India , 41st Annual National Conference Of the Indian Academy of Forensic Medicine.

References

1. Galstaun G. A study of ossification as observed in Indian subjects. *Indians J. Med. Res*.1937; 25: 267-324.
2. Hjern A, Brendler-Lindqvist M, Norredam M . Age assessment of young asylum seekers. *Acta Paediatrica*. 2012; 102(1): 4-7
3. Garamendi PM, Landa MI, Ballesteros J, Solano MA . Reliability of the methods applied to assess age minority in living subjects around 18 years old A survey on a Moroccan origin population. *Forensic Science International*. 2005; 154: 3-12
4. A. Schmeling, C. Grundmann, A. Fuhrmann, H.-J. Kaatsch, B. Knell, F. Ramsthaler, W. Reisinger, T. Riepert, S. Ritz-Timme, F. W. Rösing, K. Rötzscher, G. Geserick . Criteria for age estimation in living individuals. *International Journal of Legal Medicine*. 2008; 122(6): 457-460
5. Foti B, Lalys L, Adalian P, Giustiniani J, Maczel M, Signoli M, Dutour O, Leonetti G . New forensic approach to age determination in children based on tooth eruption. *Forensic Science International*. 2003; 132(1): 49-56. DOI: 10.1016/s0379-0738(02)00455-3
6. Mincer H. H., Harris E. F., and Berryman, H. E.. The A.B.F.O. Study of Third Molar Development and Its Use As an Estimator of Chronological Age. *Journal of Forensic Science*. 1993; 38(2): 379-390
7. Schmidt S., Koch B., Schulz R. *et al.* Comparative analysis of the applicability of the skeletal age determination methods of Greulich–Pyle and Thiemann–Nitz for forensic age estimation in living subjects. *International Journal of Legal Medicine.* 2007;121**,** 293–296 https://doi.

org/10.1007/s00414-007-0165-7

8. Bhise S.S., Nanandkar S. D. Age determination from radiological study of epiphysial appearance and fusion around elbow joint. *Journal of Forensic Medicine, Science and Law*. 2011; 20(1): 23-31
9. Garg A, Goyal N, Gorea RK,. Radiological Age Estimation from Xiphi-Sternal Joint in Living Person. *Journal of Indian Academy of Forensic Medicine*. 2011; 33(1): 27-29
10. Parikh CK. Parikh's Textbook of Medical Jurisprudence, Forensic Medicine and Toxicology. 6th ed. New Delhi, India: C.B.S. Publisher and Distributors; 2006. p. 2.10.
11. Miyazaki C S, Maranho D A, Agnollitto P.M, Nogueira-Barbosa M H . Study of secondary ossification centers of the elbow in the brazilian POPULATION. *Acta Ortop Bras*. 2017; 25(6): 279-282. doi: 10.1590/1413-785220172506170954.
12. Alwahbany S A M, Ahmed T O and Hussien A D . Radiological assessment of closure time of around elbow secondary ossification centers, in Khartoum hospital, Khartoum north hospital, Omdurman hospital, police hospital, and Umbadda hospital, in Khartoum, Sudan (December 2009-2010) . *International Journal of Orthopedic Sciences*. 2017; 3( ): 806-812. DOI: http://dx.doi.org/10.22271/ortho.2017.v3.i2i.87
13. Marak Fremingston K., Sangma W B Ch., Singh M Shyamo, Kharrubon B. A roentgenographic study for age estimation in boys and girls of North - Eastern region of India. *International Journal of Medical Toxicology & Legal Medicine*. 2008; 10(2): 38-43
14. Patel B, Reed M, Patel S. Gender-specific pattern differences of the ossification centers in the pediatric elbow. *Pediatric Radiology*. 2009; 39(3): 226-231.
15. Nyhagen, Karoline. Age Estimation in the Living Using Magnetic Resonance Imaging – A Review of Current Methods Identifying the 18-Years-Old Threshold. DOI: 10.13140/RG.2.2.25127.44966.
16. Altman DG. *Practical statistics for Medical Research*. London:Chapman and Hall;1991
17. McHugh Mary L. Interrater reliability: the kappa statistic. *Biochemia Medica*. 2012; 22(3): 276-282
18. Fleiss J L. Measuring Nominal Scale Agreement Among Many Raters. *Psychological Bulletin.*1971;76 (5): 378-382.