Pascal Adalian,[1] Ph.D.; Marie Dominique Piercecchi-Marti,[2] M.D., Ph.D.; Brigitte Bourliere-Najean,[3] M.D.; Michel Panuel,[3] M.D.; Catherine Fredouille,[4] M.D.; Olivier Dutour,[1] M.D., Ph.D.; and Georges Leonetti,[1,2] M.D., Ph.D.

Postmortem Assessment of Fetal Diaphyseal Femoral Length: Validation of a Radiographic Methodology

REFERENCE: Adalian P, Piercecchi-Marti MD, Bourliere-Najean B, Panuel M, Fredouille C, Dutour O, Leonetti G. Postmortem assessment of fetal diaphyseal femoral length: validation of a radiographic methodology. J Forensic Sci 2001:46(2):215–219.

ABSTRACT: Depending on the general condition of fetal remains, forensic specialists might face difficulties concerning age estimation. Reference tables and regression equations are helpful devices in this task, although they are generally applied for complete fetuses or fetal remains including soft tissues. However, the problem of age estimation stays for osseous remains, both for entire bones and ossified parts, since most of the reference tables come from ultrasonographic measurements, which are not easily reproducible on fetal osseous remains. Furthermore, the ultrasonographic measurements contain slight errors in comparison to the real anatomical ones. This study describes a radiographic protocol and a measurement technique that facilitate and improve bone measurements, and therefore, facilitate age estimation, too. A qualitative criterion, namely a clear-cut bony endplate, was defined and tested. Its reliability (repeatability and reproducibility) turned out to be good, showing nonsignificative differences to the threshold of 0.05, with average errors of 0.26 and 0.44 mm respectively. Moreover, concerning the test of eventual size differences between the right and left femurs showed a P value < 0.0001. The test of the qualitative criterion was based on the comparison of the radiographic in situ femur measurements and the radiographic measurements of the same bones after dissection. The results were satisfactory, since an average error of 0.58 mm was obtained, which did not give any significant differences to the threshold of 0.05. It was concluded that this methodology provides an easy and precise new measurement tool for forensic practice, and can allow us to establish some nonultrasonographic tables, which fit our population.

KEYWORDS: forensic science, forensic anthropology, fetus, femur, diaphyseal length, radiology, methodology, dissection

Estimating gestational age can be one of the forensic specialists' possible tasks. Depending on the general condition of fetal remains, the specialists can use several methods with reference to published growth standards in order to estimate age with the most accuracy. Several reference tables or regression equations are available for complete fetuses and fetal remains including soft tissues based on ultrasonographic examination. In these cases, the measurements of biparietal diameter (1–6), crown-heel length (1,6–8), crown-rump length (8–9), and foot length (10–11) can be used. Even if it was noticed that such in utero measurements contain slight errors (1 to 1.5 mm) compared to the real anatomical size (2), the tables based on them were found to be acceptable.

The second type of most frequently examined fetal remains are bones, or more precisely, the ossified portion of the cartilaginous matrix of the bone under construction. Most of the existing reference tables (2,8,12–21) for such fetal remains are of ultrasonographic origin, while the nonultrasonographic ones are old and do not fit our population. The purpose of our study was to assess a precise and easily reproducible method that can help us to establish new fetal growth standards.

Material and Methods

Our information on fetuses included in this study came from anonymous fetopathologic examination records, which were set down after various fetal deaths (spontaneous abortions, in utero death, stillborn fetuses). In view of the multidisciplinary character of the examination, which is the result of the existence of prenatal diagnosis centers established by the French law, every report was written by a specialist. Therefore, each review file contained: a radiological report, a frontal and lateral radiographic examination of the fetus, an eventual anomaly description, and a fetopathologic report that gave general information about the mother concerning her age, the number of preceding pregnancies, the number of pregnancies led to term, the present cause of fetal death, the expulsion mode, the possible diseases, and the familial medical history. Concerning the fetuses, we obtained the caryotype results, the biometric parameter values, and the description of the eventual visceral and external anomalies from the specialist.

We studied 1000 of the above-mentioned reports and selected 525 fetuses according to the following criteria (22): age between twelve and 41 weeks, absence of maternal diseases (diabetes, hypertension, infectious diseases with fetal repercussions, or congenital disease), and normal fetal caryotype. From the selected 525 reports 27 were twin pregnancies, presenting slightly different growth tendencies as unique pregnancies (15). These cases were excluded since they were liable to be mistaken with intrauterine

[1] Université de la Méditerranée, Faculté de Médecine, Unité d'Anthropologie - UMR 6578, 27 bd Jean Moulin, 13385 Marseille Cedex 05, France.
[2] Université de la Méditerranée, Faculté de Médecine, Service de Médecine Légale, 27 bd Jean Moulin, 13385 Marseille Cedex 05, France.
[3] Service de Radiologie Pédiatrique, CHU Timone, 264 rue St Pierre, 13385 Marseille Cedex 05.
[4] Service d'Anatomie Pathologique et Neuropathologique, CHU Timone, 264 rue St Pierre, 13385 Marseille Cedex 05.

Received 20 March 2000; and in revised form 12 May 2000; accepted 12 May 2000.

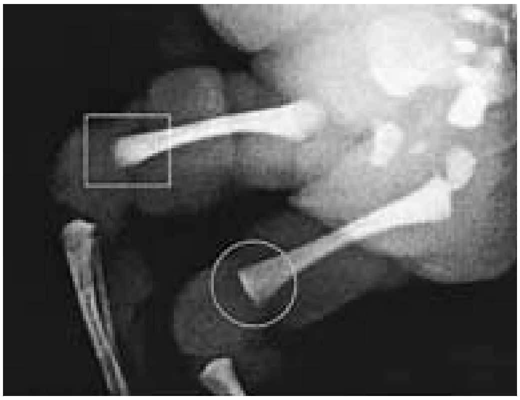

FIG. 1—*Qualitative criterion—The clear-cut bony endplate (see circle) identifies a good qualitative criterion. The other endplate (see square) is vague, and therefore corresponds to a bad qualitative criterion.*

growth retardation. Finally, the examined material involved 269 males and 229 females, totalling 498 fetuses.

Radiographic Examination Protocol

The radiographs were performed with a PHILIPS Diagnost 4 radiography table and a PHILIPS PCR/ACE treatment console with a focus-film distance of 1 m. The X-rays were taken on standard V phosphorus screen cassettes and printed on hard copy films.

Each fetus was examined from a frontal and a lateral view. They were laid directly on the table, fixing the head and limbs with adhesive ribbons. For the frontal view the limbs were stretched out, while for the lateral one the inferior limbs were half bent in order to avoid any superposition (Fig. 1). The applied radiographic parameters were 41 kV and 2 mAs.

Measurement Technique

As a consequence of the principles of radiography, and the fact that we worked on fetuses, what we measured was the ossified shaft of the developing bone. Without further precision, this portion is commonly termed only as "diaphysis" by the majority of authors, although the term "ossified shaft" includes the diaphysis (the result of the bone's primary ossification center) and the flared ends, also called metaphyses. However, in order to avoid any confusion when compared to other articles, we decided to use "diaphysis" too.

First of all, for the radiographic measurement of these diaphyses we had to take into consideration the difficulties of positioning. On the one hand, the bones had to be parallel to the film in order to avoid parallax errors. On the other hand, the bones had to lie directly on the radiographic plate in order to avoid the "cone of projection" phenomenon. As a consequence of these conditions, we decided to take the measurements on the profile radiographs, since it was sure that the bones of one side were leaning definitely on the radiographic plate. In spite of this choice, the lateralization of the limbs was not always evident. Therefore, to be sure of the measurement, we had to define a qualitative criterion. This criterion was the existence of a clear-cut endplate of the growing femur in the radiograph, which ensured the avoidance of the projection cone and the parallax phenomena, and, at the same time, we were informed about the bone lateralization (Fig. 1). After noting the lateralization, we measured right and left bones without any differentiation.

The measurements were carried out with a 0.5 mm graduated plastic ruler. First, we measured the diaphyseal length, and then the scale that was applied in the radiograph. After taking into consideration the ratio between the real and the radiographic scale each time (since this ratio was different concerning each radiograph), we could figure out proportionately the real diaphyseal sizes. The measurements were rounded up to the nearest upper 0.5 mm. We measured the maximal length of the diaphysis, taking care to remain parallel to its great axis.

Tests

As in each experimental protocol, and especially because we proposed a new methodology, we had to use several tests. Since we established a qualitative, or a subjective criterion, it was necessary to test it first.

Therefore, after the bones had been measured with the application of this criterion without considering sex, we arbitrarily chose to take the ages of 19, 26, and 33 gestational weeks in order to skip the influence of the radiographic visual amelioration due to the increasing calcification rate, which could have modified the methodology testing. We randomly selected 30 radiographs. The 60 femurs belonging to the chosen fetuses were dissected and X-rayed. They were directly laid on the radiographic plate in order to avoid the parallax and the projection cone artefacts, and, therefore the utilization of the criterion (Fig. 2). The radiographs were taken with the same previously described equipment, having the parameters of 45 kV and 8 mAs. Two radiographs were taken: one with the landmarks and the scale determined by the radiologist (Fig. 2) and the other one without.

The second test we made focused on the reliability of the measurements. Some authors were already interested in the estimation of measurement errors, notably White (23), who proposed a reduced gaps method in 1991. However, this technique does not take into account either the range of measurement, or the minimal technical error inherent to each measurement. As a consequence of these conditions, we preferred to choose the method published in 1995 by Signoli and Dutour (24). These authors proposed to check first of all the "repeatability," which is the error between two measurements taken at two different times by the same observer. Therefore, we remeasured the radiographs after a three-month interval.

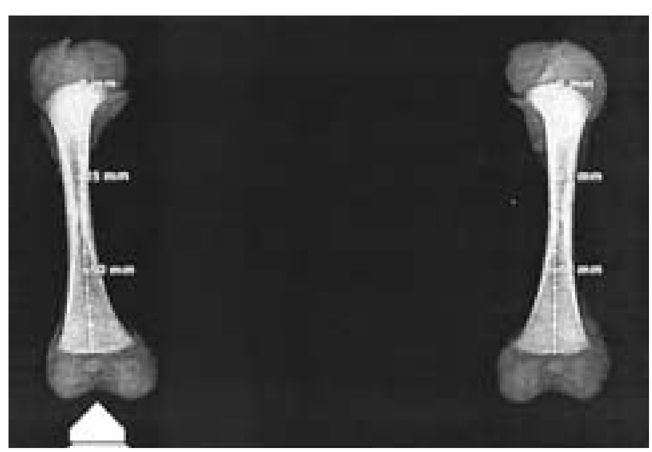

FIG. 2—*Radiograph of dissected femurs—X-ray picture of a right and a left side femur with a scale applied by the radiologist (the right side femur is indicated by a white arrow).*

The other point to be tested was the "reproducibility," which is the error attributable to the change of observers. Therefore, the second observer left out of consideration the results of the first one, and measured 30 radiographs using the same qualitative criterion and rounding up method of measurements.

Results

Difference Between In situ *and* Ex situ *Measurements*

The two radiographs of each dissected femur (with a scale and without) proved that we did not make errors in the measurement of the ossified parts (data not shown), and therefore we could create in situ and ex situ measurement tables for comparison (Table 1). The calculated average error between the two kinds of measurements was 0.58 mm. Because this error existed (even if it seemed to be weak), we checked if it resulted in significant differences.

As we did not expect any significant statistical difference between the in situ and ex situ measurements, we checked whether the pairing of the values was effective. We obtained a Spearman correlation of 0.9973, which corresponds to a highly significant pairing ($P < 0.0001$). Moreover, as we wanted to take into account the relative importance of the error as compared to the value, we applied the Wilcoxon matched pairs signed ranks test. The P value of the test was 0.7497, and therefore we could assert there was no significant difference to the threshold of 0.05 between the two series of measurements.

TABLE 1—*Difference between* in situ *and* ex situ *measurements.*

	In situ		Ex situ		
No.	LF	RF	LF	RF	Difference
1	34.8		33.9	34.5	0.9
2	33.8		33.5	33.1	0.3
3		30.8	31.3	31.3	0.5
4		28.8	29.2	28.9	0.1
5		32.4	32.2	31.9	0.5
6		31.8	31.6	31.6	0.2
7	39.6		41.2	41.2	1.6
8	33.6		33.8	33.5	0.2
9	38		37	37	1
10		38.6	37.3	37.7	0.9
11		34.3	34.7	34.7	0.4
12		43.7	44.4	44	0.3
13		30.6	30.7	31	0.4
14	50.6		51	51	0.4
15	48.9		48.2	48.2	0.7
16		43.2	42.8	42.4	0.8
17	37.1		37.1	37.3	0
18	53.1		52.5	52.5	0.6
19	59.4		57.8	58.7	1.6
20	57.6		57.9	57.5	0.3
21	63.8		64	64	0.2
22	69		70	70	1
23		64.2	64.7	65.2	1
24	63.5		63.8	63.8	0.3
25	63.5		64.2	63.7	0.7
26	73.7		73.2	72.7	0.5
27	73.7		72,8	72.8	0.9
28	71		71.5	71.1	0.5
29		77.9	77.8	77.4	0.5
30		78.7	79.3	79.3	0.6

NOTE: Using the qualitative criterion, the in situ measurements were taken directly on the profile view of the complete fetuses, while the ex situ measurements were taken after the dissection of each femur, laying them directly on the radiographic plate. (All measurements are in mm; RF stands for right femur, while LF for left femur).

Repeatability

After we had demonstrated that the chosen qualitative criterion implicated no significant differences, the second point to check in our methodology was the reliability. The radiographs were remeasured ignoring the results of the first measurement. The table of results allowed us to calculate the average error, which was 0.26 mm. The Wilcoxon nonparametric test (25) gave a Z score of -0.18 and a probability of 0.857. Therefore, we could conclude that there was no significant difference to the threshold of 0.05.

Reproducibility

A graphic representation allowed us to observe that the maximal error, which is due to a double error of measurement (an error of 0.5 mm on the bone and 0.5 mm on the scale), was 2 mm, while the average difference was 0.44 mm when compared to the first observers' measurements. The interpretation of these results by the Wilcoxon test gave a Z score of -0.55 and a probability of 0.582. Therefore, we could state that there was no significant difference to the threshold of 0.05.

Difference Between Right and Left Sides

The controversial question of the difference between right and left sides led us to test their possible differential growth. Therefore, we decided to compare pair by pair the radiographic measurements, which were taken on each dissected femur for the qualitative criterion test (comparison of in situ/ex situ measurements). Actually, it was in anticipation of this control that we dissected both femurs of the 30 fetuses. The table of measurements (Table 1) allowed us to calculate the average difference, which was 0.24 mm. In order to verify the possible statistical significance of this difference by the nonparametric Wilcoxon matched pairs signed ranks test, we wanted to confirm that the pairing was effective. We obtained a Spearman correlation value of 0.9978, which corresponded to a P value less than 0.0001. Therefore, the pairing proved to be highly effective. The Wilcoxon test gave a P value of 0.4537, allowing us to assert indisputably that there was no significant difference to the threshold of 0.05 between the right and left femoral lengths.

Discussion

Proposed Methodology

For the radiographic measurements, we proposed the utilization of a plastic ruler. This did not seem to be a source of error since the repeatability, which gave an average error of 0.25 mm, was excellent. Although the average error of the reproducibility test was slightly higher (0.44 mm), it did not result in significant differences either, therefore, changing the measure tool did not challenge us. As a consequence, we counsel the utilization of the 0.5 mm graduated plastic ruler.

The difference between in situ and ex situ measurements, which we tested to evaluate the determined qualitative criterion, gave an average error of 0.58 mm. We statistically proved that it did not result in a significant difference, therefore we could certify the criterion reliability.

We could conclude that the above described method is valid and can be used to establish new and precise growth standards by applying it on larger samples. Although in a recent study Warren (26) used similar techniques and measurements, he did not evaluate the reliability of such a method. The purpose of the present study was the validation of our methodology for its proper application in further studies carried out on the same material.

Why Radiography Rather than Ultrasonography?

One of the reasons that led us to use radiography came from the necessity to develop precise reference tables. In fact, the advent of ultrasonography three decades ago, and especially the high resolution of the real time ultrasonography, allowed the detection of developmental anomalies at increasingly precocious ages. In accordance with this, fetal pathology and a need for standards concerning precocious development became important. The situation was the same concerning radiography, of which interpretation had to be based on the knowledge of morphology and dimensions of the normal skeleton at all developmental stages. Unfortunately, even if the standards that were determined by ultrasonographic studies and realized on tens of thousands cases established a very good correlation between the ultrasonographic measurement and the gestational age, according to many authors, these measurements did not correspond to the real anatomical size. For instance, Alonso and Portman (1) stated that even if the ultrasonographic measurement of the biparietal diameter was an excellent age predictor, the femur length measured by ultrasonography was significantly smaller than the anatomical length (the coefficient of the variation was 23%) and, therefore, was not a reliable indicator of age and fetal development for a biometric study. This statement was supported by Guihard-Costa and Droulle (2), who proved that the admitted echographic measurement error was 1 to 1.5 mm.

These errors were known for a long time and in 1980 Cronck (27) already revealed some causes: dimensions that could be measured by ultrasonography depended on the different gestation periods, differed according to the image quality, which allowed one to see anatomical landmarks, and the size of the relative dimension to the image resolution. We also noted that fetal movements resulted in errors that could not be corrected. According to Cronck, the ultrasonic beam width itself is already a source of error in positioning the landmarks.

Another reason for our choice came from the observation that ultrasonographists did not use the osseous diaphyseal length very often, and, thereby, the reference tables concerning them were less numerous than those based on the biparietal diameter or the transverse thoracic width. Concerning essential applications of ultrasonography, in 1983 Cronck (27) wrote that the femur measurement is only used as a cross test with age determination based on the biparietal diameter.

Conclusion

The appropriate measurement of femur diaphyseal length on radiographs necessitated the utilization of a qualitative measurement criterion, which improved the reliability of this method in comparison to the ex situ measurements. The measurements taken on this basis presented statistically significant repeatability and reproducibility, which, besides providing a new tool for forensic specialists, offers the possibility of establishing new fetal growth standards after its application on all long bones and on bigger samples. Accordingly, we plan to create these reference tables and establish our own regression equations for age determination in a following study.

Acknowledgments

Sincerest thanks are given to Marta Maczel for editorial assistance during preparation of this manuscript.

References

1. Alonso K, Portman E. Fetal weights and measurements as determined by postmortem examination and their relation with ultrasound examination. Arch Pathol Lab Med 1995;119:179–80.
2. Guihard-Costa AM, Droulle P. Croissance du diamètre bipariétal, du diamètre abdominal transverse et de la longueur du fémur chez le fœtus. Influence du sexe. Cahiers d'Anthropologie et de Biométrie humaine 1990;VIII(1–2):49–69.
3. Hern MW. Correlation of fetal age and measurements between 10 and 26 weeks of gestation. Obstet Gynecol 1984;63:26–32.
4. Deter RL, Harrist RB, Hadlock FP, Carpenter RJ. The use of ultrasound in the assessment of normal fetal growth: a review. J Clin Ultrasound 1981 Dec;9:481–93.
5. Deter RL, Harrist RB, Hadlock FP, Poindexter AN. Longitudinal studies of fetal growth with the use of dynamic image ultrasonography. Am J Obstet Gynecol 1982;143:545–54.
6. Fazekas IG, Kosa K. Forensic fetal osteology. Budapest: Akademiai Kiado Publishers, 1978.
7. Brenner WE, Edelman DA, Hendricks CH. A standard of fetal growth for the United States of America. Am J Obstet Gynecol 1976;126: 555–64.
8. Olivier G. Précisions sur la détermination de l'âge d'un fœtus d'après sa taille et la longueur de ses diaphyses. Médecine Légale et Dommage Corporel 1974;4:297–9.
9. Daya S. Accuracy of gestational age estimation by means of fetal crown-rump length measurement. Am J Obstet Gynecol 1993;168(3–1): 903–8.
10. Mercer BM, Sklar S, Shariatmader A, Gillieson MS, D'Alton ME. Fetal foot length as a predictor of gestational age. Am J Obstet Gynecol 1987;156(2):350–5.
11. Kumar GP, Kumar UK. Estimation of gestational age from hand and foot length. Med Sci Law 1993;33(4):48–50.
12. Chitty LS, Altman DG, Henderson A, Campbell S. Charts of fetal size: 4. Femur length. Br J Obstet Gynaecol 1994;101:132–5.
13. Bagnall KM, Harris PF, Jones PRM. A radiographic study of the longitudinal growth of primary ossification centres in limb long bones of the human fetus. Anat Rec 1982;203:292–9.
14. Warda AH, Russel LD, Rossavik IK, Carpenter RJ, Hadlock FP. Fetal femur length: a critical reevaluation of the relationship to menstrual age. Obstet Gynecol 1985;66(1):69–75.
15. Chervenak FA, Skupski DW, Romero R, Myers MK, Smith-Levitin M, Rosenwaks Z, et al. How accurate is fetal biometry in the assessment of fetal age? Am J Obstet Gynecol 1998;178:678–87.
16. Hohler CW, Quetel TA. Fetal femur length: equations for computer calculation of gestational age from ultrasound measurement. Am J Obstet Gynecol 1982;143(4):479–81.
17. Miller JM, Foster TA, Brown HL, Gabert HA. Fetal anthropometry at term: effect of menstrual age and relative fetal size. J Clin Ultrasound 1989 March–Apr;17:193–6.
18. Kelemen E, Janosa M, Calco W, Fliedner TM. Developmental age estimated by bone length measured in human fetuses. Anat Rec 1984;209:547–52.
19. Hill LM, Guzick D, Hixson J, Peterson CS, Rivello DM. Composite assessment of gestational age: a comparison of institutionally derived and published regression equations. Am J Obstet Gynecol 1991;166: 551–5.
20. Yeh MN, Bracero L, Reilly KB, Murtha L, Aboulafia M, Barron BA. Ultrasonic measurement of the femur length as an indicator of fetal gestational age. Am J Obstet Gynecol 1982;144:519–22.
21. Merz E, Grußner A, Kern F. Mathematical modeling of fetal limb growth. J Clin Ultrasound 1989 March–Apr;17:179–85.
22. Rodriguez JI, Palacios J, Rodriguez S. Transverse bone growth and cortical bone mass in the human prenatal period. Biol Neonate 1992; 62:23–31.
23. White T. Human osteology. San Diego/New York (US), Academic Press Inc., 1991.
24. Signoli M, Dutour O. Biométrie de l'os tympanal. Mise au point et évaluation de nouvelles mesures. Préhistoire Anthropologie Méditerranéennes 1995;4:71–7.

25. Chenorkian R. La pratique archéologique statistique et graphique. Ed. Errance et Adam, 1996.
26. Warren MW, Radiographic determination of developmental age in fetuses and stillborns. J Forensic Sci 1999;44(4):708–12.
27. Cronck CE. Fetal growth as measured by ultrasound. Year Phys Anthropol 1983;26:65–89.

Additional information and reprint requests:
Prof. Georges Leonetti
Université de la Méditerranée, Faculté de médecine
Service de médecine légale
27 bd Jean Moulin
13385 Marseille Cedex 05, France