PRENATAL DIAGNOSIS
Prenat Diagn 2000; **20**: 210–214.

Critical reappraisal of the utility of sonographic fetal femur length in the prediction of trisomy 21

Patrizia Vergani[1], Anna Locatelli[1], Maria Giovanna Piccoli[1], Eloisa Mariani[1], Nicola Strobelt[1], John C. Pezzullo[2] and Alessandro Ghidini[2]*

[1]*Divisione di Ostetricia e Ginecologia, Istituto di Scienze Biomediche San Gerardo, Monza, Italy*
[2]*Department of Obstetrics and Gynecology, Georgetown University Medical Center, Washington, D.C., U.S.A.*

Measurement of femur length (FL) has been advocated as part of a genetic sonogram for the prediction of Down syndrome (DS). However its predictive ability has been inconsistent. We have studied the diagnostic value of this sonographic parameter in a prospective cohort of women with singleton gestations undergoing genetic sonogram between 14 and 22 weeks because of advanced maternal age or family history of aneuploidies. Genetic sonograms were performed at a mean gestational age of 17.0 weeks (range 14–22). DS was diagnosed in 30 fetuses, while 888 were euploid. Mean $\pm$SD observed/expected (O/E) values of FL (1.00 ± 0.10 versus 0.97 ± 0.01, $p = 0.07$) were not significantly different between euploid and DS fetuses. Comparison of the regression equations of FL versus biparietal diameter revealed that while the intercepts were not significantly different between euploid and DS fetuses, the difference in slopes reached significance ($p = 0.04$) suggesting that the predictive ability of FL may increase with advancing gestational age. In addition, a MEDLINE search (National Library of Medicine) was conducted for articles published between 1985 and 1998 on fetal femur length in the prediction of trisomy 21. Review of the published literature on the subject suggests that FL is not a consistent or reliable sonographic predictor of DS. Published thresholds of FL should not be used outside of the Institution from which they originated, and each Institution should establish whether this parameter has predictive ability in its own population. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: prenatal diagnosis; trisomy 21; ultrasonography

INTRODUCTION

Ultrasonography offers a non-invasive means to adjust the age-related risk of Down syndrome, and thus improve the selection of pregnant women who may be candidate for invasive prenatal diagnosis, by utilizing a series of markers that are more frequently present in aneuploid than euploid fetuses (Nadel *et al.*, 1995; Vintzileos *et al.*, 1995; Bromley *et al.*, 1997; Nyberg *et al.*, 1998). Length of fetal long bones has been proposed as a reliable indicator of trisomy 21 by several centres. Much attention has been focused on femur length, because it is the long bone most commonly measured during standard prenatal ultrasonographic examinations. The reproducibility of this marker however has been inconsistent among different institutions.

To calculate the predictive ability of ultrasonographic measurements of femur length during the early second trimester for the detection of fetal trisomy 21, we conducted a prospective cohort study in a pregnant population undergoing genetic sonogram because of advanced maternal age of 35 years or older, or family history of aneuploidies. In addition, we have reviewed the series published in the literature on the value of femur length in the prediction of trisomy 21.

*Correspondence to: A. Ghidini, Department of Obstetrics and Gynecology, Georgetown University Medical Center, 3800 Reservoir Road, N.W. – 3 HC, Washington, D.C. 20007, U.S.A. E-mail: ghidina@gusun.georgetown.edu

METHODS

During a six-year period (January 1990–December 1996), all women with singleton fetuses receiving genetic counselling because of maternal age 35 years or older at delivery, or family history of aneuploidies underwent a targeted ultrasonographic study between 14 and 22 weeks' gestation. Maternal serum biochemical screening was not used in the study population. All ultrasonographic examinations were performed by six physicians with expertise in prenatal diagnosis and without prior knowledge of the fetal karyotype. After excluding cases examined after 22 weeks' gestation ($n = 19$), voluntary terminations of pregnancy for whom karyotype analysis was not available ($n = 3$), intrauterine demises with no karyotype available ($n = 5$), preterm neonatal death with no karyotype ($n = 1$), chromosome anomalies other than autosomal trisomies ($n = 1$), and aneuploidies other than trisomy 21 ($n = 10$), 918 cases were available for analysis.

The second trimester ultrasonographic examination (Ultramark 9, Advanced Technology Laboratories, Bothell, WA, USA) included standard fetal biometry and evaluation of fetal anatomy for detection of structural anomalies. Fetal karyotype information was obtained from the results of cytogenetic examinations performed either at mid-trimester by amniocentesis, or at birth in cases clinically indicated. Paediatric assessment at birth was available in 100% of cases.

In addition, a MEDLINE search (National Library of Medicine) was conducted for articles published

Received: 30 September 1999
Revised: 9 December 1999
Accepted: 13 December 1999

between 1985 and 1998 in which fetal long bones length was used in the prediction of trisomy 21. The reference lists of all identified articles were examined to elicit additional relevant studies.

Statistical analysis

Regression analysis was performed for femur length (FL) as a function of biparietal diameter (BPD). Based on this regression equation, expected values for a given BPD were calculated. The ratios of observed to expected values were compared between euploid and trisomy 21 fetuses using the Student t-test.

RESULTS

Of the 957 women enrolled in the study, 30 had fetuses with Down syndrome. Mean gestational age at ultrasound was 17.0 weeks [standard deviation (SD) = 1.7, range 14–22 weeks]. Mean maternal age was 38.4 years (SD = 2.1, range 26–47 years). FL was available in 888 euploid and 30 Down syndrome cases. The regression line of FL based on the BPD in euploid fetuses was: FL = 0.88*BPD − 9.89 (Figure 1). Based on this regression line, we calculated the O/E values of FL for a given BPD. Mean ± SD observed/expected (O/E) FL was not significantly different between euploid and aneuploid fetuses (1.00 ± 0.10 versus 0.97 ± 0.01, $p = 0.07$). Of interest, comparison of the regression equations of FL versus BPD revealed that while the intercepts were not significantly different between euploid and aneuploid fetuses, the difference in slopes reached significance ($p = 0.04$), suggesting that the predictive ability of this sonographic marker may improve with advancing gestational age (Figure 1). The conclusions were not different when the analysis was limited to the patients at high risk because of advanced maternal age alone.

Thirty-two studies were identified which evaluated the diagnostic indices of FL (Table 1). Seven additional studies were not considered in our analysis either because the population was subsequently included in other studies from the same centre (Vintzileos *et al.*, 1995; Vintzileos *et al.*, 1996a,b; Nyberg *et al.*, 1993), or because the methodology used did not provide information on the univariate analysis comparing Down syndrome and euploid fetuses (Platt *et al.*, 1992; Deren *et al.*, 1998; Winston and Horger, 1988; Bahado-Singh *et al.*, 1998).

DISCUSSION

Long bones measurement has been advocated as a useful parameter in the diagnosis of trisomy 21 because it does not require considerable expertise. Even though fetal femur length would theoretically be an ideal component of a genetic sonogram because it is easy to obtain, we did not find it a useful parameter to discriminate between trisomy 21 and euploid fetuses. A review of the published series on the subject reveals that the regression lines of FL on BPD were not significantly different between Down syndrome and euploid fetuses in six other centres in addition to ours (LaFollette *et al.*, 1989; Lynch *et al.*, 1989; Marquette *et al.*, 1990; Shah *et al.*, 1990; Perrella *et al.*, 1988; Winston and Horger, 1988). Six additional centres found that the regression lines of FL versus BPD were statistically different between Down syndrome and euploid fetuses, suggesting that true biologic differences may exist. However, they noted that these differences were not of sufficient magnitude to be clinically useful (Peters *et al.*, 1989), or optimal thresholds could not be found to significantly dis-

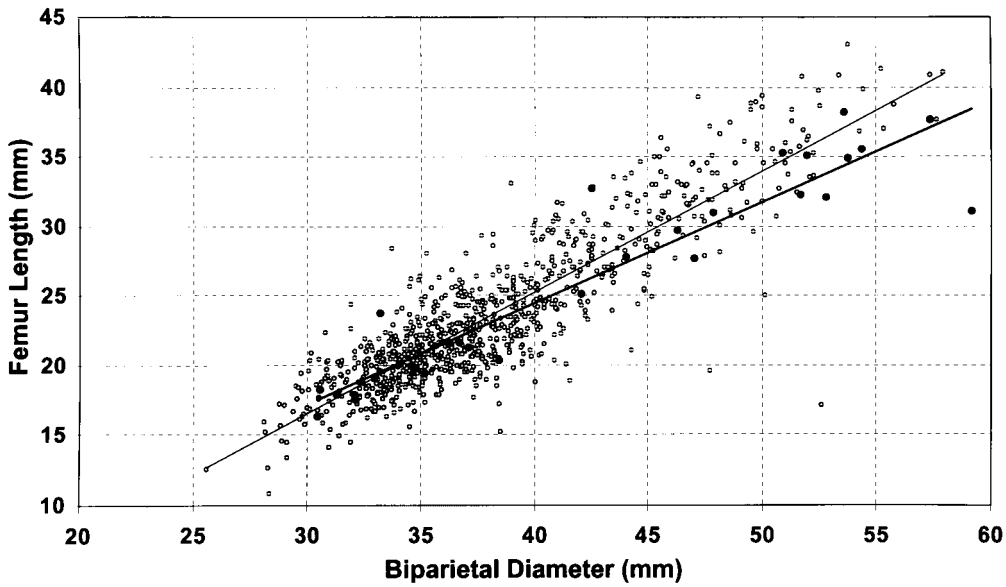

Figure 1—Relationship between biparietal diameter and femur length in Down syndrome (filled circles, bold line) and euploid (open circles, thin line) fetuses. The regression line of FL in euploid fetuses is: FL = 0.88*BPD − 9.89; in fetuses with trisomy 21 it is: FL = 0.72*BPD − 4.44

Table 1—Studies on the efficacy of femur length in the prediction of fetal trisomy 21

Author (year)	Type of study	Abnormal test definition	True positive rate	False positive	*p* value
Benacerraf *et al.* (1987)	Case-control	O/E ≤ 0.91	19/28 (68%)	4/192 (2%)	<0.001
Lockwood *et al.* (1987)	Case-control	BPD/FL > 1.5 SD	18/35 (51%)	26/349 (7%)	<0.001
Perrella *et al.* (1988)	Case-control	O/E ≤ 0.91	5/19 (26%)	29/128 (23%)	0.95
Winston and Herger (1988)	Case-control	Not calculated	Not calculated	Not calculated	0.673
Peters *et al.* (1989)	Case-control	O/E ≤ 0.91	2/16 (12%)	14/194 (7%)	0.78
Brumfield *et al.* (1989)	Case-control	BPD/FL ≥ 1.80	6/15 (40%)	1/45 (2%)	<0.001
La Follette *et al.* (1989)	Case-control	O/E ≤ 0.91	4/30 (13%)	27/229 (12%)	1
Lynch *et al.* (1989)	Case-control	O/E ≤ 0.91	5/9 (56%)	5/9 (56%)	1
Dicke *et al.* (1989)	Case-control	BPD/FL > 1.5 SD	6/33 (18%)	7/177 (4%)	0.007
Cuckle *et al.* (1989)	Case-control	O/E ≤ 0.85	9/83 (11%)	19/1340 (1%)	<0.001
Hill *et al.* (1989)	Case-control	O/E ≤ 0.91	11/22 (50%)	43/286 (15%)	<0.001
Benacerraf *et al.* (1989)	Case-control	O/E ≤ 0.91	7/20 (35%)	28/709 (4%)	<0.001
Shah *et al.* (1990)	Case-control	BPD/FL > 90%	3/17 (17%)	1/17 (6%)	0.59
Ginsberg *et al.* (1990)	Case-control	BPD/FL > 1.5 SD	5/11 (45%)	14/212 (7%)	<0.001
Marquette *et al.* (1990)	Case-control	BPD/FL > 1.5 SD	3/31 (10%)	14/155 (9%)	1
Nyberg *et al.* (1990)	Case-control	O/E ≤ 0.91	7/49 (14%)	35/572 (6%)	0.06
Rodis *et al.* (1991)	Case-control	BPD/FL > 95th	2/11 (18%)	95/1890 (5%)	0.2
Rotmensch *et al.* (1992)	Case-control	O/E < 0.90	8/43 (18%)	18/204 (9%)	0.1
Benacerraf *et al.* (1992)	Case-control	O/E ≤ 0.91	23/32 (71%)	63/588 (11%)	<0.001
Hadlock *et al.* (1992)	Cohort	BPD/FL > 90%	4/16 (25%)	202/1770 (11%)	0.19
Lockwood *et al.* (1993)	Cohort	O-E ≤ 3.4	6/41 (24%)	180/4874 (4%)	<0.002
Benacerraf *et al.* (1994)	Case-control	O/E ≤ 0.91	20/45 (44%)	4/106 (4%)	<0.001
Campbell *et al.* (1994)	Case-control	BPF/FL > 1.5 SD	3/6 (50%)	15/264 (6%)	0.003
Johnson *et al.* (1995)	Case-control	BPD/FL > 1.80	7/26 (27%)	495/2763 (18%)	<0.001
Bromley *et al.* (1997)	Case-control	O/E < 0.90	19/46 (41%)	5/149 (3%)	<0.001
Vintzileos *et al.* (1997)	Case-control	O/E ≤ 0.91	10/23 (43%)	11/581 (2%)	<0.001
Verdin and Ecoromides (1998)	Cohort	BPD/FL > 97.5%	6/11 (54%)	5/449 (1%)	<0.001
Nyberg *et al.* (1998)	Case-control	O/E < 0.89	27/142 (19%)	11/930 (1%)	<0.001

Note: O/E, ratio of observed/expected values; BPD, biparietal diameter; FL, femur length; SD, standard deviation.

criminate between trisomy 21 and euploid fetuses (Nyberg *et al.*, 1990; Rodis *et al.*, 1991; Rotmensch *et al.*, 1992; Hadlock *et al.*, 1992). The difference in sample size of Down syndrome cases between the studies that did ($n=17$, median Down sydrome cases = 30, range = 6–142) versus did not ($n=12$, median Down syndrome cases = 18, range = 9–49) find differences in FL between euploid and DS fetuses is not statistically significant ($p=0.28$), suggesting that sample size is not a critical issue for the discrepant conclusions.

A few centres did find that FL was a good predictor of trisomy 21, with the following caveats. First, there are significant differences among institutions in the regression lines of euploid fetuses (Hill *et al.*, 1989; Lockwood *et al.*, 1987). In addition, there is significant heterogeneity among different centres in O/E FL values in euploid and/or trisomy 21 fetuses (Cuckle *et al.*, 1989; Nyberg *et al.*, 1993; Lockwood *et al.*, 1987). Therefore, regression equations should be generated at each institution to calculate optimal cut-off values, and each centre should evaluate whether FL is a significant and independent predictor of aneuploidies. Relying on curves generated at outside institutions, as it has been proposed (Vintzileos *et al.*, 1996a) could adversely affect the predictive ability and accuracy of this biometric marker. Of interest, discrepancies in the predictive ability of FL have also been observed among different operators within the same centre (Grandjean *et al.*, 1995) and in different studies published by the same centres (Rodis *et al.*, 1991 versus Vintzileos *et al.*, 1996b, 1997; Lockwood *et al.*, 1987 versus Rotmensch *et al.*, 1992; Benacerraf *et al.*, 1989 versus Benacerraf *et al.*, 1992). This suggests that centres which rely on FL as a predictive marker of trisomy 21 should establish their inter-operator variability and should periodically reassess the predictive ability of this marker.

A second caveat is that each centre found statistical significance using different definitions and thresholds for abnormal FL. When institutions compared the predictive ability of different definitions within the same population, they did not find consistent results (Dicke *et al.*, 1989; Hill *et al.*, 1989; Hadlock *et al.*, 1992; Johnson *et al.*, 1995). This suggests that each institution should find the definition and threshold which optimally discriminates between euploid and trisomy 21 fetuses, even though such practice will inevitably hamper comparison among different institutions. A final caveat is that the false positive rate in the euploid population in some studies was close to or in excess of 10%, suggesting that the high detection rate of Down syndrome would lead in some series to an excessively high rate of amniocenteses in normal fetuses (Hill *et al.*, 1989; Benacerraf *et al.*, 1992; Johnson *et al.*, 1995; Grandjean and Sarramon, 1995).

The repercussion of this approach would be particularly worrisome in women without prior risk for trisomy 21, in whom the higher false positive rates would translate into a greater proportion of procedure-related losses of euploid fetuses.

Our finding that the slopes of the regression lines of FL are different between euploid and trisomy 21 fetuses suggests that as femur shortening becomes more evident with advancing gestational age, the predictive ability of FL may improve. A similar observation was also made by other investigators (Shah *et al.*, 1990; Platt *et al.*, 1992). The role of gestational age at genetic sonogram in the predictive ability of FL cannot be fully evaluated from the published literature because most studies do not report the mean/median gestational age at ultrasound examination in the study population.

In summary, FL does not appear to be a reliable and reproducible enough parameter to be proposed as a screening method for Down syndrome in all institutions. Even though recent studies propose the use of multiple genetic markers for the prediction of DS, many still incorporate FL among the markers used in a genetic sonogram. Our findings suggest that FL should not be incorporated in diagnostic algorithms or scoring systems inclusive of other parameters, such as maternal serum markers or sonographic nuchal fold thickness, which have withstood the test of comparison among different operators and institutions.

REFERENCES

Bahado-Singh RO, Deren O, Oz AU, Tan A, Hunter D, Copel J. 1998. An alternative for women initially declining genetic amniocentesis: individual Down syndrome odds on the basis of maternal age and multiple ultrasonographic markers. *Am J Obstet Gynecol* **179**: 514–519.

Benacerraf BR, Gelman R, Frigoletto FD. 1987. Sonographic identification of second trimester fetuses with Down's syndrome. *N Engl J Med* **317**: 1371–1376.

Benacerraf BR, Cnann A, Gelman R, Laboda LA, Frigoletto FD. 1989. Can sonographers reliably identify anatomy features associated with Down syndrome in fetuses? *Radiology* **173**: 377–380.

Benacerraf BR, Neuberg D, Bromley B, Frigoletto FD. 1992. Sonographic scoring index for prenatal detection of chromosomal abnormalities. *J Ultrasound Med* **11**: 449–458.

Benacerraf BR, Nadel A, Bromley B. 1994. Identification of second-trimester fetuses with autosomal trisomy by use of a sonographic scoring index. *Radiology* **193**: 135–140.

Bromley B, Lieberman E, Benacerraf BR. 1997. The incorporation of maternal age into the sonographic scoring index for detection at 14–20 weeks of fetuses with Down's syndrome. *Ultrasound Obstet Gynecol* **10**: 321–324.

Brumfield CG, Hauth JC, Cloud GA, Davis RO, Henson BV, Cosper P. 1989. Sonographic measurements and ratios in fetuses with Down syndrome. *Obstet Gynecol* **73**: 644–646.

Campbell WA, Vintzileos AM, Rodis JF, Ciarleglio L, Craffey A. 1994. Efficacy of the biparietal diameter/femur length ratio to detect Down syndrome in patients with an abnormal biochemical screen. *Fetal Diagn Ther* **9**: 175–182.

Cuckle H, Wald N, Quinn J, Royston P, Butler L. 1989. Ultrasound fetal femur length measurement in the screening for Down's syndrome. *Br J Obstet Gynaecol* **96**: 1373–1378.

Deren O, Mahoney MJ, Copel JA, Bahado-Singh RO. 1998. Subtle ultrasonographic anomalies: do they improve the Down syndrome detection rate? *Am J Obstet Gynecol* **178**: 441–445.

Dicke JM, Gray DL, Songster GS, Crane JP. 1989. Fetal biometry as a screening tool for the detection of chromosomally abnormal pregnancies. *Obstet Gynecol* **74**: 726–729.

Ginsberg N, Cadkin A, Pergament E, Verlinsky Y. 1990. Ultrasonographic detection of the second-trimester fetus with trisomy 18 and trisomy 21. *Am J Obstet Gynecol* **163**: 1186–1190.

Grandjean H, Sarramon MF. 1995. Femur/foot length ratio for detection of Down syndrome: result of a multicenter prospective study. *Am J Obstet Gynecol* **173**: 16–19.

Hadlock FP, Harrist RB, Martinez-Poyer J. 1992. Fetal body ratios in second trimester, a useful tool for identifying chromosomal abnormalities? *J Ultrasound Med* **11**: 81–85.

Hill LM, Guzick D, Belfar HL, Hixson J, Rivello D, Rusnak J. 1989. The current role of sonography in the detection of Down syndrome. *Obstet Gynecol* **74**: 620–623.

Johnson MP, Michaelson JE, Barr M, Treadwell MC, Hume RF, Dombrowski MP. 1995. Combining humerus and femur length for improved ultrasonographic identification of pregnancies at increased risk for trisomy 21. *Am J Obstet Gynecol* **172**: 1229–1235.

LaFollette L, Filly RA, Anderson R, Golbus MS. 1989. Fetal femur length to detect trisomy 21. *J Ultrasound Med* **8**: 657–660.

Lockwood C, Benacerraf B, Krinsky A, Blakemore K, Belanger K, Mahoney M. 1987. A sonographic screening method for Down syndrome. *Am J Obstet Gynecol* **157**: 803–808.

Lockwood C, Lynch L, Ghidini A, Lapinski R, Berkowitz G, Thayer B. 1993. The effect of fetal gender on the prediction of Down syndrome by means of maternal serum α-fetoprotein and ultrasonographic parameters. *Am J Obstet Gynecol* **169**: 1190–1197.

Lynch J, Berkowitz GS, Chitkara U, Wilkins IA, Mehalek KE, Berkowitz RL. 1989. Ultrasound detection of Down syndrome: is it really possible? *Obstet Gynecol* **73**: 267–270.

Marquette GP, Boucher M, Desrochers M, Dallaire L. 1990. Screening for trisomy 21 with ultrasonographic determination of biparietal diameter/femur length ratio. *Am J Obstet Gynecol* **163**: 1604–1605.

Nadel AS, Bromley B, Frigoletto FD, Benacerraf BR. 1995. Can the presumed risk of autosomal trisomy be decreased in fetuses of older women following a normal sonogram? *J Ultrasound Med* **14**: 297–302.

Nyberg DA, Resta RG, Hickok DE, Hollenback KA, Luthy DA, Mahony BS. 1990. Femur length shortening in the detection of Down syndrome. Is prenatal screening feasible? *Am J Obstet Gynecol* **162**: 1247–1252.

Nyberg DA, Resta RG, Luthy DA, Durlin EH, Williams MA. 1993. Humerus and femur length shortening in detection of Down's syndrome. *Am J Obstet Gynecol* **168**: 534–538.

Nyberg DA, Luthy DA, Resta RG, Nyberg BC, Williams MA. 1998. Age-adjusted ultrasound risk assessment for fetal Down's syndrome during the second trimester: description of the method and analysis of 142 cases. *Ultrasound Obstet Gynecol* **12**: 8–14.

Perrella R, Duerinckx AJ, Grant E, Tesler F, Tabsh K, Crandall BF. 1988. Second-trimester sonographic diagnosis of Down syndrome: role of femur-length shortening and nuchal-fold thickening. *AJR* **151**: 981–985.

Peters MT, Lockwood CJ, Miller WA. 1989. The efficacy of fetal sonographic biometry in Down syndrome screening. *Am J Obstet Gynecol* **161**: 297–300.

Platt LD, Medearis AL, Carlson DE, Falk RE, DeVore GR, Horenstein J. 1992. Screening for Down syndrome with the femur length/biparietal diameter ratio: a new twist of the data. *Am J Obstet Gynecol* **167**: 124–128.

Rodis JF, Vintzileos AM, Fleming AD, Ciarleglio L, Nardi A, Scorza WE. 1991. Comparison of humerus length with femur length in fetuses with Down syndrome. *Am J Obstet Gynecol* **165**: 1051–1056.

Rotmensch S, Luo JS, Liberati M, Belanger K, Mahoney MJ, Hobbins JC. 1992. Fetal humeral length to detect Down syndrome. *Am J Obstet Gynecol* **166**: 1330–1334.

Shah YG, Eckl CJ, Stinson SK, Woods JR. 1990. Biparietal diameter/femur length ratio, cephalic index, and femur length

measurements, not reliable screening techniques for Down syndrome. *Obstet Gynecol* **75**: 186–188.

Verdin SM, Economides DL. 1998. The role of ultrasonographic markers for trisomy 21 in women with positive serum biochemistry. *Br J Obstet Gynaecol* **105**: 63–67.

Vintzileos AM, Egan JFX. 1995. Adjusting the risk for trisomy 21 on the basis of second trimester ultrasonography. *Am J Obstet Gynecol* **172**: 837–844.

Vintzileos AM, Egan JF, Smullian JC, Campbell WA, Guzman E, Rodis JF. 1996a. Adjusting the risk for trisomy 21 by a simple ultrasound method using fetal long-bone biometry. *Obstet Gynecol* **87**: 953–958.

Vintzileos AM, Campbell WA, Rodis JF, Guzman ER, Smulian JC, Knuppel RA. 1996b. The use of second-trimester genetic sonogram in guiding clinical management of patients at increased risk for fetal trisomy 21. *Obstet Gynecol* **87**: 948–952.

Vintzileos AM, Campbell WA, Guzman ER, Smulian JC, McLean DA. 1997. Second-trimester ultrasound markers for detection of trisomy 21: Which markers are best? *Obstet Gynecol* **89**: 941–944.

Winston YE, Horger EO. 1988. Down syndrome and femur length. *Am J Obstet Gynecol* **159**: 1018.