1-1-2009

Reevaluating humeral length for the detection of fetal trisomy 21

Diana L. Gray
Washington University School of Medicine in St. Louis

Jeffrey M. Dicke
Washington University School of Medicine in St. Louis

Rachel Dickerson
Washington University School of Medicine in St. Louis

Carolyn McCourt
Washington University School of Medicine in St. Louis

Anthony O. Odibo
Washington University School of Medicine in St. Louis

Reevaluating Humeral Length for the Detection of Fetal Trisomy 21

Diana L. Gray, MD, Jeffrey M. Dicke, MD, Rachel Dickerson, MD, Carolyn McCourt, MD, Anthony O. Odibo, MD, MSCE

Objective. The purpose of this study was to analyze humeral length (HL) in a normal population and to compare that with HL in a population of fetuses with trisomy 21 to determine the most efficient discriminating parameters for diagnostic accuracy. ***Methods.*** A nested case-control study comparing HLs from a normal population and a population of fetuses with trisomy 21 was conducted. Humeral length was regressed against gestational age for a consecutive well-dated population of normal singleton gestations presenting to the Washington University School of Medicine prenatal diagnosis units over a 5-year period. A second population of well-dated pregnancies with trisomy 21 diagnosed either prenatally or postnatally was also selected on the basis of the same criteria, except that anomalous fetuses were included. Various discriminating thresholds for a short HL were compared for efficiency in the detection of trisomy 21. These included the following: observed/expected HL (≤0.89), biparietal diameter/HL greater than 1.3, 1.4, 1.5, 1.6, 1.7, 1.75, 1.8, and 1.85 SD above the mean for gestation, HL less than 0.8 or less than 0.9 multiple of the median, and HL less than the fifth percentile for gestation. ***Results.*** A total of 620 normal pregnancies and 32 with trisomy 21 were extracted from the database. A receiver operating characteristic curve revealed HL less than the fifth percentile as the optimal discriminator for trisomy 21 detection (area under the receiver operating characteristic curve = 0.80). The positive likelihood ratio (LR+) was greatest (25.0) with HL less than the fifth percentile. When HL was considered in isolation without other sonographic markers of trisomy 21, the LR+ was 6.3. ***Conclusions.*** Humeral length less than the fifth percentile was the most effective discriminator among the many studied. ***Key words:*** Down syndrome; fetal trisomy 21; humeral length; prenatal diagnosis; sonography.

Abbreviations
BPD, biparietal diameter; FL, femur length; FPR, false-positive rate; GA, gestational age; HL, humeral length; LMP, last menstrual period; LR, likelihood ratio; MoM, multiple of the median

Received April 16, 2009, from the Division of Maternal-Fetal Medicine and Ultrasound, Department of Obstetrics and Gynecology, Washington University School of Medicine, St Louis, Missouri USA. Revision requested May 7, 2009. Revised manuscript accepted for publication June 23, 2009.

Address correspondence to Diana L. Gray, MD, Department of Obstetrics and Gynecology, Washington University School of Medicine, Campus Box 8091, 660 S Euclid Ave, St Louis, MO 63110 USA.
E-mail: grayd@wustl.edu

Shortened humeri and femora are well-recognized sonographic markers for fetal trisomy 21.[1–8] Most investigations of humeral length (HL) in the detection of fetal trisomy 21 have shown superior performance of this marker compared to shortened femur length (FL). The likelihood ratios (LR) cited by Nyberg et al[9] were 5.1 for a shortened humerus compared with 1.5 for a shortened femur when used as isolated sonographic markers for trisomy 21. A study by Bromley et al[4] yielded an LR of 5.8 for a shortened humerus compared with 1.2 for a shortened femur. Various discriminating thresholds for a shortened humerus have been shown as effective markers for fetal trisomy 21, including observed to predicted HL,[2,4] multiples of the median (MoMs),[6] and HL less than the fifth percentile for gestation.[1] However, another study by Vergani et al[10] that included a meta-analysis of studies on HL in the detection of fetal trisomy 21 concluded that the utility of this marker was inconsistent. This group concluded that only institutions with

locally generated regression equations and documented predictive ability of this marker should be using it as a screening test for trisomy 21. The objective of the investigation described herein was to analyze HL in a normal population of patients at our center and to compare that with a population of fetuses with trisomy 21 to determine the most efficient discriminating parameters for diagnostic accuracy.

Materials and Methods

This was a nested case-control study comparing fetal HLs from a normal population with those from a population of fetuses with trisomy 21. The study was approved by the Washington University School of Medicine Institutional Review Board for Human Studies. Humeral length was measured for a consecutive, well-dated, normal population of 1068 singleton gestations presenting to the Washington University School of Medicine (WUSM) prenatal diagnosis units over a 5-year period from May 26, 1994, through August 6, 1999. These pregnancies were selected from the division's computerized database on the basis of the following inclusion criteria: singleton gestations dated by a known first day of the last menstrual period (LMP) with no oral contraceptive use within the 2 months before conception and positive urine pregnancy test results within 5 weeks of the LMP date, a known conception date, or a first-trimester crown-rump length consistent with 5 days of the expected gestation by LMP. Cases were excluded if the pregnancy ended in abortion or stillbirth or if a fetal anomaly was diagnosed before or after birth. These cases were excluded because it is well known that pregnancies destined for spontaneous abortion or stillbirth very frequently have abnormal growth patterns. All fetal structural anomalies identified in our prenatal sonography units, from anencephaly to minor dysmorphisms, are prospectively coded and entered into the database. If a fetal anomaly code was entered for a particular pregnancy, that pregnancy was excluded from the group used to generate the normal HL growth curve. Intrauterine growth restriction is also coded as an anomaly; thus, such pregnancies were excluded. A single sonographic examination in which HL was measured was selected for those pregnancies in which multiple examinations occurred; the process of selection involved selecting HL measurements from weeks in which the data to generate the normal curve were limited, ensuring that no overlap with a previously included pregnancy occurred. From the same period and thus the same overall population, a second group of well-dated pregnancies with trisomy 21 diagnosed either prenatally or postnatally was also selected on the basis of the same criteria, except that structurally anomalous fetuses were allowed in this study group. There were no significant differences in demographic characteristics between the two populations studied (Table 1). The HLs of the two populations were then compared, and various discriminating thresholds were studied.

Sonographic measurements in both populations were made with similar real-time state-of-the-art equipment, including Acuson 128XP (Siemens Medical Solutions, Mountain View, CA), Sonoline Elegra (Siemens Medical Solutions, Malvern, PA), Ultramark 8 (Philips Healthcare, Bothell, WA), and a GE 3000 (GE Healthcare, Milwaukee, WI). The HL was measured in the standardized fashion inclusive of the diaphyseal portion of the bone. The measurements were made by experienced sonographers or sonologists who were unaware of the fetal diagnosis or karyotype at the time of measurement.

Using least squares analysis, HL was regressed against gestational age (GA) for the normal population studied. Discriminating thresholds for a shortened HL were compared for efficiency in the detection of fetal trisomy 21 using receiver operating characteristic curves. These thresholds included the following: observed/expected HL (≤0.89), biparietal diameter (BPD)/HL greater

Table 1. Demographic Characteristics of the Study Populations

Characteristic	Trisomy 21 (n = 32)	No Trisomy 21 (n = 620)	*P*
Mean age ± SD, y	35.7 ± 5.3	36.6 ± 4.2	.37
Race, n (%)			
White	26 (84)	451 (73)	
Black	1 (3)	54 (9)	
Hispanic	1 (3)	10 (2)	.59
Asian	1 (3)	4 (1)	
Other	3 (7)	91 (15)	
Mean GA ± SD at scanning, wk	20.4 ± 5.6	18.5 ± 1.8	.06

than 1.3, 1.4, 1.5, 1.6, 1.7, 1.75, 1.8, and 1.85 SD above the mean for gestation, HL less than 0.8 or 0.9 MoM, and HL less than the fifth percentile. Statistical methods, including χ^2 and *t* tests, were used for categorical and continuous variables, respectively. All analyses were performed using Stata 9.0 SE (StataCorp, College Station TX).

Results

For the normal population, 620 pregnancies in the database met inclusion criteria. These pregnancies were well distributed from 12 to 40 weeks' gestation. The regression equation for HL versus GA with the best fit for the normal population was the following: $HL = -3.527 + 0.417\ (GA) - 0.004\ (GA)^2$. The fit was excellent, as evidenced by the R^2 value, which was 0.97 (Figure 1). Thirty-two fetuses with trisomy 21 met inclusion criteria and constituted the trisomy 21 population. There were no significant differences in demographic characteristics between the two populations studied (Table 1). Receiver operating characteristic analysis revealed that HL less than the fifth percentile was the most efficient discriminator (Figure 2). Mean and fifth percentile HLs by gestation are presented in Table 2.

When short HL was tested for effectiveness in the detection of fetal trisomy 21 in our population, it was found to be quite effective. The efficiency parameters for the most discriminating HL thresholds are found in Table 3. Humeral length less than the fifth percentile for GA was found to be the most efficient discriminator, with 12 of 32 fetuses with trisomy 21 detected, for sensitivity of 37.5%, a positive screen rate (false-positive rate [FPR]) of 1.5%, a positive predictive value of 57% (1 per 1.75), and a positive LR (LR+) of 25.0. The negative LR (LR–) was 0.6. In descending order, the next most effective thresholds were observed/expected HL of 0.89 or less, with an LR+ of 15.5; HL less than 0.9 MoM, with an LR+ of 9.3; and BPD/HL greater than 1.7 or greater than 1.75, both with LR+ values of 3.4. The less effective discriminating thresholds included all of the other variations of BPD/HL (>1.3–1.85) and HL less than 0.8 MoM, with LR+ values ranging from 0 to 2.9. All screening efficiency parameters for these thresholds are listed in Table 4.

When HL less than the fifth percentile was evaluated as an isolated finding in the population of fetuses with trisomy 21, 3 of 32 (9.4%) were found to have this finding in the absence of other sonographic markers of trisomy 21. In the normal population, 9 of 620 (1.5%) had a shortened HL. Therefore, as an isolated sonographic marker of fetal trisomy 21, the LR+ for HL less than the fifth percentile for gestation was 6.3.

Discussion

Comparing shortened HL less than the fifth percentile as a sonographic marker of fetal trisomy 21 to a previous study of shortened FL from our institution by Dicke et al,[5] we found that HL less

Figure 1. Humeral length regression curve: $HL = -3.527 + 0.417\ (GA) - 0.004\ (GA)^2$. CI indicates confidence interval.

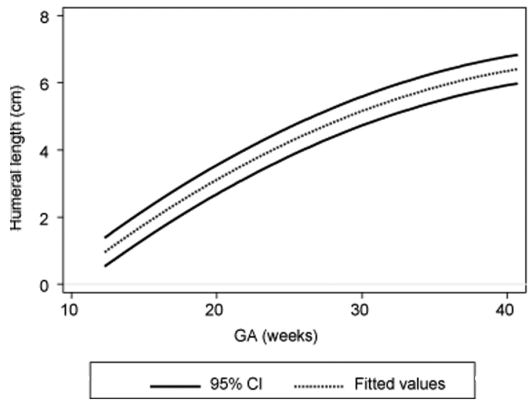
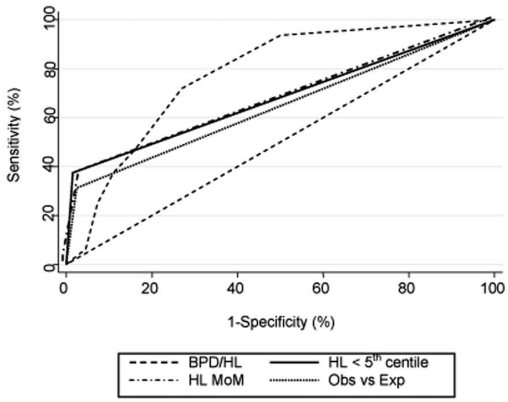


Figure 2. Receiver operating characteristic curve comparing BPD/HL greater than 1.7, 1.75, and greater than 1.8 with fifth percentile, less than 0.9 MoM, and abnormal observed (Obs)/expected (Exp) HL (≤0.89).



than the fifth percentile performed better as a predictor of fetal trisomy 21. The sensitivity of HL less than the fifth percentile of 37.5%, FPR of 1.5%, positive predictive value of 1 per 1.75, and LR+ of 25.0 were substantially better than the parameters for a shortened femur using BPD/FL greater than 1.5 SD (sensitivity of 18%, FPR of 4%, positive predictive value of 1 per 169, and LR+ of 4.5). The low FPR of 1.5% indicates that the nomogram we generated for HL was very strict in the definition of defining a shortened femur. This is not a limitation of the study because it should reduce the anxiety and further testing provoked by screening tests with higher FPRs. In our previous study of FL and trisomy 21, the fetuses with a shortened FL as the only finding were not segregated from the population of fetuses with trisomy 21 and multiple sonographic findings. Hence, the performance of HL as an isolated marker for the detection of trisomy 21 cannot be compared to FL in isolation.

Comparisons with previous studies of shortened HL can be found in Table 5. Rodis et al[1] published the first study of the effectiveness of HL as a marker for fetal trisomy 21 in 1991. This investigation also used HL less than the fifth percentile as the discriminating threshold in a small population of fetuses with trisomy 21 (n = 11) with a resulting LR+ of 12.8 and an FPR of 5%. Our study was second in size for the number of fetuses with trisomy 21 to a 1993 study published by Nyberg et al,[2] which had 45 fetuses with trisomy 21 and yielded sensitivity of 24.4%, an FPR of 4.5%, and an LR+ of 5.4. The LR+ values of 25.0 when not controlled for other markers of trisomy 21 and 6.3 when found in isolation for this study compare favorably with those from the other studies listed in Table 5. One possible limitation of the

Table 2. Humeral Length at 5th and 50th Percentiles by Gestation

GA, wk	HL, cm	
	5th	50th
15	1.4	1.8
16	1.7	2.1
17	2.0	2.3
18	2.4	2.6
19	2.6	3.0
20	2.8	3.2
21	3.1	3.4
22	3.2	3.6
23	3.5	3.9
24	3.8	4.1
25	4.0	4.4
26	4.3	4.6
27	4.6	4.8
28	4.7	4.9
29	4.8	5.0
30	4.9	5.1
31	5.0	5.2
32	5.2	5.3
33	5.4	5.6
34	5.5	5.7
35	5.6	5.9
36	5.7	6.1
37	5.8	6.2
38	6.2	6.3
39	6.3	6.4
40	6.4	6.5

Table 3. Most Effective Discriminating Thresholds

Parameter	Sens, % (95% CI) (Tri 21 = 32)	Spec, % (95% CI) (Normal = 620)	FPR, %	PPV, % (95% CI)	NPV, % (95% CI)	LR+	LR–
HL <5th	37.5 (21–56)	98.5 (97–99)	1.5	57 (34–78)	97 (95–98)	25.0	0.6
HL <5th (isolated)	9.4 (2–25)	98.5 (97–99)	1.5	25 (6–57)	95.5 (94–97)	6.3	0.9
HL <0.9 MoM	37.5 (21–56)	96 (94–98)	4.0	34 (19–52)	97 (95–98)	9.3	0.7
BPD/HL >1.7	47 (29–65)	86 (83–89)	14.0	15 (9–23)	97 (95–98)	3.4	0.6
BPD/HL >1.75	34 (19–53)	90 (87–92)	10.0	15 (8–26)	96 (95–98)	3.4	0.7
Obs/Exp HL ≤0.89	31 (16–50)	98 (96–99)	2.0	40 (21–61)	97 (95–98)	15.5	0.7

CI indicates confidence interval; NPV, negative predictive value; Obs/Exp, observed/expected; PPV, positive predictive value; Sens = sensitivity; Spec = specificity; and Tri 21, trisomy 21.

Table 4. Least Effective Discriminating Thresholds

Parameter	Sens, % (95% CI) (Tri 21 = 32)	Spec, % (95% CI) (Normal = 620)	FPR, %	PPV, % (95% CI)	NPV, % (95% CI)	LR+	LR–
BPD/HL >1.3	100 (89–100)	1.0 (0.36–2.1)	99	5.0 (3.4–6.9)	100 (54–100)	1.0	0
BPD/HL >1.4	100 (89–100)	9.2 (7.0–11.7)	90.8	5.4 (3.7–7.5)	100 (94–100)	1.1	0
BPD/HL >1.5	97 (84–99.9)	37.9 (34–41.9)	62.1	7.5 (5.1–10.4)	99.6 (97.7–100)	1.6	0.1
BPD/HL >1.6	84.4 (67.2–94.7)	64.8 (60.9–68.6)	35.2	11.0 (7.4–15.6)	98.8 (97.2–99.6)	2.4	0.2
BPD/HL >1.8	18.8 (7.2–36.4)	93.5 (91.3–95.4)	6.5	13.0 (4.9–26.3)	95.7 (93.8–97.2)	2.9	0.9
BPD/HL >1.85	3.1 (0.0–16.2)	96.6 (94.9–97.9)	3.4	4.5 (0.1–22.8)	95.1 (93.1–96.6)	0.9	1.0
HL <0.8 MoM	3.1 (0.0–16.2)	100 (99.4–100)	0	100 (2.5–100)	95.2 (93.3–96.7)	0	1.0

CI indicates confidence interval; NPV, negative predictive value; PPV, positive predictive value; Sens = sensitivity; Spec = specificity; and Tri 21, trisomy 21.

study was that most of the HL measurements were made in the second trimester between 14 and 22 weeks. The regression model included fetuses at all weeks' gestation, however. The second-trimester time frame is that in which most screening for fetal trisomy 21 occurs; therefore, the findings are reliable. As a result of this study, we will incorporate the use of HL shortening in the screening protocol for trisomy 21 when appropriate in our prenatal diagnosis units. It will be used in our clinical practice in addition to other sonographic markers for trisomy 21, such as a thickened nuchal skin fold, by factoring the HL LR+ into the a priori risk for trisomy 21. The a priori risk may have been generated on the basis of maternal age, serum screening, or first-trimester screening. As did Vergani et al,[10] we urge caution when using biometric markers as screening tools for fetal syndromes. The nomograms and abnormal thresholds for biometric markers are most accurate if they have been generated in the local center where they will be used or systematically verified as effective in that population. Hence, as is the case whenever biometric parameters are used to screen for fetal disorders, the applicability of these possible markers for general use may be limited. Such markers are best applied in a tertiary level prenatal diagnosis center that has studied their use as predictors for the fetal disorders in question.

Table 5. Comparison With Previous Studies of Short HL

Reference	Discriminator	Sens, % (Tri 21 Detected/Total Tri 21)	FPR, %	LR+
Rodis et al[1]	<5th	64 (7/11)	5	12.8
Nyberg et al[2]	Obs/Exp ≤0.89	24.4 (11/45)	4.5	5.4
Bahado-Singh et al[6]	Obs/Exp <0.92	55 (6/11)	9	6.1
Vintzileos et al[7]	Obs/Exp <0.89	48 (10/21)	2	12.0
Benacerraf et al[11]	Obs/Exp <0.9	50 (12/24)	6.25	8.0
This study	<5th	37.5 (12/32)	1.5	25.0
		9.4 (3/32) (isolated)	1.5	6.3

Obs/Exp indicates observed/expected; Sens, sensitivity; and Tri 21, trisomy 21.

References

1. Rodis JF, Vintzileos AM, Fleming AD, et al. Comparison of humerus length with femur length in fetuses with Down syndrome. Am J Obstet Gynecol 1991; 165:1051–1056.
2. Nyberg DA, Resta RG, Luthy DA, Hickok DE, Williams MA. Humerus and femur length shortening in the detection of Down's syndrome. Am J Obstet Gynecol 1993; 168:534–538.
3. Nyberg DA, Luthy DA, Resta RG, Nyberg BC, Williams MA. 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 1998; 12:8–14.
4. Bromley B, Lieberman E, Shipp TD, Benacerraf BR. The genetic sonogram: a method of risk assessment for Down syndrome in the second trimester. J Ultrasound Med 2002; 21:1087–1096.
5. Dicke JM, Gray DL, Songster GS, Crane JP. Fetal biometry as a screening tool for the detection of chromosomally abnormal pregnancies. Obstet Gynecol 1989; 74:726–729.
6. Bahado-Singh RO, Oz AU, Kovanci E, et al. New Down syndrome screening algorithm: ultrasonographic biometry and multiple serum markers combined with maternal age. Am J Obstet Gynecol 1998; 179:1627–1631.
7. Vintzileos AM, Guzman ER, Smulian JC, McLean DA, Ananth CV. Choice of second-trimester genetic sonogram for detection of trisomy 21. Obstet Gynecol 1997; 90:187–190.
8. Lockwood C, Benacerraf B, Krinsky A, et al. A sonographic screening method for Down syndrome. Am J Obstet Gynecol 1987; 157:803–808.
9. Nyberg DA, Souter VL, El-Bastawissi A, Young S, Luthhardt F, Luthy DA. Isolated sonographic markers for detection of fetal Down syndrome in the second trimester of pregnancy. J Ultrasound Med 2001; 20:1053–1063.
10. Vergani P, Locatelli A, Ghidini A, et al. Caveats for the use of humerus length in the prediction of fetal Down syndrome. Fetal Diagn Ther 2003; 18:190–195.
11. Benacerraf BR, Neuberg D, Frigoletto FD Jr. Humeral shortening in second-trimester fetuses with Down syndrome. Obstet Gynecol 1991; 77:223–227.