

New crown–rump length curve based on over 3500 pregnancies

A. PEXSTERS*, A. DAEMEN†, C. BOTTOMLEY‡, D. VAN SCHOUBROECK*, L. DE CATTE*, B. DE MOOR†, T. D’HOOGHE*, C. LEES§, D. TIMMERMAN* and T. BOURNE*¶

*Department of Obstetrics and Gynecology, University Hospitals and †Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, Leuven, Belgium, ‡Department of Obstetrics and Gynaecology, Chelsea and Westminster Hospital and ¶Imperial College London, Hammersmith Campus, London and §Division of Fetal–Maternal Medicine, Addenbrookes Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

KEYWORDS: crown–rump length; menstrual age; vaginal ultrasound

ABSTRACT

Objectives The Robinson and Hadlock crown–rump length (CRL) curves are commonly used to estimate gestational age (GA) based on the CRL of an embryo or fetus. However, the Robinson curve was derived from a small population using transabdominal sonography and the Hadlock curve was generated using early transvaginal ultrasound equipment. The aim of this study was to use transvaginal and transabdominal ultrasound to study a large population of early pregnancies to assess embryonic or fetal size, and so create a new normal CRL curve from 5.5 weeks’ gestation. We compared this with the Robinson and Hadlock CRL curves.

Methods A retrospective database study of CRL in first-trimester embryos was conducted in a fetal medicine referral center with a predominantly Caucasian population. Linear mixed-effects analysis was performed to determine the relationship between CRL and GA. After internal validation of this curve, the CRL was compared with the expected CRL at a given GA according to both the Robinson and Hadlock models based on the paired t-test. Bland–Altman plots were constructed to compare the CRL measurements obtained in our study population with those predicted according to GA by both the Robinson and Hadlock curves.

Results In total 3710 normal singleton pregnancies with a known last menstrual period were included in the study, corresponding to 4387 scans. Our data differed significantly from both the Robinson and the Hadlock curves (paired t-test, $P < 0.0001$). A mixed-effects model for CRL as a function of GA was developed on 70% of the data and internally validated with z-scores on the

remaining 30%. The new curve extended from 5.5 to 14 weeks’ gestation. Compared to our CRL curve, the Robinson curve gave a 4-day underestimation of GA at 6 weeks with a difference in CRL of 3.7 mm and a 1-day overestimation from 11 to 14 weeks with a difference in CRL of 0.9–1 mm. A comparison between our curve and the Hadlock curve showed a difference in CRL of 2.7 mm at 6 weeks, equivalent to an underestimation of 3 days, and a difference in CRL of 4.8 mm at 14 weeks, equivalent to an overestimation of 2 days. At 9 weeks all three curves were similar.

Conclusion The new CRL curve suggests differences in the range of CRL measurements compared with the Robinson and Hadlock curves. These differences are most significant at the beginning and the end of the first trimester, and may lead to more accurate estimations of GA. Copyright © 2010 ISUOG. Published by John Wiley & Sons, Ltd.

INTRODUCTION

In clinical obstetrics, gestational age (GA) is estimated as the time from the first day of the last menstruation onwards, assuming a menstrual cycle of 28 days¹. Ultrasound measurements of embryonic and fetal crown–rump length (CRL) are used in the early stages of pregnancy to estimate GA. Between 5.5 and 14 weeks’ gestation this approach has been described in studies using either static image scanners or transabdominal sonography². Early studies using the transvaginal route have produced similar results at very early gestations^{3–6}.

Correspondence to: Dr A. Pexsters, Department of Obstetrics and Gynecology, University Hospitals, Katholieke Universiteit Leuven, Herestraat 49, 3000 Leuven, Belgium (e-mail: anne.pexsters@uz.kuleuven.ac.be)

Accepted: 12 October 2009

The most commonly used method for predicting GA on the basis of a CRL measurement is the classic Robinson curve, which was derived from 214 transabdominal scans carried out 35 years ago on 80 women with known dates of the last menstrual period (LMP)^{7,8}. A limitation of this study was that Robinson included few data before 7 weeks' and after 13 weeks' gestation. In addition, MacGregor *et al.*⁹ studied women with known dates of conception following infertility treatment and suggested that there was a generalized underestimation of GA when using the Robinson curve.

Nearly 30 years ago, Hadlock *et al.*³ extended the CRL curve using data from 416 patients and starting from an embryonic length of 2 mm. In their review they concluded that their findings were in general agreement with the original Robinson model up to 12 weeks' gestation.

Owing to the methodological and equipment differences between the studies of Hadlock *et al.* and Robinson performed more than 25 years ago and today's practice, a critical re-evaluation of CRL in relation to GA is overdue. The aim of our study was to examine data derived from a large population of pregnancies at between 5.5 and 14 weeks' gestation to reassess the relationship between embryonic or fetal size and GA, thereby developing a 'new' normal range for CRL. We then compared this approach to both the Robinson and Hadlock CRL curves widely used in clinical practice.

METHODS

We conducted a retrospective database study on the CRLs of embryos and fetuses at different gestations in the first trimester of pregnancy. This was carried out in a referral center for fetal medicine with a predominantly Caucasian population. We included consecutive patients who underwent a transvaginal or transabdominal ultrasound scan in the first trimester between 2002 and 2008. Only those singleton intrauterine pregnancies that were subsequently found to be viable at the time of the nuchal scan at between 11 and 14 weeks and that had at least one registered CRL measurement were included. Only women with recorded known and certain LMP dates were included in the study, based on the electronic file of each patient, which included the following prospectively completed fields: date of LMP known vs. unknown; if known, certain vs. uncertain. We excluded pregnancies that on long-term follow-up were found to have resulted in a miscarriage, stillbirth, genetic or other congenital abnormality. Other exclusion criteria were pregnancies resulting from infertility treatment and all pregnancies with an uncertain LMP.

All the women underwent an ultrasound assessment using a transabdominal (2.5–5-MHz) or transvaginal (5–8-MHz) transducer for B-mode imaging, with either an Acuson Sequoia (Siemens-Acuson Inc., Mountain View, CA, USA), Voluson 730 (GE Medical Systems, Zipf, Austria) or ESAOTE Technos (Esaote, Genova, Italy) machine. The CRL was measured by placing the caliper at the outer side of the crown and rump of the embryo

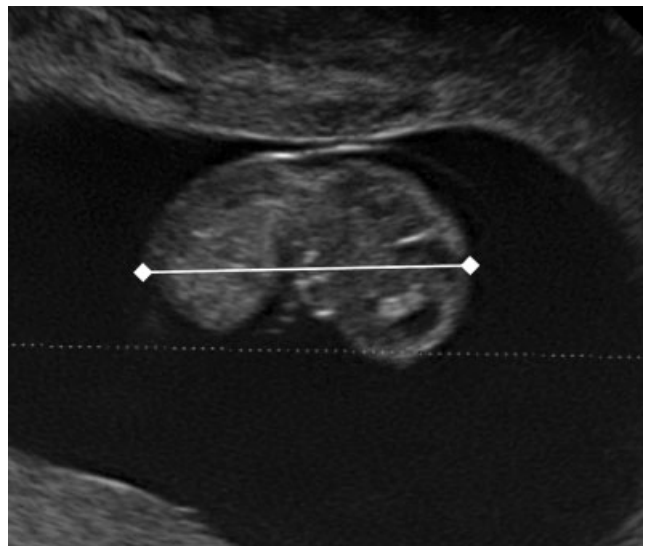


Figure 1 Embryonic crown–rump length at 6–8 weeks' gestation.



Figure 2 Fetal crown–rump length at > 8 weeks' gestation.

or fetus (greatest length) and was measured to the nearest mm (Figures 1 and 2)^{1,2}. All ultrasound assessments were carried out by gynecologists with specialist training in obstetric and gynecological sonography.

All data were recorded on a computer database (Astraia, Munich, Germany) and subsequently entered into an Excel spreadsheet for statistical analysis. Ethics committee approval was obtained at University Hospitals Leuven.

Statistical analysis

Statistical analyses were performed using SAS version 9.1 for Windows (SAS Institute Inc., Cary, NC, USA). In order to account for possible codependency of multiple measurements in the same patients, a linear mixed-effects model was used. The model was developed on a training set (70% of the pregnancies, chosen in order of ascending

hospital number) to examine the relationship between CRL and GA, with GA as an independent or explanatory variable and expanded with a polynomial term up to the power of two (GA^2) (because of evidence for a non-linear relationship between GA and CRL based on scatter plots)¹⁰. The covariance structure for the fixed effects GA and GA^2 was set to a simple structure with only the variances equal to σ^2 , while each covariance was set to zero. An exponential and Gaussian structure did not lead to an improvement in the likelihood of the model. As random effects, an intercept was included to account for within-subject variability, as well as GA and GA^2 , because growth expressed in terms of CRL is not considered to be linear¹¹. An unstructured covariance matrix was chosen for the random effects. The parameters of the model were estimated with the maximum likelihood approach. The Akaike information criterion (AIC), taking the complexity of the model into account, was calculated as a measure of goodness of fit. The curve was internally validated on the remaining 30% of pregnancies with use of the paired *t*-test. This test was also used for a comparison of all datapoints with respect to the Robinson and Hadlock curves. Because the observed CRL values had a slightly negatively skewed distribution, data were log-transformed after reflection of the distribution. $P < 0.05$ was chosen for statistical significance.

The GA included in the study ranged between 40 and 98 days. Data outside 4 SD from the expected CRL for GA according to the Robinson curve were considered as outliers.

Bland–Altman plots were constructed to compare the CRL measurements obtained in our study population with those predicted according to GA by both the Robinson and Hadlock curves¹². The 95% CIs were calculated of the differences between the observed and expected CRL according to the Robinson and the Hadlock curves, and defined as the mean difference + 2 SD and – 2 SD. These were compared on the *x*-axis to the mean of the observed and expected CRL measurements for both Robinson (Figure 5) and Hadlock (Figure 6) curves.

The 5th and 95th percentiles for the new CRL curve were calculated at each GA individually, taking the different variability in CRL across the GA range into account.

RESULTS

The initial dataset of scans in patients with known LMP contained 4698 scans from 3809 pregnancies. After exclusion of scans taken outside the GA range under consideration ($n = 37$), datapoints with a CRL outside the range mean ± 4 SD ($n = 126$), and scans with an unknown CRL ($n = 148$), the final data set contained 4387 datapoints from 3710 singleton pregnancies with one or multiple scans in early pregnancy. Of the 4387 datapoints, 3050 were derived from transabdominal scans. Of the scans carried out before 10 weeks, 96.4% were transvaginal.

The paired *t*-test was applied for the comparison of all 4387 datapoints with respect to the Robinson and

Hadlock curves. Our datapoints differed significantly from both curves ($P < 0.0001$). The mean difference after log-transformation between our data and the Robinson curve was included with 95% confidence in the interval 0.03–0.04 mm. For the Hadlock curve the 95% CI was 0.01–0.02 mm.

A linear mixed-effects model for CRL as a function of GA was developed on 70% of the data, i.e. 3064 scans from 2597 pregnancies. The correlation of CRL with GA was expressed by the equation:

$$CRL = -9.09 - (0.26 \times GA) + (0.012 GA^2)$$

with an AIC of 18 565.

Our CRL curve and the Robinson curve are shown in Figure 3.

As internal validation, the remaining 1113 pregnancies, corresponding to 1323 datapoints, were compared with respect to our CRL curve. Using a paired *t*-test did not show a significant deviation of the validation data from our curve ($P = 0.3125$, -0.003 (95% CI, -0.008 to 0.003) on a logarithmic scale). Figure 4 shows our curve with the validation data, indicating that the proposed CRL curve is applicable to new patients.

Our CRL chart was compared with the Robinson chart, and the differences in CRL for each day of gestation are shown in Table S1. At 6 weeks' gestation there was an observed difference in CRL of 3.7 mm, equivalent to an underestimation of 4 days by Robinson. From 11 to 14 weeks' gestation there was an observed difference in CRL of 0.9–1 mm, equivalent to 1 day overestimation by Robinson.

Comparison of Hadlock's curve and ours showed a difference in CRL of 2.7 mm at 6 weeks (Figure 3 and Table S2). This is equivalent to an underestimation of 3 days by Hadlock. There was also a difference in CRL of 4.8 mm at 14 weeks, equivalent to an overestimation

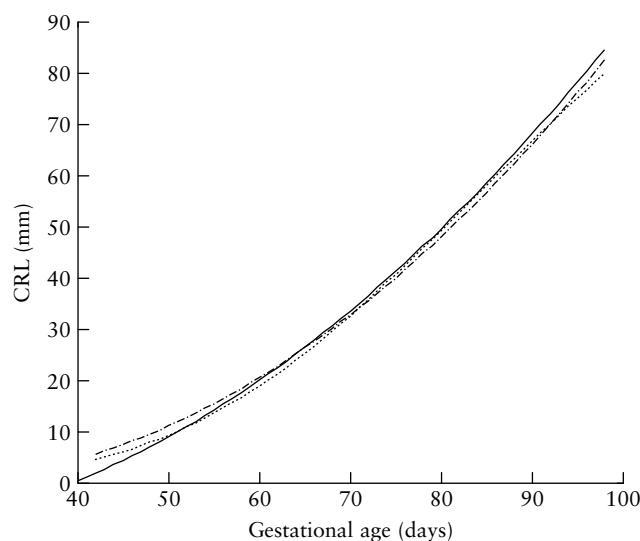


Figure 3 Comparison of our new crown–rump length (CRL) curve (—) with the Robinson⁸ (---) and Hadlock³ (.....) curves.

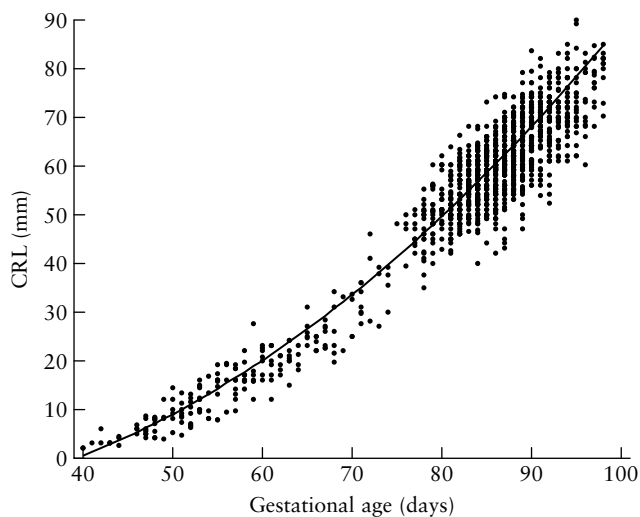


Figure 4 Our new crown–rump length (CRL) curve validated on 1113 pregnancies.

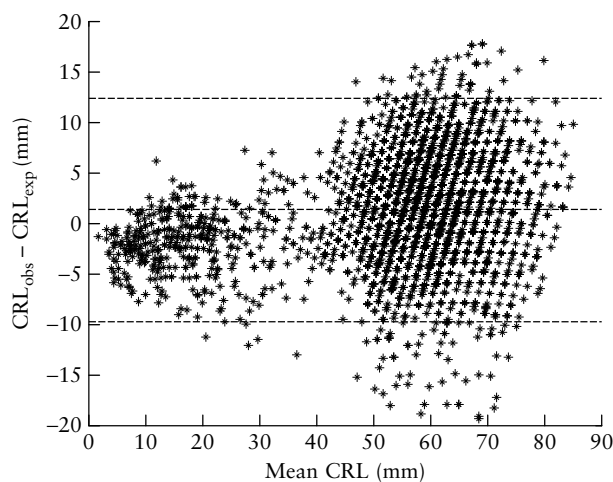


Figure 5 Bland–Altman plot: comparison of crown–rump length (CRL) observed in our study population (CRL_{obs}) with the results expected from the Robinson curve (CRL_{exp}). Lines represent mean and mean ± 2 SD.

of 2 days by Hadlock. At 9 weeks the three curves are similar (Figure 3).

In general, the observed CRL values used to construct the new CRL curve correlated well with the expected CRL values based on both the Robinson curve and the Hadlock curve ($\rho^2 = 0.95$, $P < 0.0001$). The mean difference between the observed and expected CRL was 1.3 mm for Robinson, as shown in the Bland–Altman plot (Figure 5). As the differences were normally distributed, 95% of them lay between mean -2 SD and mean $+2$ SD, equal to -9.7 mm and 12.3 mm for the Robinson data. The 95% CI for the bias in CRL value was 1.1 – 1.5 mm. The 95% CI for the mean -2 SD was -10.1 to -9.4 mm, and for the mean $+2$ SD it was 12.0 – 12.7 mm.

When compared with the Hadlock curve, the mean difference was 0.7 mm (Figure 6). The 95% CI for the bias in CRL value was 0.5 – 0.9 mm. The 95% CI for the mean -2 SD (-9.9) was -10.2 to -9.6 mm, and for the mean $+2$ SD (11.2) it was 10.9 – 11.6 mm.

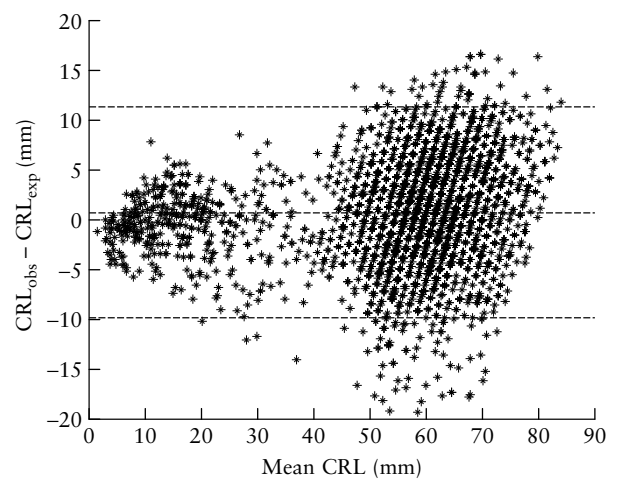


Figure 6 Bland–Altman plot: comparison of crown–rump length (CRL) observed in our study population (CRL_{obs}) with the results expected from the Hadlock curve (CRL_{exp}). Lines represent mean and mean ± 2 SD.

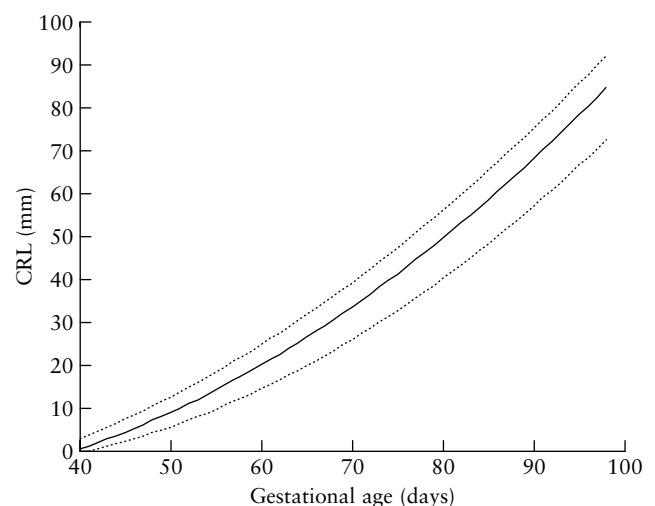


Figure 7 Our new crown–rump length (CRL) curve with 5th and 95th percentiles.

The mean CRL and 5th and 95th percentiles are given in Figure 7 and Table S3 for the new CRL curve. The 5th and 95th percentiles were calculated at each specific GA. Those percentiles were not symmetrically distributed, and the specific shape depended on the specific GA. In the lower GA range, the distribution was more positively skewed, in the higher GA range negatively skewed, while the distribution approximated a symmetrical distribution in the middle range.

DISCUSSION

We have developed a new CRL normal range on a large number of pregnancies to define the size of the embryo and fetus in early pregnancy. The new normal range is different from the commonly used Robinson curve and suggests that both the Robinson and Hadlock curves significantly underestimate size at early gestations. This

difference in early gestations is even more pronounced in our study based on pregnancies after spontaneous conception when compared with the difference observed in a study based on pregnancies with known dates of conception following infertility treatment⁹. Moreover, the new normal range overestimates size, with respect to the old curves, after 9 weeks' gestation, with a difference that is more pronounced with the Robinson curve than with the Hadlock curve. All ranges agree on size at around 9 weeks' gestation. The relationship between the datapoints used to derive the new CRL curve and the Robinson and Hadlock curves can be seen most clearly in the Bland–Altman plots. The Bland–Altman plot is the most appropriate way of comparing the differences for a known gestational age between the current 'gold standard' curves of expected CRL (those of Hadlock and Robinson), and the observed CRL measurements in our population.

Strengths of this study are that it is based on a large number of pregnancies with reported certain menstrual dates and that the data are derived from a relatively homogeneous patient population; such an extensive evaluation has not been published in the literature to date. However, we acknowledge that our study has limitations. The study is retrospective and even with certain menstrual dates there might be some uncertainty over the dating of a pregnancy because of cycle irregularity. Moreover, it has been shown that GA estimates based on certain menstrual dates can be invalid as a result of the incidence of delayed ovulation¹³. Although validated internally, our own reference curve needs to be validated externally by testing it prospectively on a different study population to assess its reproducibility. Another limitation of our study is that we do not know if our ultrasound assessment before 7 weeks' gestation is reproducible. We therefore acknowledge that a prospective study on intra- and interobserver variability of first-trimester estimates with modern ultrasound equipment would be appropriate, and have started such a study in our unit. However, a recent report on intra- and interobserver variability of early fetal growth parameters, including CRL, overall shows good reproducibility of these measurements between 9 and 14 weeks¹⁴.

The possibility exists that differences in study populations may explain the differences between our results and those of Robinson and Hadlock. Ethnic- and age-related variations in growth rates in early pregnancy have been described, and there may be other confounding factors that influence the developing embryo^{15,16}. However, as far as we can tell our predominantly Caucasian population was similar to those recruited in the studies by both Robinson and Hadlock. Future consideration may have to be given to the development of individualized growth curves in early pregnancy that reflect the demographic characteristics of a population attending any particular clinic. Otherwise, inaccurate dating may lead to abnormalities of growth being unrecognized later in pregnancy and to inappropriate intervention.

Using CRL measurements to define GA based on the Robinson reference curve has been widely used for years,

and now in the UK forms the basis of assigning the due date of a pregnancy¹⁷. The method was said to have a low interobserver variability in an early report using static image scanners¹⁸. However, the majority of ultrasound scans performed in early pregnancy are now carried out transvaginally and possible differences using this approach have not been extensively described. In 1992 Hadlock *et al.*³ used 416 women to build their model, using relatively early transvaginal ultrasound as well as transabdominal sonography. Their data were broadly in agreement with those of Robinson.

Between 11 and 14 weeks' gestation we found a 1-day overestimation by Robinson compared with our data. However for very early pregnancies – before 6 weeks – the observation that there is a 4-day difference or more may be of clinical importance. The low number of patients in Robinson's data set, the relative paucity of measurements at early gestation, and the use of mostly static-image ultrasound equipment might explain this discrepancy. Our findings are in disagreement with Hadlock's data for pregnancies before 8 weeks. An important issue to consider in this context is the accuracy of CRL measurements. Until it reaches a length of 4 mm, the embryo is straight and a measurement of the CRL is a true measurement of the greatest length². Between 4 and 18–22 mm, the greatest length of the embryo is the neck–rump length, owing to curvature of the embryo. It is likely that measurement differences may occur at this stage¹¹. As ultrasound equipment has technically improved over recent years, the images that can now be obtained may provide a more precise definition of the embryo. This improvement in image resolution could explain the differences in CRL between our study and Hadlock's. Nevertheless, our comparative analysis of the curves shows that a possible 2-day difference due to measurement errors has to be taken into account.

The greatest disparity between our curve and both Robinson's and Hadlock's is seen before 8 weeks' gestation. This can be most clearly seen from Figure 3. We found that for measurements of CRL below 20 mm, the majority of observed measurements are considerably lower than would be expected using either the Robinson or the Hadlock CRL curve. This is a clear and unambiguous finding. It is likely, given the sample size and the use of predominantly modern transvaginal ultrasound equipment, that our curve is more accurate at these relatively early gestations. This is relevant as it is becoming evident that discrepancies in embryonic size at early gestations may be associated with both short- and long-term adverse outcomes^{19–21}. Furthermore, increasing numbers of women are attending for a scan earlier in pregnancy to confirm viability. Accurate dating is important for optimal timing of first-trimester screening and delivery. While a difference of several days may not seem clinically important in normal pregnancies, it becomes relevant for clinical decision making at the extremes of viability at around 24 weeks' gestation, and when determining the appropriate time for post-term induction of labor.

The new CRL range for normal pregnancies described in this study probably gives a more accurate pregnancy assessment at earlier gestation owing to the far greater sample size than those used by Robinson and Hadlock. The range described may improve the accuracy of decisions based on GA in those pregnancies dated at an early gestation.

ACKNOWLEDGMENTS

A.D. is research assistant of the Fund for Scientific Research - Flanders (FWO-Vlaanderen). We are grateful to Pushpika Thilakarathne for his statistical contribution. This work is partially supported by (1) Research Council KUL: GOA AMBioRICS, CoE EF/05/007 SymBioSys, PROMETA, several PhD/postdoc & fellow grants; (2) Flemish Government: FWO: PhD/postdoc grants, projects G.0241.04 (Functional Genomics), G.0499.04 (Statistics), G.0232.05 (Cardiovascular), G.0318.05 (sub-functionalization), G.0553.06 (VitamineD), G.0302.07 (SVM/Kernel), research communities (ICCoS, ANMMM, MLDM); IWT: PhD Grants, GBOU-McKnow-E (Knowledge management algorithms), GBOU-ANA (biosensors), TAD-BioScope-IT, Silicos; SBO-BioFrame, SBO-MoKa, TBM-Endometriosis, TBM-IOTA3, O&O-Dsquare; (3) Belgian Federal Science Policy Office: IUAP P6/25 (BioMaGNet, Bioinformatics and Modeling: from Genomes to Networks, 2007-2011); and (4) EU-RTD: ERNSI: European Research Network on System Identification; FP6-NoE Biopattern; FP6-IP e-Tumours, FP6-MC-EST Bioptrain, FP6-STREP Strokemap.

REFERENCES

- Jurkovic D, Gruboeck K, Campbell S. Ultrasound features of normal early pregnancy development. *Curr Opin Obstet Gynecol* 1995; 7: 493–504.
- Bottomley C, Bourne T. Dating and growth in the first trimester. *Best Pract Res Clin Obstet Gynaecol* 2009; 23: 439–452.
- Hadlock FP, Shah YP, Kanon OJ, Lindsey JV. Fetal crown-rump length: reevaluation of relation to menstrual age (5–18 weeks) with high-resolution real-time ultrasound. *Radiology* 1992; 182: 501–505.
- Drumm JE, Clinch J, MacKenzie G. The ultrasonic measurement of fetal crown-rump length as a method of assessing gestational age. *Br J Obstet Gynaecol* 1976; 83: 417–421.
- Silva PD, Mahairas G, Schaper AM, Schauburger CW. Early crown-rump length. A good predictor of gestational age. *J Reprod Med* 1990; 35: 641–644.
- Kustermann A, Zorzoli A, Spagnolo D, Nicolini U. Transvaginal sonography for fetal measurement in early pregnancy. *Br J Obstet Gynaecol* 1992; 99: 38–42.
- Robinson HP. Sonar measurement of fetal crown-rump length as means of assessing maturity in first trimester of pregnancy. *BMJ* 1973; 4: 28–31.
- Robinson HP, Fleming JEE. A critical evaluation of sonar “crown-rump length” measurements. *Br J Obstet Gynaecol* 1975; 82: 702–710.
- MacGregor SN, Tamura RK, Sabbagha RE, Minogue JP, Hoffman DI. Underestimation of gestational age by conventional crown-rump length dating curves. *Obstet Gynecol* 1987; 70: 344–348.
- Verbeke G and Molenberghs G. *Linear Mixed Models for Longitudinal Data*. Springer-Verlag, Inc.: New York, 2000.
- Deter RL, Buster JE, Casson PR, Carson SA. Individual growth patterns in the first trimester: evidence for difference in embryonic and fetal growth rates. *Ultrasound Obstet Gynecol* 1999; 13: 90–98.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310.
- Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatr Perinat Epidemiol* 2007; 21 (Suppl 2) 86–96.
- Verburg BO, Mulder PG, Hofman A, Jaddoe VW, Witteman JC, Steegers EA. Intra- and interobserver reproducibility study of early fetal growth parameters. *Prenat Diagn* 2008; 28: 323–331.
- Jacquemyn Y, Sys SU, Verdonk P. Fetal biometry in different ethnic groups. *Early Hum Dev* 2000; 57: 1–13.
- Bottomley C, Daemen A, Mukri F, Papageorghiou AT, Kirk E, Pexsters A, De Moor B, Timmerman D, Bourne T. Assessing first trimester growth: the influence of ethnic background and maternal age. *Hum Rep* 2009; 24: 284–290.
- <http://www.nice.org.uk/nicemedia/pdf/CG62FullGuidelineCorrectedJune2008.pdf> (accessed 22nd February 2009).
- Daya S. Accuracy of gestational age estimation by means of fetal crown-rump length measurements. *Am J Obstet Gynecol* 1993; 168: 903–908.
- Mukri F, Bourne T, Bottomley C, Schoeb C, Kirk E, Papageorghiou AT. Evidence of early first-trimester growth restriction in pregnancies that subsequently end in miscarriage. *BJOG* 2008; 115: 1273–1278.
- Bukowski R, Smith GC, Malone FD, Ball RH, Nyberg DA, Comstock CH, Hankins GD, Berkowitz RL, Gross SJ, Dugoff L, Craigo SD, Timor-Tritsch IE, Carr SR, Wolfe HM, D’Alton ME; FASTER Research Consortium. Fetal growth in early pregnancy and risk of delivering low birth weight infant: prospective cohort study. *BMJ* 2007; 334: 836.
- Smith GC. First trimester origins of fetal growth impairment. *Semin Perinatol* 2004; 28: 41–50.

SUPPORTING INFORMATION ON THE INTERNET

Table S1 Crown-rump length (CRL) values for the Robinson curve⁸ (considered as standard curve) and the new CRL curve: gestational age-dependent prediction of CRL

Table S2 Crown-rump length (CRL) values for the Hadlock curve³ (considered as standard curve) and the new CRL curve: gestational age-dependent prediction of CRL

Table S3 Mean crown-rump length values according to our derived curve with 5th and 95th percentiles