American Journal of Obstetrics and Gynecology (2004) **190**, 1077–81

ELSEVIER

AMERICAN JOURNAL of OBSTETRICS and GYNECOLOGY

www.elsevier.com/locate/ajog

GENERAL OBSTETRICS AND GYNECOLOGY: OBSTETRICS

First trimester ultrasound screening is effective in reducing postterm labor induction rates: A randomized controlled trial

Kelly A. Bennett, MD,[a,*] Joan M. G. Crane, MD,[b] Patrick O'Shea, MD,[b] Joanne Lacelle, MD,[b] Donna Hutchens, BN,[b] Joshua A. Copel, MD[c]

Department of Obstetrics and Gynecology, Vanderbilt University, Nashville, Tenn[a], the Department of Obstetrics and Gynecology, Memorial University of Newfoundland, St. John's, Newfoundland, Canada[b], and the Department of Obstetrics and Gynecology, Yale University School of Medicine, New Haven, Conn[c]

Received for publication May 23, 2003; revised September 23, 2003; accepted September 30, 2003

KEY WORDS
Ultrasound
Postterm
Induction

Objective: This study was designed to test the null hypothesis that first trimester ultrasound crown-rump length measurement for gestational age determination will result in no difference in the rate of induction of labor for postterm pregnancy, compared with second trimester biometry alone.

Study design: Two hundred eighteen women were randomly assigned to receive either first trimester ultrasound screening or second trimester ultrasound screening to establish the expected date of confinement. Sample size was calculated by using a 2-tailed $\alpha = .05$ and power $(1\text{-}\beta) = 80\%$. Data were analyzed with χ^2 and Fisher exact tests.

Results: Of 104 women randomly assigned to the first trimester screening group, 41.3% had their gestational age adjusted on the basis of the crown-rump length measurement. Of 92 women randomly assigned to the second trimester screening group, 10.9% were corrected as a result of biometry ($P < .001$, relative risk = 0.26, 95% CI = 0.15-0.46). Five women in the first trimester screening group and 12 women in the second trimester screening group had labor induced for postterm pregnancy ($P = 0.04$, relative risk = 0.37, 95% CI = 0.14-0.96).

Conclusion: The application of a program of first trimester ultrasound screening to a low-risk obstetric population results in a significant reduction in the rate of labor induction for postterm pregnancy.

The estimated date of confinement (EDC) as calculated by menstrual age is based on the assumption that pregnancy lasts 280 days from the first day of the last menstrual period (LMP).[1] It is now widely recognized that this approach is laden with potential error because it is based on accurate recall of the first day of the LMP and it assumes that ovulation occurs on the 14th day of

Presented at the Twenty-third Annual Meeting of the Society of Maternal Fetal Medicine, Oral Concurrent Session C, San Francisco, Calif, February 3-8, 2003. (Awarded Best Paper).

* Reprint requests: Kelly A. Bennett, MD, Department of Obstetrics and Gynecology, Vanderbilt University, B1100 Medical Center North, Nashville, TN 37212.

E-mail: kelly.a.bennett@vanderbilt.edu

0002-9378/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ajog.2003.09.065

the menstrual cycle. It fails to account for the highly variable duration of the follicular phase of the menstrual cycle, which can range from 7 to 21 days.[2,3] This results in systematic error inherent in ignoring the true length of gestation, which is the interval between the time of fertilization of the ovum and the time of delivery of the fetus.

Studies of basal body temperature[4,5] and second trimester ultrasound biometry[6] indicate that menstrual dating is often inaccurate and can lead to induction of labor for a presumed but inaccurate diagnosis of postterm pregnancy. It is now widely accepted that a more precise method of determining gestational age (GA) is the first trimester measurement of the crown-rump length (CRL) of the fetus. By using this measurement, an estimate of the EDC can be determined with an error of $\pm$4.7 days.[7] Systematic overestimation of GA can be problematic because it can result in unnecessary induction, dysfunctional labor, and cesarean delivery.[8]

The primary objective of this randomized controlled trial was to determine whether application of a program of routine first trimester ultrasound screening (to precisely identify GA) to a general, low-risk obstetric population would result in a decrease in the rate of induction of labor for postterm pregnancy. The overall prevalence of diagnosis of postterm pregnancy and the prevalence of induction of labor for any indication were considered secondary outcomes.

Methods

This study was conducted at the Health Care Corporation of St. John's, St. John's, Newfoundland, Canada, from December 31, 1999, to April 11, 2002. This tertiary referral center for perinatal care, which serves an almost entirely white population of 500,000, is the site of 2,300 of the province's 5,000 annual births. Induction of labor for postterm pregnancy is typically considered after 41 weeks' GA, in keeping with the guidelines of the Society of Obstetricians and Gynaecologists of Canada (SOGC).[9] The research proposal was approved by the Human Investigation Committee of the Faculty of Medicine, Memorial University of Newfoundland.

Women between the ages of 16 and 40 years, in the first trimester of pregnancy who presented to the center during the study period were eligible for inclusion. Patients who presented with any one of the following were considered ineligible: women with an indication for or having had a prior first trimester ultrasound, having a known multiple gestation, being unwilling to participate, or being younger than 19 years with no guardian present.

All women in the first trimester of pregnancy who attended prenatal visits with one of 15 participating family physicians or one of 4 participating obstetricians were informed of the study and given written information about the protocol. Women were enrolled in the study by their attending physician who explained the protocol and its potential risks and benefits before obtaining written consent. Each woman was then randomly assigned to receive either first trimester ultrasound screening to establish EDC by using CRL measurement, or second trimester ultrasound screening to establish EDC with the use of biometry.

Opaque envelopes each containing a card indicating group allocation were prepared by an administrative staff member who used computer-generated random number tables. Allocation of each consenting woman to her ultrasound screening group was determined by opening the next sequentially numbered envelope. Those women randomly assigned to the first trimester screening group were scheduled for an initial ultrasound and a clinical pelvic examination between 8 and 12 weeks' gestation, as determined by LMP-derived dates.

All first trimester ultrasounds were performed by one of two maternal fetal medicine specialists with the ATL HDI 5000 ultrasound system (Phillips Medical Systems, Markham, Ontario, Canada) equipped with a 9-5 MHz transvaginal probe. Each woman was evaluated for the presence or absence of a gestational sac, a yolk sac and a fetal pole with detectable cardiac activity, and a CRL measurement was made.

If the estimate of EDC derived by the date of the LMP differed by 5 days or more from that derived by ultrasound measurement of CRL, the GA and EDC were revised to reflect the estimates on the basis of the ultrasound assessment. All referring physicians received written notification informing them of any change in the estimate of GA and the EDC.

Second trimester ultrasound examination is a standard part of prenatal care at the study center. Women randomly assigned to the second trimester screening group were managed according to usual care, being scheduled for second trimester ultrasounds at 19 weeks' gestation on the basis of LMP dating and pelvic examination.[10] All second trimester ultrasounds were performed by two radiologists not involved in the study using the ATL HDI 3000 and 5000 ultrasound systems (Phillips Medical Systems, Markham, Ontario, Canada) equipped with 4-2 and 5-2 MHz abdominal probes.

If the estimates of the EDC calculated from the LMP and from ultrasound biometry measurements differed by 10 days or more, the GA and EDC were revised according to the ultrasound-derived estimate. The duration of pregnancy was calculated by using ultrasound dating criteria from the corrected estimate of the first day of the LMP to the date of delivery of the newborn infant. If a study participant was undelivered at 41 weeks' gestation (greater than 287 days), the woman was referred to the maternal fetal medicine unit and offered labor induction.

Table I Demographic characteristics

Characteristic	First trimester screening (n = 104)	Second trimester screening (n = 92)
Maternal age	29.3 (±4.2)	29.1 (±4.1)
Race		
White	103 (99.0%)	92 (100%)
Other	1 (1.0%)	0 (0%)
Parity		
Nulliparous	51 (49.0%)	47 (51.1%)
Multiparous	53 (51.0%)	45 (48.9%)
Previous postterm birth	13 (12.5%)	10 (10.8%)
Previous labor induction	16 (15.3%)	16 (17.3%)
Previous preterm birth	2 (1.9%)	3 (3.3%)
Previous cesarean delivery	11 (10.5%)	7 (7.6%)

Data presented as mean (±SD) or n (%).

Decisions regarding method of induction were made at the discretion of the maternal fetal medicine specialist and in keeping with the patient's wishes. Women who were determined to have an unripe cervix received misoprostol or dinoprostone for cervical ripening, whereas those with a favorable cervix underwent amniotomy, followed by oxytocin augmentation. Decisions pertaining to labor management were made by the attending physician.

Sample size was calculated on the basis of the primary outcome, rate of induction of labor for postterm pregnancy. An intended sample size (n = 272) was calculated assuming a 50% reduction in the rate of postterm labor induction, a 2-tailed $\alpha = .05$ and $\beta = .20$ (PEPI, Version 2, 1995, Computer Programs for Epidemiologic Analysis, Epidemiologic Analysis, Salt Lake City, Utah). Recruitment to the study was stopped when the provincial task force studying a maternal serum screening program recommended first trimester ultrasound dating scans for all women interested in maternal serum screening.

Data were analyzed according to the intent-to-treat principle by parametric and nonparametric statistics, using statistical analysis software (SPSS 10.0, Analytical Software, Tallahassee, Fla). Decision levels and hypotheses to be tested were determined at the outset to minimize bias. Hypothesis testing was performed on the primary outcome measure. Statistical significance of the observed difference in the primary outcome measure was assessed with χ^2 and Fisher exact tests, using a significance threshold of $P < .05$. The significance threshold for all secondary and hypothesis-generating analyses was set at $P < .001$ to minimize biases incurred by multiple testing.

Results

Of 218 women with singleton pregnancies, 9 women randomly assigned to the first trimester screening group and 10 women randomly assigned to the second trimester screening group had pregnancy loss in the first trimester. One woman in the first trimester screening group was diagnosed with an ectopic pregnancy. Two participants in the second trimester screening group were lost to follow-up after emigration from the province. Consequently, 22 patients generated no available outcome data and could not be considered in the analysis. Analysis included 104 women in the first trimester screening group and 92 women in the second trimester screening group. Maternal preinduction demographic data are presented in Table I.

Table II Reasons for induction

Indication	First trimester screening (n = 104)	Second trimester screening (n = 92)	P	RR	95% CI
Postterm	**5**	**12**	**.04**	**0.37**	**0.14-0.96***
Preeclampsia	2	6	.10	0.29	0.07-1.28†
Oligohydramnios	1	3	.25	0.29	0.03-2.42†
Diabetes	0	1	.63	1.76	0.17-18.50†
Premature membrane rupture	1	1	.93	0.88	0.05-13.90†
Term membrane rupture	6	8	.43	0.66	0.24-1.86*
Growth restriction	1	1	.93	0.88	0.05-13.90†
Other	5	2	.32	2.21	0.46-2.21†

* χ^2 test.
† Fisher exact test.

Of 104 women, 43 (41.3%) in the first trimester screening group had their GA and EDC adjusted on the basis of the CRL measurement. In all the women in the first trimester screening group, the CRL-corrected EDC estimate was confirmed by second trimester biometry. Second trimester screening was performed in this group for the purpose of performing an anomaly screen. Only 10 of 92 women (10.9%) randomly assigned to the second trimester screening group required adjustment of their GA and EDC according to the results of second trimester biometry ($P < .001$, relative risk [RR] = 0.26, 95% CI = 0.15-0.46).

With respect to the primary outcome measure, a significant difference in postterm pregnancy induction rates was noted between the two groups. Five women (4.8%) in the first trimester ultrasound screening group underwent labor induction for postterm pregnancy, compared with 12 women (13.0%) in the second trimester screening group ($P = .04$, RR = 0.37, 95% CI = 0.14-0.96). All patients offered induction for postterm pregnancy consented to scheduled labor induction. Additional data regarding specific indications for induction of labor are presented in Table II.

A difference in the prevalence of the diagnosis of postterm pregnancy was noted between groups. In the first

trimester screening group, 7 of 104 women (6.7%) delivered at a GA of 287 days or greater, compared with 15 of 92 (16.3%) in the second trimester screening group ($P = .03$, RR = 0.41, 95% CI = 0.18-0.94). No significant difference in birth route was observed between the two groups. In the first trimester screening group, there were 73 vaginal deliveries and 31 cesarean births (11 elective repeat). In the second trimester screening group, there were 71 vaginal deliveries and 21 cesarean births (7 elective repeat). No significant differences in neonatal outcomes were observed between the two groups.

Comment

The true prevalence of pregnancies lasting longer than 41 or 42 weeks' gestation has been difficult to establish because many such diagnoses occur as a result of an inability to accurately establish time of conception. Saito et al[4] studied prolonged pregnancy and time of ovulation using basal body temperature profiles in 129 Japanese women reporting that delay of ovulation was the major contributing cause of apparent prolongation of pregnancy beyond 295 days. Boyce et al[5] confirmed these findings in a study of basal body temperature profiles of 317 French women, reporting that 68% of women who completed 42 postmenstrual weeks had a less-advanced GA when calculated on the basis of their ovulation date.

Kramer et al[6] compared the estimate of GA based on the LMP to estimates determined by measurement of the biparietal diameter from second trimester ultrasound in 11,045 women, reporting that the positive predictive value of the diagnosis of postterm pregnancy based on the mothers' LMP was only 12%. Pregnancy has traditionally been considered postterm at 294 days from the first day of the LMP, with a reported frequency ranging from 3% to 15%.[9] The frequency of pregnancies completing 41 weeks' gestation (calculated as 287 days from the LMP) may be as high as 27%.[11,12]

First trimester ultrasound dating is accurate within days ($\pm$4-5 days), but does not allow for anatomically detailed examination, because of the fetus's small size and the early stage of development of the anatomic structures of interest. An additional benefit is that it does allow for early diagnosis of missed abortion, ectopic pregnancy and multiple gestations. Second trimester ultrasound is only accurate in dating a pregnancy within weeks ($\pm$7-14 days), but it does allow for diagnosis of congenital anomalies.

Bukowski et al[12] studied 3588 women undergoing first trimester ultrasound as part of the multicenter First and Second Trimester Evaluation for Aneuploidy Trial. Gestational age determination using the CRL as opposed to LMP significantly affected the proportion of pregnancies considered greater than 41 weeks (8.2% vs 22.1%, $P < .001$, RR = 0.37, 95% CI = 0.33-0.40, RR reduction [RRR] 61%). These results compare well with our own. The proportion of pregnancies considered greater than 41 weeks was decreased in the first trimester screening group compared with the second trimester screening group (6.7% vs 16.3%, $P = .03$, RR = 0.41, 95% CI = 0.18-0.94, RRR 63 %.)

A search of MEDLINE using the search terms "crown rump length" and "labor induction" from 1970 to 2002 revealed no published studies that addressed the question of first trimester ultrasound measurement of CRL as a strategy to reduce labor induction rates. Three large randomized trials have studied the effect of second trimester ultrasound biometry on induction rates as secondary outcomes.[13-15] All 3 demonstrated a decreased rate of labor induction after 42 weeks' gestation in the ultrasound screening group.

Recruitment to this study was stopped before the predetermined sample size was reached when the provincial task force studying a maternal serum screening program recommended first trimester ultrasound dating scans for all women interested in maternal serum screening. Despite this, the study had sufficient power to detect a significant difference in the primary outcome. The results of this trial suggest that dating by CRL measurement in the first trimester is superior to second trimester biometry alone in providing an accurate estimation of the expected date of confinement. The application of a program of routine first trimester ultrasound to a general, low-risk obstetric population reduces systematic error in estimating GA and significantly reduces postterm labor induction rates.

References

1. Treloar AE, Behn BG, Cowan DW. Analysis of gestational interval. Am J Obstet Gynecol 1967;99:34-45.
2. Munster K, Schmidt L, Helm P. Length and variation in the menstrual cycle—a cross sectional study from a Danish country. BJOG 1992;99:422-9.
3. Rawlings EE, Moore BA. The accuracy of methods of calculating the expected date of delivery for use in the diagnosis of postmaturity. Am J Obstet Gynecol 1970;106:676-9.
4. Saito M, Keijiro Y, Akinori H, Takahiro K, Nozumu N, Kohei K. Time of ovulation and prolonged pregnancy. Am J Obstet Gynecol 1972;112:31-8.
5. Boyce A, Mayaux M, Schwartz. Classical and true gestational postmaturity. Am J Obstet Gynecol 1976;125:911-3.
6. Kramer MS, McLean F, Boyd ME, Usher RH. The validity of gestational age estimation by menstrual dating in term, preterm, and postterm gestations. JAMA 1988;260:3306-8.
7. Robinson HP, Fleming JEE. A critical evaluation of sonar CRL measurements. BJOG 1975;82:702.
8. Boyd ME, Usher RH, McLean FH, Kramer MS. Obstetric consequences of postmaturity. Am J Obstet Gynecol 1988;158:334-8.
9. Society of Obstetricians and Gynaecologists of Canada. Clinical practice guidelines. Committee Opinion. Guidelines for Postterm Pregnancy. London, Ontario (Canada): Rodgers Publishing; 1997. No. 15.

10. Society of Obstetricians and Gynaecologists of Canada. Clinical practice guidelines. Committee Opinion. Guidelines for ultrasound as part of routine prenatal care. London, Ontario (Canada): Rodgers Publishing; 1999. No.78.
11. Beischer NA, Evans JH, Townshed L. Studies in prolonged pregnancy: the incidence of prolonged pregnancy. Am J Obstet Gynecol 1969;103:476-82.
12. Bukowski R, Saade G, Malone F, Hankins G, D'Alton M. A decrease in postdate pregnancies is an additional benefit of first trimester screening for aneuploidy. Am J Obstet Gynecol 2001;185(suppl):S148.
13. Waldenstrom U, Axelsson O, Nilsson S, Gunnar E, Fall O, Lindeberg S, et al. Effects of routine one-stage ultrasound screening in pregnancy: a randomized controlled trial. Lancet 1988;2:585-8.
14. Saari-Kemppainen A, Karjalainen O, Ylostalo P, Heinonen OP. Ultrasound screening and perinatal mortality: controlled trial of systematic one-stage screening in pregnancy. The Helsinki Ultrasound Trial. Lancet 1990;336:387-91.
15. Ewigman BG, Crane JP, Frigoletto FD, LeFevre ML, Bain RP, McNellis D. Effect of prenatal ultrasound screening on perinatal outcome. N Engl J Med 1993;329:821-7.