Published in final edited form as:

Contraception. 2009 September ; 80(3): 287–291. doi:10.1016/j.contraception.2009.03.008.

Unexpected heaping in reported gestational age for women undergoing medical abortion

Irving Sivin^{1,*}, James Trussell^{2,3}, E Steve Lichtenberg⁴, Mary Fjerstad⁵, Kelly Cleland³, and Vanessa Cullins⁵

¹Reproductive Health Program, Population Council, New York NY 10065, USA ²The Hull York Medical School, University of Hull, Hull HU6 1RX, England ³Office of Population Research, Princeton University, Princeton NJ 08544, USA ⁴Northwestern University Feinberg School of Medicine, Chicago IL 60611, USA ⁵Planned Parenthood Federation of America, New York NY 10001, USA

Abstract

Background—Planned Parenthood Federation of America (Planned Parenthood) conducted an extensive audit in August 2006 of first trimester medical abortion with oral mifepristone plus buccal misoprostol through 56 days of gestation so that patients could be given accurate information about the success rate of the new regimen.

Objectives—We sought to evaluate the effectiveness of this buccal misoprostol regimen and examine correlates of its success during routine service delivery.

Methods—Audits at 10 large urban service points were conducted in 2006 to estimate success rates of the buccal regimen. Success was defined as medical abortion without vacuum aspiration.

Results—We discovered unexpected heaping of reported gestational age on days divisible by 7.

Conclusion—Such heaping, which has not been reported in the literature, would make it more difficult to detect a modest trend in declining effectiveness with increasing gestational age if there were one. High coefficients of variation of sac size and crown-rump length characterize the early gestational weeks. We suspect, but are unable to prove, that the source of the heaping found in our investigation is a tendency for operators of ultrasound machines at some sites to simplify reporting by rounding a portion of the results to a date corresponding to the nearest complete gestational week. We believe that immediate supervisory awareness and feedback may reduce the extent of the problem. However, the problem may persist in multiple-site studies given the underlying variability of the ultrasound measurements with differently calibrated machines and different rules for recording data, some of which may permit acceptance of an estimate based on the stated date of the last menses, if it differs by no more than 2 or 3 days from the ultrasound result.

Keywords

Gestational age; heaping; abortion trials

© 2013 Elsevier Inc. All rights reserved.

*Corresponding author: Irving Sivin, Reproductive Health Program, Population Council, New York NY 10065, USA, isivin@popcbr.rockefeller.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

From April 2006, patients choosing medical abortion at Planned Parenthood Federation of America (Planned Parenthood) health centers with gestational age ≤ 56 days were offered 200 mg of mifepristone followed 24–48 h later by 800 mcg of buccal misoprostol (administered by the woman herself, at home). The buccal misoprostol regimen requires patients to place two pills (200 mcg each) in each cheek for 30 min, and then swallow any remaining pill fragments. We conducted an evaluation of the effectiveness of this mifepristone-buccal misoprostol regimen.

2. Methods

The study design is described in more detail elsewhere [1]. Briefly, a geographically stratified sample of 10 Planned Parenthood health centers was selected to provide information on characteristics and a team including one of the authors (MF) and three experienced Planned Parenthood nurse practitioners under her supervision conducted medical records audits at 10 metropolitan sites. The study protocol and design were submitted to and approved by Allendale Investigational Review Board.

Gestational age (GA) was determined or confirmed by transvaginal ultrasound in all cases where a gestational sac was visible (those cases seen approximately >35 days after the start of the last menses). At 8 centers, the ultrasound machine provided estimates of GA in standard weeks-and-days format (e.g., 7 weeks and 3 days from the start of the LMP) and hereafter referred to as the standard format. GA was recorded in the chart in the same format. At one center, one machine provided estimates in the standard format while the other provided estimates in a range (e.g., 5–6 weeks); however, one of the authors (MF) reviewed the static images and calculated gestational age using mean sac diameter in mm or embryonic length in mm [2,3]. In the final center, some machines provided estimates in the standard format and others in day format (e.g., 52 days from the start of last menses).

In every case, one of the authors (MF) examined each ultrasound image in detail to verify that placement of the calipers was correct (more precisely, not obviously incorrect), to confirm that measurements of mean sac diameter or of embryonic length were accurate, and to recalculate the GA with the Rossavik or Goldstein formula [2,3].

Although some providers use the patient-reported first day of the LMP to estimate GA if it closely approximates GA by ultrasound, this is not standard practice at Planned Parenthood health centers, and was not the practice at any of the 10 sites that were audited. Instead, in each case, the GA was determined by ultrasound, and was recorded in the Planned Parenthood medical record. One of the authors verified that the GA documented in the medical record was correct, or, in a few cases in which a range of GA was recorded, calculated the GA into the standard format using the Rossavik or Goldstein formula. The verified or re-calculated GA was relied upon for the purposes of this analysis. Only if no sac was seen on ultrasound was GA estimated from the date of the start of the LMP.

If no gestational sac could be seen, three possibilities arose: 1) the woman was not pregnant; 2) she was pregnant, but because the pregnancy was ≤ 35 days the sac could not be visualized or 3) the woman had an ectopic pregnancy. To determine her status and eligibility for medical abortion, we proceeded as follows: a woman was deemed eligible only if she met several criteria two of which were that the GA, based on the onset of the LMP, was ≤ 35 days and a urine pregnancy test sensitive to hCG at 25 IU/L, was positive (A negative test here indicated the woman was not pregnant and not eligible for a medical abortion.) Other eligibility criteria required that a bimanual exam reveal small uterine size consistent with early pregnancy; that neither tenderness nor adnexal masses were present and a second

urinary hCG test sensitive to levels above 2000 IU/L was negative. Levels of hCG >2000 IU/L imply almost certain visualization of a gestational sac by transvaginal ultrasound of a normal pregnancy. Absence of the sac in a woman who had conceived and had hCG levels above 2000 IU/L would point to a work-up for ectopic pregnancy) If the patient met these eligibility conditions, a quantitative serum beta hCG was drawn, the patient ingested mifepristone at the clinic and was given misoprostol to take home. Signs and symptoms of ectopic pregnancy were discussed, and she returned two days after taking misoprostol. A history was then taken to assess whether she had experienced bleeding, and a second quantitative beta hCG test was drawn. If the most recent beta hCG dropped by 50% or more the pregnancy was deemed to have aborted [4,5]. In these cases, the GA was estimated as the date of examination minus the onset date of the LMP.

If on presentation a patient's estimated GA was >35 days based on a combination of ultrasound findings, urine hCG tests and bimanual pelvic exam as described above, and no sac was visualized, she was deemed ineligible for medical abortion and was not included in our dataset. She received a workup for ectopic pregnancy instead.

Descriptive statistics of key sample characteristics were computed. Univariate and multivariate logistic regression analyses were used to examine whether individual or clinic characteristics were related to the success rates (defined as no surgical intervention required) of the abortion procedure. The Kolmogorov-Smirnoff test was used to test for the uniform distribution of gestational age in days. All calculations were preformed in SAS (SAS Institute Inc, Cary North Carolina) or Cytel Studio 7 (Cytel Inc, Boston Massachusetts).

Although the standard format for reporting GA is in weeks and days (so that weeks begin on days divisible by 7), the conventional way of grouping days into weeks for statistical analysis is for weeks to end on days divisible by 7; we hereafter use the term conventional gestational week to refer to this grouping.

3. Results

It is evident in Fig. 1 that there is heaping on gestational ages divisible by 7, ages that mark the end of a conventional gestational week (e.g., days 28, 35, 42, 49, and 56) rather than other days. Further to this point, 25.3% of cases are reported with these gestational ages marking the end of a gestational week. The heaping is so pronounced that a formal statistical test is hardly needed. Nevertheless, we did test to see whether the distribution is uniform across the seven days in a gestational week. To ensure the same number of possible days of the week, we limited the analysis to days 28–55; the p-value is < 0.0001. We obtained the same result when we limited the analysis to days 28–48 and days 28–41. The last case produces a very conservative test, since as can be seen in Fig. 1, the caseload is rising over that range, so other things equal, we would expect to find a deficit of days divisible by 7 since these are the *first* days of a gestational week (in these analyses); in fact, there is a surplus of first days (20.6%). One might object to including the week beginning with day 28 because gestational age cannot be estimated well by ultrasound. So we repeated the analyses for days 36–56 and days 43–56. In both cases the p-value was <0.0001. The last case produces a very conservative test, since as can be seen in Fig. 1, the caseload is declining over that range, so other things equal, we would expect to find a deficit of days divisible by 7 since these are the *last* days of a conventional gestational week; in fact, there is a surplus of last days (17.5%).

One possible source of this heaping is that there is a tendency to round to the nearest gestational age in weeks; in this case, 7 weeks and 3 days would be reported as 7 weeks, while 7 weeks and 5 days would be reported as 8 weeks. Fig. 2 invites that interpretation. If

this hypothesis were true, then mean reported gestational age in days would equal the true mean, since rounding down (days 1, 2 and 3) and rounding up (days 4, 5 and 6) would cancel each other exactly if gestational ages overall were uniform in distribution. However, it is also possible that instead there is a tendency for gestational age to be reported in completed weeks (the days are dropped); in this case, all fractions of 7 weeks, even 7 weeks and 6 days, would be reported as 7 weeks. Fig. 3 invites the interpretation that there is a tendency to report completed weeks only. In this case, mean reported gestational age in days would be biased downward.

The overall success rate of medical abortion using buccal misoprostol was 98.3% (Table 1). Logistic regression results showed no association between success and gestational age in days (Table 1). We explored further whether there was a trend toward decreased effectiveness with increased gestational age by examining effectiveness by gestational age in weeks. If reported gestational age is rounded to the nearest completed week, with days 1, 2 and 3 being reported as 7 weeks while days 4, 5 and 6 are reported as 8 weeks, then Fig. 1 suggests that a more accurate classification by week would be obtained if weeks were *centered* on days divisible by 7 than if weeks are classified in the conventional way, *ending* in days divisible by 7. If reported gestational age is reported in completed weeks (days are dropped), then Figure 2 suggests that a more accurate classification by week would be obtained if weeks *begin* on days divisible by 7 than if weeks are classified in the conventional way, *ending* in days divisible by 7. Results for all three classifications are shown in Table 1. P-values for a linear trend when analyzed by week are greatly reduced in the classifications when weeks are *centered* on days divisible by 7, indicating the more likely existence of a true trend. Indeed, when weeks are *centered* on days divisible by 7, effectiveness statistically significantly decreases with increasing gestation ($p=0.045$).

Another possible source of bias is that clinicians may have wished to extend service to patients whose pregnancies were a few days beyond the allowable limit of 56 days of gestation by adjusting the gestational age downward and reporting it as 56 days. If this were true, heaping would have been evident at the 56-day limit of gestation. However, the distribution of days when the maximum gestational age is 53 days yields a pattern virtually identical to those in Fig. 1 and 2 (results not shown). Consequently there is no evidence of more heaping on day 56, the maximum gestational age for which the buccal regimen was permitted than on earlier days divisible by 7; indeed, when weeks are grouped in the conventional way *ending* on days divisible by 7, only 14.7% of cases in week 8 are on day 56, as would be expected if there were no heaping artifact created to include women over the gestational age limit of 56 days.

When we examined heaping at only the 8 centers in which the ultrasound machines provided estimates in the standard weeks-and-days format and gestational age was recorded in the chart using that standard format, the pattern was similar, but the heaping on days divisible by 7 was somewhat blunted (20.9% of cases versus 25.3% for all centers); nevertheless, the distribution was decidedly not uniform ($p<0.0001$).

4. Discussion

We found clear evidence of heaping of reported gestational age to integers divisible by 7. We initially identified three hypotheses for the source of the observed heaping. The first is that the ultrasound machines themselves produce the heaping by virtue of a tendency of their embedded software programs to round off gestational ages to integers. The second is that clinicians sometimes do not record verbatim in the charts the detailed estimates (in weeks plus days or total gestational days) provided by the machines; on such occasions, clinicians either round to the nearest week or record only completed weeks (by dropping the days), or

both. The third is that the source is neither the machines nor the clinicians but the investigators who *extracted* the data from the charts. Of course, the source could be a combination of the three possibilities.

We find the second and third of these explanations implausible: one of the authors (MF) and her collaborators conducted a review of all charts and verified that all data were transferred verbatim from ultrasound images to patient charts and thence to the logs used to transmit these data to the investigators. The first hypothesis poses that the ultrasound machines produced a systematic artifact. However, when we excluded data from ultrasound machines that did not report gestational ages in the standard weeks-and-days format, the heaping on days divisible by 7 (20.9% of cases) remained. How could the machines produce heaping if they report data in weeks and days, or alternately in total gestational days? Thus, this notion seems far-fetched.

Perhaps the answer is the human element involved in taking the image. The machine produces results in accordance with where the sonographer places calipers that measure sac size or the longitudinal axis of the somite (early embryo). Any technician (or nurse or mid-level practitioner or doctor) performing the original ultrasound in the clinic can affect the outcome by placing the two ultrasound calipers a little closer or a little farther from each other, depending on how the sonographer does the measurement and whether s/he has a bias, whether conscious or not. The bias could be a desire to qualify a 59-day gestational age embryo for a procedure permissible to women only up to 56 days, though we found no evidence for this. The more likely bias is a tendency toward simplicity when it does not affect health center protocols. Under this hypothesis, the sonographer may (consciously or not) adjust the calipers to produce a simple result like 7 weeks plus 0 days instead of the more complicated result of 7 weeks and 1 day. There is nothing devious in this approximation and nothing unethical; it is merely a desire to simplify and expedite the processing of the patient (it matters not at all to the patient whether she is 7 weeks plus 0 days or 7 weeks plus 1 day). Although not the case in our Planned Parenthood data, another variation of this rounding-off artifact could result when a health center has a formal or informal policy of calculating gestational age by using the obstetrical concept of “best obstetrical estimate” to date a pregnancy. Using this technique, first day of the last menstrual period is the default for gestational age dating as long as gestational age derived from first day of last menstrual period is consistent with dating by examination and sonography; i.e., dating by last menstrual period rules if it does not differ by more than 2 or 3 days from the on-site ultrasound on that day.

Regardless of the source, such heaping, which has not been previously reported, would be expected to diminish the ability to detect a modest declining trend in success with increased gestational age if there were one. We found that for buccal administration of misoprostol, when gestational age was grouped into weeks *centered* on a day divisible by 7, there is a significant decline in success with increasing gestational age using multivariate analysis (p-value 0.045), whereas such a statistically significant trend could not be detected when gestational age was measured either in days or in weeks ending or beginning in a day divisible by 7 (Table 1). The quantitative extent of heaping in our data set, resulting in a p-value of <0.0001 in a test for a uniform distribution, makes it extremely unlikely that this phenomenon appeared solely by chance.

Many medical abortion trials have reported declines in efficacy with advancing gestational age, particularly in the week from 57 to 63 days of gestation [6–8]. Since declining efficacy with advancing gestational age is an important clinical finding, future studies should examine the reporting of gestational age at the study sites before undertaking the study. They should set rules designed to diminish the presence of gestational age heaping,

including establishment of quality control in reporting of gestational age data and prompt review of these data. If heaping is evident, adjustment for this effect by centering weeks on days divisible by 7, instead of beginning or ending them on these days may help to reveal the true magnitude of declining efficacy with advancing gestational age, if there is such a trend. Similarly, it is useful to have sample sizes of sufficient quantity to permit evaluation of the difference between the upper and middle terciles, quartiles or quintiles.

To our knowledge, no previous medical abortion efficacy trials have included testing for evidence of heaping in reported gestational age data. Heaping may mask the true effects of moderate declining efficacy of some medical abortion regimens in future studies and indeed may have done so in previous ones.

Acknowledgments

There was no external funding for this study or of writing of this manuscript.

References

1. Fjerstad M, Sivin I, Lichtenberg ES, Trussell J, Cleland K, Cullins V. Effectiveness of medical abortion with mifepristone and buccal misoprostol through 59 gestational days. *Contraception*. in press.
2. Rossavik IK, Torjusen GO, Gibbons WE. Conceptual age and ultrasound measurements of gestational sac and crown-rump length in vitro fertilization pregnancies. *Fertil Steril*. 1988; 49:1012–1017. [PubMed: 3286286]
3. Goldstein SR, Wolfson R. Endovaginal ultrasonographic measurement of early embryonic size as a means of assessing gestational age. *J Ultrasound Med*. 1994; 13:27–31. [PubMed: 7636950]
4. Callen, PW. *Ultrasonography in Obstetrics and Gynecology*. Philadelphia: W.B.Saunders Company; 2000.
5. Schaff E, Fielding SL, Eisinger S, Stadlius L. Mifepristone and misoprostol for early abortion when no gestational sac is present. *Contraception*. 2001; 63:251–254. [PubMed: 11448464]
6. Aubény E, Peyron R, Turpin CL, et al. Termination of early pregnancy (up to 63 days of amenorrhea) with mifepristone and increasing doses of misoprostol. *Int J Fertil Menopausal Stud*. 1995; 40(Suppl 2):85–91. [PubMed: 8574255]
7. Spitz IM, Bardin CW, Benton L, Robbins A. Early pregnancy termination with mifepristone and misoprostol in the United States. *N Engl J Med*. 1998; 338:1241–1247. [PubMed: 9562577]
8. von Hertzen H, Honkanen H, Piaggio G, et al. WHO multinational study of three misoprostol regimens after mifepristone for early medical abortion I: Efficacy. *Br J Obstet Gynaecol*. 2003; 110:808–818.

Fig. 1.
Buccal misoprostol-mifepristone abortion caseload by gestational age in days

Reported Gestational Age by Day

Fig. 2.

Evidence of heaping in reporting of gestational age in days

Note:

Day 0 is a gestational age in days divisible by 7 (e.g., 28, 35, 42, 49, 56).

Day 1 is a gestational age in days one greater than a day divisible by 7 (e.g., 29, 36, 43, 50, 57).

Day -1 is a gestational age in days one short of a day divisible by 7 (e.g., 27, 34, 41, 48, 55).
Other days are categorized analogously.

Reported Gestational Age by Day

Fig. 3.

Evidence of heaping in reporting of gestational age in days

Note:

Day 0 is a gestational age in days divisible by 7 (e.g., 28, 35, 42, 49, 56).

Day 1 is a gestational age in days one greater than a day divisible by 7 (e.g., 29, 36, 43, 50, 57).

Day 2 is a gestational age in days two greater than a day divisible by 7 (e.g., 30, 37, 44, 51, 58). Other days are categorized analogously.

Table 1

Success rate (%) by week of gestation of buccal misoprostol-mifepristone abortions

Week/Days	N	% success	P-value for trend
All Days	1,349	98.3	0.079
Weeks grouped in days <i>ending</i> in multiples of 7			
Week 4 (days 22–28)	28	96.4	0.142
Week 5 (days 29–35)	198	99.0	
Week 6 (days 36–42)	376	99.2	
Week 7 (days 43–49)	444	97.5	
Week 8 (days 50–56)	300	98.3	
Weeks grouped in days <i>centered</i> on multiples of 7			
Week 4 (days 25–31)	77	97.4	0.045
Week 5 (days 32–38)	276	99.6	
Week 6 (days 39–45)	416	99.0	
Week 7 (days 46–52)	408	97.1	
Week 8 (days 53–59)	171	97.7	
Weeks grouped in days <i>beginning</i> in multiples of 7			
Week 4 (days 28–34)	149	99.3	0.231
Week 5 (days 35–41)	342	98.8	
Week 6 (days 42–48)	467	98.1	
Week 7 (days 49–55)	342	98.3	
Week 8 (days 56–59)	46	95.7	

Notes:

1. P-value for a linear trend in days or weeks resulting from a logistic regression model, which includes the woman's age, buccal caseload/site, and the presence of multiple fetuses as well as gestation in days or weeks.
2. Weeks with three or fewer cases are not shown; therefore, the sum of the number of cases grouped by weeks does not add up to the total number of cases. However, all cases are included in the logistic regression analyses used to assess the p-value for a linear trend in weeks.