



Published in final edited form as:

Paediatr Perinat Epidemiol. 2010 January ; 24(1): 31–34. doi:10.1111/j.1365-3016.2009.01089.x.

Accuracy of Reporting Bleeding During Pregnancy

Reem Hasan^{a,b}, Michele L. Jonsson Funk^b, Amy H. Herring^{c,d}, Andrew F. Olshan^{b,d}, Katherine E. Hartmann^{e,f}, and Donna D. Baird^a

^aEpidemiology Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC

^bDepartment of Epidemiology, University of North Carolina, Chapel Hill, NC

^cDepartment of Biostatistics, University of North Carolina, Chapel Hill, NC

^dCarolina Population Center, University of North Carolina, Chapel Hill, NC

^eInstitute for Medicine and Public Health, Vanderbilt University Medical Center, Nashville, TN

^fDepartment of Obstetrics and Gynecology, Vanderbilt University Medical Center, Nashville, TN

Summary

Vaginal bleeding during pregnancy has been considered a marker of an at-risk pregnancy, but the accuracy of reported bleeding has not been assessed. We sought to evaluate the agreement in vaginal bleeding reports based on prospective daily diary and retrospective recall at first trimester interview and to investigate predictors of reporting accuracy. Participants recruited prior to pregnancy for a community-based pregnancy cohort (n=153) completed web-based daily diaries beginning prior to pregnancy up through the end of the first trimester. A comprehensive first-trimester interview was conducted, and the bleeding data from diary and interview were compared. Kappa statistics were used to quantify agreement. Log-linear models were used to investigate maternal age, prior miscarriage, and current pregnancy outcome as potential predictors of agreement. We found that bleeding characteristics (number of bleeding episodes, bleeding heaviness, duration, and gestational timing) from the diary and interview were reported with high levels of agreement. Kappas ranged from 0.77-0.84. Retrospective report of any bleeding had a sensitivity of 0.80 and specificity of 1.0; however, sensitivity was lower when examined within smaller time intervals. Important predictors of agreement were not identified in this analysis, but the sample was small. Overall, the presence of vaginal bleeding, a common and potentially alarming symptom of early pregnancy, may be assessed by interview later in pregnancy with reasonable accuracy.

Introduction

Vaginal bleeding is common in early pregnancy, and has been associated with adverse outcomes.¹⁻⁴ Most studies of vaginal bleeding rely on recalled data, which may be influenced by outcome and other pregnancy events.^{5,6} No studies have evaluated the accuracy of vaginal bleeding reports in pregnancy.

This analysis was undertaken to compare retrospective bleeding data collected from interview with prospective data obtained from daily diaries in a sample of pregnant women. We assessed the extent of agreement between interview and diary.

Methods

Study Population

Right From the Start (RFTS) is an ongoing community-based pregnancy cohort that began enrollment of pregnant women in several states (North Carolina, Tennessee, Texas) of the United States (US) in 2000.⁷ Briefly, participants were at least 18 years old, English- or Spanish-speaking, had not used assisted reproductive technologies to conceive, and intended to carry the pregnancy to term. Women who were not yet pregnant but attempting to conceive could pre-enroll prior to pregnancy and were followed until formal enrollment at the time of a positive pregnancy test. A subset of pre-enrolled women completed a web-based daily diary during the pre-enrollment period and throughout the first trimester. The diary included information about common symptoms and signs of pregnancy, including vaginal bleeding and spotting. A comprehensive telephone interview was completed (median gestational week 13) which collected detailed information about the first trimester, including information about bleeding symptoms. Informed, signed consent was obtained from each study participant in compliance with all Institutional Review Board procedures.

This analysis is based on the 153 participants enrolled in RFTS who completed both the daily diary during the first trimester and completed the first trimester interview.

Bleeding Episodes

This analysis evaluated first trimester bleeding episodes, regardless of pregnancy outcome. Bleeding episodes that terminated within four days of miscarriage were not included.

First-trimester interview—Participants reported the total number of episodes experienced during the first trimester, and detailed information was collected about the timing, heaviness, and duration of bleeding for the first three reported episodes. Participants provided the date on which an episode began, and episode duration was reported in days of bleeding. A ‘spotting’ episode was one that was only noticed when wiping; a ‘bleeding’ episode included at least one day of light or heavy bleeding.

Daily diary—In the daily diary, bleeding and spotting were queried separately. Episodes were defined in a manner similar to that in the interview. All episodes were classified according to their timing (date began), duration (number of days with any bleeding or spotting), and heaviness (bleeding or spotting).

Statistical Analysis

Interview and diary reporting were compared for the occurrence of any episode, the total number of episodes (no report, 1 episode, or 2+ episodes), the total duration of all episodes combined (no report, 1 day, or 2+ days), the timing of the first episode (no report, before 7 weeks gestation, 7-9 weeks gestation, 10+ weeks gestation), and the heaviness of the heaviest episode (no report, spotting only, bleeding) for all women. Sensitivity, specificity, and kappa for presence of any episode were calculated, and a weighted kappa statistic was calculated for all other comparisons. For the weighted kappa, categories were equally spaced and weighted according to squared differences. Sensitivity and specificity were also calculated for episodes reported in the diary and the interview, in one- and two-week intervals of the first trimester.

Log-linear models were used to evaluate maternal age, prior miscarriage, and current pregnancy outcome as possible predictors of agreement for the number of episodes reported and heaviness of bleeding. This method models the distribution of observations in a

contingency table, accounting for agreement due to chance and beyond-chance agreement in the data.⁸⁻¹⁰

Stata (version 9.2) and StatXact (version 6) were used for all analyses.

Results

The majority of participants were white, married, and had at least a college education (Table 1). About half of all women with prior pregnancies reported having had a miscarriage. Sixty-five women (42%) reported at least one episode of spotting or bleeding in the diary; fifty-two of these women reported at least one episode in the first-trimester interview (sensitivity=0.80). No participants reported episodes in the interview without reporting at least one episode in the diary (specificity=1.0). More spotting episodes were reported in the diary compared to interview. The thirteen women who had episodes recorded in the diary and did not report any episodes in the interview recorded only spotting episodes in the diary. All participants who recorded bleeding episodes in the diary reported at least one episode (spotting or bleeding) in the interview, although some misclassification in heaviness was present.

The extent of agreement between information reported in the diary and interview was evaluated by calculating Cohen's kappa and weighted kappa statistics. The kappa for overall agreement of bleeding or spotting reported at any time of gestation was 0.82 (95% CI 0.73, 0.91). Results from the diary and interview were also reported with high levels of agreement for specific bleeding characteristics, with kappas ranging from 0.77 to 0.84 (Table 2).

When all reported bleeding episodes were examined by week of gestation, the sensitivity of interview compared with diary was generally low for a given week (range 0.08, 1.0) but specificity remained high (≥ 0.94 for all weeks). When gestational age at reported bleeding was categorized in two-week intervals, the sensitivity increased (range 0.48, 0.73).

None of the factors that we examined as potential predictors of agreement (maternal age, prior miscarriage, or miscarriage in current pregnancy) were important predictors of agreement. Estimates were imprecise with wide confidence intervals (data not shown).

Discussion

Our overall measures of agreement were high, suggesting that early pregnancy bleeding is accurate when recalled at the end of the first trimester. Our results also show that exact timing is not well reported. Narrow categorization of bleeding episodes into one-week intervals yielded low week-specific sensitivity. However, the increase in sensitivity of episode reports based on two-week intervals indicates that women can retrospectively approximate the timing of their bleeding.

Overall, bleeding episode information obtained from the diary was more detailed compared to interview data. The number, duration, and heaviness of episodes reported in the diary were attenuated when reported in the interview. More episodes were reported in the diary compared to the interview, but these were spotting episodes, suggesting that spotting is more easily forgotten than bleeding.

The major limitation of this analysis is the homogeneity of our study population. Participants were highly educated women who planned their pregnancy and volunteered for our study. A third had had a prior miscarriage. These participants are likely to be highly aware of their pregnancy-related symptoms and motivated to report episodes both in the daily diary and in the first trimester interview. The proportion of women reporting any

bleeding in interview (approximately 34%) was higher in this analysis compared to a related analysis of the entire cohort, in which 27% of participants reported any bleeding. The act of filling out the diary may improve reporting in the interview for this subgroup of women. Additionally, participants were asked to recall events over a relatively short time period. Thus, our results may be viewed as a best-case scenario, based on a select population of women whose recall may be better than the general population.

To conclude, we found high levels of agreement for reports of vaginal bleeding episodes obtained from daily diary and recalled interview.

Acknowledgments

The field research was supported in part by grants from the National Institute of Child and Human Development (5R01HD043883 and 5R01HD049675) and the American Water Works Association Research Foundation (2579). Additional funds were provided by the National Institute of Environmental Health Sciences (Intramural Research Program and P30ES10126).

We thank Lea Cupul Uicab and Matt Longnecker for helpful comments and Sue Edelman for assistance with graphics. We also acknowledge the contributions of *Right From the Start* participants and staff.

References

1. Ananth CV, Savitz DA. Vaginal bleeding and adverse reproductive outcomes: a meta-analysis. *Paediatr Perinat Epidemiol.* 1994; 8:62–78. [PubMed: 8153019]
2. Yang J, Hartmann KE, Savitz DA, Herring AH, Dole N, Olshan AF, Thorp JM Jr. Vaginal bleeding during pregnancy and preterm birth. *Am J Epidemiol.* 2004; 160:118–25. [PubMed: 15234932]
3. Yang J, Savitz DA, Dole N, Hartmann KE, Herring AH, Olshan AF, Thorp JM Jr. Predictors of vaginal bleeding during the first two trimesters of pregnancy. *Paediatr Perinat Epidemiol.* 2005; 19:276–83. [PubMed: 15958150]
4. Hasan R, Baird DD, Herring AH, Olshan AF, Funk MLJ, Hartmann KE. Association between first-trimester vaginal bleeding and miscarriage. *Obstet Gynecol.* 2009 in press.
5. Rockenbauer M, Olsen J, Czeizel AE, Pedersen L, Sorensen HT. Recall bias in a case-control surveillance system on the use of medicine during pregnancy. *Epidemiology.* 2001; 12:461–6. [PubMed: 11416783]
6. Savitz DA, Chan RL, Herring AH, Howards PP, Hartmann KE. Caffeine and miscarriage risk. *Epidemiology.* 2008; 19:55–62. [PubMed: 18091004]
7. Promislow JH, Makarushka CM, Gorman JR, Howards PP, Savitz DA, Hartmann KE. Recruitment for a community-based study of early pregnancy: the Right From The Start study. *Paediatr Perinat Epidemiol.* 2004; 18:143–52. [PubMed: 14996255]
8. Tanner MA, Young MA. Modeling agreement among raters. *J Am Stat Assoc.* 1985; 80:175–180.
9. Agresti A. Modelling patterns of agreement and disagreement. *Stat Methods Med Res.* 1992; 1:201–18. [PubMed: 1341658]
10. Velema JP, Blettner M, Restrepo M, Munoz N. The evaluation of agreement by means of log-linear models: proxy interviews on reproductive history among floriculture workers in Colombia. *Epidemiology.* 1991; 2:107–15. [PubMed: 1932307]

Table 1

Participants who completed the daily diary and first trimester interview (n=153), *Right From the Start* (2000-2008)

| | n | (%) |
|-----------------------------|-----|--------|
| Maternal age (years) | | |
| ≤ 30 | 95 | (62.1) |
| >30 | 58 | (37.9) |
| Race | | |
| White | 132 | (86.8) |
| Other | 20 | (13.2) |
| Missing | 1 | |
| Education | | |
| Less than college | 16 | (10.4) |
| College or more | 137 | (89.5) |
| Marital status | | |
| Married, cohabiting | 152 | (99.3) |
| Single | 1 | (0.7) |
| Smoking in pregnancy | | |
| No | 148 | (96.7) |
| Yes | 5 | (3.3) |
| Gravidity | | |
| Primigravida | 54 | (35.3) |
| 1 or more prior pregnancy | 99 | (64.7) |
| History of miscarriage | | |
| Primigravida | 54 | (35.3) |
| No | 50 | (32.7) |
| Yes | 49 | (32.0) |
| History of induced abortion | | |
| Primigravida | 54 | (35.3) |
| No | 85 | (55.6) |
| Yes | 14 | (9.2) |
| Outcome | | |
| Miscarriage | 19 | (15.4) |
| Live birth/stillbirth | 104 | (84.6) |
| Missing | 30* | |

* These participants were still pregnant at time of analysis

Table 2
 Comparison of bleeding characteristics reported in diary and interview (n=153), *Right From the Start*, (2000-2008).

| | Interview | | Diary | | | Percent agreement | Kappa [95% CI] |
|----------------------------------------------|-----------|-----------|-----------|------------|------------|-------------------|-------------------|
| | None | Spot | Bleed | Total | Total | | |
| Heaviness of heaviest episode | | | | | | | |
| None | 88 | 13 | 0 | 101 | | 0.84 | 0.82 [0.75, 0.88] |
| Spot | 0 | 33 | 5 | 38 | | | |
| Bleed | 0 | 6 | 8 | 14 | | | |
| Total | 88 | 52 | 13 | 153 | | | |
| Number of episodes | | | | | | | |
| None | 88 | 11 | 2 | 101 | | 0.84 | 0.83 [0.76, 0.91] |
| 1 | 0 | 20 | 9 | 29 | | | |
| 2+ | 0 | 3 | 20 | 23 | | | |
| Total | 88 | 34 | 31 | 153 | | | |
| Total duration of all episodes (days) | | | | | | | |
| None | 88 | 8 | 5 | 101 | | 0.86 | 0.84 [0.76, 0.92] |
| 1 | 0 | 9 | 7 | 16 | | | |
| 2+ | 0 | 2 | 34 | 36 | | | |
| Total | 88 | 20 | 46 | 153 | | | |
| Timing of first episode (weeks) | | | | | | | |
| None | 88 | 7 | 4 | 2 | 101 | 0.84 | 0.77 [0.65, 0.89] |
| ≤6 | 0 | 26 | 3 | 2 | 31 | | |
| 7-9 | 0 | 4 | 7 | 0 | 11 | | |
| 10+ | 0 | 0 | 3 | 7 | 10 | | |
| Total | 88 | 37 | 17 | 11 | 153 | | |