

9 2 0 0 '00 FEB 15 A9:50

February 11, 2000

Dockets Management Branch
U.S. Food and Drug Administration
Room 10-61
5630 Fishers Lane
Rockville, MD 20852

Re: Docket 97P-0329

Enclosed are two scientific studies and a *New England Journal of Medicine* editorial on the effects of caffeine in pregnant women. The first study is a meta-analysis that pooled 32 studies exploring the effects of caffeine on rates of spontaneous abortion and low birth weight.¹ That study found a significant increase in the rates of miscarriage and low-birth-weight babies in pregnant women who consumed more than 150 milligrams of caffeine per day. The odds ratio for spontaneous abortion was 1.4 with a 95 percent confidence interval between 1.3 and 1.5. For low birth weight, the odds ratio was 1.5 with a 95 percent confidence interval between 1.4 and 1.6.

The second study found an association between caffeine consumption (more than 6 cups of coffee per day) with spontaneous abortion in women.² Instead of using self-reported data on caffeine consumption, the investigators measured the serum concentration of the caffeine metabolite paraxanthine to determine caffeine exposure. They relied on a single serum paraxanthine measurement to determine the level of exposure throughout pregnancy.

In an accompanying editorial, Brenda Eskenazi, Ph.D., from the University of California Berkeley School of Public Health, critiqued the paraxanthine study and discussed the policy

¹ Fernandes, O., Sabharwal, M., Smiley, T., Pastuszak, A., Koren, G., and Einarson, T., Moderate to heavy caffeine consumption during pregnancy and relationship to spontaneous abortion and abnormal fetal growth: a meta-analysis. *Reproductive Toxicology* 1998;12: 435-444.

² Klebanoff, M.A., Levine, R.J., DerSimonian, R., Clemens, J.D., Wilkins, D.G., Maternal serum paraxanthine, a caffeine metabolite, and the risk of spontaneous abortion. *New England Journal of Medicine* 1999;341: 1639-1644.

97P-0329

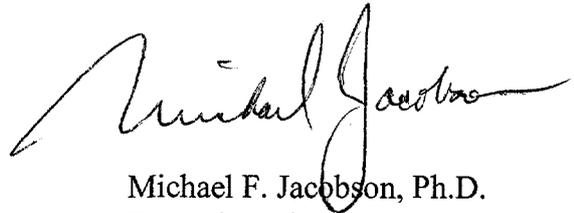
C238

implications of the meta-analysis.³ She contended that a single measure of paraxanthine could not reliably estimate the consumption level in pregnant woman. She also expressed concern that certain subgroups of the population may be particularly sensitive to the effects of caffeine and that a safety factor of 10 must be used to determine a safe level of exposure to protect those people. Her editorial called for labeling of caffeine content and continued education of pregnant women to limit caffeine intake.

Sincerely,



Patricia Lieberman, Ph.D.
Staff Scientist



Michael F. Jacobson, Ph.D.
Executive Director

enclosures

³ Eskenazi, B., Caffeine -- Filtering the facts. *New England Journal of Medicine* 1999; 341: 1688-1689.



● *Original Contributions*

**MODERATE TO HEAVY CAFFEINE CONSUMPTION DURING
PREGNANCY AND RELATIONSHIP TO SPONTANEOUS ABORTION AND
ABNORMAL FETAL GROWTH: A META-ANALYSIS**

OLAVO FERNANDES,* MONA SABHARWAL,* TOM SMILEY,* ANNE PASTUSZAK,†
GIDEON KOREN,† and THOMAS EINARSON*†‡

*Doctor of Pharmacy Program, Faculty of Pharmacy, University of Toronto; †The Motherisk Program, The Division of Clinical Pharmacology and Toxicology, Department of Pediatrics and Research, The Hospital for Sick Children; and ‡Department of Health Administration, Faculty of Medicine, University of Toronto, Toronto, Canada

Abstract — The objective was to determine the association of moderate to heavy caffeine consumption during pregnancy on spontaneous abortion and birth weight in humans. Data sources used included a computerized literature search of MEDLINE (1966–July 1996); EMBASE (1988–November 1996); Psychlit I (1974–1986); Psychlit II (1987–1996); CINAHL (1982–May 1996) and manual search of bibliographies of pertinent articles. Inclusion criteria were: English language research articles; pregnant human females; case control or cohort design; documented quantity of caffeine consumption during pregnancy; control group with minimal or no caffeine consumption (0 to 150 mg caffeine/d); documented data regarding spontaneous abortion and/or fetal growth. The exclusion criteria were: case reports; editorials; review papers. The methods section of each study was examined independently by two blinded investigators with a third investigator adjudicating disagreements. Two independent investigators extracted data onto a standardized form. A third investigator adjudicated discrepancies. We compared a caffeine-exposed group (>150 mg/d) and controls (0 to 150 mg/d), using Mantel–Haenszel pooling. Of the 32 studies meeting inclusion criteria, 12 had extractable data (6 for spontaneous abortion, 7 for low birth weight, 1 common study). Mantel–Haenszel odds ratio (CI_{95%}) was 1.36 (1.29–1.45) for spontaneous abortion in 42,988 pregnancies. The overall risk ratio was 1.51 (1.39–1.63) for low birthweight (<2500 g) in 64,268 pregnancies. Control for confounders such as maternal age, smoking, and ethanol use was not possible. We concluded that there is a small but statistically significant increase in the risks for spontaneous abortion and low birthweight babies in pregnant women consuming >150 mg caffeine per d. A possible contribution to these results of maternal age, smoking, ethanol use, or other confounders could not be excluded. © 1998 Elsevier Science Inc.

Key Words: caffeine; pregnancy; spontaneous abortion; congenital malformations.

INTRODUCTION

In 1980, the United States Food and Drug Administration issued a warning regarding the use of caffeine during pregnancy (1). While conclusions about human teratogenicity could not be definite at that time, the FDA suggested that as a precautionary measure, pregnant women should be advised to avoid or limit their consumption of food or drugs containing caffeine. Due to the large worldwide consumption of caffeinated beverages (e.g., coffee, tea, cola) it is important to know whether

such a warning is actually warranted. Should caffeine consumption during pregnancy be linked to adverse effects such as spontaneous abortion or fetal growth retardation, that finding would have important implications for public health. Furthermore, the potential impact of that association is underscored by the fact that low birth weight is associated with high mortality and morbidity in neonates.

Caffeine clearance from the body is essentially unchanged during the first trimester of pregnancy. However, a significant delay in elimination occurs in the second and third trimester, as the half life of caffeine extends to 10.5 h from a normal half life of 2.5 to 4.5 h in the nonpregnant woman (2). Caffeine is known to readily cross the placenta. Substantial quantities pass into

Address correspondence to Dr. Koren at the Hospital For Sick Children, 555 University Avenue, Toronto, ON, M5G 1X8 Canada.

Received 24 October 1997; Final revision received 1 April 1998; Accepted 4 April 1998.

the amniotic fluid, umbilical cord blood, and the urine and plasma of neonates. In addition, the human fetus and neonate have low levels of the enzymes necessary for caffeine metabolism.

Several mechanisms for caffeine to produce adverse outcomes have been postulated. For example, caffeine increases cellular cyclic adenosine monophosphate (cAMP) through phosphodiesterase inhibition. The rise in cAMP may interfere with fetal cell growth and development (3).

Animal studies of caffeine and pregnancy outcomes have reported considerable variability in results. Some studies have suggested a link between caffeine and teratogenesis, fetal resorption, and decreased fetal weight (4,5). An increase in the malformation rate, specifically cleft palate and ectrodactyly, was demonstrated in rats and mice at caffeine doses of 100 mg/kg/d or more (5). This effect was not seen at doses of 50 mg/kg/d. It is important to note that humans ingest caffeine at significantly lower doses of 1.7 to 4.5 mg/kg/d (5).

Epidemiologic studies have produced incomplete or conflicting results concerning the effects of caffeine exposure during pregnancy. To date, we are unaware of a formal meta-analysis quantifying the potential risks. Therefore, the present meta-analysis was conducted to determine the association of moderate to heavy caffeine consumption during pregnancy on spontaneous abortion and birth weight in humans.

MATERIALS AND METHODS

Original research studies investigating the effects of moderate to heavy caffeine consumption during pregnancy on spontaneous abortion and fetal growth in humans were examined using meta-analysis based on methods described by Einarson et al. (6). *Spontaneous abortion* was defined for this analysis as expulsion from the uterus of products of conception before the fetus is viable (approximately 20 weeks of gestation). Included in the definition were fetal loss, fetal death, and miscarriage. *Fetal growth* was defined by standard measures that included birth weight, birth weight for gestational age, birth weight by percentile, body length, and head circumference. *Low birth weight* was defined as birth weight less than 2500 g and intrauterine growth retardation (IUGR) was defined as birth weight less than the tenth percentile for gestational age (7).

Values for *caffeine content* of beverages and foods were recorded as defined by the caffeine content in milligrams outlined in each individual study. If the caffeine content was not presented in milligrams (mg), then the following standard conversions were used: one cup of coffee was equivalent to 74 mg of caffeine and one cup of tea was equivalent to 27 mg of caffeine (8).

Due to the nature of caffeine consumption, a group of pregnant women with absolute non-exposure to caffeine is difficult to find. We recognize that the term *control group* has several different meanings in epidemiology. For the purposes of this meta-analysis, the *control* was defined as a group exposed to minimal or no caffeine (0 to 150 mg caffeine/d). *Moderate caffeine consumption* was defined as 151 to 300 mg caffeine/d and *heavy caffeine consumption* was >300 mg caffeine/d.

Search strategy

A computerized literature search was completed using the following databases: MEDLINE (1966–December 1996); EMBASE (1988–November 1996); Psychlit I (1974–1986) and Psychlit II (1987–1996); CINAHL (1982–May 1996). Articles examining the relationship between caffeine consumption and pregnancy were identified using the search terms “pregnancy and caffeine” along with “pregnancy outcome and caffeine.” Search terms were initially kept as broad as possible in order to ensure that articles that were not indexed strictly by the desired outcomes (i.e., spontaneous abortion and fetal growth) were not missed. All abstracts retrieved from the computer search were independently reviewed by at least two investigators to identify articles relating to the desired outcomes. Additional references were identified from bibliographies of retrieved articles and selected reviews (4,5,9,10).

Study selection

An independent review was conducted by removing all identifiers and having two investigators independently evaluate the methods sections of all retrieved articles using a checklist of the inclusion and exclusion criteria. The inclusion criteria for the meta-analysis were: English language articles; pregnant human females; case control or cohort design; caffeine exposure during pregnancy; documented quantity of caffeine consumption; control group with minimal or no caffeine consumption (0 to 150 mg caffeine/d); and documented data regarding spontaneous abortion and/or fetal growth. Case reports or case series, editorials, and review articles were excluded. Reviewers were blinded to journal names, author names and study results. In the event that agreement could not be reached between the two reviewers or if sufficient information was not provided in the methods sections, a third investigator was consulted who served as the adjudicator. If the adjudicator could not reach a decision based on the methods section alone, the entire paper was reviewed by all judges.

Data extraction

Data extraction was performed on all included articles independently by two investigators who were

Table 1. Excluded studies and reasons for rejection

Barr et al. (35) Pastore and Savitz (37) Peacock et al. (42) Rosenberg et al. (36) Wisborg et al. (44)	Do not contain desired outcomes as defined by our inclusion criteria
Fried and O'Connell (34) Godel et al. (41) Kline et al. (40) Larroque et al. (38)	Control group not identified as defined by our inclusion criteria
Shu et al. (26) Watkinson and Fried (13) Wilcox et al. (28)	Data not combinable according to our caffeine stratification criteria (rejected on attempt to extract data)
Barr and Streissguth (14) Munoz et al. (43) Peacock et al. (25) Vandenberg (39) Weathersbee et al. (15)	Data not extractable
Beaulac-Baillargeon and Desrosiers (18) Desrosiers (18) Furuhashi et al. (33) Olsen et al. (24)	Control group not defined by study

blinded to the journal and authors' names. Data were extracted using a standardized form that recorded study characteristics, sample characteristics, caffeine content stratification, confounding factors, and outcome results in both caffeine and control groups. Extracted quantitative data for spontaneous abortion and fetal growth were entered in 2×2 tables for control and caffeine groups. Data extraction forms were reviewed for agreement by a third investigator, who conducted an individual assessment of the study if there was a disagreement in values or information.

Statistical methods

Odds ratios were calculated for individual case control studies and risk ratios were calculated for individual cohort studies along with associated 95% confidence intervals. A combined Mantel-Haenszel odds ratio (11) was calculated for each outcome comparison and an overall 95% confidence interval was calculated by the method described by Miettinen (12). In the event that all

the combined studies were cohort studies, an overall Mantel-Haenszel relative risk ratio was calculated. A Q value (χ^2) and P -value for homogeneity of samples was calculated using standard statistical methods. A level of $P < 0.05$ (two-tailed) was considered significant for all statistical tests. For each outcome, the main analysis comprised of a comparison of a caffeine exposure group (>150 mg caffeine/d) to a control group (0 to 150 mg caffeine/d). Subgroup and sensitivity analyses were performed to investigate comparisons among moderate, heavy, control, and zero caffeine consumption levels; research design (i.e., cohort and case control studies); the effect of large studies; and the effect of adding studies that did not meet the meta-analysis caffeine content stratification criteria to identify and evaluate possible changes to odds ratios and relative risk ratios.

RESULTS

Study selection

Over 275 abstracts of articles dealing with caffeine exposure in pregnancy published between 1966 and 1996 were identified by the initial search strategy. Upon examination of these abstracts and review articles, 32 papers were identified as potentially eligible and were entered into the study selection process (13-44). After the blinded independent study selection process, 21 articles (13-33) met the inclusion criteria and were potentially eligible for the meta-analysis. Interobserver agreement was 87% after the initial application of the inclusion criteria. However, full consensus was reached after adjudication. The data extraction process was performed on the 21 articles. During this process, an additional nine studies were excluded. Table 1 outlines excluded studies and reasons for rejection. A total of 12 studies were accepted into the analysis process (six studies for the spontaneous abortion outcome (16,17,27, 29-31) and seven studies (19-23,30,32) for the fetal growth outcome, which includes one study (30) that was accepted for both outcomes. At this time, two additional studies (23,30) were excluded from the main analysis and used only for sensitivity analyses since their caffeine stratification did not meet the exact caffeine stratification

Table 2. Relationship between caffeine exposure and spontaneous abortion: Summary of study characteristics

Article	Study design	<i>n</i>	Sources of caffeine identified	Caffeine conversion factors (mg/cup unless specified)
Armstrong et al. (17)	cohort, retrospective	35,848	coffee	not specified
Dominguez-Rojas et al. (31)	cohort, retrospective	691	coffee	coffee: 140
Fenster et al. (27)	case control, retrospective	852 cases 1618 controls	coffee, tea, cola	coffee: 107; tea: 34; cola: 47 mg/can
Infante-Rivard et al. (29)	case control, retrospective	331 cases 993 controls	coffee, tea, cola	coffee: 107; tea: 34; cola: 47 mg/can
Mills et al. (30)	cohort, prospective	423	coffee, tea, cola, cocoa, medications	coffee: 100; tea: 40; cocoa: 30; cola: 40 mg/can; decaffeinated coffee: 1.5

Table 3. Relationship between caffeine exposure and fetal growth: Summary of study characteristics

Article	Study design	n	Sources of caffeine identified	Caffeine conversion factors (mg/cup unless specified)
Caan and Goldhaber (23)	case control, retrospective	131 cases 136 controls	coffee, tea, cola,	not specified
Fenster et al. (21)	cohort, retrospective	1,230	coffee, tea, cola,	coffee: 107; tea: 34; cola: 47 mg/can;
Fortier et al. (19)	cohort, retrospective	7,025	coffee, tea, cola, chocolate	coffee: filtered or percolated 109 mg; instant 74 mg; espresso 168 mg; tea: bag 49 leaves or instant 30 mg; cola: 29 mg; chocolate: 56 mg
Linn et al. (32)	cohort, retrospective	12,205	coffee, tea	not specified
Martin and Bracken (20)	cohort, prospective	3,891	coffee, tea, cola, medications	coffee: 107 mg; tea: 34; cola: 47 mg/serving
McDonald et al. (22)	cohort, retrospective	40,445	coffee	not specified
Mills et al. (30)	cohort, prospective	423	coffee, tea, cola, cocoa, medications	coffee: 100; tea: 40; cola: 40 mg/can; cocoa: 30 mg; decaffeinated coffee: 1.5

definitions for combinability. Tables 2 and 3 summarize characteristics of accepted studies.

Spontaneous abortion

Table 4 presents odds ratios and risk ratios for individual studies for spontaneous abortion. In the main analysis, comparing spontaneous abortions in the caffeine exposure group (>150 mg caffeine/d) to the control group, a total of five studies (16,17,27,29,31) were included (three cohort and two case control studies) involving a total of 42,889 patients. The combined odds ratio (CI_{95%}) was 1.36 (1.29–1.45) with $Q = 21.21$ ($P < 0.001$) for heterogeneity of outcome. Table 5 outlines the type and results of sensitivity analysis, which did not greatly alter the odds ratio. However, the removal of one study (31) greatly improved the homogeneity of the analysis ($Q = 3.52$, $P < 0.318$). Due to caffeine stratification of the accepted studies, a comparison of the control to "zero" caffeine consumption was not possible.

Fetal growth

All seven accepted studies measured fetal growth according to low birth weight (<2500 g). Two of those studies also evaluated intrauterine growth retardation

(IUGR). Table 6 outlines individual odds ratios and risk ratios for individual studies used in the low birth weight outcome comparison. In the main analysis, comparing low birth weight babies in the caffeine exposure group (>150 mg caffeine/d) to the control group, a total of five studies (19–22,32) were included (all cohort designs) involving a total of 64,268 patients (Table 7). The combined relative risk was 1.51 (1.39–1.63) with $Q = 8.72$ ($P = 0.068$) for heterogeneity of outcome. Table 7 outlines the type and results of sensitivity analysis, which did not greatly alter the summary relative risk. Further subgroup analyses on the low birth weight outcome are also displayed in Table 7. The risk ratios (CI_{95%}) for comparisons of moderate caffeine consumption to control was 1.33 (1.21–1.47), 1.81 (1.61–2.04) comparing heavy caffeine consumption to control, and 1.38 (1.20–1.60) comparing heavy caffeine consumption to moderate caffeine consumption. A risk ratio of 1.06 (1.00–1.13) resulted from our comparison of the control to "zero" caffeine consumption. As this risk ratio included unity, it validated our choice of control group (0 to 150 mg caffeine/d). A combined risk ratio of 1.56 (1.34–1.82) was calculated for two studies (19,21) that inves-

Table 4. Results of individual studies comparing spontaneous abortions in caffeine exposure (>150 caffeine mg/d) to control groups (0 to 150 caffeine mg/d)

Article	Caffeine exposure >150 mg/d		Controls		OR/RR (95% CI)
	Spontaneous abortion	No spontaneous abortion	Spontaneous abortion	No spontaneous abortion	
Armstrong et al. (17)	1,577	4,564	6,183	23,524	RR = 1.23 (1.18–1.29)
Dominguez-Rojas et al. (31)	146	329	23	193	RR = 2.89 (1.92–4.35)
Fenster et al. (27)	152	256	455	1,028	OR = 1.34 (1.07–1.69)
Infante-Rivard et al. (29)	92	186	239	807	OR = 1.67 (1.25–2.22)
Mills et al. (30) ^a	43	291	16	70	RR = 1.44 (0.86–2.44)
Srisuphan and Bracken (16)	27	852	41	2,215	RR = 1.69 (1.05–2.73)
Summary odds ratio					OR = 1.36 (1.29–1.45)

OR/RR = odds/risk ratio calculated with the Mantel-Haenszel formula; 95% CI = 95% confidence interval for odds ratio.

^anot included in summary odds ratio as study's caffeine stratification of groups did not meet the exact caffeine stratification definitions or combinability in the meta-analysis (assumptions made to test data as a sensitivity analysis as described in Methods).

Table 5. Combined results of studies comparing spontaneous abortion in caffeine exposure (>150 mg caffeine/d) to control groups (0 to 150 mg caffeine/d)

Analysis	n	Summary Ratio	95% CI	Test for homogeneity		
				Q	DF	P
Primary analysis (Studies 16, 17, 27, 29, 31)	42,889	OR = 1.36	1.29–1.45	21.21	4	<0.001
Cohort studies only	39,674	RR = 1.26	1.20–1.33	13.52	2	0.001
Case control studies only	3,215	OR = 1.46	1.22–1.74	1.36	1	0.244
Sensitivity: removing study #17		OR = 1.69	1.45–1.98	14.48	3	0.002
Sensitivity: adding in study #30		OR = 1.37	1.29–1.45	21.38	5	<0.001
Sensitivity: removing outlier study #31		OR = 1.33	1.26–1.42	3.52	3	0.318

DF = degrees of freedom.

tigated IUGR, which supports our data for low birth weight.

DISCUSSION

Epidemiologic studies involving caffeine consumption by pregnant women have resulted in differing results concerning adverse fetal outcomes. This meta-analysis was designed to examine the risk of spontaneous abortion and fetal growth retardation, as these two outcomes are sources of significant morbidity, mortality, and societal burden in terms of costs.

This meta-analysis indicates a modest but statistically significant relationship between moderate to heavy caffeine consumption in pregnancy and the risk for spontaneous abortion and low birth weight. In order to reasonably assess the implications of these findings, the results of subgroup and sensitivity analyses must be examined along with the limitations of the accepted studies.

Defining an appropriate reference group was a challenge in this analysis. Most of the studies did not explicitly differentiate low-exposed and unexposed groups. The control group (0 to 150 mg caffeine/d) was chosen based on the large number of studies that utilized this categorization as "light" caffeine users (<150 mg/d). Results of the sensitivity analysis comparing "zero"

caffeine intake to the control group (0 to 150 mg/d) validated the assumption that less than 150 mg of caffeine consumption per d constituted an appropriate control group for our meta-analyses. Srisuphan et al. (16) reported similar findings in their research, which explored caffeine consumption and the risk of spontaneous abortion. They postulated a possible "threshold effect" of around 150 mg/d, reasoning that intake below this level would not be enough caffeine to affect the fetus in terms of cell growth, cell division, or uteroplacental circulation.

Spontaneous abortion

The analysis indicates a small but significant relationship between moderate to heavy caffeine consumption and the risk for spontaneous abortion. The study by Armstrong and colleagues (17) contributed heavily to the sample size of this analysis. When that study was removed from the main analysis (Table 5), the odds ratio increased to 1.69 (95% CI, 1.45–1.98). The overall analysis for spontaneous abortion exhibited a large degree of heterogeneity ($Q = 21.21$, $P < 0.001$). We identified the study by Dominguez-Rojas and coworkers (31) as an outlier by examining its contribution to the variance of Q for homogeneity. We performed a sensitivity analysis to determine the effect of this study on

Table 6. Results of individual studies comparing low birth weight caffeine exposure (>150 mg caffeine/d) to control groups (0 to 150 mg caffeine/d)

Article	Caffeine exposure >150 mg/d		Controls		OR/RR (95% CI)
	LBW	Birth weight >2500 g	LBW	Birth weight >2500 g	
Caan, Goldhaber (23) ^{†a}	34	27	96	108	OR = 1.42 (0.79–2.52)
Fenster et al. (21)	26	217	61	926	RR = 1.73 (1.12–2.68)
Fortier et al. (19)	79	1,156	242	5,251	RR = 1.45 (1.14–1.86)
Linn et al. (32)	116	1,152	839	10,098	RR = 1.19 (0.99–1.49)
Martin and Bracken (20)	32	987	38	2,603	RR = 2.18 (1.37–3.47)
McDonald et al. (22)	455	5,001	1,837	33,152	RR = 1.59 (1.44–1.75)
Mills et al. (30) ^a	5	82	16	320	RR = 1.21 (0.46–3.20)
Summary odds ratio					OR = 1.51 (1.39–1.63)

OR/RR = odds/risk ratio from individual studies; 95% CI = 95% confidence interval for odds ratio. LBW = low birth weight defined as birth weight < 2500 g.

^anot included in summary odds ratio as study's caffeine stratification of groups did not meet the exact caffeine stratification definitions for combinability in the meta-analysis (assumptions made to test data as a sensitivity analysis as described in Methods).

Table 7. Combined results of studies examining low birth weight.

Analysis	Summary			Homogeneity		
	n	Ratio	95% CI	Q	DF	P
Primary analysis (Studies 19-22, 32) ^a	64,268	RR = 1.51	1.39-1.63	8.72	4	0.068
Sensitivity: removing study #22		RR = 1.38	1.20-1.57	6.53	3	0.088
Sensitivity: adding study #30		RR = 1.50	1.39-1.63	8.91	5	0.113
Sensitivity: adding study #23		OR = 1.55	1.42-1.69	9.55	5	0.089
Combined results comparing LBW in moderate ^b caffeine exposure to control ^c group (19-22, 32)	61,374	RR = 1.33	1.21-1.47	6.97	4	0.137
Combined results analysis comparing LBW in heavy ^d caffeine exposure to control ^c group (19-22, 32)	58,013	RR = 1.81	1.61-2.04	10.98	4	0.029
Combined results analysis comparing LBW in heavy ^d caffeine exposure to moderate ^b group (19-22, 32)	9,221	RR = 1.38	1.20-1.60	2.76	4	0.598
Combined results analysis comparing LBW in "zero" ^e caffeine exposure to control ^c group (19-22, 32)	82,640	RR = 1.06	1.00-1.13	3.07	4	0.546

LBW = low birth weight defined as birth weight <2500 g.

^aMain analysis compares caffeine exposure (>150 mg/d) to control group.

^b150 to 300 mg caffeine/d

^c0 to 150 mg caffeine/d.

^d>300 mg caffeine/d.

^eGroups designated as having no caffeine intake.

both homogeneity and the summary odds ratio. After its removal, the summary odds ratio was reduced slightly to 1.31, and the heterogeneity statistics became nonsignificant ($Q = 3.52$, $P = 0.318$). That study was analyzed to determine a possible explanation for its divergent results. The study had been conducted in Madrid, Spain and

espresso coffee was the most common form of caffeine consumption reported. This coffee was found to contain about 140 mg caffeine/cup, which is almost twice as strong as a cup of coffee consumed in North America (8). When examining the results of the expectant mothers who had consumed more than 420 mg caffeine per d, it was found that 61 out of 87 (71%) pregnancies resulted in spontaneous abortion. In comparison, the study by Armstrong et al. (17), conducted in Montreal, Canada, investigated a group of women who had consumed over 700 mg of caffeine/d in the form of regular coffee. The rate of spontaneous abortion in those women was 30.9%. It is interesting to note the difference in the concentration of caffeine in the two studies and the resultant rates of spontaneous abortion. It may be that the consumption of beverages containing highly concentrated caffeine over short periods has a more significant effect on fetal development than conventional beverage consumption, or that espresso coffee has other unmeasured ingredients that may also be contributing to outcomes. These theories warrant consideration in future research.

In summary, the sensitivity and subgroup analyses performed within the spontaneous abortion analysis indicated that no group of studies (cohort or case control) or individual investigation influenced, to any degree, the overall main analysis of the summary odds ratio. The study by Dominguez-Rojas and coworkers (31) contributed to the heterogeneity of the analysis, however, the change in summary odds ratio was negligible, and nonheterogeneity was achieved when that study was removed from the analysis.

Table 8. Potential confounders as identified by included studies

1. Smoking
2. Alcohol
3. Maternal age
4. Cannabis
5. Previous abortion
6. Gravidity
7. Parity
8. Employment status
9. Education
10. Body type
11. Infection
12. Family history
13. Race
14. Drug use
15. Married status
16. Menarcheal age
17. Prior gynecologic surgery
18. Interval from previous pregnancy less than 6 months
19. Pregnancy induced hypertension
20. Uterine abnormality
21. Previous stillbirth
22. Insurance coverage
23. Use of tap water
24. Nausea during pregnancy
25. Weight extremes
26. Hours of physical activity/week
27. Previous low birth weight newborn
28. Previous preterm births
29. Weight gain in pregnancy

Table 9. Comparison of confounding factors of accepted spontaneous abortion studies

Study	Confounders identified ^a	Reported RR/OR (95% CI)	Adjusted RR/OR (95% CI)	Comments
Armstrong et al. (17)	1,2,3,5,6,8,9,13	OR = 1.17 (1.03–1.32) for zero vs. 375–675 mg caffeine consumption	not reported	After adjusting for confounders, reported risk of spontaneous abortion increases by factor of 1.017 per cup of coffee consumed per day ($P = 0.01$)
Dominguez-Rojas et al. (31)	1,2,3,5,15,16	RR = 1.87 (1.12–3.14) for 141–280 mg, RR = 5.15 (2.70–9.82) for 291–420 mg, RR = 20.47 (10.85–38.65) for >420 mg	OR = 2.20 (1.22–3.96) for 141–280 mg, OR = 4.81 (2.28–10.14) for 291–420 mg, OR = 5.43 (7.34–32.43) for >420 mg	Outlier study (see discussion) Spanish hospital workers Expresso coffee (160 mg/cup)
Fenster et al. (27)	1,2,3,5,6,8,9,13,15,21,22,23,24	OR = 1.24 (0.90–1.71) 151–300 mg, OR = 1.55 (1.04–2.31) >300 mg	OR = 1.17 (0.84–1.62) 151–300 mg, OR = 1.22 (0.8–1.87) >300 mg	55% increased likelihood of heavy caffeine consumption for cases as compared to controls (association increases with dose)
Infante-Rivard et al. (29)	1,2,3,6,9,19,20	not reported	OR = 1.95 (1.29–2.93) for 0–48 vs 163–321 mg caffeine	Pattern of increased risk for fetal loss with increasing quantity of caffeine intake OR increased by a factor of 1.22 for each 100 mg caffeine intake per day.
Mills et al. (30)	1,2,3,5,6,7,8,9,11,12,13,14	not stated	OR = 1.15 (0.89–1.49)	Study used for sensitivity purposes only
Srisuphan and Bracken (16)	1,2,3,4,5,6,7,13,16,17,18	RR = 1.95 [P (for significance) = .07 for control] (1–150 mg vs >150 mg)	RR = 1.73; P (for significance) = 0.03	Positive association between caffeine use and smoking
Wilcox et al. (28)	1,2,3,4,5,6,9,10	RR = 2.4 (0.80–7.00) control vs >150 mg	N/A	Sample too small ($n = 171$) for extensive multivariate analysis

^aSee Table 8 for confounder numbering scheme.

Low birth weight

The risk ratio calculated for the main analysis of studies (control versus > 150 mg caffeine/d) was 1.51 (1.43–1.69). The study by McDonald et al. (22) was substantially larger than the others. Upon removal of that study from the main analysis for low birth weight, the relative risk was reduced to 1.38 (1.20–1.57). Adding the study by Mills and colleagues (30) or Caan and Goldhaber (23), which used different controls and consumption groups than our criteria, had virtually no effect on the outcome of the summary risk ratio or confidence intervals calculated in the main analysis (Table 7).

In order to explore a possible dose–effect relationship between caffeine consumption and risk of outcome, we calculated summary risk ratios for control versus moderate (150 to 300 mg caffeine/d) and control versus heavy (>300 mg caffeine/d) groups. These increasing ratios suggest an increased risk of low birth weight neonates in relation to the amount of caffeine consumed above 150 mg/d.

Finally, two studies (19,21) were combined that defined IUGR as birth weight less than the tenth percentile for gestational age, to determine whether there was an association. The summary risk ratio calculated by combining these two studies was almost identical to the

overall analysis for low birth weight, thus supporting our original findings.

Study limitations

When combining studies addressing the reproductive risks of caffeine, one has to acknowledge the limitations inherent to this research.

Measurement of caffeine

All of the studies that were accepted into the meta-analyses depended on the recall of the mother or expectant mother with regards to her level and sources of caffeine consumption. Ability to accurately recall and report the amount of caffeine ingested partly depends on whether the research was done prospectively or retrospectively. To study this type of recall bias, Fenster et al. (45) tested the recall of women who had been asked to report their caffeine consumption six months earlier. They showed that the women were able to reproduce their answers from six months earlier, within 1 cup of coffee, 77% of the time. As presented in Tables 2 and 3, four of the five studies in each of the two overall meta-analysis groups were researched retrospectively. Although recall bias may be a source of error in estimation, Fenster et al. showed that a habitual beverage

Table 10. Comparison of confounding factors of accepted fetal growth studies

Study	Confounders identified ^a	Reported RR/OR (95% CI)	Reported Adjusted RR/OR (95% CI)	Comments
Caan et al. (23)	1,2,3,6,13,29	OR = 2.30 (CI not stated) for zero vs >300 mg caffeine consumption per day	OR = 3.53 (1.05–11.81) for zero vs >300 mg caffeine consumption per day	Case control study. Small study size.
Fenster et al. (21)	1,2,3,5,7,8,9,13,15,19,21,22	OR = 2.36 (1.17–4.93) for zero vs >300 mg/d caffeine consumption (LBW)	OR = 2.05 (0.86–4.88) zero vs. >300 mg caffeine consumption (LBW)	Reported OR adjusted for age, parity, race, hypertension, cigarette, and alcohol consumption.
Fortier et al. (19)	1,2,3,7,8,9,15,21,25,26,27,28	OR = 1.86 for 0–10 mg vs 151–300 mg in IUGR (CI not stated)	OR = 1.4 (1.05–1.85) for 0–10 mg vs 151–300 mg in IUGR	OR for IUGR were for full term births.
Linn et al. (32)	1,2,3,5,6,7,8,9,13,15,21,25	OR = 1.45 for >300 mg vs control (CI not stated)	OR = 1.17 (0.85–1.61) for control vs >300 mg	Frequency of smoking was over 3 times greater in coffee drinkers 300 mg/d
Martin and Bracken (20)	1,2,3,4,5,7,8,13,15,21,25,29	RR = 4.0 (1.9–8.6) for zero vs >300 mg/d	RR = 4.6 (2.0–10.5) for zero consumption vs >300 mg/d	Analysis included term births only RR actually increased after adjusting for confounders
McDonald et al. (22)	1,2,3,5,6,7,8,9,13	not stated as crude OR (estimated by logistic regression in next column)	OR = 1.34 (1.10–1.65) for low birth weight adjusted for gestational age; zero vs 375–750 mg per d	Women who drank >10 cups of coffee per day OR = 1.43 (1.02–2.02).
Mills et al. (30)	1,2,3,5,6,7,8,9,11,12,13,14	not stated	not stated	Birth weight percentile was lower in caffeine users after adjustment. <i>P</i> (for significance) = 0.06

^aSee Table 8 for confounder numbering scheme.

such as coffee (which was the main source of caffeine in all studies) is often consumed in daily patterns that may not be very difficult to predict.

A second possible type of error introduced into the issue of caffeine measurement is encountered upon estimating the amount of caffeine contained in specific servings. Most studies utilized an educated "guess" by taking averages of various analyzed samples obtained from their study population. In order to examine the variation in caffeine content of beverages, conversion factors used for estimating caffeine content in each study are presented in Tables 2 and 3.

A third potential error involving caffeine intake estimation involves the lack of identification of all sources of caffeine consumed. Tables 2 and 3 summarize the sources of caffeine included in each study. Varying protocols exist, with three studies taking only coffee into account in their estimation of consumption. Although coffee is the most common source of caffeine, this systematic error would generally lead to a degree of underestimation of caffeine use. It is assumed that this underestimation would occur to the same extent in the control and caffeine consumption groups.

Combinability of results

In order to perform a meta-analysis of studies one needs data that can be validly combined. Some studies were excluded on the basis of stratifications of caffeine consumption that differed from those used in the meta-

analysis. Furthermore, additional studies were excluded because they lacked valid controls.

A formal quality assessment was not performed on individual studies. However, the inclusion/exclusion criteria, along with a systematic data extraction process, served as an inherent quality assessment mechanism.

Undetected spontaneous abortions

Wilcox and associates (28) showed that approximately 25% of biochemically detected pregnancies ended before being clinically detected. One might assume that early loss of pregnancy would follow the patterns of late spontaneous abortion among the various stratifications of caffeine consumption. The extent to which this assumption is valid determines the amount of error introduced into the meta-analysis.

Confounders

Various confounding factors were identified in the accepted articles (Table 8). Tables 9 and 10 illustrate how individual studies handled the important issue of confounders and how they have adjusted risk ratios accordingly. The most important common confounders appear to be concurrent smoking, alcohol use, maternal age over 35, and previous spontaneous abortion. Most other confounding factors would be equally distributed among the various stratifications of caffeine consumption. However, levels of smoking, alcohol use, and maternal age have been shown to have a positive

correlation with levels of caffeine consumption (9). Certainly, adjustment for these confounders by multivariate analysis would be a desirable component to interpreting the influence of confounders on summary ratios. Unfortunately, due to the nature of data presentation in the individual studies, adjustment for these confounders was not possible. For example, when these studies reported the number of spontaneous abortions in caffeine-exposed and non-exposed groups, they did not specifically delineate the number of smokers in each caffeine stratification group. Consequently, we were unable to adjust the overall summary ratios for either outcome of the meta-analysis for smoking or any of the other confounders.

Risk of spontaneous abortion increases as the quantity of cigarettes smoked/d increases (9). In most of the five studies in the main analysis for spontaneous abortion, the odds ratios were not altered significantly even after the researchers adjusted for smoking and other confounders (as reported in each study). It would have been interesting to have been able to quantify the interrelationship between caffeine and smoking on spontaneous abortion.

The issue of smoking as a confounder in the fetal growth analysis is an interesting one. It has been postulated that the negative effects of caffeine on fetal growth occur in the third trimester, as this is the time when the greatest rate of growth occurs (5). Nicotine is known to increase the rate of metabolism of caffeine in humans. Pregnant women have only one-third the capacity to metabolize caffeine in the third trimester of pregnancy (2). This reduction, combined with the fact that the fetus is unable to metabolize caffeine, is thought to be an important contributing factor in the proposed fetal growth retardation. The issue is complex because smoking is known to retard fetal growth by a different mechanism (9). Our accepted studies illustrate these conflicting results as two of the studies (20,23) actually showed higher risk ratios for the relationship of caffeine consumption with low birth weight after adjustment for smoking. In direct contrast to this observation, the study by Fortier and colleagues (19) showed that the risks from caffeine consumption and smoking were more than additive in their contributions to the risk of fetal growth retardation. Although the results of any particular study involved in the analyses were not significantly altered by adjustment for confounders, it should be noted that the summary effect of the confounder may have inflated the results of the meta-analyses.

Our results suggest a small but statistically significant increase in the risks for spontaneous abortion and for low birth weight babies in pregnant women consuming more than 150 mg of caffeine per d. Pregnant women should be encouraged to be aware of dietary caffeine

intake and to consume less than 150 mg of caffeine/d from all sources throughout pregnancy. Future research should include reliable methods of caffeine measurement, standardized control and caffeine exposure groups, and a standard approach to control for confounders.

REFERENCES

- Goyan JE. Food and Drug Administration news release no.P80-36. Washington, DC: September 1980;4.
- Knutti R, Rothweiler H, Schlatter C. The effect of pregnancy on the pharmacokinetics of caffeine. *Arch Toxicol.* 1982; 5(Suppl): 187-92.
- Weathersbee PS, Lodge R. Caffeine: its direct and indirect influence on reproduction. *J Reprod Med.* 1977;19:55-63.
- Thayer PS, Palm PE. A current assessment of the mutagenic and teratogenic effects of caffeine. *CRC Crit Rev Toxicol.* 1975;3:345.
- Wilson JG, Scott WJ. The teratogenic potential of caffeine in laboratory animals. In: Dews PB, ed. *Caffeine: perspectives from recent research.* Berlin: Springer-Verlag; 1984:165-87.
- Einarson TR, Leeder JS, Koren G. A method for meta-analysis of epidemiologic studies. *Drug Intell Clin Pharm.* 1988;22:813-24.
- Arbuckle TE, Sherman GJ. An analysis of birth weight by gestational age in Canada. *Can Med Assoc J.* 1989;140:157-60, 165.
- Gilbert RM, Marshman JA, Schwieder M, Berg R. Caffeine content of beverages as consumed. *Can Med Assoc J.* 1976;114: 205-8.
- Golding J. Reproduction and caffeine consumption—a literature review. *Early Hum Dev.* 1995;43:1-14.
- Martin JC. An overview: maternal nicotine and caffeine consumption and offspring outcome. *Neurobehav Toxicol Teratol.* 1982;4: 421-7.
- Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst.* 1959;22: 719-48.
- Miettinen O. Estimability and estimation in case-referent studies. *Am J Epidemiol.* 1976;103:226-35.
- Watkinson B, Fried PA. Maternal caffeine use before, during and after pregnancy and effects upon offspring. *Neurobehav Toxicol Teratol.* 1985;7:9-17.
- Barr HM, Streissguth AP. Caffeine use during pregnancy and child outcome: a seven year prospective study. *Neurotoxicol Teratol.* 1991;13:441-448.
- Weathersbee PS, Olsen LK, Lodge JR. Caffeine in pregnancy. *Postgrad Med.* 1977;62:64-9.
- Srisuphan W, Bracken M. Caffeine consumption during pregnancy and association with late spontaneous abortion. *Am J Obstet Gynecol.* 1986;154:14-20.
- Armstrong BG, McDonald AD, Sloan M. Cigarette, alcohol, and coffee consumption and spontaneous abortion. *Am J Public Health.* 1992;82:85-7.
- Beaulac-Baillargeon L, Desrosiers C. Caffeine-cigarette interaction on fetal growth. *Am J Obstet Gynecol.* 1987;157:1236-40.
- Fortier I, Marcoux S, Beaulac-Baillargeon L. Relation of caffeine intake during pregnancy to intrauterine growth retardation and preterm birth. *Am J Epidemiol.* 1993;137:931-40.
- Martin TR, Bracken M. The association between low birth weight and caffeine consumption during pregnancy. *Am J Epidemiol.* 1987;126:813-21.
- Fenster L, Eskenazi B, Windham GC, Swan SH. Caffeine consumption during pregnancy and fetal growth. *J Public Health.* 1991;81:458-61.
- McDonald AD, Armstrong BG, Sloan M. Cigarette, alcohol, and coffee consumption and prematurity. *Am J Public Health.* 1992: 87-90.
- Caan BJ, Goldhaber MK. Caffeinated beverages and low birth-weight: a case-control study. *Am J Public Health.* 1989;79:1299-1300.

24. Olsen J, Overvad K, Frische G. Coffee consumption, birthweight and reproductive failures. *Epidemiology*. 1991;2:370-374.
25. Peacock JL, Bland JM, Anderson HR. Effects on birthweight of alcohol and caffeine consumption in smoking women. *J Epidemiol Community Health*. 1991;45:159-63.
26. Shu XO, Hatch MC, Mills J, Clemens J, Susser M. Maternal smoking, alcohol drinking, caffeine consumption, and fetal growth: results from a prospective study. *Epidemiology*. 1995;6:115-20.
27. Fenster L, Eskenazi B, Windham GC, Swan SH. Caffeine consumption during pregnancy and spontaneous abortion. *Epidemiology*. 1991;2:168-74.
28. Wilcox AJ, Weinberg CR, Baird DD. Risk factors for early pregnancy loss. *Epidemiology*. 1990;1:382-5.
29. Infante-Rivard C, Fernandez A, Gauthier R, David M, Rivard GE. Fetal loss associated with caffeine intake before and during pregnancy. *JAMA*. 1993;270:2940-3.
30. Mills JL, Holmes LB, Aarons JH, et al. Moderate caffeine use and the risk of spontaneous abortion and intrauterine growth retardation. *JAMA*. 1993;269:593-7.
31. Dominguez-Rojas V, Juanes-Pardo JR, Astasio-Arbiza P, Ortega-Molina P, Gordillo-Florenico E. Spontaneous abortion in a hospital population: are tobacco and coffee intake risk factors? *Eur J Epidemiol*. 1994;10:665-8.
32. Linn S, Schoenbaum SC, Monson RR, Rosner B, Stubblefield PG, Ryan KJ. No association between coffee consumption and adverse outcomes of pregnancy. *N Engl J Med*. 1982;306:141-5.
33. Furuhashi N, Sato S, Suzuki M, Hiruta M, Tanaka M, Takahashi. Effects of caffeine ingestion during pregnancy. *Gynecol Obstet Invest*. 1985;19:187-91.
34. Fried PA, O'Connell. A comparison of the effects of prenatal exposure to tobacco, alcohol, cannabis and caffeine on birth size and subsequent growth. *Neurotoxicol Teratol*. 1987;9:79-85.
35. Barr HM, Streissguth AP, Martin DC, Herman CS. Infant size at eight months of age: relationship to maternal use of alcohol, nicotine, and caffeine during pregnancy. *Pediatrics*. 1984;74:336-41.
36. Rosenberg L, Mitchell AA, Shapiro S, Slone D. Selected birth defects in relation to caffeine-containing beverages. *JAMA*. 1982;247:1429-32.
37. Pastore LM, Savitz DA. Case-control study of caffeinated beverages and preterm delivery. *Am J Epidemiol*. 1995;141:61-9.
38. Laroque B, Kaminski M, Lelong N, Subtil D, Dehaene P. Effects on birth weight of alcohol and caffeine consumption during pregnancy. *Am J Epidemiol*. 1993;137:941-50.
39. Van den berg BJ. Epidemiologic observations of prematurity: effects of tobacco, coffee, and alcohol. In: Reed DM, Stanley FJ, eds. *The epidemiology of prematurity*. Baltimore: Urban and Schwarzenberg Inc.; 1977:157-76.
40. Kline J, Levine B, Silverman J, et al. Caffeine and spontaneous abortion of known karyotype. *Epidemiology*. 1991;2:409-17.
41. Godel JC, Pabst HF, Hodges PE, Johnson KE, Froese GJ, Joffres MR. Smoking and caffeine and alcohol intake during pregnancy in a northern population: effect on fetal growth. *Can Med Assoc J*. 1992;147:181-8.
42. Peacock JL, Bland JM, Anderson HR. Preterm delivery: effects of socioeconomic factors, psychological stress, smoking, alcohol, and caffeine. *BMJ*. 1995;311:531-6.
43. Muñoz LM, Lönnerdal B, Keen CL, Dewey KG. Coffee consumption as a factor in iron deficiency anemia among pregnant women and their infants in Costa Rica. *Am J Clin Nutr*. 1988;48:645-51.
44. Wisborg K, Henriksen TB, Hedegaard M, Secher NJ. Smoking during pregnancy and preterm birth. *Br J Obstet Gynaecol*. 1996;103:800-5.
45. Fenster L, Swan SH, Windham GC, Neutra RR. Assessment of reporting consistency in a case-control study of spontaneous abortions. *Am J Epidemiol*. 1991;133:477-88.

MATERNAL SERUM PARAXANTHINE, A CAFFEINE METABOLITE, AND THE RISK OF SPONTANEOUS ABORTION

MARK A. KLEBANOFF, M.D., M.P.H., RICHARD J. LEVINE, M.D., M.P.H., REBECCA DERSIMONIAN, Sc.D., JOHN D. CLEMENS, M.D., AND DIANA G. WILKINS, Ph.D.

ABSTRACT

Background Whether the consumption of caffeine during pregnancy increases the risk of spontaneous abortion is controversial. Prior studies have determined caffeine consumption by questionnaire. We used a biologic marker, serum paraxanthine, a metabolite of caffeine, to measure the dose of caffeine.

Methods In a nested case-control study, we measured serum paraxanthine in 591 women who had spontaneous abortions at less than 140 days' gestation and in 2558 matched women from the same clinic who gave birth to live infants at 28 weeks' gestation or later and who had serum drawn on the same day of gestation as the women who had abortions. The women were enrolled in the Collaborative Perinatal Project during the period from 1959 to 1966, and serum paraxanthine was measured over 30 years later.

Results A total of 487 women who had spontaneous abortions (82 percent) and 2087 controls (82 percent) had quantifiable serum paraxanthine concentrations. However, the mean serum paraxanthine concentration was higher in the women who had spontaneous abortions than in the controls (752 vs. 583 ng per milliliter, $P < 0.001$). The odds ratio for spontaneous abortion was not significantly elevated in the women who had serum paraxanthine concentrations of 1845 ng per milliliter or lower, corresponding to the 95th percentile of the matched women. However, the adjusted odds ratio for spontaneous abortion among women with serum paraxanthine concentrations higher than 1845 ng per milliliter, as compared with women who had concentrations below 50 ng per milliliter, was 1.9 (95 percent confidence interval, 1.2 to 2.8).

Conclusions Only extremely high serum paraxanthine concentrations are associated with spontaneous abortion. This suggests that moderate consumption of caffeine is unlikely to increase the risk of spontaneous abortion. (N Engl J Med 1999;341:1639-44.)

©1999, Massachusetts Medical Society.

WHETHER consumption of caffeine during pregnancy increases the risk of spontaneous abortion is controversial. Several studies have indicated that even moderate caffeine consumption is associated with a risk of fetal loss that is more than double the risk in women who do not consume caffeine.^{1,2} Others have reported that the risk is elevated only for women who consume three or more cups of coffee per day³ or those who consume large amounts of caffeine and

who are nauseated during pregnancy.⁴ Still other studies have found no increase in risk even among women who consume large amounts of caffeine.⁵⁻⁸

Possible reasons for these discrepant results include small samples, particularly with respect to the groups of women who consumed large amounts of caffeine; retrospective or prospective ascertainment of caffeine consumption, spontaneous abortion, or both; changes in caffeine consumption during pregnancy; differences in the adequacy of statistical control for nausea during pregnancy; and problems of assessing caffeine intake on the basis of responses to a questionnaire.⁹ Wide variations in individual rates of caffeine metabolism make it difficult to translate even an accurately reported intake into serum concentrations of caffeine and its metabolites.¹⁰ We measured serum paraxanthine, the primary metabolite of caffeine, to determine whether the consumption of caffeine is associated with spontaneous abortion.

METHODS

Study Subjects

Our sample consisted of women enrolled in the Collaborative Perinatal Project, a prospective study of pregnancy, labor, and child development conducted at 12 sites in the United States from 1959 to 1966. The women in that study were enrolled when they presented for prenatal care and were followed for the remainder of their pregnancy. There were approximately 55,000 births to 42,000 women.¹¹ Although no information was collected on the consumption of coffee, tea, or soft drinks, serum was obtained approximately every two months during pregnancy, at delivery, and six weeks after delivery. Information about vomiting was obtained at enrollment and at each prenatal visit, and gestational age was estimated on the basis of the reported first day of the last menstrual period.

A total of 830 women had early fetal losses (less than 140 days after the first day of the last menstrual period); serum was obtained during the pregnancy from 704 of these women. The relatively small number of women with early fetal losses was due to the late gestational age at which many women were enrolled in the study.¹¹ The women were stratified according to the clinical center and the day of gestation on which the earliest serum sample was obtained. For the women with early fetal losses in each stratum, we selected four times the number of women at the same center who gave birth to live infants after at least 28 weeks of gestation and who had serum drawn on the same day of gestation (con-

From the Division of Epidemiology, Statistics, and Prevention Research, National Institute of Child Health and Human Development, Bethesda, Md. (M.A.K., R.J.L., R.D., J.D.C.); and the Center for Human Toxicology, University of Utah, Salt Lake City (D.G.W.). Address reprint requests to Dr. Klebanoff at the Division of Epidemiology, Statistics, and Prevention Research, National Institute of Child Health and Human Development, National Institutes of Health, 6100 Bldg., Rm. 7B03, MSC 7510, Bethesda, MD 20892-7510, or at mk90h@nih.gov.

trols). Serum was obtained at an unknown time of day during clinic visits and during hospitalization for delivery.

Serum caffeine and paraxanthine were assayed. In a pilot study,¹² serum caffeine and paraxanthine concentrations were positively correlated with the caffeine consumption reported by pregnant women, but the serum paraxanthine concentration was more closely correlated with caffeine consumption than was the serum caffeine concentration, particularly among smokers. The serum concentration of paraxanthine is less sensitive than that of caffeine to very recent caffeine intake.¹⁰ Accordingly, our primary objective was to test the hypothesis that the mean serum paraxanthine concentration was higher in women who had spontaneous abortions than in women who delivered live infants. An additional objective was to determine whether there was a threshold above which the serum paraxanthine concentration was associated with spontaneous abortion.

Biochemical Assays

Serum caffeine and paraxanthine were measured with the use of high-performance liquid chromatography.¹² The limit of quantitation was established at 50 ng per milliliter for caffeine and paraxanthine; the limit of detection was 25 ng per milliliter. The intraassay and interassay coefficients of variation were less than 6.9 percent at 200, 800, and 2000 ng per milliliter. The laboratory personnel who performed the assays were unaware of the outcome of each pregnancy. Serum samples from the women who had spontaneous abortions and from the matched controls were analyzed in the same batch; the order of the samples varied from batch to batch. Since this analysis involved previously collected specimens from which identifying information had been removed, the Office of Human Subjects Research found it to be exempt from the requirement for approval by an institutional review board.

Statistical Analysis

Continuous variables were compared with use of Student's *t*-test or analysis of variance, and categorical variables were compared with use of the chi-square test. The standard deviation for the serum paraxanthine concentration was proportional to the mean, violating the assumptions of the *t*-test and analysis of variance. Log transformation of the serum paraxanthine values solved this problem. Since the results with the use of log-transformed data did not differ substantially from the results with the use of untransformed data, only the latter are reported here. The association between the serum paraxanthine concentration and spontaneous abortion was analyzed by conditional logistic regression.¹³

RESULTS

There were 704 women who had early fetal losses and 2816 controls. Since the Collaborative Perinatal Project had only one code for all early fetal losses, the original study records were reviewed to identify the women who had spontaneous abortions. Forty-six of the women with early fetal losses had induced abortions, ectopic pregnancies, or iatrogenic termination of pregnancy or died during pregnancy. For the group of 658 women in whom fetal loss was due to spontaneous abortion, it was not possible to determine from a review of the records whether serum drawn on the day of spontaneous abortion was obtained before or after the event, so the 57 women in whom the serum sample had been obtained on the day of abortion were excluded from the analysis. In an additional 10 women who had spontaneous abortions, insufficient serum was available for analysis. The exclusion of these 113 women required the exclusion of 208 matched controls, and in 50 additional controls, in-

sufficient serum was available for analysis. The final study group thus comprised 591 women who had spontaneous abortions and 2558 matched controls.

We compared the group of 591 women with spontaneous abortions whose serum samples were available for analysis with the group of 193 women with spontaneous abortions for whom serum samples were not available. The median date of enrollment was January 1963 for the former group and December 1960 for the latter ($P < 0.001$), suggesting that study procedures improved over time, and the two groups of women were enrolled on day 76 and day 80 of gestation, respectively ($P = 0.02$). On average, 24 days elapsed from enrollment to the spontaneous abortion for women for whom serum was available, as compared with 12 days for women for whom serum was not available ($P < 0.001$). The proportion of women from whom serum was obtained varied significantly among the study sites, ranging from 70 to 100 percent.

The characteristics of the women who had spontaneous abortions and the controls are shown in Table 1. Serum was drawn on the same day of gestation in the two groups. The mean duration of pregnancy was slightly more than 14 weeks among the women who had spontaneous abortions and was 39 weeks among the controls. The median interval from the collection of serum to abortion was 17 days. The women who had spontaneous abortions were significantly older than the controls ($P < 0.001$), more likely to smoke ($P < 0.001$), and less likely to have vomited ($P < 0.001$) or to have taken medications containing caffeine ($P = 0.02$) during pregnancy.

The serum paraxanthine concentrations are shown in Table 2 according to the outcome of pregnancy and maternal characteristics. In both the group of women who had spontaneous abortions and the control group, higher serum paraxanthine concentrations were associated with increasing age, white race, smoking, and the absence of vomiting during pregnancy. The serum paraxanthine concentration was positively associated with the level of education only in the control group. In almost every category of each of these characteristics, the serum paraxanthine concentration was higher in the women who had spontaneous abortions than in the controls (Table 2).

A total of 487 women who had spontaneous abortions (82 percent) and 2087 controls (82 percent) had quantifiable serum paraxanthine concentrations ($P = 0.64$). However, the mean serum paraxanthine concentration was significantly higher in the abortion group than in the control group (752 vs. 583 ng per milliliter, $P < 0.001$). The odds ratios for spontaneous abortion according to the serum paraxanthine concentration, with the women who had unquantifiable serum paraxanthine concentrations (< 50 ng per milliliter) used as the reference group and with adjustment for smoking status, age, and race or ethnic group, are

TABLE 1. CHARACTERISTICS OF THE WOMEN WHO HAD SPONTANEOUS ABORTIONS AND THOSE WHO GAVE BIRTH TO LIVE INFANTS (CONTROLS).

CHARACTERISTIC	WOMEN WITH SPONTANEOUS ABORTIONS (N=591)		CONTROLS (N=2558)	
Length of gestation when blood sample obtained (days)				
Mean	78		78	
Median	76		76	
Length of gestation at time of abortion or delivery (days)				
Mean	100		274	
Median	99		278	
Interval between blood sample and abortion (days)				
Mean	22		—	
Median	17			
Mean age at enrollment (yr)	27		25	
Smoker (%)*	46		39	
Vomiting since last menstrual period (%)†	38		56	
Education (%)‡				
<12 yr	50		40	
12 yr	31		37	
≥13 yr	19		23	
Race or ethnic group (%)				
White	61		65	
Black	33		29	
Other or unknown§	6		6	
Diagnosis of diabetes mellitus before pregnancy (%)	3		2	
Use of medications containing caffeine (%)				
During month serum sample was obtained	6		9	
During month before serum sample was obtained	4		6	

*Data were available for 591 women who had spontaneous abortions and 2542 controls.

†Data were available for 515 women who had spontaneous abortions and 2544 controls.

‡Data were available for 447 women who had spontaneous abortions and 2518 controls.

§Eleven percent of the women in this category were Asian, 85 percent were Puerto Rican, and 4 percent were unclassified with respect to race or ethnic group.

shown in Figure 1. Data on vomiting during pregnancy and educational level were missing for a substantial number of women. However, adjustment for these factors did not substantially change the odds ratios (data not shown). The increased risk of spontaneous abortion was almost entirely restricted to women with serum paraxanthine concentrations higher than 1845 ng per milliliter, corresponding to the 5 percent of controls with the highest concentrations (adjusted odds ratio, 1.9; 95 percent confidence interval, 1.2 to 2.8). For the remainder of the analyses, the women were grouped according to their serum paraxanthine concentrations (<50 ng per milliliter, 50 to 1845 ng per milliliter, and >1845 ng per milliliter, corresponding roughly to <20th percentile of serum paraxanthine values in the controls, 20th to 95th percentile, and >95th percentile).

The association between serum paraxanthine con-

TABLE 2. MEAN SERUM CONCENTRATIONS OF PARAXANTHINE ACCORDING TO THE OUTCOME OF PREGNANCY AND MATERNAL CHARACTERISTICS.

CHARACTERISTIC	WOMEN WITH SPONTANEOUS ABORTIONS		CONTROLS	
	SERUM PARAXANTHINE	P VALUE	SERUM PARAXANTHINE	P VALUE
	ng/ml		ng/ml	
Age		<0.001*		<0.001*
<20 yr	447		359	
20-24 yr	512		498	
25-29 yr	835		681	
30-34 yr	1068		713	
≥35 yr	1024		870	
Education		0.74*		0.01*
<12 yr	714		515	
12 yr	748		593	
≥13 yr	680		612	
Smoker		<0.001		<0.001
Yes	899		762	
No	626		474	
Vomiting		<0.001		<0.001
Yes	585		536	
No	849		646	
Race or ethnic group		<0.001		<0.001
White	931		679	
Black	435		357	
Other or unknown	652		659	

*The P value is for trend.

centrations and spontaneous abortion according to other factors is shown in Table 3. All odds ratios were adjusted for maternal age, smoking status, and race or ethnic group. For women with very high serum paraxanthine concentrations, the odds ratio for spontaneous abortion did not differ significantly according to whether the abortion occurred at 100 or more days of gestation or earlier (100 days was the median interval), whether the serum sample had been obtained 17 or fewer days before spontaneous abortion or more than 17 days earlier (17 days was the median interval), or whether the woman had or had not vomited since her last menstrual period.

To determine whether differences in desiccation over time affected the results, we measured serum sodium in 3057 samples, using direct potentiometry with ion-selective electrodes. The mean (\pm SD) serum sodium concentration was 137 ± 27 mmol per liter. After standardization of the serum paraxanthine concentration to a serum sodium concentration of 135 mmol per liter, the results shown in Figure 1 were largely unchanged.

DISCUSSION

Our results indicate that the serum concentration of paraxanthine, the primary metabolite of caffeine, is higher in women who have spontaneous abortions than in women who give birth to live infants. However, the risk of spontaneous abortion is not in-

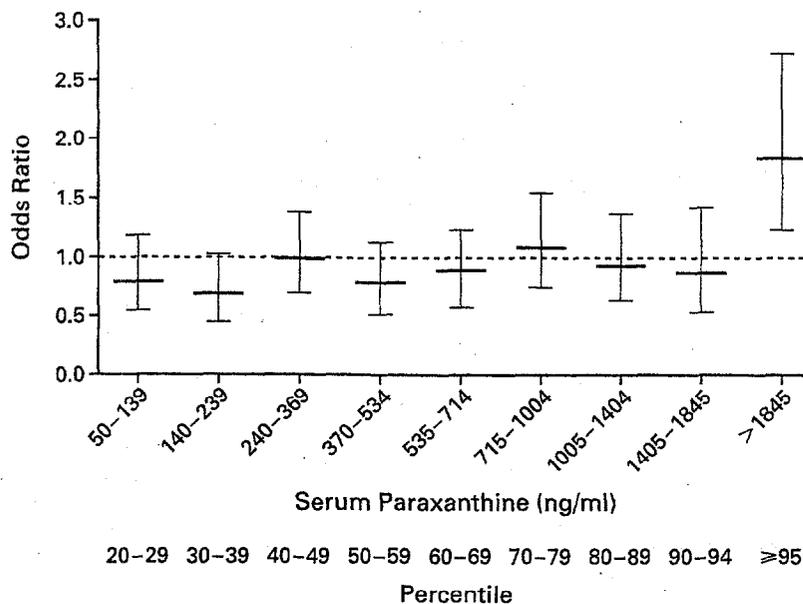


Figure 1. Adjusted Odds Ratios and 95 Percent Confidence Intervals for Spontaneous Abortion According to the Serum Concentration of Paraxanthine.

The reference category is values of less than 50 ng per milliliter. The odds ratios have been adjusted for smoking status, age, and race or ethnic group. The percentiles are for the serum paraxanthine values in the controls.

creased until extremely high serum paraxanthine concentrations are reached. Our results support previous studies showing that the consumption of large amounts of caffeine is associated with an increased risk of spontaneous abortion¹⁻⁴ but that moderate consumption does not increase the risk.³⁻⁸

Since there is no precise way to equate a serum paraxanthine concentration with an amount of caffeine intake, our results cannot directly answer the question of how much caffeine is safe during pregnancy. However, there may be indirect ways to answer this question. Our pilot study¹² involved women who had participated in the Birmingham, Alabama, study of infant growth in the mid-1980s.¹⁴ The highest caffeine intake in that cohort was 1530 mg per day (equivalent to approximately 15 cups of coffee). The highest measured serum paraxanthine concentration was 1165 ng per milliliter, which was substantially lower than the value at the 95th percentile in this study (1845 ng per milliliter). Even with allowance for volume loss during storage, the 95th percentile of serum paraxanthine in this study is higher than the highest value in the Birmingham study. Extrapolating from our pilot data, a 60-kg woman who did not smoke and who consumed 600 mg of caffeine (about 6 cups of coffee) per day or a 60-kg woman who smoked and who consumed 1100 mg of caffeine (about 11 cups of coffee) per day would have an estimated serum paraxanthine concentration of 1845 ng per milliliter.

Additional information to equate serum paraxanthine concentrations with caffeine intake comes from the California Child Health and Development Studies,¹⁵ involving a prospective cohort of pregnant women in the 1960s. In that study, women were asked about their intake of coffee and tea. Assuming that a cup of tea contains half the caffeine of a cup of coffee, the 95th percentile of caffeine intake was equivalent to 8.5 cups of coffee per day, which is consistent with the extrapolated data from our pilot study and conservatively suggests that the 95th percentile of caffeine intake in the current study was the equivalent of more than 5 cups of coffee per day.

Several caveats should be noted. First, the women in the Collaborative Perinatal Project were enrolled relatively late in gestation, and the majority of abortions occurred in the second trimester. Furthermore, karyotype analyses were not performed for any of the aborted fetuses. Fetuses aborted early in gestation are more likely to have chromosomal abnormalities than are fetuses aborted later.¹⁶ However, abortion of chromosomally normal fetuses is a more sensitive indicator of exogenous risk factors.¹⁶ The association between caffeine intake and spontaneous abortion has been reported to be similar for chromosomally normal and abnormal fetuses,⁸ suggesting either that caffeine increases the risk of loss for both types of fetuses or that the association is not causal.

Second, the Collaborative Perinatal Project recorded data on vomiting during pregnancy but not on

TABLE 3. ADJUSTED ODDS RATIOS FOR SPONTANEOUS ABORTION ACCORDING TO THE SERUM PARAXANTHINE CONCENTRATION AND ADDITIONAL FACTORS.

FACTOR	WOMEN WITH SPONTANEOUS ABORTIONS CONTROLS* TOTAL			ADJUSTED ODDS RATIO (95% CI)†
	number			
>17 Days between collection of serum and spontaneous abortion‡				
Serum paraxanthine, <50 ng/ml	50	240	290	1.0
Serum paraxanthine, 50-1845 ng/ml	206	879	1085	1.0 (0.6-1.5)
Serum paraxanthine, >1845 ng/ml	33	62	95	2.0 (1.2-3.6)
≤17 Days between collection of serum and spontaneous abortion‡				
Serum paraxanthine, <50 ng/ml	54	231	285	1.0
Serum paraxanthine, 50-1845 ng/ml	212	1078	1290	0.8 (0.5-1.1)
Serum paraxanthine, >1845 ng/ml	36	68	104	1.8 (1.0-3.1)
Spontaneous abortion at ≥100 days' gestation‡				
Serum paraxanthine, <50 ng/ml	51	255	306	1.0
Serum paraxanthine, 50-1845 ng/ml	205	1042	1247	1.0 (0.7-1.4)
Serum paraxanthine, >1845 ng/ml	33	79	112	1.7 (1.0-3.0)
Spontaneous abortion at <100 days' gestation‡				
Serum paraxanthine, <50 ng/ml	53	216	269	1.0
Serum paraxanthine, 50-1845 ng/ml	213	915	1128	0.8 (0.6-1.1)
Serum paraxanthine, >1845 ng/ml	36	51	87	2.1 (1.2-3.6)
Vomiting since last menstrual period§				
Serum paraxanthine, <50 ng/ml	49	190	239	1.0
Serum paraxanthine, 50-1845 ng/ml	221	860	1081	1.1 (0.7-1.6)
Serum paraxanthine, >1845 ng/ml	47	73	120	2.2 (1.2-4.0)
No vomiting since last menstrual period§				
Serum paraxanthine, <50 ng/ml	49	277	326	1.0
Serum paraxanthine, 50-1845 ng/ml	133	1087	1220	0.7 (0.5-1.0)
Serum paraxanthine, >1845 ng/ml	16	57	73	1.8 (0.8-3.8)

*Controls were women who gave birth to live infants after at least 28 weeks of gestation, who were at the same clinic as the women who had spontaneous abortions, and who had serum drawn on the same day of gestation as the women who had spontaneous abortions.

†Odds ratios have been adjusted for maternal age, smoking status, and race or ethnic group. The reference group is women in both study groups who had serum paraxanthine concentrations of less than 50 ng per millimeter. CI denotes confidence interval.

‡Controls were randomly matched with women who had spontaneous abortions in this stratum.

§Data on vomiting were available for 515 women who had spontaneous abortions and 2544 controls.

nausea. Nausea is thought to be a marker for a healthy pregnancy, and nausea and food aversions may cause women to reduce their consumption of coffee and other foods with strong aromas.¹⁷ If so, then even the elevated risk of spontaneous abortion among women with extremely high serum paraxanthine concentrations may simply reflect the fact that a viable pregnancy causes a woman to reduce her intake of caffeine. Since nauseated women consume less caffeine than women without nausea and also have a reduced risk of spontaneous abortion, the likely effect of incomplete data on nausea and vomiting would be to overestimate the level of risk associated with high levels of caffeine consumption.

Third, although unlike the serum caffeine concentration, the serum paraxanthine concentration does not fluctuate greatly during the day, the serum half-

lives of the two substances are similar: approximately 5 hours during the first trimester and 10 hours during the second trimester.^{18,19} Therefore, serum paraxanthine is a marker only of short-term caffeine intake. Although we are unaware of any data that confirm this observation, the likelihood that caffeine intake is relatively constant from day to day provides support for the use of serum paraxanthine as a biologic marker of caffeine intake.

Fourth, the serum samples we used had been stored for over 30 years. The stability of paraxanthine during long-term storage at -20°C is unknown. In our pilot study,¹² we found that the paraxanthine concentration in serum samples stored for eight years at -70°C was closely correlated with the reported caffeine intake, suggesting that paraxanthine remains stable under these conditions. We quantified paraxan-

thine in 82 percent of the serum samples from the Collaborative Perinatal Project and detected it below the limit of quantitation in an additional 4 percent. In the California Child Health and Development Studies, 13 percent of the women reported that they consumed neither coffee nor tea. This finding is consistent with our 85 percent detection rate and suggests that marked deterioration of paraxanthine was unlikely to have occurred.

Our study has several strengths. The serum samples were collected in the 1960s, when few pregnant women were advised to reduce their intake of caffeine. Per capita coffee consumption in the United States peaked in 1962 and then declined, particularly among people less than 40 years old.²⁰ The Collaborative Perinatal Project is therefore likely to have enrolled many women who consumed large amounts of caffeine. Most investigators have found it difficult to enroll sufficient numbers of women who consumed large quantities of caffeine.^{1,4,6,7}

In conclusion, if caffeine causes spontaneous abortion, it does so only at serum paraxanthine concentrations, and presumably levels of caffeine intake, that were uncommonly high in the 1960s, and these high levels are probably even less common now.

Supported by a contract (NO1-HD-7-3262) from the National Institutes of Health.

REFERENCES

1. Infante-Rivard C, Fernandez A, Gauthier R, David M, Rivard G-E. Fetal loss associated with caffeine intake before and during pregnancy. *JAMA* 1993;270:2940-3.
2. Srisuphan W, Bracken MB. Caffeine consumption during pregnancy and association with late spontaneous abortion. *Am J Obstet Gynecol* 1986;154:14-20.
3. Dlugosz L, Belanger K, Heilenbrand K, Holford TR, Leaderer B, Bracken MB. Maternal caffeine consumption and spontaneous abortion: a prospective cohort study. *Epidemiology* 1996;7:250-5.
4. Fenster L, Eskenazi B, Windham GC, Swan SH. Caffeine consumption during pregnancy and spontaneous abortion. *Epidemiology* 1991;2:168-74.
5. Armstrong BG, McDonald AD, Sloan M. Cigarette, alcohol, and coffee consumption and spontaneous abortion. *Am J Public Health* 1992;82:85-7.
6. Mills JL, Holmes LB, Aarons JH, et al. Moderate caffeine use and the risk of spontaneous abortion and intrauterine growth retardation. *JAMA* 1993;269:593-7.
7. Fenster L, Hubbard AE, Swan SH, et al. Caffeinated beverages, decaffeinated coffee, and spontaneous abortion. *Epidemiology* 1997;8:515-23.
8. Kline J, Levin B, Silverman J, et al. Caffeine and spontaneous abortion of known karyotype. *Epidemiology* 1991;2:409-17.
9. Eskenazi B. Caffeine during pregnancy: grounds for concern? *JAMA* 1993;270:2973-4.
10. Lelo A, Miners JO, Robson R, Birkett DJ. Assessment of caffeine exposure: caffeine content of beverages, caffeine intake, and plasma concentrations of methylxanthines. *Clin Pharmacol Ther* 1986;39:54-9.
11. Niswander KR, Gordon M. The women and their pregnancies: the collaborative perinatal study of the National Institute of Neurological Diseases and Stroke. Philadelphia: W.B. Saunders, 1972.
12. Klebanoff MA, Levine RJ, Dersimonian R, Clemens JD, Wilkins DG. Serum caffeine and paraxanthine as markers for reported caffeine intake in pregnancy. *Ann Epidemiol* 1998;8:107-11.
13. SAS/STAT software: changes and enhancements, release 6.07. Technical report P-229. Cary, N.C.: SAS Institute, 1992:433-81.
14. Bergsjø P, Hoffman HJ, Davis RO, et al. Preliminary results from the Collaborative Alabama and Scandinavian Study of Successive Small-for-Gestational Age Births. *Acta Obstet Gynecol Scand* 1989;68:19-25.
15. van den Berg BJ, Christianson RE, Oechsli FW. The California Child Health and Development Studies of the School of Public Health, University of California at Berkeley. *Paediatr Perinat Epidemiol* 1988;2:265-82.
16. Kline J, Stein Z, Susser M. Conception to birth: epidemiology of prenatal development. New York: Oxford University Press, 1989:81-117.
17. Stein Z, Susser M. Miscarriage, caffeine, and the epiphenomena of pregnancy: the causal model. *Epidemiology* 1991;2:163-7.
18. Yesair DW, Branfman AR, Callahan MM. Human disposition and some biochemical aspects of methylxanthines. In: Spiller GA, ed. The methylxanthine beverages and foods: chemistry, consumption, and health effects. New York: Alan R. Liss, 1984:215-33.
19. Aldridge A, Bailey J, Neims AH. The disposition of caffeine during and after pregnancy. *Semin Perinatol* 1981;5:310-4.
20. Gilbert RM. Caffeine consumption. In: Spiller GA, ed. The methylxanthine beverages and foods: chemistry, consumption, and health effects. New York: Alan R. Liss, 1984:185-213.

with Pap smears. Unfortunately, blanket screening for HPV in sexually active young women will probably do more harm than good because of the high prevalence of HPV and the tendency of the infection and lesions to regress.¹⁰ The key is to identify persistent, type-specific infection, but surprisingly, there is no consensus on what constitutes persistent infection.

It seems rational to define persistent infection as present when the same type of HPV DNA is detected at least twice over a period of one or more years. Since 20 percent of new HPV infections persist for at least one year in young women, and since the longer a young woman is infected the more likely she is to have continued persistent infection,¹⁰ this definition should be clinically useful in identifying women at high risk for high-grade cervical disease (i.e., cervical intraepithelial neoplasia of grade 2 or worse). Screening for HPV will detect prevalent infection and identify the women with a persistent type-specific infection who are likely to be at high risk for continued persistence of HPV and cervical cancer. The detection of persistent infection in older women should be even more useful as a method of identifying women with high-grade cervical neoplasia. Additional studies are needed to define the natural history of HPV among women of various ages and ethnic and racial backgrounds and those at various degrees of risk and to assess the value of these methods of detecting lesions with a high likelihood of progression.

In sum, cervical cancer often begins with the sexual transmission of HPV to a woman who is susceptible to persistent infection. Over time, the HPV lesion progresses to invasive cervical cancer. The prevention of HPV infection, through vaccination, public health measures, or identification through cytologic or molecular screening, is needed to fulfill the criteria of Hill for definitive evidence of causation and to eliminate cervical cancer.

ROBERT D. BURK, M.D.

Albert Einstein College of Medicine
Bronx, NY 10461-1602

REFERENCES

- zur Hausen H. Papillomavirus in anogenital cancer: the dilemma of epidemiologic approaches. *J Natl Cancer Inst* 1989;81:1680-2.
- IARC monographs on the evaluation of carcinogenic risks to humans. Vol. 64. Human papillomaviruses. Lyon, France: International Agency for Research on Cancer, 1995:409.
- Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295-300.
- Herrero R, Hildesheim A, Conception B, et al. A population-based study of human papillomavirus infection and all grades of cervical neoplasia in rural Costa Rica. *J Natl Cancer Inst* (in press).
- Bosch FX, Manos MM, Munoz N, et al. Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. *J Natl Cancer Inst* 1995;87:796-802.
- Wallin K-L, Wiklund E, Ångström T, et al. Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer. *N Engl J Med* 1999;341:1633-8.
- Ho GY, Burk RD, Klein S, et al. Persistent genital human papillomavirus infection as a risk factor for persistent cervical dysplasia. *J Natl Cancer Inst* 1995;87:1365-71.
- Liaw KL, Glass AG, Manos MM, et al. Detection of human papillomavirus DNA in cytologically normal women and subsequent cervical squamous intraepithelial lesions. *J Natl Cancer Inst* 1999;91:954-60.
- Sun X-W, Kuhn L, Ellerbrock TV, Chiasson MA, Bush TJ, Wright TC Jr. Human papillomavirus infection in women infected with the human immunodeficiency virus. *N Engl J Med* 1997;337:1343-9.
- Ho GY, Bierman R, Beardsley L, Chang CJ, Burk RD. Natural history of cervicovaginal papillomavirus infection in young women. *N Engl J Med* 1998;338:423-8.

©1999, Massachusetts Medical Society.

CAFFEINE — FILTERING THE FACTS

THE article by Klebanoff and colleagues¹ in this issue of the *Journal* adds to the growing body of literature examining the association of caffeine intake with adverse pregnancy outcomes and developmental problems. This study reports an association between spontaneous abortions, primarily in the second trimester, and unusually high levels of consumption of caffeine, equivalent to more than six cups of coffee a day.

Unlike previous studies, in which exposure was estimated from the women's reports of the amount of caffeinated beverages they consumed, Klebanoff et al. used a biologic marker of caffeine intake, the levels of the caffeine metabolite paraxanthine in serum, to estimate exposure. Estimation of exposure on the basis of women's own reports is problematic, both because of potentially inaccurate or biased reports and because the amount of caffeine in a cup of tea or coffee differs greatly, depending on the method of preparation.² The use of a biologic marker of exposure may help to reduce misclassification. Nevertheless, a single serum measurement, as used in the study by Klebanoff et al., may not accurately reflect a woman's exposure during the critical period of fetal development, since in early pregnancy the half-life of caffeine is short (approximately three to seven hours) and caffeine intake may vary markedly during pregnancy because of nausea and food aversion.

Given the results of the study by Klebanoff et al.,¹ can we now conclude that the consumption of caffeine at usual levels during pregnancy is safe? The weight of the evidence still suggests otherwise. A recent meta-analysis³ concluded that there is a small increase in the crude risk of both spontaneous abortion (odds ratio, 1.4; 95 percent confidence interval, 1.3 to 1.5) and low birth weight (odds ratio, 1.5; 95 percent confidence interval, 1.4 to 1.6) in women who consume more than 150 mg of caffeine, or roughly one to two cups of coffee, per day. Furthermore, on the basis of studies in animals, spontaneous abortion and low birth weight may not be the most sensitive end points to use in determining the in utero effects of caffeine consumption.

The most obvious effects of caffeine in nonpregnant adults are cardiovascular and neurobehavioral.

Knowledge of the effects of caffeine on the nervous system dates back centuries, to the time when Ethiopian shepherds noticed that their sheep stayed awake all night after grazing on wild coffee cherries. In fact, caffeine is the most widely consumed behaviorally active substance in the world, with the U.S. per capita consumption at nearly 3.5 kg of coffee per year, or more than 150 mg of caffeine per day.⁴ Caffeine, like nicotine, albeit to a lesser extent, meets some of the criteria of the World Health Organization and the American Psychiatric Association for a drug of dependence and acts on the dopaminergic system in the same way as amphetamines and cocaine.⁵ Caffeine and its metabolites are known to cross the blood-brain barrier readily in adults and fetuses alike. They act by blocking adenosine A₁ and A_{2A} receptors, which leads to secondary effects on many classes of neurotransmitters.

Intake of caffeine during pregnancy or the early postnatal period would be expected to have similar or more profound cardiovascular and neurobehavioral effects on fetuses and infants than on the caffeine-consuming mothers. This is because of caffeine's ready passage through the placenta, its presence in breast milk, its increasing half-life during pregnancy (up to 11 hours late in pregnancy) and in infants (up to 100 hours), the smaller body mass of fetuses and infants, and the inability of the fetus and neonate to detoxify caffeine.² Changes in fetal heart rate and breathing patterns have been noticed even when maternal intake of caffeine is moderate and when it has no apparent effects on the mother.⁶

Studies in animals provide supporting evidence of anatomical, neurochemical, and neurobehavioral changes in the developing nervous system after in utero or early postnatal exposure to caffeine.⁷ In animals, in utero or early postnatal exposure to moderate-to-high doses of caffeine has been associated with decreased brain weight and alterations in brain development and in learning and memory, and it has been associated with an increased incidence of apnea in adult rats.⁷⁻⁹ Unfortunately, the results of the few studies of the neurobehavioral effects in children after in utero or early postnatal exposure are conflicting.

There may be subgroups of pregnant women and children who are particularly susceptible to the effects of caffeine. Usually, legal regulation of allowable levels of exposure to a chemical includes a safety factor of 10 to 100 when data are extrapolated from animals to humans, in order to protect the most susceptible. However, many of the effects of caffeine in animals occur after doses that are within the range of human consumption, once metabolic differences among species are taken into account.⁷ Several factors alter susceptibility to the effects of caffeine, including exposure to other substances. For example, the half-life of caffeine is halved in smokers and doubled in women taking oral contraceptives. Caffeine

may potentiate the teratogenic and other adverse effects of substances such as ethanol, nicotine, and certain drugs, such as vasoconstricting agents.⁷ Premature infants may be more susceptible than term infants, given the longer half-life of caffeine in their bodies. Also, infants with poor nutritional status may be particularly vulnerable to the neurochemical and concomitant neurobehavioral effects of caffeine.⁷

Given these considerations, what is the appropriate course of action for clinicians and policy makers? In 1981, the Food and Drug Administration (FDA) advised pregnant women to "avoid caffeine-containing foods and drugs, if possible, or consume them only sparingly."¹⁰ This advisory has not percolated down to the general population. And no advisories have targeted breast-feeding women or parents of young children. In part because of our own dependence on our morning cup of coffee, and because of our inability to find strong associations with effects on health in humans, we have accepted that more than 75 percent of pregnant women consume caffeinated beverages. In 1997, the Center for Science in the Public Interest petitioned the FDA to require that caffeine content be disclosed on food labels, but Congress has not yet acted. Food labeling is an appropriate first step in educating the public and should be undertaken. In the meantime, health care providers should continue to counsel women who are pregnant or breast-feeding to limit their caffeine intake.

BRENDA ESKENAZI, PH.D.

University of California School of Public Health
Berkeley, CA 94720

REFERENCES

1. Klebanoff MA, Levine RJ, DerSimonian R, Clemens JD, Wilkins DG. Maternal serum paraxanthine, a caffeine metabolite, and the risk of spontaneous abortion. *N Engl J Med* 1999;341:1639-44.
2. Eskenazi B. Caffeine during pregnancy: grounds for concern? *JAMA* 1993;270:2973-4.
3. Fernandes O, Sabharwal M, Smiley T, Pastuszak A, Koren G, Einerson T. Moderate to heavy caffeine consumption during pregnancy and relationship to spontaneous abortion and abnormal fetal growth: a meta-analysis. *Reprod Toxicol* 1998;12:435-44.
4. Fredholm BB, Bättig K, Holmén J, Nehlig A, Zvartau EE. Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacol Rev* 1999;51:83-133.
5. Nehlig A. Are we dependent upon coffee and caffeine? A review on human and animal data. *Neurosci Biobehav Rev* 1999;23:563-76.
6. Salvador HS, Koos BJ. Effects of regular and decaffeinated coffee on fetal breathing and heart rate. *Am J Obstet Gynecol* 1989;160:1043-7. [Erratum, *Am J Obstet Gynecol* 1989;161:669.]
7. Nehlig A, Debray G. Potential teratogenic and neurodevelopmental consequences of coffee and caffeine exposure: a review on human and animal data. *Neurotoxicol Teratol* 1994;16:531-43.
8. Tye K, Pollard I, Karlsson L, Scheibner V, Tye G. Caffeine exposure in utero increases the incidence of apnea in adult rats. *Reprod Toxicol* 1993;7:449-52.
9. Fisher S, Guillet R. Neonatal caffeine alters passive avoidance retention in rats in an age- and gender-related manner. *Brain Res Dev Brain Res* 1997;98:145-9.
10. Caffeine in pregnancy. Rockville, Md.: Food and Drug Administration, 1981. (FDA 81-1081.)

©1999, Massachusetts Medical Society.

CSPI Center for
Science in the
Public
Interest

sher of ***Nutrition Action Healthletter***
1875 Connecticut Avenue, N.W., Suite 300
Washington D.C., 20009-5728

**Do Not
Fold**

Dockets Management Branch
U.S. Food and Drug Administration
Room 10-61
5630 Fishers Lane
Rockville, MD 20852

