



## Practice of Epidemiology

### A Little Bit Pregnant: Modeling How the Accurate Detection of Pregnancy Can Improve HIV Prevention Trials

Courtney A. Schreiber, Mary Sammel, Sharon L. Hillier, and Kurt T. Barnhart

*Initially submitted January 3, 2008; accepted for publication October 2, 2008.*

The prevalence of unplanned pregnancies contributes to the methodological challenges of human immunodeficiency virus (HIV) prevention trials. In this paper, the authors discuss the incidence of pregnancy, including chemical pregnancy, and how the different methods of pregnancy diagnosis could affect the statistical power and calculated outcomes of HIV prevention trials. Study sample size inflation factors are estimated to aid in the design of clinical trials. The authors used published data of women attempting pregnancy as well as data from HPTN 055 ([www.HPTN.org/research\\_studies/hptn055.asp](http://www.HPTN.org/research_studies/hptn055.asp)) to estimate the percentage of early study discontinuation that would be associated with 3 diagnostic methods for pregnancy in a hypothetical clinical trial. They classified chemical pregnancies as false-positive pregnancy tests and showed the sample size adjustment that would be necessary in clinical trial design because of the early discontinuations associated with pregnancy. There is a greater than 3-fold difference in the number of falsely positive pregnancy tests that will be detected, depending upon the diagnostic method used. The number of incident pregnancies may render HIV prevention trial sample sizes inadequate by as much as 50%. Pregnancy prevention and precise pregnancy diagnosis are critical to the statistical power and integrity of HIV prevention trials.

clinical trials as topic; contraception; HIV; pregnancy; primary prevention

Abbreviations: hCG, human chorionic gonadotropin; HIV, human immunodeficiency virus.

Rates of human immunodeficiency virus (HIV) infection in women are steadily increasing; according to the World Health Organization, 17.5 million women worldwide were infected with HIV in 2005. Vaginal microbicides and other pharmaceuticals such as preexposure prophylaxis regimens and vaccines are in development to curb HIV acquisition in vulnerable populations, including sexually active women of reproductive age (1). Pregnancy is another outcome that can result from sexual activity in this same population of women. Despite the fact that plans to become pregnant are exclusionary, the pregnancy rate in HIV prevention trials ranges from 23 to 70 per 100 woman-years (2). The prevalence of pregnancy in these trials has exceeded scientists' expectations and has the potential to undermine study results because most trials require that women who become pregnant stop using investigational drugs to avoid fetal exposure. In part because there is no "gold standard" for

screening women at high risk of pregnancy, and in part because both HIV and pregnancy originate from the same vector, incident pregnancies have materialized as a significant concern within HIV prevention trials. In this paper, we discuss the problem that the diagnosis of pregnancy within HIV prevention trials presents, the different methods of diagnosing pregnancy, and the impact that the different methods of pregnancy diagnosis have on trial outcomes.

Diagnosis of an incident pregnancy during an HIV prevention trial has repercussions beyond the known social, health, and economic consequences that unplanned pregnancies have in high-risk populations outside the clinical trial setting. The pregnant state may affect investigational drug use or efficacy and the virulence of HIV, which in turn may complicate trial outcomes (3, 4). The decreased use of investigational product that accompanies these pregnancies results in differential product/placebo exposure in the 2 trial

Correspondence to Dr. Courtney A. Schreiber, 3701 Market Street, Suite 810, Philadelphia, PA 19104 (e-mail: [cschreiber@obgyn.upenn.edu](mailto:cschreiber@obgyn.upenn.edu)).

arms, increased bias, and decreased statistical power. This decrement in statistical power comes from 2 sources: 1) the censored observations reduce the total number of participants left at risk for the event under study, and 2) in an intention-to-treat analysis, the effect size decreases because pregnant women who stop using the investigational product have a likelihood of seroconversion comparable to that of the control group. Some microbicides have contraceptive efficacy, so participants may expect reduced fertility while using a microbicide. The comparator arm may not have the same, or any, contraceptive efficacy, which could lead to different durations of study participation and even different rates of seroconversion between the 2 groups (3–6). The diagnosis of pregnancy during a clinical trial is not trivial, so how is it best made?

Several tools are available to diagnose pregnancy: the level of human chorionic gonadotropin (hCG) in the serum, the presence of hCG in the urine (through either high-sensitivity or low-sensitivity tests), pelvic ultrasound (not routinely available in the developing world), and a missed menstrual period. These diagnostic methods may detect pregnancy at different stages of progression, and not all pregnancies progress to delivery of a baby.

Clinical investigators have options regarding how to best diagnose pregnancy. The choice of test might be based on availability or cost, but the impact that the timing and accuracy of the pregnancy diagnosis may have on study participation should also be considered. With the hope of minimizing theoretically poor outcomes, many trials institute frequent pregnancy testing with the goal of early detection of pregnancy so that participants can stop using the investigational product early and for the duration of pregnancy. While this approach would appear to be appropriately cautious, the reality is that the results of many of these positive pregnancy tests do *not* indicate the beginning of a true clinical pregnancy but instead are chemical pregnancies characterized by a transient rise in the pregnancy hormone level only.

The term *chemical pregnancy* is used to describe a transiently positive hCG level not associated with the development of an embryo or even a gestational sac. With the advent of high-sensitivity urine pregnancy tests in the 1980s, early diagnosis of pregnancy is now widely possible. A pregnancy test can be positive as early as the first days of the approximate time of implantation or when traces of hCG are detectable in the maternal serum. It has been established that as many as 25% of pregnancies fail even before the woman has any subjective indication that she is pregnant, that is, before she misses her menstrual period (7–9) or has symptoms of pregnancy (10).

In the general population, most chemical pregnancies go unrecognized. Chemical pregnancies are diagnosed under active monitoring for pregnancy only if hCG levels are tested *prior* to a missed menstrual period. The American Society of Reproductive Medicine and the Society for Assisted Reproductive Technology distinguish chemical pregnancies from clinical pregnancies, which include spontaneous abortions. The transient rise in hCG that characterizes a chemical pregnancy is distinct from the widely recognized outcomes of a clinical pregnancy, which include

**Table 1.** Frequency of Pregnancy Outcomes in Population A: Women Attempting Pregnancy According to 3 Longitudinal Trials in the United States and China<sup>a</sup>

Pregnancy Outcome	Wilcox et al. (7)		Zinaman et al. (8)		Wang et al. (9)		Total	
	No.	%	No.	%	No.	%	No.	%
Total conceptions <sup>b</sup>	198		116		587		901	
Chemical pregnancies <sup>c</sup>	44	22	15	13	152	26	210	23
Abortions	18	9	21	18	49	8	89	10
Deliveries	135	68	79	68	373	63	587	65
Other <sup>d</sup>	1	0.5	1	0.8	12	2	14	2

<sup>a</sup> Percentages may not add to 100% because of dropout/loss to follow-up or an undetermined pregnancy outcome at the study's conclusion.

<sup>b</sup> Total conceptions with a known outcome.

<sup>c</sup> The term *chemical pregnancy* describes a transiently positive human chorionic gonadotropin level not associated with the development of an embryo or even a gestational sac.

<sup>d</sup> Includes ectopic pregnancies, molar pregnancies, and induced abortions.

spontaneous and induced abortions, ectopic pregnancy, and delivery. In the absence of routine use of ultrasound, a chemical pregnancy could be defined by the combination of a low peak in hCG (<100 mIU/mL), rapid fall in urinary or serum hCG concentration, and lack of substantial delay in onset of the next menstrual period to help differentiate this entity from a clinical pregnancy (7, 9, 11).

In this paper, we demonstrate that, given the prevalence of chemical pregnancies, differential detection of chemical and clinical pregnancies has important implications for clinical trials. To raise consciousness about this issue, and to aid in future protocol development, we modeled the ways in which different methods of pregnancy diagnosis could affect the statistical power and integrity of HIV prevention trials. We then estimated inflation factors for study sample size to provide a metric by which to compare the impact of the different diagnostic methods on clinical trials.

## MATERIALS AND METHODS

We utilized 2 different sources to calculate the prevalence of the different possible outcomes from a positive hCG test. Three previously published longitudinal studies of women attempting pregnancy (7–9) provided information on a total of 939 women, 2,700 cycles, and 901 total conceptions with known outcomes (population A). The combined data for the different pregnancy outcomes from these trials are shown in Table 1. In summary, all women in these trials collected and stored daily urine samples while attempting pregnancy, and they were instructed to present for care if a pregnancy was suspected. The urine samples were later assessed by using quantitative hCG tests, and outcomes were correlated with the presence or absence of clinical pregnancy during each cycle. These 3 trials illustrate the concept of chemical pregnancies by showing that hCG can be detected in urine during cycles when a pregnancy was never suspected.

**Table 2.** Frequency of Pregnancy Outcomes in Population B: Women Avoiding Pregnancy in the HPTN 055 Trial in South Africa

Pregnancy Outcome	No.	%
Total conceptions	105	
Chemical pregnancies <sup>a,b</sup>	18	17
Abortions <sup>c</sup>	55	53
Deliveries <sup>d</sup>	32	30

<sup>a</sup> The term *chemical pregnancy* describes a transiently positive human chorionic gonadotropin level not associated with the development of an embryo or even a gestational sac.

<sup>b</sup> Pregnancy test positive for  $\leq 30$  days.

<sup>c</sup> Pregnancy test positive for 31–180 days.

<sup>d</sup> Pregnancy test positive for  $\geq 181$  days.

Estimates of pregnancy outcomes for women avoiding pregnancy (population B) were derived from the HPTN 055 HIV Prevention Preparedness Trial (12) and are shown in Table 2. HPTN 055 is a prospective cohort study initiated to prepare sites for implementation of HPTN 035, a phase 3 safety and effectiveness study of 2 vaginal microbicides to prevent HIV. The study was conducted at 2 HPTN study sites in South Africa. Women were followed for up to 12 months, and HIV seroconversion was the primary outcome. Study participants were asked to avoid pregnancy for the trial duration even though no drug was tested. Details regarding the specific populations enrolled in this longitudinal cohort have been published previously (13). Most women (83% by self-report) were using a contraceptive method. The most common were injectable methods (32.7%) and male condoms (32.4%), followed by oral contraceptives (15.0%) and withdrawal or the rhythm method (13.6%) (2). Participants reported to the study site monthly for sexually transmitted infection and pregnancy testing. Pregnancy data were recorded as the number of consecutive months of positive urine hCG concentrations (or “pregnancy period”).

We chose 3 protocols for diagnosing pregnancy: 1) monthly urine pregnancy tests, 2) monthly pregnancy tests followed by a confirmatory pregnancy test at a 1-week interval should the initial test be positive, and 3) pregnancy tests following missed menses only.

- Protocol 1 tests each study participant monthly by calendar day, regardless of timing in the menstrual cycle. Monthly pregnancy tests will detect a positive hCG concentration as early as 26 days after the last menstrual period to as late as 42 days after the last menstrual period (the most advanced gestation detectable given a negative test in the prior month) depending upon menstrual cycle length and the timing of conception.
- Protocol 2 adds a follow-up test at a 1-week interval to protocol 1. We chose this option based on how the literature helps to distinguish between a chemical and clinical pregnancy (7, 10, 14–16). On the basis of longitudinal data, very few chemical pregnancies still show detectable hCG levels 1 week after the expected menstrual period, and the vast majority of clinical pregnancies can be de-

tected by 1 week after an expected menstrual period is missed. Of 40 women who had a chemical pregnancy (7), only 8 (20%) experienced a menstrual cycle lasting longer than 35 days (approximately 1 month plus 1 additional week). Although data on menstrual cycle lengths demonstrate variability both within and among women (15, 17), 28 days is considered the population-average cycle length, and when women present with an hCG result of 100 mIU/mL or less, hCG levels decline by an average of 88%–90% by 7 days from initial detection (11). This decrease would result in an hCG level of 10 mIU/mL by the second hCG test—a level too low to be detected by high-sensitivity urine pregnancy tests. Therefore, we extrapolated that the majority of women (80%) with a chemical pregnancy will have a negative urine hCG test result or will have resumed their menstrual cycle by 1 week after an initially positive pregnancy test.

- Protocol 3 uses a woman’s missed menstrual period as the indication for pregnancy testing. Since 97% of clinical pregnancies will have implanted by 7 days after the first day of the next expected menstrual period (14), a negative pregnancy test at 1 week after a missed menstrual period will exclude most clinical pregnancies.

Using the trends of hCG positivity for both chemical and clinical pregnancy outcomes (7, 14), we calculated the positive predictive values for detecting a clinical pregnancy for these 3 diagnostic algorithms. In protocol 1, the positive predictive value is equal to the proportion of clinical pregnancies detected divided by the total positive hCG tests. In protocol 2, the total number of positive hCG tests detected is minus 80% of the chemical pregnancies, while the clinical pregnancies detected remain constant. The denominator of the positive predictive value in protocol 3 is calculated by subtracting 3% of the clinical pregnancies (since 3% of clinical pregnancies are not yet diagnosed by 1 week after the expected missed menstrual period) and adding the 20% of chemical pregnancies that would still be detected by testing at this time.

We then modeled the annual rates of the reproductive outcomes for women in a hypothetical study population. Combining the 2 different population-based pregnancy data estimates, we calculated the proportion of pregnancies that would be detected by the 3 diagnostic tests by estimating the proportion of positive hCG results that would be detected by that test out of the total known positive hCG results in the studies. Table 3 shows the percentage of chemical pregnancies, “false positives” in this setting, that would be detected in the population by each diagnostic test. We assumed that annual pregnancy rates would range from 10% to 60% depending on whether or not contraception is used.

We computed an inflation factor, shown in Table 3, that can be utilized to ensure adequate statistical power of the trial. For a prespecified pregnancy rate, denoted as  $\Pr(\text{pregnancy})$ , the proportion of these pregnancies that would be detected by each protocol (protocols 1–3), described as  $\Pr(\text{Detecting the Pregnancy}|\text{protocol } j)$ , is used to determine what fraction of the total number of study subjects would be

**Table 3.** Increase in Sample Size Necessary to Adjust for the Combined Effect of Pregnancy Incidence and Method of Pregnancy Diagnosis

Diagnostic Test	Pregnancies Detected <sup>a</sup>		Inflation Factor <sup>b</sup> for a Given Pregnancy Rate		
	No.	%	10% Pregnancy Incidence	25% Pregnancy Incidence	60% Pregnancy Incidence
Three trials (7–9)					
Monthly	901	100	1.10	1.33	2.50
+1 week	733	81	1.09	1.25	1.95
Menses	712	79	1.09	1.25	1.95
HPTN 055 trial sites					
Monthly	105	100	1.10	1.33	2.50
+1 week	91	86	1.09	1.27	2.07
Menses	88	83	1.09	1.26	1.99

Abbreviation: HPTN, HIV Prevention Trials Network.

<sup>a</sup> Number of positive human chorionic gonadotropin (hCG) concentrations detected with this test/known positive hCG levels in this population.

<sup>b</sup> Inflation factor =  $1/[1 - (\text{proportion of pregnancies detected}) \times (\text{proportion of true pregnancies})]$ .

censored because of the diagnosis of pregnancy,  $C_j$ . For protocol  $j = 1, 2, 3$ , the proportion censored is

$$C_j = \Pr(\text{Detecting the Pregnancy} | \text{protocol } j) \\ \times \Pr(\text{Pregnancy}).$$

The inflation factor is then computed as  $1/(1 - C_j)$  and can be used to multiply by the originally calculated sample size

to assure that the trial accounts adequately for attrition due to pregnancy. This method of sample size adjustment conservatively adjusts for periods of time when pregnant participants are not using the investigational product by assuming that they are censored from the study for its duration.

Estimates of the total days of avoidable fetal exposure shown in Table 4 are conservative assuming the lowest

**Table 4.** Estimated Exposure Time in a Hypothetical Study of 1,000 Women, Half Randomized to the Active Treatment Arm of a Human Immunodeficiency Virus Prevention Trial

Diagnostic Protocol	% hCG Positive That Are Clinical Pregnancies (True Positives)	% hCG Positive That Are Chemical Pregnancies <sup>a</sup> (False Positives)	Avoidable Time not Using the Investigational Product/False-Positive hCG (Weeks)	Total Person-Weeks not Using the Investigational Product <sup>b</sup>	Estimated Amount of Avoidable Fetal Exposure (Days/Fetus)
Assuming a pregnancy prevalence of 10% or 50 conceptions					
1: Monthly	77–83	17–23	4	48	0
2: +1 week	94–96	4–6	4	12	7
3: Menses	98–99	1–2	4	4	≤12
Assuming a pregnancy prevalence of 25% or 125 conceptions					
1: Monthly	77–83	17–23	4	116	0
2: +1 week	94–96	4–6	4	32	7
3: Menses	98–99	1–2	4	12	≤12
Assuming a pregnancy prevalence of 60% or 300 conceptions					
1: Monthly	77–83	17–23	4	276	0
2: +1 week	94–96	4–6	4	72	7
3: Menses	98–99	1–2	4	24	≤12

Abbreviation: hCG, human chorionic gonadotropin.

<sup>a</sup> The term *chemical pregnancy* describes a transiently positive human chorionic gonadotropin level not associated with the development of an embryo or even a gestational sac.

<sup>b</sup> Total person-weeks not using the investigational product = number hCG positive × proportion false positive × time not using the product assuming retention in the trial with monthly visits (4 weeks).

**Table 5.** Positive Predictive Values<sup>a</sup> for Diagnosing Pregnancy

Population	Test Type Used to Diagnose Pregnancy		
	Protocol 1: Monthly hCG <sup>b</sup>	Protocol 2: Monthly/Weekly <sup>c</sup>	Protocol 3: Missed Menses <sup>d</sup>
Literature-based estimates of women attempting pregnancy	691/901 = 77%	691/733 = 94%	691/712 = 97%
HPTN 055 trial	87/105 = 83%	87/91 = 96%	87/88 = 99%

Abbreviation: HPTN, HIV Prevention Trials Network.

<sup>a</sup> Positive predictive value = clinical pregnancies detected with the specified diagnostic method/total human chorionic gonadotropin (hCG)-positive tests captured by using that diagnostic method.

<sup>b</sup> Monthly hCG tests based on calendar day irrespective of menstrual cycle.

<sup>c</sup> Monthly hCG tests with a confirmatory test in 1 week if the initial test is positive.

<sup>d</sup> An hCG test conducted 1 week after missed menses only.

estimate for false-positive tests (17%), which translates into the largest possible number of clinical pregnancies. The estimates for person-time not using the investigational product due to chemical pregnancies assumes the highest (23%) estimated rate.

## RESULTS

The positive predictive values, shown in Table 5, demonstrate the likelihood that a test will diagnose a clinical pregnancy and not misclassify a chemical pregnancy. The positive predictive values depend on the prevalence of clinical pregnancies in the population.

The modeled HIV prevention trial is shown in Figure 1. The algorithm demonstrates how the study outcomes are affected by 1) the number of pregnancies detected by each diagnostic tool, 2) the prevalence of the type of pregnancy (chemical vs. clinical) detected, and 3) the population's use or nonuse of contraceptives. In this figure, the percentage of pregnancies detected refers to those detected at that visit by that diagnostic tool. Any clinical pregnancies not detected would be noted at a later visit. The number of pregnancies detected by the different diagnostic methods varies because of possible resolution of chemical pregnancies, normal variation in menstrual cycle lengths, and variation in the rate of decline of hCG concentration. Therefore, although nearly 100% of pregnancies are detected by performing monthly pregnancy tests (protocol 1), a significant proportion (17%–23%) would be chemical pregnancies. In contrast, fewer false-positive (chemical) pregnancies are identified when the monthly pregnancy test is followed up with a confirmatory test after 1 week (protocol 2). With this testing strategy, the percentage of chemical pregnancies would range from 2% to 4%. When pregnancy tests are performed only after the expected menses is missed (protocol 3), proportionally more of these pregnancies are considered clinical because the percentage of chemical pregnancies is reduced to 1%–2%.

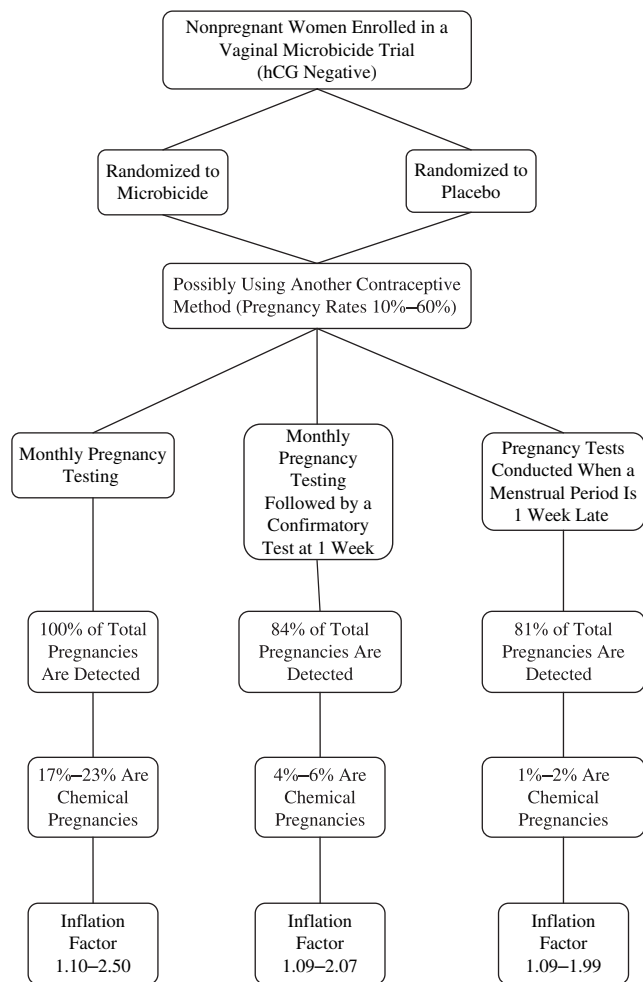
Table 3 shows how a clinical trial will be affected by varied pregnancy prevalences. The sample size inflation factor is the factor by which the initial sample size would have to be multiplied to achieve the appropriate statistical power given the pregnancy prevalence associated with each

diagnostic tool. The inflation factor is designed to help correct for the effect of censoring pregnant participants. For example, in a protocol that uses monthly pregnancy tests and requires enrollment of 1,000 women to demonstrate a difference in HIV seroconversion between the intervention and control arms of a trial, the inflation factor would increase the target enrollment to 1,100–2,500 once pregnancy prevalence is accounted for. Table 4 illustrates the effect that pregnancy prevalence and method of diagnosis have on exposure time in our modeled trial of 1,000 women. For the sake of comparison, we display, for each pregnancy diagnosis method, the number of days that each fetus would be exposed to an investigational drug.

## DISCUSSION

Despite the age-old saying, the diagnosis of pregnancy is not binary: women *can* be a little bit pregnant. In this analysis, we classified detection of a chemical pregnancy as a false-positive pregnancy test—something that one would prefer to avoid in clinical trials. Early detection of clinical pregnancies, which we classified as true-positive tests, should be optimized in clinical trials. The data put forth in this paper underscore the varied pregnancy outcomes that can occur because of the natural history of a positive hCG test.

The use of highly sensitive pregnancy tests has highlighted, in effect generated, the high prevalence of chemical pregnancies in the population. In the HIV prevention trial setting, a positive pregnancy test usually results in a woman either being withdrawn from the remainder of the trial or discontinued from the investigational product until the pregnancy has ended, which has ethical consequences. The diagnosis of pregnancy is not untroubling for many women, and a false-positive pregnancy test may result in unnecessary stress. The methodological consequence of such diagnoses is a decrease in exposure-time to the investigational product, resulting in a decreased effect-size in an intention-to-treat analysis and/or censoring, which leads to a reduction in the statistical power of the trial to demonstrate a difference in HIV seroconversion between the intervention and control arms. The population



**Figure 1.** Three different methods of pregnancy diagnosis and the corresponding impact on the trial sample size required to demonstrate an effect. The chemical pregnancy percentages vary because a spot high-sensitivity urine pregnancy test can detect a pregnancy that resulted from ovulation any time between 10 and 25 days prior, with a more advanced pregnancy being less likely to be chemical. The range in each inflation factor is due to the variation in pregnancy rates, as shown in Table 3. The term *chemical pregnancy* describes a transiently positive human chorionic gonadotropin (hCG) level not associated with the development of an embryo or even a gestational sac. Inflation factor =  $1/[1 - (\text{proportion of pregnancies detected}) \times (\text{proportion of true pregnancies})]$ .

urgently needs methods of HIV prevention: we cannot afford falsely negative studies.

Our analysis has several limitations. Using the longitudinal trials that defined the entity of chemical pregnancies, we estimated the chemical pregnancy rate to be 23% in the general population, ranging from 15% to 25%. Although the definition of a chemical pregnancy is somewhat different in each of the 3 trials, we chose to combine the data sets to achieve more stable estimates. To moderate this limitation, we have provided ranges of the impact of the different pregnancy outcomes. Given the estimated chemical pregnancy rate of 17% in HPTN 055, we have at best over-

estimated the impact of chemical pregnancies in clinical trials, which results in more conservative inflation factors. Another limitation is that our estimates of the pregnancy outcomes in HPTN 055 depend on the accuracy of the reported “pregnancy periods.” These periods were calculated by considering the number of consecutive months with positive hCG tests. Finally, we used data on the rate at which hCG rises and falls to estimate the prevalence of detectable chemical pregnancies at different time points. Doing so required the use of point estimates, which may overestimate or underestimate the positive predictive values of our diagnostic tests.

Even with these limitations in mind, we have shown that the manner in which pregnancy is defined has implications for clinical trials. Because of overlap in the populations at risk of HIV and the populations at risk of pregnancy, the diagnosis of pregnancy within an HIV prevention trial has many possible ramifications. If such a diagnosis is made outside of the usual mode of clinical care (e.g., random monthly testing), women may begin to mistakenly associate the investigational product with the diagnosis of a miscarriage instead of a chemical pregnancy (an event that would not have been evident to that woman in more common life circumstances). By taking the pregnancy diagnosis out of the context of the menstrual cycle and using monthly tests, we are in fact redefining the diagnosis of pregnancy, which will introduce new bias.

The effect that pregnancy has on the HIV prevention regimen is unknown with respect to safety, acceptability, or effectiveness. It is possible that becoming pregnant itself results in a differential risk of HIV acquisition. A disparate risk of pregnancy in the arms of the trial, compounded with a possible differential risk of HIV acquisition in pregnant women, could result in biased trial results. A woman’s attitude toward her pregnancy may modify her sexual risk behaviors as well as her adherence and compliance behaviors within the clinical trial.

Clearly, the best way to avoid the problems associated with incident pregnancies in clinical trials is to provide highly effective contraceptives for women considering enrollment in an HIV prevention trial, and throughout the trial (18). The data in Tables 3 and 4 underscore how effective contraception helps to keep the impact of pregnancy low. Although the most effective methods of contraception approximate success rates of 100% (19), pregnancies may still occur in these trials if less-effective methods are used or if methods are used incorrectly. Furthermore, routine use of specific contraceptive methods may not be a practical requirement of these trials given the communities in which they are being conducted. If highly effective methods are not culturally acceptable in the regions where HIV is most prevalent, investigators may have little choice but to sacrifice contraceptive efficacy in favor of high rates of seroconversion.

Short of eliminating pregnancy altogether in these trials, uniform definitions of pregnancy can be met and incorporated into the protocol. Baseline information about community pregnancy outcomes would provide useful data to extrapolate effects of the drug, or even study participation, on pregnancy outcomes. To help curtail the number of false-positive pregnancies, we advocate a diagnostic method that avoids monthly pregnancy tests and instead adheres to

a clinical standard of care by using a missed menstrual period or symptoms of pregnancy as the trigger to perform a test. This method may result in delayed pregnancy diagnosis and more time using the investigational product for pregnant women; however, given the fact that organogenesis does not begin until 5–6 weeks after the last menstrual period, the increased exposure of 12 or fewer days may be acceptable in the setting of reassuring reproductive toxicity studies. If the specifics of the study population suggest that clinical parameters are too unreliable, perhaps the scheduled monthly pregnancy test can be confirmed by a second test a week later and the woman not diagnosed as being pregnant until that time. As shown in Table 4, this approach would have the advantage of minimizing exposure of the developing fetus to the investigational product.

In this paper, we present estimates from the literature of the different pregnancy outcomes for women who were attempting pregnancy and compare them to those who are supposed to be avoiding pregnancy in an HIV prevention preparedness trial. In addition to the prevalence of chemical pregnancies, we also call attention to the striking number (53%) of pregnancies in HPTN 055 that ended before 25 weeks of gestation. Whether these abortions are spontaneous or induced, this trial highlights the poor pregnancy outcomes that this population suffers. Improving reproductive outcomes, in addition to combating the HIV epidemic, is of paramount public health importance.

#### ACKNOWLEDGMENTS

Author affiliations: Department of Obstetrics and Gynecology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania (Courtney A. Schreiber, Kurt T. Barnhart); The Center for Clinical Epidemiology and Biostatistics, Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania (Mary Sammel, Kurt T. Barnhart); Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania (Sharon L. Hillier); Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania (Sharon L. Hillier); and Reproductive Infectious Disease Research, Magee-Women's Hospital of the University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania (Sharon L. Hillier).

This work was supported by the National Institutes of Health and the National Institute of Allergy and Infectious Diseases (grant 1 U01-068633-01). Funding for this project was provided in part by the BIRCIW K12-HD43459 grant "Career Development in Women's Health Research" Award.

Conflict of interest: none declared.

#### REFERENCES

1. Zaba B, Gregson S. Measuring the impact of HIV on fertility in Africa. *AIDS*. 1998;12(suppl 1):S41–S50.
2. Kelly C, Ramjee G, Mâsse B. Self-reported contraceptive use and pregnancy in a site preparedness study (HPTN 055) [abstract]. Presented at the Microbicides 2006 Conference, Cape Town, South Africa, April 23–26, 2006.
3. Gray RH, Li X, Kigozi G, et al. Increased risk of incident HIV during pregnancy in Rakai, Uganda: a prospective study. *Lancet*. 2005;366(9492):1182–1188.
4. Morrison CS, Wang J, Van Der Pol B, et al. Pregnancy and the risk of HIV-1 acquisition among women in Uganda and Zimbabwe. *AIDS*. 2007;21(8):1027–1034.
5. Morrison CS, Richardson BA, Mmiro F, et al. Hormonal contraception and the risk of HIV acquisition. *AIDS*. 2007; 21(1):85–95.
6. Stephenson JM. Systematic review of hormonal contraception and risk of HIV transmission: when to resist meta-analysis. *AIDS*. 1998;12(6):545–553.
7. Wilcox AJ, Weinberg CR, O'Connor JF, et al. Incidence of early pregnancy loss. *N Engl J Med*. 1988;319(4):189–194.
8. Zinaman MJ, Clegg E, Brown CC, et al. Estimates of human fertility and pregnancy loss. *Fertil Steril*. 1996;65(3): 503–509.
9. Wang X, Chen C, Wang L, et al. Conception, early pregnancy loss, and time to clinical pregnancy: a population-based prospective study. *Fertil Steril*. 2003;79(3):577–584.
10. Sayle AE, Wilcox AJ, Weinberg CR, et al. A prospective study of the onset of symptoms of pregnancy. *J Clin Epidemiol*. 2002;55(7):676–680.
11. Chung K, Sammel M, Zhou L, et al. Defining the curve when initial levels of human chorionic gonadotropin in patients with spontaneous abortion are low. *Fertil Steril*. 2006;85(2): 508–510.
12. HPTN 055 HIV Prevention Preparedness Study. ([http://www.hptn.org/research\\_studies/hptn055.asp](http://www.hptn.org/research_studies/hptn055.asp)).
13. Ramjee G, Kapiga S, Weiss S, et al. The value of site preparedness studies for future implementation of phase 2/III/III HIV prevention trials: experience from the HPTN 055 study. *J Acquir Immune Defic Syndr*. 2008;47(1):93–100.
14. Poikkeus P, Hiilesmaa V, Tiitinen A. Serum hCG 12 days after embryo transfer in predicting pregnancy outcome. *Hum Reprod*. 2002;17(7):1901–1905.
15. Wilcox AJ, Baird DD, Dunson D, et al. Natural limits of pregnancy testing in relation to the expected menstrual period. *JAMA*. 2001;286(14):1759–1761.
16. Wilcox AJ, Dunson DB, Weinberg CR, et al. Likelihood of conception with a single act of intercourse: providing benchmark rates for assessment of post-coital contraceptives. *Contraception*. 2001;63(4):211–215.
17. Creinin MD, Keverline S, Meyn LA. How regular is regular? An analysis of menstrual cycle regularity. *Contraception*. 2004;70(4):289–292.
18. Raymond EG, Taylor D, Cates W Jr, et al. Pregnancy in effectiveness trials of HIV prevention agents. *Sex Transm Dis*. 2007;34(12):1035–1039.
19. Trussell J. Contraceptive failure in the United States. *Contraception*. 2004;70(2):89–96.