

## **SHOULD BIOMARKER ESTIMATES OF HIV INCIDENCE BE ADJUSTED?**

**Ron Brookmeyer<sup>a, b</sup>**

**From the <sup>a</sup> Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, New Jersey, USA and the <sup>b</sup> Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA.**

**Correspondence to Ron Brookmeyer, PhD, Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Baltimore, MD 21205, USA**

**Tel +1 410 955 3519; email [rbrook@jhsph.edu](mailto:rbrook@jhsph.edu)**

A final version of this manuscript has been published in [AIDS](#), 23(4):485-491, 2009.

Key Words: BED, biomarker, epidemiology, HIV incidence, statistics

## **Abstract**

**Objective:** To evaluate adjustment procedures that have been proposed to correct HIV incidence rates derived from cross-sectional surveys of biomarkers (BED). These procedures were motivated by some reports that the biomarker BED approach overestimates incidence when compared to cohort studies.

**Design:** Considered the Hargrove and McDougal adjustment procedures that adjust biomarker estimates of HIV incidence rates for misclassification with respect to the timing of infections.

**Methods:** Performed mathematical and statistical analysis of the adjustment formulas. Evaluated sources of error in cohort studies of incidence that could also explain discrepancies between cohort and biomarker estimates.

**Results:** The McDougal adjustment has no net effect on the estimate of HIV incidence because false positives exactly counterbalance false negatives. The Hargrove adjustment has a mathematical error that can cause significant underestimation of HIV incidence rates especially if there is a large pool of prevalent long standing infections.

**Conclusion:** The two adjustment procedures of biomarker incidence estimates evaluated here that purport to correct for misclassification, do not increase accuracy and in some situations can introduce significant bias. Instead, the accuracy of biomarker estimates can be increased through improvements in the estimates of the mean window period of the populations under study and the representativeness of the cross-sectional samples. Cohort estimates of incidence are also subject to important sources of error and should not blindly be considered the gold standard for assessing the validity of biomarker estimates.

## **Introduction**

HIV incidence, the rate at which new infections are occurring in populations, measures the current growth of the epidemic. Knowledge of incidence rates leads to more effective and targeted intervention efforts. Incidence rates are important when designing studies to evaluate prevention programs, for example in sample size determinations, and indeed, the lack of accurate estimates has been identified as one reason for the failure of several large scale HIV prevention trials [1]. There are enormous challenges to obtaining reliable estimates of HIV incidence in populations. Recent progress in the use of novel biomarkers in cross-sectional samples is an exciting methodological development that offers the promise of addressing these challenges. The basic idea is to use a biomarker to identify persons recently infected. The approach requires only a cross-sectional sample of persons from whom serum specimens are collected at a single point in time which is in contrast to the cohort approach that requires multiple specimens from participants collected longitudinally over time.

The biomarker approach was first suggested in connection with an assay for P24 antigenemia [2]. This particular biomarker necessitates large sample sizes to obtain statistically reliable estimates of incidence because the durations that persons are P24 antigenemic are relatively short. Subsequently, the Serologically Testing Algorithm for Recent HIV Seroconversion (STARHS) was developed [3] which consists of a dual antibody testing system in which persons are tested with a standard HIV antibody assay and those that are positive are then tested with a second less sensitive assay. Persons who are positive on the first and also negative on the second assays are said to be in the “window” period. The second assay that is currently in widespread use is the BED HIV-1 capture enzyme immunoassay [4]. A person is in the “window period” if they are confirmed positive on the standard enzyme immunoassay (the first assay) and negative on the BED assay (the second assay).

The BED approach was the basis for the recently reported figure that approximately 56,000 new infections occurred in the U.S. in 2006 which was higher than the previous estimate of 40,000[5,6]. That statistical revision raises questions whether the increase is due to changes in the methodology or changes in the underlying incidence, and whether the new (BED) method is any more accurate than previous methods? UNAIDS had actually issued a cautionary statement that the BED assay should neither be used for estimating incidence nor for monitoring trends because of reports from African and Asian settings that the BED approach overestimate rates [7].

The reports that the BED approach overestimates HIV incidence when compared to cohort studies have led to proposals for statistical adjustments [8, 9]. These adjustments purportedly correct for “misclassification” that arises because BED does not always accurately distinguish recent from past infection; of particular concern is that some long

standing infections have been found in the window period. Yet, such statistical adjustment procedures have not been universally adopted [10]. For example, the U.S. incidence estimate was not adjusted for misclassification which raises the question whether it was an overestimate because it wasn't adjusted? This article evaluates the appropriateness of statistical adjustments of biomarker incidence estimates for misclassification with respect to the timing of infection.

### **Statistical Framework**

The biomarker approach to estimating incidence from a cross-sectional survey relies on the epidemiological relationship that the prevalence of a condition is equal to the incidence multiplied by the mean duration of that condition. Here the “condition” refers to the “window period”. The window period is not a fixed number but rather has a probability distribution. We use the symbol  $\mu$  for the mean (or expected value) of the window period. Persons can have window periods either shorter or longer than the mean  $\mu$ . The probability that the window period is greater than  $t$  days, the survival function, is called  $S(t)$ . The biomarker approach assumes that the expected number of new infections occurring in the population per unit time is approximately constant for the most recent period of time up to the maximum possible window period  $M$ ; there have been no reports that the BED window period is longer than about  $M=3$  years. If  $g$  is the probability density of incident infection, (that is, the expected number of new infections per day divided by the population size), which is assumed constant over the past  $M$  years, then the probability that an individual is in the window period at the time of recruitment into the cross-sectional sample is:

$$P(W) = g \int_0^M S(t) dt \quad (1)$$

The mean window period is mathematically equal to the integrated survival function, that is  $\mu = \int_0^M S(t) dt$ , which is a relationship we will use several times in this article. Thus, the right side of equation (1) is equal to  $g \mu$ . Using that fact and dividing equation (1) by the proportion of persons who are either uninfected or in the window period, we obtain:

$$I \approx \frac{p}{\mu} \quad (2)$$

where  $I$  is the incidence rate, and  $p$  is the proportion of persons that are in the window period from among persons who are either negative on the standard enzyme immunoassay or in the window period. If the mean  $\mu$  is expressed in days, then  $I$  is the fraction of the uninfected population that becomes infected per day. Equation (2) is the basis of the biomarker approach for estimating incidence which has also been termed a snapshot approach because it requires only a single cross-sectional sample [11]. The approach relies on an unbiased estimate of the mean window period. Statistical

approaches have been proposed for incorporating uncertainty in the mean window period into confidence intervals for the incidence rate including an analytic procedure [12] or a Monte Carlo procedure [13]. It is assumed that persons with advanced HIV disease (AIDS) are removed from the cross-sectional sample before calculating incidence from equation 2 [10].

### **Evaluation of Statistical Adjustments**

Two adjustment procedures have been proposed to account for misclassification that purportedly arise if persons not infected “recently” are in the window period (the false positives), or if persons infected “recently” are not in the window period (the false negatives). These adjustment procedures define “recent infection” as an infection that occurred within the past  $\mu$  days, where  $\mu$  is the mean window period. The procedures use sensitivity and specificity corrections to adjust the numbers in the window to obtain a “corrected” number infected within the past  $\mu$  days.

#### **The McDougal Adjustment**

McDougal and colleagues derive a correction factor (see equation 2 in [8]) which is the ratio  $P(T_0)/P(W)$  where  $P(T_0)$  is the probability a person was infected within the past  $\mu$  days and  $P(W)$  is the probability of being in the window. An important point is that the McDougal approach makes a distinction between  $P(T_0)$  and  $P(W)$ .

However, it is a mathematical fact that the frequency (probability) of false negatives is exactly equal to the frequency of false positives. We prove this remarkable fact in Figure 1 which shows two circles that represent the probability of being in the window period,  $P(W)$ , and the probability of being infected within the past  $\mu$  days,  $P(T_0)$ . The two regions in which the circles don't overlap are the false negative and false positive probabilities and these two regions are of exactly equal size implying that the two circles themselves are of exactly equal size, that is,  $P(W)=P(T_0)$  (see figure legend for proof). While such a result does not hold in general, it is true in this context in which the “screening test” is the window period and that screening test is being used to classify persons as to whether they were infected within the past  $\mu$  days or not.

The fact that  $P(W)$  is equal to  $P(T_0)$  can also be proved simply and directly as follows. The probability a person was infected within most recent  $\mu$  days is just the rate infections occur in the population,  $g$ , multiplied by the duration of the time interval which is  $\mu$ ; that argument implies that  $P(T_0)$  is equal to  $g\mu$ . But,  $P(W)$  is also equal to  $g\mu$  (see equation (1)), thus proving that  $P(W)=P(T_0)$ .

Thus, the McDougal correction factor, which is the ratio  $P(T_0)/P(W)$ , is theoretically equal to 1. The adjustment for false negatives exactly counterbalances the adjustment for false positives. The net effect of the McDougal adjustment is zero.

A subtle point is that although the probabilities  $P(W)$  and  $P(T_0)$  are exactly equal, the particular persons who comprise the group in the window period are not exactly the same as the persons who comprise the group infected in the last  $\mu$  days. In other words, although the sizes of the groups are exactly the same, the members of the groups are not. Accordingly, the assay for the window period should not be used to make *individual* determinations (diagnoses) as to whether or not a particular person was infected recently; rather it is used to obtain an aggregate number of persons in the window period which is then used to estimate the incidence rate in the population.

### **The Hargrove Adjustment**

Hargrove and colleagues [9] developed a simplified version of the McDougal adjustment using an additional assumption that the sensitivity of the BED is equal to its specificity among persons infected between  $\mu$  and  $2\mu$  days earlier (see equation 3 in [9]). They develop an adjustment that has an input factor  $\varepsilon$  which is the probability of being in the window period if infected at least  $2\mu$  days earlier; Hargrove [9] use  $\varepsilon = 0.053$ . Larger (smaller) values of  $\varepsilon$  lead to larger (smaller) downward adjustments to the incidence rate.

However, we show in the Appendix that the mathematical implication of the Hargrove assumptions is that  $\varepsilon$  is 0. That is, there is a fundamental mathematical inconsistency in the Hargrove formula with any nonzero value for  $\varepsilon$  and that inconsistency can produce very anomalous results.

We illustrate the anomalies of the Hargrove adjustments with a numerical example. Consider two communities  $A$  and  $B$  with the same *current* rates of HIV incidence and population sizes. Suppose the HIV incidence rates in  $A$  were 0 up until 3 years ago. The epidemic in  $B$  occurred in two waves. In the first wave, there was a burst of infections that occurred 5 years earlier resulting in 20% of the population of  $B$  becoming infected and thereby creating a pool of prevalent long standing infections. In the second wave, the HIV incidence rates in  $B$  were exactly the same as in community  $A$  during the past 3 years. Cross-sectional samples of 10,000 individuals are taken in both communities. Table 1 shows the numbers of persons in the sample that were uninfected, infected, and in the window period, along with the Hargrove adjusted and unadjusted rates (see table 1 footnote for explanation). We assume that no persons from  $B$  who were infected 5 years earlier are currently in the window period (because there are no data suggesting window periods are longer than about 3 years). The unadjusted incidence rates were 3.6% per year in both communities. The Hargrove adjusted incidence rates in  $A$  and  $B$  were 3.5 and 0.7% per year respectively. Thus, the Hargrove procedure yields an anomalous result because the adjusted rate in  $A$  is about 5 times greater than  $B$  even though the current HIV incidence rates are equal in the two communities.

Why does the Hargrove adjustment produce such an anomalous result? The Hargrove adjustment is subtracting off a fraction of the long standing HIV prevalent infections from the numbers in the window period. But because these long standing infections are not in the window period (they were infected 5 years earlier, and there is no data suggesting window periods extend that long) the procedure is incorrectly deflating the

numbers in the window period. As the pool of long standing prevalent infections becomes larger, the Hargrove adjusted incidence will become increasingly more biased downwards. In fact, the Hargrove adjustment procedure produces *negative* estimates of the HIV incidence rates in many plausible settings. For example, the Hargrove adjustment (with  $\varepsilon = 0.053$ ) will give a negative estimate of incidence if the true HIV incidence rate is 1% per year and if the percentage of the population with long standing prevalent infections of greater than 3 years duration is 10% or more. In effect, the Hargrove adjustment reduces to assuming that a fraction of infected persons ( $\varepsilon$ ) remain in the window period indefinitely. Yet there are no data to indicate that the window period extends beyond approximately 3 years.

### **Why Cohort and Biomarker Incidence Estimates May Not Agree**

There have been several reports especially from Africa that cohort estimates of incidence are lower than biomarker (BED) estimates [7,14]. These reports raise questions as to why cohort estimates do not agree with the biomarker estimates, and whether cohort estimates should be regarded as the gold standard for assessing the accuracy of biomarker estimates?

We consider two sources of error that could result in cohort studies yielding biased estimates of incidence rates of a population. The first concerns *selection bias* into cohorts that result if persons who would agree to participate in cohort studies, the “compliers,” have different risks of infection than other persons. For example, compliers may be less mobile than the non-compliers and mobility may be associated with HIV risks. The second source of error, we call *adherence effects*, occurs if the follow-up visits themselves have an effect on HIV incidence perhaps through repeated exposure to counseling (such as condom promotion or other prevention messages) among persons who adhere to the schedule of visits. We numerically investigated the sensitivity of cohort estimates of incidence to selection bias and adherence effects. Table 2 shows the ratio of the incidence rate in a population to that from a cohort study. For example, if 50% of a population are compliers and the relative risks associated with selection bias and adherence effects are both 0.67, then the population incidence rate would be 1.86 times greater than estimated from a cohort study. While we do not have direct data on the actual magnitude of the selection and adherence effects, we do know that a significant fraction of persons who were asked to participate in several major cohort studies, either refused to participate or did not complete the follow-up schedule. For example among studies cited as examples where the BED estimates were higher than the cohort estimates, only 61% [15] and 71% [9] of persons at baseline gave follow-up blood specimens.

There are also important sources of error in the biomarker approach for estimating population incidence. A main source of error in the biomarker approach is the mean window period. The mean should be calculated from a representative sample of window periods of the population. The mean window period may depend on the circulating HIV strain, and other co-infections in the population (10). Furthermore, neither short nor long window periods should be systematically excluded from the calculation. If long

window periods are excluded, then the mean will be underestimated, potentially causing an overestimation of incidence. For example, the reported mean window period for clade C HIV-1 virus in southern Africa was 187 days (.512 years) [9]. However, that estimate systematically excluded censored data (cases still in the window period at last follow-up) which resulted essentially in the exclusion of all window periods greater than 12 months (see figure 1 in [9]). Excluding censored observations from the statistical analysis will result in downwardly biased estimates of the mean window period.

We illustrate the impact of excluding long window periods or censored observations when calculating the mean window period. Suppose 85% of observations have window periods 365 days or less, and the mean of those observations is 187 days; and 15% of observations have window periods greater than 365 days and their mean is 620 days. The effect of excluding the 15% of observations with windows greater than 365 days is to produce a mean window period that is 74% of what it should be, and a biomarker incidence estimate that is 1.35 times larger than what it should be.

It is also worth reiterating that equation (2) requires the *mean* window period and not the *median*; this is a potentially important distinction because the distribution of window periods is right skewed in which case the median would be smaller than the mean. Accordingly, if the median was used instead of the mean in equation 2, the incidence rate would be overestimated. The median has been used in some analyses [4].

In summary, discrepancies between biomarker and cohort incidence estimates, even by a factor of 2 or more, could well be explained by a combination of reasons including underestimation of the mean window period due to exclusion of censored observations, and the impact of selection bias and adherence effects in cohort studies.

## **Discussion**

Reports of discrepancies between biomarker and cohort incidence rates have motivated adjustment procedures to correct biomarker estimates. These procedures attempt to correct for misclassification with regard to the timing of HIV infection. Our analysis shows that use of two of these adjustment procedures are generally misguided. The McDougal adjustment has no numerical effect because mathematically the false positives exactly counter balance the false negatives. The Hargrove adjustment has a mathematical error that can cause significant underestimation of the HIV incidence rates especially when there is a pool of long standing HIV prevalent infections in the population.

Our analysis of the adjustment procedures assumed a maximum possible window period ( $M$ ), and published analyses are consistent with that assumption. Indeed, there are no data documenting any window period longer than approximately three years (see for example figure 7 in [4], and figure 1 in [9].) If, however, a proportion of HIV positive persons are identified who remain in the window period *indefinitely* (that is, for periods of time

considerably greater than 3 years), then, under such circumstances, an adjustment would be necessary.

Although published analyses of the distribution of window periods have not reported persons re-entering the window period after having exited, we recognize that such a phenomenon is a possibility. Under such circumstances,  $\mu$  in equation 2 refers to the *total* expected time spent in the window period (the initial time in the window period plus any subsequent revisits; see [11] for theoretical justification .) Further studies are needed to refine our understanding of the duration of window periods including whether persons re-enter the window. It is imperative that these studies not exclude persons with long window periods. Persons still in the window period at last follow-up should be treated as censored data using survival analysis techniques.

How can we improve the accuracy of biomarker estimates of population HIV incidence? It is critically important that cross-sectional samples be representative of the target population with respect to HIV risks. For example, specimens that are drawn principally from antenatal clinics, sexually transmitted disease clinics or voluntary testing and counseling centers may not be representative of the broader population. In such cases, weighting of the sample will be necessary to appropriately adjust the incidence to reflect the target population (see [5] for an example of such adjustments). Adjustments to improve the representativeness of cross-sectional surveys are very different than the adjustments for misclassification discussed in [8] and [9] and evaluated in this article. Persons with advanced HIV disease should be excluded from the cross-sectional samples. The accuracy of biomarker estimates can also be increased through improved estimates of the mean window period across HIV subtypes and diverse populations.

The inconsistency between cohort and biomarker estimates of incidence reported in some studies is likely due to a multitude of reasons. These include errors in the mean window period of the biomarker as well as selection bias and adherence effects in cohort studies. In view of the potential errors with both cohort and biomarker approaches, cohort estimates should not blindly be considered the gold standard for assessing the validity of biomarker estimates.

## Appendix

### Inconsistency of the Hargrove Adjustment

Hargrove [9] define the sensitivity (*sen*) as the conditional probability  $P(W | T_0)$  of being in the window ( $W$ ) if the person was infected within the last  $\mu$  days ( $T_0$ ):

$$sen = P(W | T_0) = \frac{\int_0^{\mu} gS(t)dt}{g\mu} = \frac{\int_0^{\mu} S(t)dt}{\mu}$$

Hargrove [9] define *spec 1* to be the specificity among persons infected between  $\mu$  and  $2\mu$  days ago. Let  $T_1$  represent the event that an individual was infected between  $\mu$  and  $2\mu$  days ago. We have

$$P(W | T_1) = 1 - spec1 = \frac{\int_{\mu}^{2\mu} gS(t)dt}{g\mu} = \frac{\int_{\mu}^{2\mu} S(t)dt}{\mu}$$

Let  $T_2$  represent the event that an individual was infected between  $2\mu$  and  $M$  days ago, then the probability such an individual is in the window is

$$P(W | T_2) = \frac{\int_{2\mu}^M gS(t)dt}{g(M - 2\mu)} = \frac{\int_{2\mu}^M S(t)dt}{(M - 2\mu)}. \quad (A1)$$

We can express the mean as a sum of integrals over three intervals because  $\mu = \int_0^M S(t)dt$ :

$$\int_0^{\mu} S(t)dt + \int_{\mu}^{2\mu} S(t)dt + \int_{2\mu}^M S(t)dt = \mu$$

Dividing the above equation by  $\mu$ , we have

$$sen + (1 - spec1) + \frac{1}{\mu} \int_{2\mu}^M S(t)dt = 1 \quad (A2)$$

The Hargrove adjustment assumes that  $sen = spec 1$ , which implies from equation (A2)

$$\int_{2\mu}^M S(t)dt = 0,$$

It follows (from A1) that  $P(W | T_2) = 0$  which implies  $\varepsilon = 0$  ( $\varepsilon$  is the conditional probability of being in the window period if infected more than  $2\mu$  days ago.) Thus, the Hargrove assumption that  $sen = spec = 1$  implies  $\varepsilon = 0$ . Any nonzero value for  $\varepsilon$  is mathematically incompatible with the Hargrove assumptions. The Hargrove adjustment uses  $\varepsilon = 0.052$ .

**Author Contribution:** Dr. Brookmeyer takes responsibility for the concept, analysis, interpretation and drafting of the manuscript.

**Acknowledgement:** The author thanks Peter Ghys and journal reviewers for helpful comments on an earlier draft of the manuscript.

## References

1. Lagakos S, Gable A, editors. *Methodological Challenges in HIV Prevention Trials*. Institute of Medicine, National Academy Press, 2008.
2. Brookmeyer R, Quinn TC. Estimation of current human immunodeficiency virus incidence rates from a cross sectional survey using early diagnostic tests. *Am J Epidemiol*, 1995; 141(2):166-172.
3. Janssen RS, Satten GA, Stramer S et al. New testing strategy to detect early HIV-1 infection for use in incidence estimates and for clinical and prevention purposes. *JAMA* 1998; 280(1):42-48.
4. Parekh BS, Kennedy MS, Dobbs T, Pau CP, Byers R, Green T, et al. Quantitative Detection of Increasing HIV Type 1 antibodies after seroconversion: a simple assay for detecting recent HIV infection and estimating incidence. *AIDS Res Hum Retroviruses* 2002; 18:295-307.
5. Hall IH, Song R, Rhodes P, Prejean J, An Q, Lee LM, et al. Estimation of HIV incidence in the United States. *JAMA*, 2008; 300(5):520-529.
6. Karon JM, Song R, Brookmeyer R, Kaplan E, Hall HI. Estimating HIV incidence in the United States from HIV/AIDS surveillance data and biomarker HIV test results. *Statistics in Medicine* 2008 4617-4633.
7. UNAIDS Reference Group for Estimates, Modelling and Projections. *Statement on the Use of the BED-assay for the Estimation of HIV-1 Incidence for Surveillance or Epidemic Monitoring. Report of a meeting of the UNAIDS Reference Group for Estimates, Modeling and Projections*. Athens, Greece, December 13-15, 2005. Geneva:UNAIDS; 2005. <http://www.epidem.org/Publications/BED%20statement.pdf>.
8. McDougal JS, Parekh BS, Peterson ML, Branson BM, Dobbs T, Ackers M, et al. Comparison of HIV-1 incidence observed during longitudinal follow-up with incidence estimated by cross-sectional analysis using the BED capture enzyme immunoassay. *AIDS Res Hum Retroviruses* 2006; 22:945-952
9. Hargrove JW, Humphrey JH, Mutasa K, Parekh BS, McDougal JS, Ntozini R. Improved HIV-1 incidence estimates using BED capture enzyme immunoassay. *AIDS* 2008, 22:511-518.
10. Using the BED HIV-1 Capture EIA Assay to estimate incidence using STARHS in the context of surveillance in the United States, October 2007. Centers for Disease Control and Prevention website. <http://www.cdc.gov/hiv/topics/surveillance/resources/factsheets/BED.htm>.

11. Kaplan E Brookmeyer R Snapshot estimators of recent HIV incidence. *Operations Research* 1999; 47(1): 29-37
12. Brookmeyer R. Accounting for follow-up bias in estimation of Human Immunodeficiency virus incidence rates. *J R. Statist. Soc A*, 1997; 160:127-140.
13. Cole SR, Chu R, Brookmeyer R. Confidence Intervals for Biomarker-based Human Immunodeficiency Virus Incidence Estimates and Differences using Prevalent Data. *American Journal of Epidemiology* 2007; 165(1):94-100.
14. Karita E, Price M, Hunter E, Chomba E, Allen S, Fei L, et al Investigating the utility of the HIV-1 BED capture enzyme immunoassay using cross-sectional and longitudinal seroconverter specimens from Africa. *AIDS* 2007; 21:403-408.
15. Kamali A, Quigley M, Nakiyingi J, Kinsman J, Kengeya-Kayondo J et al. Syndromic management of sexually transmitted infections and behavior change interventions on transmission of HIV-1 in rural Uganda: a community randomized trial. *The Lancet*, 2003, 361:645-652.

**Table 1: Numerical illustration of Hargrove adjustment procedure in samples of 10,000 from two communities with the same current HIV incidence rates.<sup>a</sup>**

	<u>Community A</u>	<u>Community B</u>
<u>Sample Sizes<sup>b</sup></u>		
Uninfected	9,657	7,726
Total Infected	343	2,274
infected & in window period	180	144
Total Sample Size	10,000	10,000
<u>Incidence Rates (% per year)<sup>c</sup></u>		
Hargrove Adjustment	3.5	0.7
Unadjusted	3.6	3.6

<sup>a</sup>The HIV incidence rates in community A and B are identical for the last most recent 3 years. There were no infections in community A prior to 3 years ago. In community B, exactly 20% of the population was infected 5 years earlier; the remaining 80% who were not infected 5 years earlier are subjected to the same rates of infection as community A beginning 3 years ago.

<sup>b</sup>Sample sizes in community A were chosen to match the data given in [9, p514] and then scaled to give a sample size of 10,000. Sample sizes in community B were obtained as follows: number uninfected in B=0.8 x number uninfected in A; number in window period in B=0.8 x number in window period in A; number infected in B=0.8 x numbers infected in A + 0.2 x 10,000. It is assumed that none of those who were infected 5 years earlier in community B are currently in the window period (because there have been no reports that the distribution of window periods extends beyond about 3 years).

<sup>c</sup>The calculations assumed a mean window period of .512 years (187 days)[9]. The Hartgrove adjustment formula (see equation 3 in [9]) is

$$(R - \epsilon P) / (R + .512N - \epsilon T)$$

where  $R$  is the number in the window (180 in A, 144 in B),  $\epsilon=.052$  [9],  $P$  is the total number HIV infected (343 in A, 2274 in B),  $N$  is the total uninfected (9,657 in A, 7726 in B), and  $T$  is the total sample size (10,000 in both A and B). The unadjusted rates from equation 2 are

$$\frac{180}{.512(9657+180)} = .036 \text{ in community A and } \frac{144}{.512(7726+144)} = .036 \text{ in community B.}$$

**Table 2: Sensitivity of HIV incidence rates in cohort studies to selection bias and adherence effects: ratio of the true incidence rate in a population to that from a cohort study.<sup>a</sup>**

$f$	$R$		
	0.5	0.67	0.75
.75	2.5	1.67	1.44
.50	3.0	1.86	1.55
.25	3.5	2.04	1.67

---

<sup>a</sup> The fraction of the population who are compliers (agree to participate in a cohort study if approached) and non-compliers are  $f$  and  $1-f$  respectively. Suppose the ratio of the HIV incidence rate among compliers to non-compliers is  $R_s$ . Then, the HIV incidence rate in the population is

$$I_p = fR_s I_0 + (1-f)I_0$$

where  $I_0$  is the incidence rate among the non-compliers. Suppose the adherence effects that result from the scheduled follow-up visits is to multiply rates by the factor  $R_a$ . Then, the incidence among compliers who adhere to the schedule of follow-up visits in a cohort study is  $R_a R_s I_0$ . The ratio of the incidence rate in the population ( $I_p$ ) to that in a cohort study ( $R_a R_s I_0$ ), is:

$$\frac{f}{R_a} + \frac{1-f}{R_a R_s}.$$

Table 2 gives values of that ratio when  $R_s=R_a=R$ .

**Figure 1: The probability of being in the window period,  $P(W)$ , and the probability of being infected within the last  $\mu$  days,  $P(T_0)$ , are equal.** The false negative probability ( $FN$ ) is the probability of being infected within the last  $\mu$  days and not being in the window period. The false positive probability ( $FP$ ) is the probability of being in the window period and having been infected more than  $\mu$  days ago. The figure illustrates

that  $FN=FP$  which is true because  $FP = g \int_{\mu}^M S(t)dt$  and

$$FN = g \int_0^{\mu} [1 - S(t)]dt = g \left[ \mu - \int_0^{\mu} S(t)dt \right] = g \int_{\mu}^M S(t)dt = FP .$$

