ON THE PERSON YEARS CONCEPT IN EPIDEMIOLOGY AND DEMOGRAPHY

MINDEL C. SHEPS

The need to estimate a risk of some kind in a sample of individuals who are followed for a period of time is common to many different kinds of studies. This paper will report an investigation into the behavior, under defined conditions, of the "exposure time" or "person years" indices that are often used to estimate risks in such studies. One example, the Pearl Index[1], has been common in studies of the effectiveness of contraceptives.[2] Analogous indices are not infrequently calculated in therapeutic and epidemiological investigations. It will be the contention of this paper that the use of this type of index should be confined to rather limited circumstances. In particular, if the individuals in a sample cannot be assumed to experience constant and equal risks, this class of indices does not constitute an appropriate measure.

The particular form of the Pearl Index to be discussed here, the *conception rate,* may be defined as the number of women in a group who conceive for the first time during a period of observation, divided by the total number of months during which they are exposed to the risk of pregnancy.[3] Similarly, in some medical follow-up studies, the number of years during which each patient is observed are added together and the observations expressed as the number of deaths or of relapses per person-year of observation. The pur-

pose of the calculation in each case is to estimate a risk, i.e. the probability of an occurrence per unit time, while allowing for the time during which the risk has operated on each person.

Clearly, such estimates are important in many situations. Thus, the conception rate is intended to estimate an important component of reproduction rates, namely, the monthly chance of conception (fecundability). Estimates of this risk may be crucial to evaluation of contraceptive methods and of family planning programs. Similarly, comparisons of death rates, of morbidity rates, or of relapse rates may be crucial in epidemiological investigations and in evaluations of therapy.

It seems desirable to investigate the properties of an index that is often used for important purposes and that apparently has considerable appeal. The calculations obviously treat 20 years of observation in one person as equivalent to one year in 20 persons or to six months in 40 persons. Intuitively, such an assumption seems unrealistic; its effects under different circumstances merit some study.

An alternative method of measuring risk, the actuarial or life table method, is in the opinion of many a better way of studying events that may occur during a period of time.[4,5,6] The focus here is, however, the behavior of the exposure-time index under several sets of assumptions. Interest will be directed to the expected value of the index and its sampling distribution under those assumptions, with particular attention to:

1. The changes effected in the index by increasing the duration of observation or increasing the number of subjects.

2. The effect of unequal durations of observations for individual subjects. Often, subjects are under observation for unequal periods of time: they enter at different times but the study has a set closing date, they move away, or cease to report for other reasons.[4] Subjects not available for observation throughout the complete period of the study will be referred to as withdrawing from observation or as losses.

3. The effects of heterogeneity among the subjects, i.e. of unequal

risks with respect to the occurrence being studied, and also of unequal probabilities of withdrawing for different subjects.

The report will be presented in three parts: a statement of the problem and a summary of previous work, a description of the methods used in this investigation, and a report of the findings with numerical illustrations. An appendix summarizes some of the theoretical results whose derivation is presented elsewhere.[7] For concreteness, most of the ensuing discussion will be in terms of conceptions and the conception rate, although it applies equally well to many rates calculated in the same fashion. For more general purposes, we may refer to an occurrence-time index, an occurrence rate or an occurrence probability.

THE INDEX

Definition

The index may be calculated in a number of different ways. With respect to fecundability, for example, the numerator may be defined variously as all first conceptions occurring during the period of observation,[2] as all first conceptions that lead to live births,[8] or otherwise. In the presence of varying periods of observation for individuals in the sample, questions arise as to the best method of allowing for losses in both the numerator and the denominator.[4,9,10] In the particular example of the conception rate, what is known as the anniversary method is often used. Thus, if the maximum period of observation for an individual is x months plus a fraction, the risk is estimated during x months only. In effect, losses are considered to occur at the beginning of a month and not to be "at risk" during the month. This rule is applied equally to persons who do experience the event (such as conception) during their last fractional month of observation and to those who do not. If other assumptions are made, leading to modifications in the calculations, the model to be presented would also have to be modified,[7] with corresponding but nonessential changes in details of the results.

Review of the Problem and Previous Work

Since the index amounts to an average over time, an inherent assumption is involved that the risk being measured remains constant for any subject during the time of observation. This assumption may be very far from the truth in epidemiological or medical studies and in studies of contraceptive effectiveness. In the latter case, for example, a group of women may be followed for a number of months after they first start using a new contraceptive. Learning may produce more effective use, or, on the contrary, the desire to avoid conception may weaken as the months go by. This fact alone might be considered a sufficient reason for discarding the index in favor of a method (such as a life table) that is better adapted for observing changes in risk with time.

Even if we make the assumption of constant risk per individual, however, (which does not seem unreasonable in some circumstances), all the difficulties do not necessarily disappear.[11,12] First, it is necessary to distinguish between a hypothetical homogeneous population, with an identical risk for each subject, and the more likely situation where the population is heterogeneous, i.e. with fecundability varying among couples, or mortality risk varying among different individuals.

In a homogeneous sample, the index is a consistent maximum likelihood estimator of the occurrence probability. The estimator is positively biased, the bias decreasing with increased sample size.[7,10] Furthermore, under defined circumstances, an analogous exposure-time estimator of the probability of survival for the total duration of a study was found by Littell[13] to have some advantages, in a homogeneous sample subjected to a constant risk of mortality, over the usual actuarial (life table) estimator.

The behavior of the index in samples of moderate size has received little study even for homogeneous samples. For example, it has at times been assumed that the important determinant of the reliability of the index is its denominator, the exposure time. Thus this quantity is the denominator in Pearl's equation for the variance of the index, derived by treating it as a binomial variable

and equating the number of woman-months with sample size in the formula.[14] It can be shown[7] that this formulation is consistent with large sample (maximum likelihood) theory for homogeneous samples. Several questions then arise: Since the exposure-time can be increased by increasing either the duration of observations or the sample size, do these two procedures have the same effect on the index and its distribution? In general, what are the small sample properties of the index and the variance in homogeneous populations?

The behavior of the index when applied to heterogeneous populations has received some study for the case of conception rates. Thus, it has been shown that, given constant fecundability per couple, but variation from one couple to another, the proportion of those remaining at risk expected to conceive in any month decreases with the duration of follow up.[15,16] The value of the index has been shown to decrease as the period of observation grows longer, both empirically and theoretically.[17,18] In fact, after a sufficiently long period, the mean time to conception or, more generally, the mean exposure time, is expected to equal the reciprocal of the harmonic mean of the risk,[15] and hence the index will approximate the *harmonic* mean of the probability rather than its *arithmetic* mean. Accordingly, as Gini[11] and Potter[18] have emphasized, comparisons of rates calculated for groups followed for different periods of time become highly questionable.

Studies have also been made of the index in artificial sampling experiments on real data. Thus, on sampling from a set of histories, Beebe[19] found that the observed variance of the index in his samples exceeded the variance estimated from Pearl's formula. Potter and Sagi,[20] in similar sampling experiments, extended the examination of the observed variance to include comparison with an estimate calculated according to the well known approximation to the variance of a ratio. They concluded that this approximation was fairly satisfactory and Pearl's formula less so for their sampling scheme. They also found appreciable departures from the normal distribution, the index tending to be skewed to the right. Since

there is no reason to assume that the data sampled in these studies were generated by homogeneous populations, the findings apply to heterogeneous groups in particular.

The sampling schemes described assumed an infinite population in which each woman had a fixed exposure time and a fixed outcome. The variation that they estimated, therefore, consisted of differences among repeated samples from this fixed population. We might also consider, however, that for any individual, the outcome and the exposure time are random variables. A scheme of this kind, using random numbers, was reported by Littell[13] for a related problem in the paper already cited.

METHODS

The Model

For the present report, the following model is postulated. In conformity with the definition of the anniversary method given earlier, it is assumed that all conceptions or events under study occur in mid-month and all losses from observation exactly at the end of a month (i.e. at the beginning of the following month).

Assume that N individuals are followed for a maximum of m months (from the beginning of month 1 to end of month m). The sample is composed of k indistinguishable subgroups. The ith subgroup (where i refers to any value 1, 2, . . . , k) consists of n_i individuals (or couples) who are characterized by two probabilities that remain constant over time for each individual, while varying among subgroups:

1. The risk or conditional probability of the occurrence, e.g. the probability that a woman in the ith group will first conceive in any month, given that she is still under observation. The risk experienced by an individual in the ith group is denoted by p_i.

2. The conditional probability that an individual in the ith group will be lost from observation at the end of any month, given that she is still at risk, denoted by λ_i.

In the "fixed sample" case to be discussed first, the numbers n_i

are considered fixed and the subject of study is the behavior of the index in identical samples. In the "fully stochastic" case, the model is extended to include also the sampling variation involved in choosing a sample, by considering an infinite population of whom the proportion P_i is characterized by the probabilities p_i and λ_i. A random sample of size N is drawn from this population and the index calculated.

Derivation of Results

Four principal approaches were used to derive the results:

1. Under the model described, the exact probability distribution of the number of losses, occurrences (A), and exposure times (B) in homogeneous and heterogeneous populations was derived, leading to the exact distribution of the index $R = A/B$.[7] Numerical results for a few homogeneous parameters were obtained with a computation program on an IBM 7090. For heterogeneous samples, however, these computations quickly become very time consuming, and have not at present been pursued to any extent.

2. In addition to these exact results, the cumulants of the ratio in small samples from homogeneous populations were obtained according to the methods of Haldane and Smith,[21] and the behavior of the cumulants studied from the results.

3. A Monte Carlo program was written to simulate the process on an electronic computer for samples of fixed composition. In the simplest case, for example, it was assumed that a homogeneous sample of individuals (or couples) all with risk p, and no losses was followed for a maximum of m months. A random number was selected and its value determined whether the event (conception) was considered to have occurred to the first individual during month t, where t could equal 1, 2, . . . , m. The result was stored and the process repeated. This sampling scheme thus allowed a random process to act separately in each hypothetical woman, her particular exposure time and the absence or presence of conception being a sample of the values that could be obtained for her.

After all members of the hypothetical sample had been taken through the process, the index was calculated for the sample. A total of 1000 samples were drawn for various combinations of N,

m and p and the results were analyzed. Extensions to include losses and heterogeneous samples are easily made. The results obtained included the distributions and moments of the number of conceptions, the exposure time and the conception rate.

4. A computer program was written to calculate, from formulas given in the appendix, expected values for the number of occurrences, the total exposure time and the variances and covariance of these variables. With these results, values for approximations to the mean and variance of the index could be calculated in the same program.

FINDINGS

Homogeneous Samples Without Losses

Theoretical results show[7] that in *homogeneous* samples without losses, under the model described, the small positive bias of the index *increases as the period of observation* (m) *is prolonged* and decreases with increasing sample size (N). As m becomes very large, the absolute value of the bias is expected to approach p(1-p)/N, and its relative value (1-p)/N. For a moderate value of p such as 0.2, the bias may therefore approach .8/N and for small values such as 0.001, it may approach .999/N. Even in this case, however, the bias is not large, unless N is quite small.

The variance of the index is reduced by increases in either m or N. The rate of decrease with increasing duration of observations (m) depends on the risk (p). On the other hand, the reduction with increasing sample size is more consistent, since it does not depend on the value of p.

The index, in small samples, is skewed to the right for the usual levels of risk (i.e. $p<0.5$), the coefficient of skewness varying as $1/\sqrt{N}$. For fixed values of N, prolonging the duration of the study frequently tends to increase the skewness of the index and hence its departure from the normal distribution. This tendency also varies with p. The kurtosis is an even more complex function of the parameters and of the duration of observations, the coefficient of kurtosis varying as 1/N.

Numerical values for exact probability distributions of the index were calculated for small samples without losses. Examples of such results for homogeneous samples of 25 persons are shown in Table 1. In these samples, the bias in the index is under four per cent but, as expected, it increases with increasing duration of observation (m). Considerable increases are shown in the coefficients of skewness and kurtosis as m grows larger.

The reduction in the bias and the coefficients of variation, skewness, and kurtosis that is associated with increased sample size may be illustrated by comparing the values in the first part of Table 2, where N = 50, with the first part of Table 1.

TABLE I. EXACT RESULTS FOR INDEX IN HOMOGENEOUS SAMPLES WITHOUT LOSSES ($N = 25$)

	Duration of Observation in Months	*Mean*	*Characteristics of Index Coefficients of: Variation*	*Skewness**	*Kurtosis**
Fecundability (p)					
0.20	1	.2000	.40	.30	.01
	4	.2034	.23	.41	.32
	8	.2048	.24	.43	.60
	12	.2056	.19	.61	.75
	16	.2060	.18	.65	.82
	20	.2063	.18	.67	.85
	∞	.2064	.19	.67	.85
0.15	1	.1500	.48	.40	.13
	6	.1530	.24	.47	.45
	12	.1540	.20	.59	.74
	18	.1546	.20	.67	.90
	24	.1549	.19	.70	.97
	30	.1551	.19	.72	.99
	∞	.1551	.20	.72	.91
0.10	1	.1000	.60	.53	.31
	6	.1019	.28	.47	.42
	12	.1025	.23	.56	.65
	18	.1029	.21	.64	.91
	24	.1031	.20	.69	.97
	30	.1033	.20	.72	1.06
	∞	.1036	.20	.74	1.04

(The values for $m \to \infty$ were calculated from approximate formulas.[7,21])

* The measure of skewness is μ_3/σ^3, and of kurtosis, $\mu_4/\sigma_4 - 3$.

TABLE 2. THE EFFECTS OF LOSSES ON THE INDEX IN A HOMOGENEOUS POPULATION. SAMPLE SIZE = 50, AND P = 0.20

Simulation Program (The values given for finite m were obtained from a simulation program.)

Duration of Observations in Months	*Characteristics of index*			
	Mean	*Variation*	*Coefficients of Skewness*	*Kurtosis*
Monthly probability of loss = 0				
1	.2000	.283	.21	<.01
4	.2012	.167	.29	.16
8	.2017	.142	.31	.30
12	.2020	.135	.43	.38
16	.2025	.132	.44	.41
20	.2026	.132	.47	.42
∞	.20320	.130	.47	.42
Monthly probability of loss = 0.1				
1	.2000	.283	.21	<.01
4	.2028	.165	.22	+.06
8	.2027	.143	.31	−.21
12	.2032	.135	.38	−.10
16	.2035	.133	.41	−.08
20	.2038	.132	.43	−.06
∞	.20316	.132	.46	+.45
Monthly probability of loss = .05				
1	.2000	.283	.21	<.01
4	.2007	.169	.24	.23
8	.2015	.151	.43	.72
12	.2018	.147	.45	.85
16	.2023	.144	.50	.99
20	.2024	.144	.50	1.00
∞	.20304	.140	.43	+.39

The Effect of Losses on Homogeneous Samples

The effect of losses on estimates made for homogeneous samples was studied both in theoretical results and by means of the simulation program. For moderately long periods of observation, the effect of introducing losses or making them more frequent (i.e. of increasing λ) may vary with the levels of p, λ and m. As m becomes very large, however, the bias in the index is expected to be decreased by the presence of losses, the relative bias approaching the value $(1-\lambda)(1-p)/N$. The effect of losses on the higher moments of R is also not a simple, monotonic function of λ. For large m, however,

and small p, the presence of losses will usually increase the relative variation.

Table 2 presents some illustrative numerical results of the simulation program, for samples of size 50 with fecundability equal to 0.20, the monthly probability of loss taking on values of 0, 0.01 and 0.05. With a monthly probability of loss equal to 0.01, the bias in the mean, i.e. in E(R), was increased for moderate values of m. The coefficients of variation and skewness showed little change, the former tending to be a little higher and the latter a little lower than in the cases without losses. The coefficient of kurtosis became negative, then decreased in absolute value with increasing m, and was expected to be positive as m became very large.

In the simulation experiments with a monthly probability of loss equal to 0.05, the mean value of R was lower, showing a smaller bias than when $\lambda = 0$, at all values of m calculated. The variance and the coefficients of variation and kurtosis were all greater than in the case without losses. The coefficient of skewness in this case showed an inconsistent relation to that in the case without losses. For very large m, however, and $p = 0.2$, the skewness is expected to be reduced by losses.

Heterogeneous Samples (Fixed Sample Case)

The expected value of the index in heterogeneous samples is a complex function of the degree of heterogeneity, the various probabilities, and the duration of observations.[7] As already mentioned, *in a heterogeneous group without losses* the *expected value decreases* with increasing m: from the arithmetic mean fecundability of the group in the first month toward the harmonic mean. If the sample is small, however, the tendency for the index to decrease in value is counteracted in part by positive terms in the expected value of the ratio, such as are responsible for the positive bias of the index in homogeneous samples.

Before turning to small samples and to other complications, it should be emphasized that even for large samples without losses and with equal periods of observations, occurrence-time indices may not serve as valid comparisons, because arithmetic and harmonic

means do not necessarily vary in the same direction. Thus, it is easy to imagine that one of two samples has a higher arithmetic mean risk but a lower harmonic mean than the other. In such a case, indices calculated after a short period of observation would indicate that this first sample had the higher risk; after a longer period of observation, the conclusion would be reversed.

Almost all the numerical calculations performed for heterogeneous groups utilized approximate formulas for E(R) and its variance.[22] Comparison between these approximations and exact results, in a few instances, showed that the approximations gave lower values than the exact results, the agreement improving with increasing sample size. In general, the approximations are therefore conservative, that for the variance being more so than that for the mean.

TABLE 3. THE EFFECT OF SAMPLE SIZE ON THE EXPECTED VALUE OF THE INDEX [E(R)] IN A HETEROGENEOUS MODEL WITHOUT LOSSES

	Ratio of the Expected Number of Occurrences to the Expected Exposure	*Approximate E(R) Sample Size*	
m	*Time*	*N = 15*	*N = 300*
1	.1667	.1667	.1667
2	.1655	.1681	.1656
3	.1643	.1680	.1645
4	.1632	.1675	.1634
5	.1621	.1669	.1624
6	.1612	.1662	.1614
7	.1602	.1656	.1605
8	.1594	.1650	.1597
9	.1586	.1644	.1589
10	.1578	.1639	.1581
11	.1571	.1634	.1574
12	.1565	.1629	.1568
13	.1559	.1625	.1562
14	.1553	.1622	.1557
15	.1549	.1618	.1552
∞	.1500	.1598	.1505

Arithmetic mean p = .1667
Harmonic mean p = .1500
Variance p = .002222

A numerical example of the effect of sample size on the expected value of the index for heterogeneous groups without losses is shown in Table 3, where it is assumed that the occurrence probability for two-thirds of each sample is equal to 0.2 and for the remaining one-third it is equal to 0.1. The first column in Table 3 shows the ratio of the expected cumulative number of occurrences (conceptions) to the expected cumulative exposure time at each month. This ratio, which is not affected by sample size, decreases with time, from the initial value of 0.1667 toward 0.1500 (the harmonic mean probability). For a sample of size 15, even the approximation for E(R), which is a little lower than the exact value, is considerably higher than the ratio just discussed. In fact, from months two to five inclusive, it is higher than the first month's value of 0.1667. It remains well above the ratio of the two means at all times. This positive discrepancy increases with m, the asymptotic value for very large m being more than six per cent higher than the ratio of the expectations. When the sample size is increased to 300, however, as in the last column, the expected value of the index closely approaches the ratio of the two means in the first column.

If probabilities of loss are included in the model, the Pearl Index for a heterogeneous sample can take on a rather wide range of values depending on the postulated probabilities of this outcome, even if they be equal for all subgroups. This phenomenon is to be expected since, at any time after the beginning of the study, the composition of those remaining is a function of time, the occurrence risks, and the probability of loss. In the context of conception rates, for example, a larger proportion of the more fecund women than of the less fecund ones will have conceived at any time after the beginning of the study, but the relative proportions change from one time to the next. Consequently, the composition of the remaining sample keeps changing and the composition of the losses must change as well. Late losses will include relatively more of the less fecund women than will early losses. Naturally, the effect of losses on the composition of the group depends also on the rate at which they are lost, i.e. on the value of λ.

Numerical examples of these effects are seen in Figure 1, for

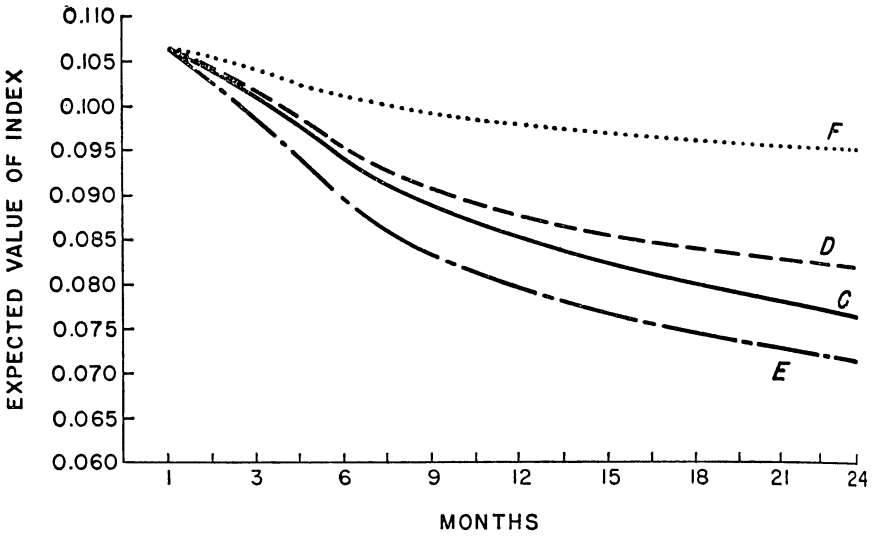

Figure 1. Approximate expected values of the Index according to duration of the study. Samples consisting of 250 individuals with monthly risk equal to 0.05 and 150 with monthly risk equal to 0.20. Monthly probabilities of leaving study are:

	Set C	Set D	Set E	Set F
Less fecund	0	.05	.01	.10
More fecund	0	.05	.10	.01

heterogeneous samples with N = 400. The composition of the samples for these four sets of calculations was identical with respect to fecundability. In Set C, the probability of loss was put equal to zero, and in Set D this probability was assumed equal to 0.05 for all individuals in the sample. This assumption produces an increase in the expected value of R, the increment increasing with the passage of time. When individuals from the more fecund group are lost at a higher rate than those from the less fecund (Set E), the expected values of the index fall below those in Set C. Conversely, if the rate of loss is higher for the less fecund group, the expected value of the index is relatively high (Set F). While these results are not surprising, they demonstrate the problems involved in using the index for heterogeneous samples containing individuals followed for unequal periods of time.

The index behaved more erratically in the examples of Figure 2, which relate to samples of 1000 consisting of a mixture of four subgroups, in all of which the distributions of fecundability were again identical. The expected values cannot be characterized consistently for different assumptions about the probabilities of loss. Again, if this probability is assumed to be identical for all individuals in the sample, the expected value of the index is higher than in the absence of losses (Set H versus Set G). For Set J, it was assumed that individuals in the sample were subject to unequal probabilities of loss, which were positively correlated with fecundability. The expected values of the index in the earlier months were below the corresponding values in Set G, from which losses were excluded. During the third year, however, the values in Set J became higher than those in Set G. Consequently, comparisons between two such samples (which have identical distributions of fecundability) would indicate higher conception rates for Set G if both samples were followed for two years or less, and lower rates for

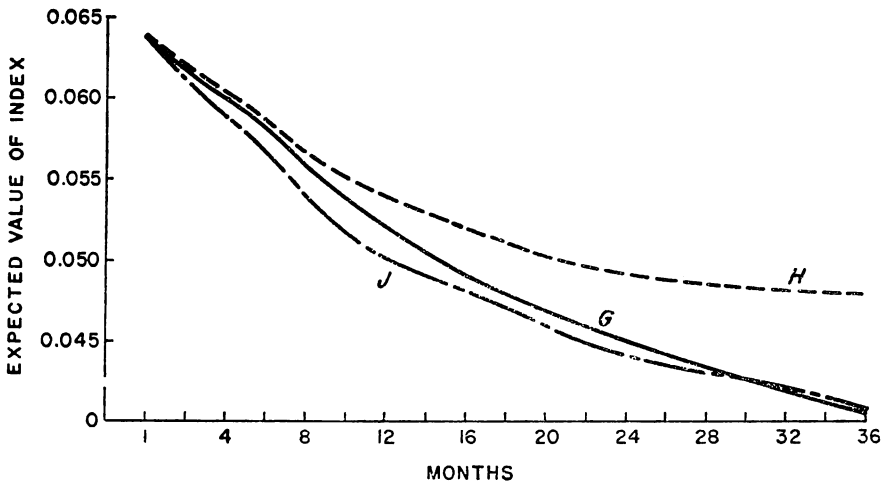

Figure 2. Approximate expected value of the Index according to duration of the study. Heterogeneous samples consisting of 4 subgroups. In all cases, the mean risk is .0645, its variance .0027, and the harmonic mean is .0214. The probability of loss is zero for set G, equal for all individuals in set H and positively correlated with risk in set J.

Set G if both samples were followed for more than three years. Thus, even keeping the duration of follow-up equal for two such samples would *not* provide a valid comparison of their fecundabilities.

Heterogeneous Population (Fully Stochastic Case)

The foregoing has been concerned with the effects of stochastic variation only in samples of fixed composition. If we add to our assumptions by providing for variation among samples drawn from a heterogeneous population in the way described earlier, it is found that both the index itself and its variance may be expected to be higher than in the fixed sample case.[7]

A numerical example is shown in Table 4, where the fixed sample has a composition identical with that of the population in the stochastic case. The example is relatively simple in that it excludes the possibility of losses and specifies only two subgroups in the population. It can be seen, however, that the variance of the index is considerably higher if the formula is based on the more appropriate assumption that includes sampling variation. More important than the variance, however, is the related effect on the expected value of the index, which departs increasingly from the changing values of the ratio of the two expectations in column 1.

Provision for losses in this model would result in values even more discouraging than those shown in Figure 2.

TABLE 4. A COMPARISON OF THE "FIXED SAMPLE" MODEL AND THE "FULLY STOCHASTIC" MODEL IN A HETEROGENEOUS POPULATION* WITHOUT LOSSES

Month of observation (1)	*Expected No. of Occurrences/ Expected Exposure Time* (1)	*Approximations for Ratio* — *N Times the Increment to be Added to Column (1) for the Expected Value of the Ratio*		*N Times the Variance*	
		Fixed Sample	*Stochastic Case*	*Fixed Sample*	*Stochastic Case*
1	.04375	0	0	.0399	.0418
4	.04093	.0137	.0164	.0088	.0105
8	.03760	.0136	.0192	.0037	.0053
12	.03473	.0124	.0200	.0022	.0036
16	.03228	.0113	.0204	.0015	.0027
20	.03020	.0106	.0206	.0011	.0022
24	.02844	.0101	.0208	.0009	.0018
∞	.01509	.0214	.0279	.00032	.00042

* P	Fecundability (p)
.625	.01
.375	.10

Arithmetic mean p = .04375
Harmonic mean p = .01509
Variance p = .001898

SUMMARY AND CONCLUSIONS

The behavior of the index "occurrences per person year of exposure" has been studied through a mathematical model. Numerical illustrations of the results, obtained by exact calculations, by a simulation program, and from approximate expressions, have been presented here. In all cases, it has been assumed that the risk per unit time is constant during the period of observation for any individual. Otherwise, the index is obviously not a suitable measure of risk.

In circumstances where it may be reasonable to assume that the sample is homogeneous, i.e. that all individuals have the same constant risk, the index may be treated as an estimate of that risk provided that the sample is large. Generally, the disadvantages are diminished when the size of samples is increased, but aggravated when the duration of observation is prolonged. In such homogeneous samples, the index has a small positive bias which increases with the duration of follow up and decreases with increased sample size. The variance decreases both with increasing sample size (N) and with the duration of observations (m). In general, the index is skewed, the skewness being to the right if the risk is less than $\frac{1}{2}$. This asymmetry increases with m and decreases when N is increased. The effect of changing probabilities of loss on these measures varies at different levels of risk. In many of the situations of interest, however, losses will reduce the bias and increase the variance of the index, according to the postulated model.

On the more realistic assumption that the risks or probabilities under study are not identical for all individuals in the sample, the index is of questionable value. In particular, if it is assumed also that losses may occur (this term including entry into the sample at different times), the behavior of the index may become highly erratic as the duration of follow up is prolonged. Consequently, the index cannot be recommended as a basis for comparing two or

more groups *even if they are observed for the same length of time,* in circumstances where it is not reasonable to assume that each sample is homogeneous. Other measures of risk that evaluate changes occurring with the passage of time, such as life table attrition rates,[6] seem clearly preferable.

The findings reported here suggest that conclusions that have been based on this index in various studies may be subject to some question. Investigators may find it necessary to re-evaluate data on which such conclusions have been based.

APPENDIX

The following symbols, some of which are defined in the paper, will be used:

N the number in the sample

n_i the number of individuals from subgroup i in a fixed sample. $\sum_{i=1}^{k} n_i = N$

p_i the risk (fecundability) of individuals in the ith group

$q_i = 1 - p_i$

λ_i the monthly conditional probability of loss for individuals in the ith group

$\gamma_i = 1 - \lambda_i$

m duration of the study

P_i the proportion of a population that belongs to subgroup i. $\sum_{i=1}^{k} P_i = 1$

$\pi_i = 1 - (\gamma_i q_i)^m$

$\nu_i = 1 - \gamma_i q_i$

A = the total number of occurrences (conceptions) observed

B = the total number of person months of observation

$R = A/B$ = the Pearl index, i.e. the number of occurrences per person-month of exposure

$E(X)$ and $V(X)$ — the mean and variance, respectively, of variable X

The theoretical results, given the assumptions in the model, are:

For homogeneous samples of size N,

$$E(R) = p + p\{(1-\nu)\pi - m\nu(1-\pi)\}/N\pi^2 + O(N^{-2}) \tag{1}$$

$$V(R) = \frac{pq\nu}{N\pi} + \frac{pq}{N^2\pi^2}\left[1 - \gamma(1-2p)(1-3\nu) - \nu\left\{\pi + \frac{2m(1-\pi)}{\pi}(2\nu + 3\gamma p) - \frac{\nu m^2 p(1-\pi)(5-3\pi)}{q\pi^2}\right\}\right] + O(N^{-3}). \tag{2}$$

Generally,

$$E(R) = \frac{E(A)}{E(B)}\left[1 + \frac{Var(B)}{\{E(B)\}^2} - \frac{Cov(A,B)}{E(A)E(B)}\right] + O(N^{-2}) \tag{3}$$

$$V(R) = \left[\frac{E(A)}{E(B)}\right]^2\left[\frac{Var(A)}{\{E(A)\}^2} + \frac{Var(B)}{\{E(B)\}^2} - \frac{2Cov(A,B)}{E(A)E(B)}\right] + O(N^{-2}) \tag{4}$$

In heterogeneous samples, the quantities needed in (3) and (4) are found to be:

For the fixed case:

E(A)	$\Sigma n_i p_i \pi_i / \nu_i$
E(B)	$\Sigma n_i \pi_i / \nu_i$
V(A)	$\Sigma n_i (p_i \pi_i)(\nu_i - p_i \pi_i)/\nu_i^2$
V(B)	$\Sigma n_i (1-\nu_i)[1-(1-\delta_i)\{1-\pi_i+(2m-1)\nu_i\}]/\nu_i^2$
Cov(A,B)	$\Sigma n_i p_i (1-\pi_i)(\pi_i - m\nu_i)/\nu_i^2$

For the stochastic case:

E(A)	$N\Sigma P_i p_i \pi_i / \nu_i$
E(B)	$N\Sigma P_i \pi_i / \nu_i$
Var(A)	$N[\Sigma P_i p_i \pi_i / \nu_i][1 - \Sigma P_i p_i \pi_i / \nu_i]$
Var(B)	$N[\Sigma P_i\{(1+\gamma_i q_i)\pi_i - 2m(1-\pi_i)\}/\nu_i^2 - \{\Sigma P_i \pi_i / \nu_i\}^2]$
Cov(A,B)	$N[\Sigma P_i p_i\{\pi_i - m\nu_i(1-\pi_i)\}/\nu_i^2 - \{\Sigma P_i p_i \pi_i/\nu_i\}\{\Sigma P_i \pi_i/\nu_i\}]$

REFERENCES

[1] Pearl, Raymond, Factors in Human Fertility and their Statistical Evaluation, *Lancet,* 225, 607–611, September 9, 1933.

[2] Tietze, Christopher and Lewit, Sarah, Recommended Procedures for the Study of Use-Effectiveness of Contraceptive Methods, *International Planned Parenthood Federation Handbook, Part I,* London, International Planned Parenthood Federation, 1962, pp. 59–72.

[3] Other forms of the Pearl Index, which attempt to estimate an average lifetime risk and include more than one pregnancy per woman, will not be considered here.

[4] Elveback, Lila, Actuarial Estimation of Survivorship in Chronic Disease, *Journal of the American Statistical Association,* 53, 420–440, June, 1958.

[5] Chiang, Chin Long, A Stochastic Study of the Life Table and its Applications: I, II, III, *Biometrics,* 16, 618–635, December, 1960; *Human Biology,* 32, 221–238, September, 1960; *Biometrics,* 17, 57–58, March, 1961.

[6] Potter, Robert G., Additional Measures of Use-Effectiveness of Contraception, *Milbank Memorial Fund Quarterly,* 41, 400–418, October, 1963.

[7] Sheps, Mindel C., Characteristics of a Ratio Used to Estimate Failure Rates, in press.

[8] Chandrasekaran, C. and Freymann, Moye W., Evaluating the Effects of Community Efforts to Modify Family Size, in *Public Health and Population Change: Current Research Issues,* M. C. Sheps and J. C. Ridley (eds)., Cambridge, Massachusetts, Schenkman Publishing Co., Inc., 1965.

[9] Dorn, Harold F., Methods of Aanalysis for Follow-up Studies, *Human Biology,* 22, 238–248, December, 1950.

[10] Seal, Hilary, The Estimation of Mortality and Other Decremental Probabilities, *Skandinavisk Aktuarietidskrift,* 37, 137–162, 1954.

[11] Gini, Corrado, Sur la mesure de l'efficacité des pratiques anticeptionalles, *Revue de l'Institut International de Statistique,* 10, 1–35, 1942.

[12] Sheps, Mindel C., Applications of Probability Models to the Study of Patterns of Human Reproduction, in *Public Health and Population Change: Current Research Issues,* M. C. Sheps and J. C. Ridley, editors, op. cit.

[13] Littell, Arthur S., Estimation of the T-year Survival Rate from Follow-up Studies Over a Limited Period of Time, *Human Biology,* 24, 87–116, May, 1952.

[14] Pearl, Raymond, On the Frequency of the Use of Contraceptive Methods, and Their Effectiveness as Used, By a Sample of American Women, *Bulletin de l'Institut International de Statistique,* 27, 208–224, 1933.

[15] Henry, Louis, Fondéments théoriques des mesures de la fécondité naturelle, *Revue de l'Institut International de Statistique,* 21, 135–151, 1953.

[16] Sheps, Mindel C., On the Time Required for Conception, *Population Studies,* 18, 85–97, July, 1964.

[17] Tietze, Christopher, Differential Fecundity and Effectiveness of Contraception, *Eugenics Review,* 50, 230–237, January, 1959.

[18] Potter, Robert G., Length of the Observation Period as a Factor Affecting the Contraceptive Failure Rate, *Milbank Memorial Fund Quarterly,* 38, 140–152, April, 1960.

[19] Beebe, G. W., *Contraception and Fertility in the Southern Appalachians,* Baltimore, The Williams and Wilkins Co., 1942, pp. 222–223.

[20] Potter, Robert G. and Sagi, P. C., Some Procedures for Estimating the Sampling Fluctuations of a Contraceptive Failure Rate, in *Research in Family Planning,* C. V. Kiser, editor, Princeton, N. J., Princeton University Press, 1962, pp. 309–405.

[21] Haldane, J. B. S. and Smith, S. M., The Sampling Distribution of a Maximum Likelihood Estimate, *Biometrika,* 43, 96–103, June, 1956.

[22] Hansen, Morris H., Hurwitz, W. N. and Madow, W. G., *Sample Survey Methods and Theory,* Vol. 2, New York, John Wiley & Sons, Inc., 1953.

ACKNOWLEDGMENTS

A large part of the investigations on which this report is based were carried out at the University of Pittsburgh. Computer programs for the results were written by Helen Chun and Joel Williamson; Lynn Doney and Florence Baseman assisted with calculation.

This investigation was supported in part by Public Health Service Research Grant GM 13436 (formerly 11134) from the National Institute of General Medical Sciences and by computer grant G 11309 from the National Science Foundation.